// NO.25-052 | 10/2025

DISCUSSION PAPER

// MASSIMO G. COLOMBO, LENA FÜNER, MASSIMILIANO GUERINI, HANNA HOTTENROTT, AND DANIEL SOUZA

The State of European Entrepreneurship: Trends in Quantity and Quality in France, Germany, and the UK (2009–2023)

Leibniz Association

ZEW

The State of European Entrepreneurship: Trends in Quantity and Quality in France, Germany, and the UK (2009–2023)

Massimo G. Colombo[1], Lena Füner[2,3,4], Massimiliano Guerini[1], Hanna Hottenrott[2,3], and Daniel Souza*[1]

[1]Department of Management, Economics and Industrial Engineering, Polytechnic University of Milan, Via Raffaele Lambruschini, 4/B, 20156, Milan, Italy

[2]Department Economics of Innovation and Industrial Dynamics, Leibniz Centre for European Economic Research (ZEW), L 7 11, 68161, Mannheim, Germany

[3]School of Management, Technical University Munich, Arcisstr. 21, 80333, Munich, Germany

[4]IAB Institute for Employment Research, Regensburger Str. 104, 90478, Nuremberg, Germany

October 14, 2025

Abstract

This paper replicates and extends the framework of Guzman and Stern (2020) to examine the evolution of entrepreneurial activity in Europe, focusing on France, Germany, and the United Kingdom between 2009 and 2023. Using harmonized national business registry data, we construct measures of both the quantity and quality of entrepreneurship across regions. In particular, we adapt the Entrepreneurial Quality Index (EQI), the Regional Entrepreneurship Cohort Potential Index (RECPI), and the Regional Entrepreneurial Acceleration Index (REAI) to capture the number of new ventures, their ex-ante growth potential, and the extent to which ecosystems translate this potential into realized outcomes. Our findings support the generalizability of this framework in the European context while revealing substantial heterogeneity across countries and regions. Major metropolitan centers such as Paris, London, and Munich combine high rates of entry with high entrepreneurial quality, but smaller knowledge- and research-intensive regions – including Cambridge, Oxford, Bonn, and Heidelberg – also emerge as important hubs. With respect to ecosystem performance, France and the UK initially exceeded expectations but later experienced steady declines, whereas Germany maintained relatively stable performance, with notable overperformance between 2012 and 2016. Moreover, we find a stronger positive correlation between entrepreneurial quantity and quality in Europe, suggesting that ecosystems capable of generating more start-ups are also more likely to produce high-quality firms. This study provides important insights for the comparative analysis of entrepreneurial ecosystems and builds a foundation for designing policies aimed at fostering high-quality, innovation-driven entrepreneurship in Europe.

Keywords: Entrepreneurial Quality; Entrepreneurial Ecosystem; High-Growth Firms; Regional Innovation

JEL Classification: G24; G32; L25; L26; M13; R12

---

*Corresponding author. Email address: danielfernando.desouza@polimi.it

This study was funded by the European Union - NextGenerationEU, Mission 4, Component 2, in the framework of the GRINS -Growing Resilient, INclusive and Sustainable project (GRINS PE00000018 – CUP D13C22002160001). The views and opinions expressed are solely those of the authors and do not necessarily reflect those of the European Union, nor can the European Union be held responsible for them.

1 Introduction

Entrepreneurial activities are widely recognized to be important drivers of innovation and economic development (Acs and Armington, 2004; Glaeser et al., 2015; Content et al., 2020; Audretsch et al., 2021). During recent decades, increasing academic and policy interest has led to significant advances in understanding how regional characteristics promote or hinder the emergence and growth of new firms (Dahl and Sorenson, 2012; Decker et al., 2016; Gourio et al., 2016; Haltiwanger, 2022).

However, a central challenge in studying entrepreneurship and its implications for innovation and regional development remained the substantial heterogeneity in the quality of entrepreneurial firms. Not all new companies have the same potential to generate and diffuse innovation (Hébert and Link, 2006; Gicheva and Link, 2016). Understanding the quality of start-ups – defined as the potential of a startup to achieve significant growth outcomes – is therefore essential to the study of entrepreneurship and its contribution to regional and national ecosystems. The shift in focus from quantity to the quality of entrepreneurship reflects the broader recognition that entrepreneurial outcomes are highly skewed: A small number of high-growth companies in specific locations typically generate a disproportionate share of innovation, employment, and value creation (Kortum and Lerner, 2000; Samila and Sorenson, 2011; Decker et al., 2016). Accounting for both the number of startups and their quality is therefore crucial when understanding entrepreneurial ecosystems and their performance (Stam and Van De Ven, 2021; Guerrero and Siegel, 2024).

Recently, significant progress has been made in measuring the quality of entrepreneurial activity. In particular, Guzman and Stern (2020) developed a novel predictive analytics model using early-stage startup characteristics, such as legal form, founder naming patterns, and IP filings, to estimate the probability that a firm will achieve a high growth event (e.g. IPO or major acquisition). This novel approach allowed for the creation of quality-adjusted entrepreneurship metrics and hence an improved mapping of entrepreneurial activities across the United States. Building on this framework, Andrews et al. (2022) significantly extended the methodology by generating fine-grained, regionally disaggregated ecosystem statistics across the United States from 1988 to 2016.

With this research, our objective is to demonstrate new empirical evidence, specifically, to assess both the quantity and quality of entrepreneurial activity across regions in Europe's three largest economies: the United Kingdom, Germany, and France. This study seeks to replicate[1] the Guzman and Stern (2020) predictive model and the Andrews et al. (2022) ecosystem framework for three major European economies. By collecting comparable business registration data and observable startup features in these countries, we aim to validate the predictive model in a European context and generate ecosystem-level indicators, that is, the Entrepreneurial Quality Index (EQI), the Regional Entrepreneurial Cohort Potential Index (RECPI), and the Regional Entrepreneurial Acceleration Index (REAI) for subnational regions. Combined, these

[1] We define replication in line with Hamermesh (2007) and Duvendack et al. (2017) as scientific replication or quasi-replication, that is, the re-examination of previous results under similar – though not identical – conditions or in different populations, with the aim of validating methodologies, prior findings and assessing their generalizability.

three measures provide a comprehensive framework for assessing both the ex-ante quality of entrepreneurial activity and the capacity of regional ecosystems to translate potential into realized entrepreneurial quality. Despite growing interest in the topic, much empirical work has focused on the United States, leaving entrepreneurial ecosystems in Europe relatively understudied - particularly with regard to their capacity to produce high-growth firms. This gap is particularly relevant, given that the lack of high-growth startups is frequently cited as one reason behind Europe's lag in innovation-led economic performance. Thus, addressing this issue has important policy implications.

Our analysis is based on business registry data from the UK, Germany, and France covering the period from 2009 to 2023. We focus on limited-liability business entities, as these are more likely to experience successful growth and represent common legal structures across the three countries. For the UK, we use the Companies House database, which provides demographic information on active firms and additional data such as insolvency filings, takeovers, and changes in capital structure. For Germany, we rely on the Mannheim Enterprise Panel, which builds on the German Business Registry. For France, we use data from SIRENE, which records the identity of all private-sector firms in Metropolitan France, and complement it with information from BODACC, covering a variety of corporate events and procedures. Overall, our data allows mapping the entrepreneurial activity in the selected countries and to identify more than 9 million firm births in the period 2009-2023.

This study contributes to filling a gap in empirical analyses of entrepreneurial ecosystem performance by providing novel, large-scale evidence on the evolution of entrepreneurial activity in Europe over the past decade. We do so by assessing both the quantity and quality of entrepreneurial activity across regions, where quantity is defined as the number of newly registered businesses, and quality as the extent to which these business are growth-oriented, exhibiting a higher potential for expansion at or near founding. In doing so, this study advances the mapping and characterization of entrepreneurial ecosystem outcomes in different institutional contexts and levels of granularity, both within and across countries.

Our analyses contribute to two main strands of literature. First, it advances research on entrepreneurial dynamics and firm growth by providing the first large-scale, systematic estimation of entrepreneurial quality in Europe. It extends recent work on firm-level growth prediction (Andrews et al., 2022; Guzman and Stern, 2020; Tartari and Stern, 2021) and complements long-standing contributions on the role of entrepreneurship in regional development (Glaeser et al., 2010; Samila and Sorenson, 2011). By applying a harmonized methodology to regions accounting for roughly 52 percent of the EU's GDP, this work fills a critical gap in a literature that remains predominantly focused on the US context.

Second, the study contributes to the literature on regional innovation and entrepreneurial ecosystems (Feldman, 2001; Lerner, 2009; Stam, 2015; Colombo et al., 2019; Ghio et al., 2019; Audretsch et al., 2021; Rocha et al., 2021) by studying inter-regional differences in the distribution of high-quality start-ups. By measuring entrepreneurial quality with greater granularity and precision, we advance the study of the microfoundations of entrepreneurial ecosystems, providing insights into how micro-level factors influence ecosystem-level outcomes. These measures provide a quantification of relative ecosystem performance, which can inform policy recommendations both at the micro and macro level.

Our findings for the European countries closely mirror the patterns reported by G&S for the US, with most predictors showing the same sign and statistical significance. Moreover, the overall predictive power is comparable to G&S, which confirms that the approach can be extended to European contexts. Moreover, our findings reveal substantial heterogeneity in entrepreneurial quality within and between the three countries, as well as over time. Our analyses highlight that entrepreneurial performance is shaped both by the supply of startups and the capacity of national ecosystems to translate high growth potential into realized outcomes. Interpreting the REAI as a reflection of the "ecosystem" as the broader economic and entrepreneurial environment in which each cohort of startups operates, our findings illustrate distinct patterns with all three countries not fully exploiting their potential. For France, for example, this implies that a startup founded in 2009, conditional on its quality, was about 2.1 times more likely to scale successfully than one founded in 2017. Digging deeper into national ecosystems, further findings reveal significant spatial variation, with certain subnational ecosystems clearly outperforming others. In contrast to the weak correlation reported for US regions by (Andrews et al., 2022), the European evidence shows a much stronger positive association between entrepreneurial quantity per capita and quality. Major metropolitan centers such as Paris, London, and Munich combine high entry per capita and high entrepreneurial quality. At the same time, also smaller knowledge- and research-intensive regions stand out, such as Cambridge and Oxford in the UK, and Bonn and Heidelberg in Germany, which achieve exceptionally high EQI despite their relatively small absolute size. This highlights the role of universities and specialized research institutions as anchors of high-quality entrepreneurship (Guerrero et al., 2015; Herrera et al., 2018) and confirms earlier insights that innovative activity in Europe is not only confined to large metropolitan agglomerations (Fritsch and Wyrwich, 2021).

This study has significant implications for both research and policymaking by informing the debate on Europe's innovative capabilities and competitiveness. The results highlight the significant heterogeneity across regions and significant divergent country-level trends. France shows a relatively smooth and steady increase in quality-adjusted entrepreneurial activity, the United Kingdom exhibits sharp fluctuations and greater volatility, while Germany follows a declining long-term trajectory. When comparing realized to expected growth outcomes, the French and British ecosystems underperform in more recent years, whereas Germany largely sustains realized growth in line with expectations. Interestingly, there seems to be a stronger quantity-quality link in Europe in the sense that entrepreneurial ecosystems capable of producing more startups also tend to generate higher-quality firms.

Finally, ecosystem structures differ substantially with France being centralized, with Paris dominating but with secondary metropolitan hubs (e.g., Toulouse, Marseille, Bordeaux) gaining importance over time. The United Kingdom is more polycentric, with London, Cambridge, and Oxford as the major hubs. Germany is distinctive in its decentralization: While Munich ranks highly with some distance to other regions, several smaller functional urban areas (e.g., Heidelberg, Bonn, Jena, Dresden) also perform strongly, reflecting a more distributed pattern of high-quality entrepreneurship. These findings support the idea that National Systems of Entrepreneurship (Acs et al., 2014) within Europe differ substantially and require a mix of policy approaches for fostering high-quality entrepreneurship. Importantly, this study quantifies significant heterogeneity in regional ecosystems (Feldman, 2001; Qian et al., 2013), which allows for better international comparisons within and beyond Europe.

The remainder of the paper is structured as follows. Section 2 outlines the framework of Guzman and Stern (2020). Section 3 details the data, applies the predictive analytics model, and discusses its validity in the European context. Section 4 examines the evolution of entrepreneurial quantity and quality across the three countries and presents regional-level statistics on entrepreneurial ecosystems. Finally, Section 5 concludes.

2 Measurement of Entrepreneurial Quality and Performance

Following the framework of Guzman and Stern (2020) (hereafter G&S), we distinguish between two dimensions of entrepreneurship: quantity, i.e. the number of entrants, and quality, defined as the start-up growth potential. Let N_{rt} denote the number of new firm registrations in region r and year t. While N_{rt} provides a measure of entrepreneurial activity, it is uninformative with respect to the underlying heterogeneity in firm growth potential. To address this limitation, G&S estimate firm-level entrepreneurial quality as the ex ante probability of attaining a growth outcome conditional on observable characteristics at the time of founding.

Formally, for firm i, founded in region r at time t, let $g_{ir(t+s)} \in \{0, 1\}$ denote an indicator variable that equals one if the firm experiences a growth outcome within s years of founding, and zero otherwise. Let H_{irt} be a vector of observable characteristics of the firm chosen by the entrepreneur at the time of entry. We define the entrepreneurial quality of firm i as the conditional probability of experiencing a meaningful growth outcome:

$$\theta_{irt} = \Pr\left(g_{ir(t+s)} = 1 \mid H_{irt}\right). \tag{1}$$

The parameter θ_{irt} is not an observed outcome but a latent propensity that summarizes, in a single probability, the *ex-ante* growth potential of firm i. Intuitively, θ_{irt} can be interpreted as the expected contribution of firm i to the distribution of future high-growth firms, given only its founding characteristics. It thus represents a forward-looking measure of entrepreneurial quality that is conceptually distinct from realized growth.

In empirical implementation, we approximate θ_{irt} by estimating a predictive function that maps founding characteristics into probabilities of growth:

$$\theta_{irt} = f(\beta H_{irt}), \tag{2}$$

where $f(\cdot)$ denotes the functional form of the prediction rule and β is the vector of parameters to be estimated.

2.1 Entrepreneurial Quality Index

Estimation of equation 2 yields $\hat{\theta}_{irt}$, the predicted growth probability for firm i, that serves as a forward-looking measure of entrepreneurial quality at the firm level. Aggregating across all entrants in a given region-year, G&S define the Entrepreneurial Quality Index (EQI) as the average predicted quality of new ventures:

$$EQI_{rt} = \frac{1}{N_{rt}} \sum_{i \in I_{rt}} \hat{\theta}_{irt}, \tag{3}$$

where I_{rt} is the set of new firms established in region r at time t, and N_{rt} is the corresponding number of entrants.

2.2 Regional Entrepreneurial Cohort Potential Index

The EQI captures entrepreneurial quality by averaging estimated growth potential across entrants. However, regional entrepreneurial potential also depends on quantity, i.e. the scale of entry. To jointly incorporate quality and quantity, G&S define the Regional Entrepreneurial Cohort Potential Index (RECPI) as:

$$RECPI_{rt} = EQI_{rt} \times N_{rt}. \tag{4}$$

The RECPI provides a measure of the quality-adjusted quantity of entrepreneurship in a given region-year. By construction, it is defined as the product of the average probability of growth across entrants and the total number of entrants. Thus, the RECPI is interpretable as the expected number of growth events generated by the cohort of firms founded in region r at time t.

2.3 Regional Entrepreneurial Acceleration Index

The RECPI provides a measure of the expected number of growth events in a given region-year. To assess the extent to which realized outcomes diverge from this expectation, G&S introduce the Regional Entrepreneurial Acceleration Index (REAI). The REAI is defined as the ratio of realized to expected growth events:

$$REAI_{rt} = \frac{\sum_{i \in I_{rt}} g_{irt}}{RECPI_{rt}}, \tag{5}$$

where the numerator is the realized number of growth events in region r and year t, while the denominator reflects expected events from the predictive model. An index above 1 indicates that the region enables high-potential firms to scale more than expected, whereas a value below 1 signals underperformance.

Taken together, the EQI, RECPI, and REAI provide a comprehensive framework for assessing both the *ex-ante* quality of entrepreneurial activity and the capacity of regional ecosystems to translate potential into realized growth. In what follows, we compute these indices for France, the UK, and Germany, closely following the G&S methodology.

3 Data and Entrepreneurial Quality Estimation

3.1 Business Registry Data in Europe

A central element of the G&S framework is the use of business registry data to estimate entrepreneurial quality. Business registries in Europe are governed by national legislation, with each country maintaining its own system. Unlike the United States, where business registries are primarily decentralized at the state level, European registries can be either centralized nationally or organized across regional jurisdictions. In France, for most of the period analyzed, an entrepreneur must submit a business registration application to the local commerce chamber[2] (*Chambre de Commerce et d'Industrie*), after which it is reviewed by a court clerk at a local commercial court (*Greffe du Tribunal de Commerce*). Upon approval, the company is recorded in the national registrar, the *Registre du Commerce et des Sociétés* (RCS), announced in the official public record *Bulletin Officiel des Annonces Civiles et Commerciale* (BODACC), and transmitted to the INSEE (*Institut National de la Statistique et des Études Économiques*), the French national statistical institute, where it is assigned a unique registry number and entered into the publicly available at the *Système d'identification du répertoire des entreprises* (SIRENE) database. In the United Kingdom, company registration is fully centralized through Companies House, the national registrar, where applications are reviewed by officials and, once approved, a unique registry number is assigned to the company, a notice is issued at the official public record (The Gazette) and recorded in the publicly accessible registry. In Germany, the registration process is decentralized at the state level, similar to the United States. Companies are registered with the local commercial register (*Handelsregister*) through the regional chamber of commerce and confirmed by its respective district court (*Amtsgericht*). Upon approval, the company is formally entered into the *Handelsregister* and its information published in the federal public register (*Bundesanzeiger*).

Our analysis relies on data from the SIRENE database and the BODACC for France, Companies House for the United Kingdom enriched by proprietary data[3], and, for Germany, the Mannheim Enterprise Panel (MUP) - a dataset derived from the German business registry and compiled and maintained jointly by Creditreform, Germany's largest credit rating agency, and ZEW, the Leibniz Centre for European Economic Research. The Data Appendix of the Supplementary Materials provides a detailed overview of these datasets.

We restrict our analysis to the complete population of limited liability, for-profit business entities, covering the period 2009–2023 for France and Germany and 2010–2020 for the United Kingdom[4]. Sole proprietorships and microbusinesses operated by solo entrepreneurs are excluded

[2]Since January 1, 2023, all formalities for company formation, modification, and cessation of activity in France must be completed online via the *Guichet des Formalités des Entreprises*. This centralized and online system managed by the *Institut National de la Propriété Industrielle* (INPI) replaces the former *Centres de Formalités des Entreprises* (CFE), which have been abolished, and applies to all businesses regardless of legal form or activity.

[3]Companies House excludes dissolved companies from the public registry. To address this limitation, we complement the dataset with information from the StatBooks® CHS platform (https://statbooks.co.uk/), which provides Companies House-derived data products covering all active and many dissolved entities. Further details are provided in the Data Appendix of the Supplementary Materials.

[4]The shorter time window for the United Kingdom reflects the unavailability of data on dissolved companies in the dataset compiled for this work.

wherever possible. These restrictions are motivated by several considerations. First, these structures are prevalent across all three countries, enabling consistent cross-country comparisons. Second, limited liability reduces personal financial risk for founders, making these entities more likely to attract external financing, undertake investment, and pursue sustained growth. Third, excluding sole proprietorships and solo-entrepreneur microbusinesses removes a large population of firms that are typically small-scale, low-capital ventures with limited growth ambitions, often established for self-employment rather than expansion. Finally, non-limited liability firms and micro-businesses that aim to grow often re-register as limited liability entities, as managers seek to scale operations, formalize governance structures, and access external financing. By focusing on limited liability firms, our analysis not only captures the segment of the business population most relevant for studying the dynamics of high-growth, high-risk entrepreneurial activity but also includes those growing firms that transition from non-limited liability or micro-business status through re-registration.

We further refine this definition, as each country recognizes multiple legal structures that satisfy our legal structure conditions, but these forms are not fully comparable across jurisdictions. In France and Germany in particular, several hybrid forms combine characteristics of corporations, limited liability companies (LLCs), and partnerships. Moreover, the legal landscape has evolved over time, with changes in the prevalence and regulation of these hybrid forms. To ensure cross-country comparability, we restrict our analysis to the most common limited liability structures in each country. Importantly, we exclude partnerships from our analysis[5]. While some partnership types may provide limited liability, they remain structurally quite distinct from corporations and LLCs and often embedded in complex hybrid forms that differ markedly across countries. Their inclusion would introduce unnecessary heterogeneity and weaken cross-country comparability. By excluding them, we ensure that our sample captures the organizational forms most relevant to entrepreneurial scaling.

Our final dataset comprises 3,007,644 firms registered in the selected time window for France, 4,958,264 for the United Kingdom, and 1,576,436 for Germany. The variation in the number of registered firms across countries reflects not only differences in the scale of their economies but also in the legal frameworks and costs associated with business creation. In the United Kingdom, for example, there is no distinct legal form for sole proprietorships, making it impossible to identify and exclude them solely from our data, and company registration costs as little as £50. By contrast, France recognizes more than 260 distinct legal forms, with registration fees ranging from zero to several thousand euros, while in Germany establishing a private limited liability company involves costs between €2,000 and €3,000, in addition to a required minimum share capital of €25,000. This institutional heterogeneity needs to be kept in mind in the cross-country comparisons of firm counts, entry rates, and quality measures. Nevertheless, by maintaining a consistent definition of our firm population within each country, we ensure internal validity in our analysis and can meaningfully compare within-country dynamics and relative changes across countries.

——————— Insert Table 1 Here ———————

[5] G&S include corporations, LLCs, and partnerships in their analysis.

Table 1 reports the distribution of legal forms among firms registered during the period covered by our analysis. In France, the firm population is almost evenly divided between Sociétés par actions simplifiées (SAS), which represent just over half of all entities (50.35%), and Sociétés à responsabilité limitée (SARL), which account for nearly the other half (49.52%). Other forms, such as Sociétés anonymes with either a board of directors or a management board, as well as the Société européenne, together make up less than 1% of firms. In the United Kingdom, the concentration is even stronger. Private Limited Companies (the closest analogue to LLCs in the US) constitute virtually the entire population (99.92%), with Public Limited Companies (the closest equivalent to US corporations) and all other categories appearing only in negligible numbers. Germany shows a somewhat more diverse landscape, though it remains highly concentrated. The Gesellschaft mit beschränkter Haftung (GmbH) accounts for more than four-fifths of firms (84.44%), followed by the hybrid GmbH & Co. KG at 13.7%. All remaining types, including joint-stock companies (AG), foreign firms, professional associations, and various partnerships, collectively represent less than 2%. Taken together, these figures highlight that in all three countries, the business population is overwhelmingly dominated by a narrow set of limited liability forms, while alternative structures remain marginal.

3.2 Measuring Startup Characteristics

Our empirical strategy closely follows Guzman and Stern (2020), who examine the relationship between observable startup characteristics at (or near) registration and subsequent growth outcomes. Replicating this approach requires combining business registry data with external sources that allow us to track firm growth and to construct comparable measures of initial characteristics across countries.

Growth. To measure growth, we rely on data from the Moody's Analytics Orbis M&A database (formerly Zephyr, Bureau van Dijk) to identify whether a startup undergoes an initial public offering (IPO) or is acquired within six years of its registration date[6]. Over the respective coverage periods for each country, we observe 1,322 growth events in France (82 IPOs), 4,562 in the United Kingdom (671 IPOs), and 2,069 in Germany (132 IPOs). These figures are lower than those reported in G&S for the United States (13,406 growth events, including 1,378 IPOs and 12,028 acquisitions), reflecting differences in the time period studied, size of the economies, and the financial market context, where capital markets and venture capital activity play a smaller role in corporate finance.

Startup Characteristics Based on Business Registry Data. We construct measures from registry data that are aligned with G&S methodology while maintaining cross-country comparability. These include indicators of legal structure and governance, name-based measures, and industry identifiers.

[6] Orbis M&A is the largest comparable dataset on private European companies. As with other Moody's Analytics datasets used in this study, coverage is not universal and varies over time. Nevertheless, there is evidence that the database provides relatively good coverage of high-value acquisitions, similar to Thomson Reuters SDC data used in the original Guzman and Stern (2020) study. Any limitations in coverage are likely to induce attenuation bias, consistent with the original analysis.

For legal structure, we define *Public* as a dummy for public limited liability companies, defined as joint-stock companies whose shares can be freely traded on public markets. In France, this category includes the Société anonyme à conseil d'administration, Société anonyme à directoire, and Société européenne. In the United Kingdom, it encompasses Public Limited Companies, the European Public Limited-Liability Company (SE), and the United Kingdom Societas (the domestic transposition of the SE after Brexit). In Germany, the definition corresponds to the Aktiengesellschaft (AG) legal structure.

We also construct a variable *Popular Business Structure*, identifying the most prevalent legal structure in each country: Société par actions simplifiée (SAS) in France, Private Limited Company (including Section 30 companies) in the United Kingdom, and Gesellschaft mit beschränkter Haftung (GmbH) in Germany. These forms account for the vast majority of limited liability firms and constitute the default organizational choice for entrepreneurs. Crucially, they are functionally analogous across jurisdictions, enabling our analysis to capture both the modal legal structure within each country and a set of broadly equivalent structures across Europe and the United States.

We replicate the two business name–based measures introduced by G&S. The first, *Short name*, is an indicator equal to 1 if the firm's registered name contains three or fewer words. While linguistic differences - such as compound words in French and German - pose some challenges, the variable is constructed in close alignment with the original definition. The second, *Eponymous*, is a dummy equal to 1 if the firm's name includes the first, middle, or last name of one of its top managers or founders. Constructing this variable is more complex, since registry data on founders and managers are inconsistently recorded across countries and the definition of "top manager" is itself heterogeneous. When information on managers at the time of registration is available, we use it directly; if this information is missing, we rely on data on current managers. In cases where no managerial data are available, we approximate by checking whether the firm name contains one of the most common personal names in the respective country[7]. The rationale is that firm names incorporating personal names are typically eponymous and often reference the founder. While this approximation is necessarily imperfect, manual validation indicate that it performs reasonably well. To the extent that remaining misclassification is unsystematic, it should attenuate rather than inflate our estimates. Full details on construction and validation are provided in the Data Appendix of the Supplementary Materials.

Finally, we assign firms to industries using the harmonized Eurostat NACE Rev. 2 system, which is required for business registry applications across all European Union (EU) countries[8]. This greatly simplifies cross-country comparability and avoids the need to infer industry from firm names, as in G&S. Our primary measure is the firm's two-digit NACE code. In some countries, such as the UK, firms may register under multiple codes, and because these codes are self-reported, missing values, inconsistencies, and updates over time are common. Whenever possible, we use the first code recorded at incorporation. If this is unavailable, we substitute the most recent code; if still missing, we classify the firm as having no sectoral information at

[7]Name frequencies are based on individual entrepreneurs in registry data for France and Germany, and on social media–based name lists for the UK. See Data Appendix in the Supplementary Materials for details.

[8]The United Kingdom uses UK the SIC 2007 classification, which is identical to NACE Rev. 2 up to the 4-digit level and still mandatory after Brexit.

founding. We interpret non-reporting of industry at incorporation as potentially informative, since incomplete registration may signal lower entrepreneurial quality. Additional details on variable construction by country are provided in the Data Appendix of the Supplementary Materials.

Startup Characteristics from External Data. To complement registry-based measures of entrepreneurial activity, we enrich our dataset with indicators of startup quality derived from intellectual property (IP) records, which provide forward-looking signals of firms' technological capabilities and market orientation. We draw on two external datasets: Moody's Analytics ORBIS Intellectual Property (ORBIS IP) and the Fraunhofer Institute for Systems and Innovation Research (ISI) Trademark Data Collection (ISI-TM). ORBIS IP links patent applications and grants from major patent offices (EPO, USPTO, WIPO, and selected national offices) to corporate entities in the ORBIS universe, enabling systematic attribution of patenting activity to legally registered firms. The ISI Trademark Data Collection (ISI-TM), developed by Fraunhofer ISI, records firm-level trademark filings at the EUIPO[9] and USPTO (Neuhäusler et al., 2021).

From these sources, we construct two firm-level binary indicators of early-stage startup quality. First, *Patent*, which takes the value of 1 if a firm files or holds a patent within the first year after incorporation. Second, *Trademark*, which equals 1 if a firm applies for a trademark at EUIPO or USPTO within the first year of registration. Both indicators capture early commitments to technological innovation (patents) and market development or brand differentiation (trademarks), thereby complementing registry-based measures of entrepreneurial quantity and quality with externally validated proxies of innovative and commercial potential.

——————— Insert Table 2 Here ———————

Table 2 reports summary statistics for the key variables. Some notable cross-country differences emerge. The incidence of growth events is very low in all countries, but with clear variation: 0.13% in Germany, 0.09% in the United Kingdom, and 0.04% in France. These rates are of the same order of magnitude as those reported by G&S for an earlier period and a different set of firms (0.07%). Name-based observables display sharper contrasts between European countries and the US baseline in G&S. Short names are markedly more common in Europe, ranging from 57% in the UK to nearly 90% in France, compared to 46% in G&S. This gap likely reflects cultural and linguistic conventions in firm naming practices. Intellectual property indicators are relatively rare but broadly similar across the three countries and close to the G&S benchmark. In contrast, industry composition differs substantially across countries, reflecting underlying structural differences in national economies[10]. A complete description of the construction of each variable is provided in the Data Appendix of the Supplementary Material.

[9]While ISI-TM does not cover trademarks registered solely at national offices, EU-level trademarks are can be regarded as stronger indicators of firm quality and market orientation. See Data Appendix for further details.

[10]Industry measures are summarized at the NACE 1-digit level to parsimoniously capture broad sectoral patterns. More detailed distributions at the NACE 2-digit level, as well as alternative aggregation levels, are provided in the Data Appendix of the Supplementary Material.

3.3 Estimation of Entrepreneurial Quality

We estimate entrepreneurial quality following the procedure of G&S, regressing *Growth* - defined as IPOs or acquisitions within six years of founding - on the set of startup characteristics using a logit specification. Figure 1 illustrate the evolution of firm entry and subsequent growth outcomes for the sample of startups included in the regression analysis, namely those for which post-entry equity growth can be observed within six years (2009–2017 for France and Germany, and 2010–2014 for the United Kingdom). In each panel, the left axis shows the intensity of firm births relative to GDP, while the right axis reports the frequency of growth events scaled by GDP. The patterns reveal marked cross-country differences: in France (Panel A), entry and growth outcomes peak in 2010–2011, dip in 2012–2013, and subsequently rise again; in the UK (Panel B), entry relative to GDP steadily increases, while growth outcomes per GDP consistently decline; and in Germany (Panel C), both entry and growth outcomes relative to GDP decrease over the period[11].

——————— Insert Figure 1 Here ———————

In Table 3, we report a series of univariate logit regressions of *Growth* on each characteristic separately, using the sample of firms for which post-entry equity growth can be observed within six years (2009–2017 for France and Germany, and 2010–2014 for the United Kingdom). Consistent with G&S, results are presented as odds ratios to facilitate interpretation, alongside the McFadden pseudo R^2 as a measure of explanatory power.

——————— Insert Table 3 Here ———————

The univariate regressions indicate that all categories of startup founding characteristics are predictive of subsequent growth. Among the name-based measures, *Short name* is positively and statistically significantly associated with growth in all three countries, with the strongest effects in France and Germany, where startups with short names are 50–60% more likely to grow. While sizable, these effects are smaller than those reported in G&S, where short names are associated with a threefold increase in growth probability. In line with G&S, *Eponymous* is negatively correlated with growth, particularly in Germany, where the odds of achieving a growth event are 86% lower for such firms. Corporate form variables exhibit the largest magnitudes. Being registered as a public (joint-stock) company is strongly predictive of growth in all countries, while *Popular business structures* displays heterogeneous patterns—positive in France and Germany but strongly negative in the UK, reflecting the limited choice of legal forms there. Intellectual property indicators are also highly predictive. Both *Patents* and *Trademarks* raise the likelihood of growth by more than an order of magnitude, with patents showing the largest effects. In France, for example, patent-holding startups at founding are 33 times more likely to grow. Nevertheless, these effects are weaker than in the US sample of G&S, where patents and trademarks increase the probability of growth by factors of 90 and 45, respectively. Industry patterns reveal sharp heterogeneity. Information and communication stands out as the most predictive sector across all three countries, with very large odds ratios, while construction,

[11] We replicate this exercise using GDP growth over five years for each cohort in the Supplementary Material, finding similar dynamics and cross-country variation in the economic context faced by each firm cohort.

accommodation and food services, and other traditional sectors are strongly negatively associated with growth. Some industries exhibit country-specific effects—for example, financial and insurance activities are highly predictive in the UK but much weaker in France and Germany. Finally, firms with missing industry codes are systematically less likely to grow, suggesting that incomplete registry information itself signals lower entrepreneurial quality. Overall, the univariate regressions confirm that observable firm characteristics carry substantial predictive power for identifying high-growth startups, consistent with the patterns documented by G&S.

We now turn to the multivariate regression analysis of entrepreneurial quality for each country. Following G&S, we sequentially introduce groups of predictors: corporate form measures in column (1), name-based measures in column (2), and intellectual property measures in column (3). Column (4) reports a specification restricted to business registry information - corporate form measures, name-based measures, and industry dummies (based on 2-digit NACE codes) - which we refer to as the nowcasting model. Column (5) presents our preferred specification, incorporating all predictors, including intellectual property measures derived from external data sources.

——————— Insert Table 4 Here ———————

Table 4 reports the results for France. Across specifications, corporate form variables exhibit the largest effects. Firms incorporated as public companies are more than 16 times as likely to experience a growth event, while those adopting the most common business structure are about three times more likely, relative to the baseline. Name-based characteristics are also predictive. Firms with short names are significantly more likely to grow, though with more modest magnitudes (around 25% higher odds in the full model), while eponymous names are strongly and negatively associated with growth, reducing the probability of growth by half. Intellectual property indicators have particularly pronounced effects: holding a patent at birth increases the probability of growth by a factor of ten, while filing a trademark raises the odds nearly fivefold. The inclusion of industry fixed effects substantially improves model fit, with the pseudo R^2 rising from 0.039 in the corporate governance baseline to 0.129 in the full model. Overall, results are qualitatively similar to G&S predictive models.

——————— Insert Table 5 Here ———————

Table 5 reports the predictive model results for the United Kingdom. Corporate form again emerges as the dominant predictor of growth: incorporation as a public company increases the odds of experiencing a growth event by roughly 90 times in the full model. The *Popular business structures* variable is omitted for the UK, as only two legal forms (public and private limited companies) are observed, leading to collinearity. Name-based measures are comparatively weaker predictors: short names have no significant effect in the full specification, while eponymous names reduce the likelihood of growth by approximately 25%. Intellectual property indicators remain highly predictive, with patents at founding increasing the odds of growth tenfold and trademarks raising them by about five times. Including industry fixed effects in the nowcasting and full models improves model fit, with the pseudo R^2 increasing from 0.062 in the corporate form baseline to 0.125 in the full specification. Compared to the G&S results for the US, name-based measures - particularly short names - carry less predictive power for entrepreneurial quality in the UK.

——————— Insert Table 6 Here ———————

Table 6 reports the predictive model results for Germany. As in France and the UK, corporate form is the strongest predictor of growth: incorporation as a public company increases the probability of experiencing a growth event by roughly 21 times in the full model, while the popular business structure is also positively associated with growth, though with a more moderate effect of around 3 times. Name-based measures are significant and display similar patterns to those observed in France: short names increase the likelihood of growth by approximately 20%, whereas eponymous names are strongly negatively associated, reducing growth odds by nearly 83%. Intellectual property indicators remain highly predictive: holding a patent at founding increases the probability of growth sixfold, and filing a trademark quadruples it. Including industry fixed effects in the nowcasting and full models substantially improves model fit, with the pseudo R^2 rising from 0.031 in the corporate form baseline to 0.119 in the full specification. Overall, these results align with patterns observed in France, the UK, and the G&S results for the US, though magnitudes and significance levels of name-based measures differ slightly across countries.

Overall, the results for the European countries closely mirror the patterns reported by G&S for the US, with most predictors showing the same sign and statistical significance. The main exception concerns name-based variables, particularly *Short Name*, which reflects cross-country differences firm naming conventions and linguistic structures. The predictive power of the models is also broadly comparable across countries, with pseudo R^2 values of 0.129 for France, 0.125 for the UK, and 0.119 for Germany. While these values remain below the 0.187 achieved by G&S, it is important to note that their US sample spans a longer period (20 years versus 9 years for France and Germany and 5 years for the UK) and includes many more observations (approximately 19 million compared to approximately 1.5 million for France, 2.1 million for the UK, and 0.9 million for Germany). In addition, G&S incorporate state fixed effects[12] and rely on US-specific features, such as Delaware incorporation, which are strong predictors of growth outcomes but unavailable for European startups. Despite these differences and the use of a parsimonious, cross-country comparable specification, our predictive models perform remarkably similarly to G&S, highlighting the robustness of their approach in a European context.

3.4 Model Validation and Robustness Checks

To assess the predictive quality of our entrepreneurial quality model, we conduct a 5-fold cross-validation procedure using firm-level data from each country[13]. Cross-validation is a standard technique in predictive modeling that involves partitioning the data into subsets, or "folds",

[12] We include NUTS-1 regional fixed effects in robustness checks reported in our Supplementary Material. This increases pseudo R^2 slightly, but coefficients remain similar in magnitude, direction, and significance, indicating that regional institutional differences in European business registration are less pronounced than in the US.

[13] We use 5 folds instead of the 10-fold approach of G&S due to smaller sample sizes, which can increase variability across folds. We also tested 10-fold cross-validation and stratified folds by growth variable and year; all approaches yield similar results.

repeatedly training the model on a portion of the data while testing it on the remaining fold. This process is repeated ten times, ensuring that each data point is used both for training and validation. The goal is to estimate how well the model generalizes to unseen data, thus providing a robust measure of out-of-sample performance.

——————— Insert Figure 2 Here ———————

Figure 2 reports the out-of-sample distribution of realized growth events across different portions of the entrepreneurial quality index for each country. The overall pattern is consistent with G&S: firms in the top 1% of the predicted quality distribution account for a disproportionate share of growth outcomes, approximately 17% in France (panel A), 19% in the UK (panel B) and 17% in Germany (Panel C), compared with 36% in the US sample. At the same time, the minimum and maximum shares across folds are much wider than in G&S, reflecting the smaller number of growth events in European samples and the resulting greater sampling variability. This dispersion indicates that while the concentration of growth is evident at the top of the distribution, fold-level estimates are less stable, making the overall relationship noisier than in the US case. Although the concentration of growth at the very top of the distribution is evident, a larger share of events also occurs in lower percentiles compared to the model validation for the US, where the top 5% and 10% of firms capture 53% and 64% of growth events, respectively. This suggests that while the models successfully identify high-potential startups, growth dynamics in Europe may be somewhat less sharply skewed toward the very top of the quality distribution.

We also test the robustness of our predictive models by considering alternative outcome measures based on employment-defined growth. This adjustment is particularly relevant in the European context, where capital markets are less developed and national economies are often smaller than some individual US states. Using employment data from the same Moody's Analytics Orbis dataset employed for equity-based outcomes and intellectual property information, we construct two additional indicators of success: *Employment Growth 500* and *Employment Growth 1000*, which equal 1 if a firm reaches more than 500 or 1,000 employees, respectively, within six years of registration. For reference, these thresholds represent approximately two and four times the employee count threshold that the European Commission uses to define a large firm[14].

Table 7 presents robustness checks for France using the growth outcomes based on employment level. The results confirm that the predictors identified in the equity-based models remain strong correlates of entrepreneurial success when measured by subsequent firm size. Corporate form continues to play a central role: incorporation as a public company increases the likelihood of exceeding 500 employees by nearly nineteen times and of surpassing 1,000 employees by more than eight times. Popular business structures are also consistently associated with higher growth probabilities, with incidence rate ratios around three to four. Name-based variables display mixed effects however: short names are positively associated with equity-based growth but negatively correlated with surpassing the 500-employee threshold, while eponymous names remain a robust negative predictor across specifications. Intellectual property indicators retain their

[14] In the Data Appendix of the Supplementary Material, we report robustness checks using alternative size definitions; the results are qualitatively similar to those presented here.

strong predictive power, with patents and trademarks substantially increasing the odds of high employment growth. Notably, the magnitude of patent effects rises with stricter employment thresholds, reaching odds ratios above 24 for firms exceeding 500 or 1,000 employees. Model fit improves under these alternative definitions of success, with pseudo R^2 values rising from 0.129 in the equity-based baseline to 0.197 and 0.215 when using the 500- and 1,000-employee thresholds, respectively, suggesting that startup characteristics provide an even sharper signal of employment outcomes in France.

——————— Insert Table 7 Here ———————

We now turn to the United Kingdom results for employment-based growth outcomes, reported in Table 8. Overall, the patterns largely align with those observed in the equity-based model, while revealing some nuances specific to employment thresholds. Incorporation as a public company remains the strongest predictor of high employment growth, though the effect is smaller compared to IPO or acquisition outcomes: the odds of exceeding 500 or 1,000 employees are roughly 13.5 and 14.8 times higher, respectively, versus a nearly 90-fold increase for equity-based growth. Name-based measures continue to show negative effects for eponymous names across both employment thresholds, while short names are now negatively associated with high employment growth - similar to the pattern observed in France - indicating that the predictive relevance of naming conventions varies by growth metric. Intellectual property measures remain highly predictive: patents increase the likelihood of reaching 500 or 1,000 employees by about 20 and 21 times, respectively, with trademarks showing similarly substantial effects. Model fit improves modestly for employment-based outcomes, with pseudo R^2 rising from 0.125 in the baseline to 0.163 for the 1,000-employee threshold.

——————— Insert Table 8 Here ———————

Finally, for Germany, Table 9 presents the results for employment-based growth outcomes. Overall, the patterns are broadly consistent with those observed in France, the UK, and the baseline equity-based model, with some differences across employment thresholds. Incorporation as a public company remains the strongest predictor of high employment growth, though the effect is smaller than for IPO or acquisition outcomes: the odds of surpassing 500 or 1,000 employees are approximately 10.6 and 13.2 times higher, respectively, compared to a 21-fold increase for equity-based growth. Popular business structures are positively associated with employment growth, but with more moderate effects. Name-based measures display a mixed pattern: eponymous names consistently reduce the likelihood of reaching high employment thresholds, whereas short names have a weak or slightly negative effect, especially at the 500-employee threshold. Intellectual property measures, patents remain a strong predictor, increasing the likelihood of reaching 500 or 1,000 employees by roughly 17 times, while surprisingly trademarks coefficients are not significant anymore. Model fit improves modestly for employment-based outcomes, with pseudo R^2 rising from 0.119 in the baseline to 0.149 at both employment thresholds. Unlike France and the UK, the model fit for the 1,000-employee threshold does not exceed that of the 500-employee threshold, suggesting that additional predictive power does not increase for the highest employment tier in Germany.

——————— Insert Table 9 Here ———————

Overall, the results for France, the UK, and Germany indicate that the predictors identified in the equity-based models continue to retain substantial explanatory power when applied to employment-based measures of high-growth firms. Corporate form and intellectual property indicators consistently emerge as strong predictors, and the models achieve reasonable fit across all three countries, suggesting that the baseline entrepreneurial quality framework remains informative even under alternative growth metrics. However, the magnitude and significance of some predictors vary across countries and employment thresholds, particularly for name-based measures and common business structures. These differences suggest that the quality measures capture slightly different dimensions of firm quality in Europe, and indicate that in the US context, IPO and M&A outcomes appear more closely linked to employment growth as a measure of success.

4 The State of European Entrepreneurship

4.1 Country-Level Trends in France, Germany, and the UK

Using the prediction models, we analyze the RECPI, weighted by annual GDP, and the REAI at the country level. We report two versions of the RECPI: one based on the nowcasting model (excluding intellectual property measures) and one based on the full model (including them).

——————— Insert Figure 3 Here ———————

Figure 3 presents RECPI-to-GDP trends across countries. The two versions of the index (nowcasting vs. full model) track each other closely for most of the sample period but diverge toward the end, with the nowcasting specification consistently producing higher values than the full model. In France (Panel A), the index shows a steady upward trajectory: after a plateau between 2010 and 2012, RECPI rises almost continuously until peaking in 2021, followed by a slight decline. In the UK (Panel B), the pattern is more volatile: RECPI climbs sharply from 2010 to 2014, dips briefly, surges again to a peak in 2017, and then falls sharply until 2020. In Germany (Panel C), the index declines markedly from its 2009 peak, partially recovers to a second peak in 2021, and resumes its downward trend thereafter. These patterns reveal notable cross-country heterogeneity. Compared with the firm entry dynamics shown in Figure 1, France's RECPI displays a much smoother upward trend up to 2017. In the UK, RECPI growth is more volatile than firm births: it rises more steeply early in the period but less so toward the end. In Germany, by contrast, the RECPI broadly mirrors the dynamics of firm entry relative to GDP.

——————— Insert Figure 4 Here ———————

We now turn to the analysis of entrepreneurial ecosystem performance using the REAI. By construction, the REAI captures whether a country's ecosystem generates more growth events than would be expected based on the quality-adjusted quantity of entrants. Following G&S, we interpret the "ecosystem" as the broader economic and entrepreneurial environment in which each cohort of startups operates. In this sense, the REAI reflects factors beyond founding characteristics - such as funding availability, legal and administrative conditions, institutional quality, and local market dynamics - that influence whether high-potential startups successfully scale.

Figure 4 presents the evolution of REAI during the training period, when growth outcomes are observable. In France (Panel A), the country ecosystem initially overperforms, with REAI reaching 1.42 between 2009 and 2011, before dropping sharply in 2012. This decline coincides with the fall in business registrations and growth events documented in Figure 1, Panel A. From 2014 onward, the index remains below unity, signaling systematic underperformance, and bottoms out at 0.68 in 2017. The United Kingdom (Panel B) follows a similar trajectory: REAI starts at 1.43 in 2010 but declines steadily, slipping below unity in 2012 and reaching 0.78 by 2014. This pattern aligns with Figure 1, Panel B, where firm creation continues to rise but growth events diverge downward. By contrast, Germany (Panel C) exhibits a distinct path. The ecosystem underperforms in 2009 (REAI = 0.82) but converges toward parity by 2010. Between 2012 and 2017, the index consistently exceeds unity, peaking at 1.15 in 2015, before dipping slightly below parity in 2017 (REAI = 0.89). This trajectory corresponds to the sharper decline in growth events relative to firm births observed in Figure 1, Panel C.

Taken together, these results draw different picture for each country. For France, we see an increase trend in the creation of startup with high potential to grow measured by the RECPI, but the capacity of the French ecosystem to make this realize growth seems to be decreasing overtime, as measured by the RECPI. A startup founded in 2009, conditional on its quality, was about 2.1 times more likely to scale successfully than one founded in 2017 (REAI ratio = 1.42/0.68). In the United Kingdom, although the RECPI rises until 2017, ecosystem performance deteriorates steadily: a startup of given quality in 2010 was roughly 1.8 times more likely to achieve a growth event than one founded in 2014 (REAI ratio = 1.43/0.78). Germany presents a different case. Despite a downward trend in high-growth potential, the ecosystem remains comparatively robust: even at the end of the period, scaling probabilities are only modestly lower, with startups founded in 2015 being about 1.3 times more likely to scale than those of similar quality founded in 2017 (REAI ratio = 1.15/0.89). Overall, these patterns highlight that entrepreneurial performance is shaped not only by the supply of startups but also by the capacity of national ecosystems to translate high growth potential into realized outcomes.

4.2 Regional Variation in Entrepreneurial Quantity and Quality

While national-level trends are informative, they mask the fact that entrepreneurial activity is inherently uneven across space. Economic forces such as agglomeration economies, localized knowledge spillovers, market access, and institutional variation generate substantial regional heterogeneity in both the quantity and quality of new firms. Understanding this spatial variation is crucial: high-potential startups tend to cluster in certain places, and the capacity of local ecosystems to support scaling may differ markedly even within the same country. A key strength of the G&S framework is its flexibility: it can be applied at different levels of geographic granularity, providing insights that are highly valuable for both academic research and policy. Following Andrews et al. (2022), we exploit this feature by generating regional-level statistics from our prediction models to examine variation in entrepreneurial quantity and quality across

France, the United Kingdom, and Germany.

Table 10 reports summary statistics for France (Panel A), the United Kingdom (Panel B), and Germany (Panel C) at different levels of regional aggregation: NUTS 1, NUTS 2, NUTS 3, Local Administrative Units (LAU), and Functional Urban Areas (FUAs)[15]. For each regional level, we report the mean and standard deviation of entrepreneurial quantity, entrepreneurial quality (EQI, rescaled by 1,000), quality-adjusted entrepreneurial quantity (RECPI), and equity growth (share of firms undergoing an IPO or acquisition). The bottom rows indicate the number of regions and total region-year observations at each level of aggregation. At more aggregate levels (such as NUTS 1), the statistics appear relatively smooth, masking substantial heterogeneity within countries. For example, in France the average EQI at NUTS 1 is 0.66 with limited dispersion (SD = 0.18), whereas at the LAU level the standard deviation rises to 0.92, revealing much greater variation across local areas, some generating, on average, very high-quality startups, while others produce substantially lower-quality firms.

——————— Insert Table 10 Here ———————

Figure 5 plots the relationship between entrepreneurial quality (EQI) and quantity per capita across FUAs in France (Panel A), the United Kingdom (Panel B), and Germany (Panel C). Each marker represents a FUA, with its size proportional to the FUA's population. In contrast to the weak correlation reported for US regions by Andrews et al. (2022), the European evidence shows a much stronger positive association between entrepreneurial quantity per capita and quality. Major metropolitan centers such as Paris, London, and Munich combine high entry per capita and high entrepreneurial quality. At the same time, knowledge- and research-intensive regions stand out: Cambridge and Oxford in the UK and Jena in Germany achieve exceptionally high EQI despite their relatively small absolute size, underlining the role of universities and specialized research institutions as anchors of high-quality entrepreneurship.

——————— Insert Figure 5 Here ———————

Next, in Tables 11, 12 and 13 we report the top 15 FUAs ranked by entrepreneurial quality for France, the United Kingdom, and Germany, respectively. To capture dynamics over time, we split observations into two distinct periods. In France (Table 11), Paris dominates the rankings throughout, confirming its role as the country's primary hub of high-quality entrepreneurship. In the earlier period, the top positions also included smaller FUAs such as Annemasse–Geneva and Annecy. In the later period, however, larger FUAs such as Bordeaux and Marseille enter the top 15, and Toulouse advances several positions, suggesting a growing importance of agglomeration effects in major urban centers. This evolution indicates that while Paris remains dominant, other large metropolitan areas have strengthened their ecosystems over time, combining scale with increasing entrepreneurial quality. In the United Kingdom (Table 12), Cambridge and Oxford

[15] The Nomenclature of Territorial Units for Statistics (NUTS) classification is the standard European system for defining territorial units (https://ec.europa.eu/eurostat/web/regions-and-cities/overview), ranging from large regions (NUTS 1) to smaller administrative areas (NUTS 2 and NUTS 3). LAUs provide a finer-grained division, corresponding roughly to municipalities or local districts. In addition, we make use of FUAs, which capture economically integrated urban regions defined by commuting flows (https://www.oecd.org/en/data/datasets/oecd-definition-of-cities-and-functional-urban-areas.html) and represent a more meaningful unit for analyzing entrepreneurial ecosystems.

consistently rank at the top across both periods, despite their relatively small populations. This highlights the role of world-class universities and research-intensive environments in generating high-quality entrepreneurship. London also maintains a strong position, reflecting its concentration of finance, global connectivity, and access to venture capital. Over time, several mid-sized FUAs, such as Reading, Cheltenham, and Coventry, enter or climb the rankings. The continued prominence of Cambridge and Oxford, together with the diversification of FUAs in later years, suggests that the UK exhibits a more polycentric pattern of high-quality entrepreneurship than France, where Paris remains overwhelmingly dominant. In Germany (Table 13), Munich, Heidelberg, and Bonn consistently rank at the top. Berlin, despite being the country's largest startup center by volume, does not dominate the quality rankings, suggesting that its ecosystem is broader but less concentrated in high-growth ventures. Notably, several smaller FUAs, such as Jena, Greifswald, and Dresden, achieve very high EQI scores. This pattern highlights a distinctive feature of the German landscape: high-quality entrepreneurship is not confined to the largest metropolitan areas but is distributed across a diverse set of regional ecosystems.

——————— Insert Table 11 Here ———————

——————— Insert Table 12 Here ———————

——————— Insert Table 13 Here ———————

Finally, Figure 6 plots the evolution of the EQI for selected FUAs over time. In France (Panel A), the entrepreneurial quality of major entrepreneurial ecosystems follows an upward trend, with Paris not only maintaining its dominance but also widening the gap with other FUAs. In the United Kingdom (Panel B) and Germany (Panel C), by contrast, EQI values are more stable, with leading FUAs sustaining consistently high levels of quality but showing less divergence over time.

——————— Insert Figure 6 Here ———————

5 Conclusion

This paper provides the first large-scale, comparative application of the G&S framework for measuring entrepreneurial quality and ecosystem potential in Europe. Using detailed firm-level registry data for France, Germany, and the United Kingdom from 2009 to 2023, we construct forward-looking indicators of entrepreneurial ecosystem performance by taking into account both the quantity and the quality of entrepreneurial activity.

Our findings confirm that the predictive analytics approach developed in the United States performs well across institutional boundaries. Founding characteristics, such as corporate form, intellectual property filings, and, to some extent, name-based features, retain strong predictive power for high-growth outcomes in European economies. At the same time, we uncover meaningful cross-country differences. Name-based indicators, particularly the use of short or eponymous firm names, show less consistent effects, suggesting that cultural and linguistic conventions shape entrepreneurial signaling. Moreover, the magnitude of predictive coefficients varies across countries, reflecting differences in institutional environments, capital market structures, and regulatory regimes. Despite these nuances, the robustness of the models highlights the generalizability of the G&S framework beyond the United States.

Our analysis of entrepreneurial activity across France, Germany, and the United Kingdom yields three broad insights. First, country-level trends diverge markedly. France shows a relatively smooth and steady increase in quality-adjusted entrepreneurial activity (RECPI), the United Kingdom exhibits sharp fluctuations and greater volatility, while Germany follows a declining long-term trajectory. When comparing realized to expected growth outcomes (REAI), the French and British ecosystems underperform in later years, whereas Germany largely sustains realized growth in line with expectations. Second, regional patterns highlight substantial heterogeneity within countries. Entrepreneurial quantity and quality are positively correlated across European FUAs, in contrast to the weak association found in the United States (Andrews et al., 2022). This suggests that in Europe, entrepreneurial ecosystems capable of producing more startups also tend to generate higher-quality firms. Third, cross-country comparisons reveal different ecosystem structures. France remains highly centralized, with Paris dominating but secondary metropolitan hubs (e.g., Toulouse, Marseille, Bordeaux) gaining importance over time. The United Kingdom is more polycentric, with London, Cambridge, and Oxford leading but mid-sized regions emerging as dynamic contributors. Germany is distinctive in its decentralization: Munich, Heidelberg, and Bonn rank highly, but several other - and smaller FUAs in terms of the overall number of business formations (e.g., Jena, Dresden, Greifswald) also perform strongly, reflecting a more distributed pattern of high-quality entrepreneurship.

In all three countries, the quality of entrepreneurial ecosystems appears to be closely associated with the presence of entrepreneurial universities (Guerrero et al., 2014) and strong research capacity (Link and Siegel, 2005; Fudickar and Hottenrott, 2019). We identify leading ecosystems in each country, reflecting a pattern similar to that observed in the United States. However, unlike the US case, we also find evidence of high-quality entrepreneurship in smaller urban areas with a strong knowledge base anchored in scientific institutions, confirming earlier insights on the spatial distribution of patenting in Europe (Fritsch and Wyrwich, 2021). More broadly, we observe a stronger positive correlation between the quantity and quality of entrepreneurship in Europe, particularly in France and the United Kingdom, where entrepreneurial activity is more concentrated in the capital cities. With respect to ecosystem performance relative to potential (REAI), France and the UK initially outperformed its potential but subsequently experienced a steady decline in the extent to which this potential was realized. By contrast, Germany's entrepreneurial ecosystem demonstrated relatively stable performance throughout the observation period, with notable overperformance between 2012 and 2016.

Future research should extend this framework to additional European countries, examine longer-term outcomes beyond equity events and employment thresholds, and investigate the causal mechanisms linking ecosystem characteristics to realized firm growth. By making the European entrepreneurial landscape more transparent and comparable, this study provides a foundation for both scholarly inquiry and evidence-based policy. These indicators could be used to evaluate recent policy initiatives such as the Tibi initiative in France and Germany's Growth and Innovation Capital for Germany (WIN) initiative.

At the same time, this approach has limitations. Our results show that measuring quality through equity growth in Europe captures a different dimension than quality measured by high employment, unlike in the United States where IPOs are more prevalent. A more detailed analysis of employment dynamics would offer deeper insights into the labor market implications of high quality entrepreneurship, including the differences between ecosystems that perform strongly and those that lag behind. Moreover, analyzing the nature of the employment created by young firms would provide additional insights on their role in regional development.

These findings suggest that the proposed framework can serve as a valuable tool for policy evaluation. By shedding light on the mechanisms underlying entrepreneurial ecosystem performance, it enables the design of more targeted and effective policies. Given the substantial variation in regional and national ecosystems across Europe, fostering high-quality entrepreneurship requires a tailored mix of policy approaches. Moreover, the framework facilitates more meaningful international comparisons, both within Europe and beyond.

References

Acs, Z. and C. Armington (2004). "Employment Growth and Entrepreneurial Activity in Cities." *Regional Studies* 38(8), 911–927.

Acs, Z. J., E. Autio, and L. Szerb (2014). "National Systems of Entrepreneurship: Measurement issues and policy implications." *Research Policy* 43(3), 476–494.

Andrews, R., C. Fazio, J. Guzman, Y. Liu, and S. Stern (2022). "The Startup Cartography Project: Measuring and mapping entrepreneurial ecosystems." *Research Policy* 51(2), 104437.

Audretsch, D. B., M. Belitski, and N. Cherkas (2021). "Entrepreneurial ecosystems in cities: The role of institutions." *PLOS ONE* 16(3), e0247609.

Colombo, M. G., G. B. Dagnino, E. E. Lehmann, and M. Salmador (2019). "The governance of entrepreneurial ecosystems." *Small Business Economics* (2), 419–428.

Content, J., N. Bosma, J. Jordaan, and M. Sanders (2020). "Entrepreneurial ecosystems, entrepreneurial activity and economic growth: new evidence from European regions." *Regional Studies* 54(8), 1007–1019.

Dahl, M. S. and O. Sorenson (2012). "Home sweet home: Entrepreneurs' location choices and the performance of their ventures." *Management Science* 58(6), 1059–1071.

Decker, R. A., J. Haltiwanger, R. S. Jarmin, and J. Miranda (2016). "Where has all the skewness gone? The decline in high-growth (young) firms in the U.S." *European Economic Review* 86, 4–23.

Duvendack, M., R. Palmer-Jones, and W. R. Reed (2017). "What Is Meant by "Replication" and Why Does It Encounter Resistance in Economics?" *American Economic Review* 107(5), 46–51.

Feldman, M. P. (2001). "The Entrepreneurial Event Revisited: Firm Formation in a Regional Context." *Industrial and Corporate Change* 10(4), 861–891.

Fritsch, M. and M. Wyrwich (2021). "Is Innovation (Increasingly) Concentrated in Large Cities? An International Comparison." *Research Policy* 50(6), 104237.

Fudickar, R. and H. Hottenrott (2019). "Public Research and the Innovation Performance of New Technology-Based Firms." *The Journal of Technology Transfer* 44(2), 326–358.

Ghio, N., M. Guerini, and C. Rossi-Lamastra (2019). "The creation of high-tech ventures in entrepreneurial ecosystems: exploring the interactions among university knowledge, cooperative banks, and individual attitudes." *Small Business Economics* (2), 523–543.

Gicheva, D. and A. N. Link (2016). "On the economic performance of nascent entrepreneurs." *European Economic Review* 86, 109–117.

Glaeser, E. L., S. P. Kerr, and W. R. Kerr (2015). "Entrepreneurship and Urban Growth: An Empirical Assessment with Historical Mines." *The Review of Economics and Statistics* 97(2), 498–520.

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

Gourio, F., T. Messer, and M. Siemer (2016). "Firm entry and macroeconomic dynamics: a state-level analysis." *American Economic Review* 106(5), 214–218.

Guerrero, M., J. A. Cunningham, and D. Urbano (2015). "Economic impact of entrepreneurial universities' activities: An exploratory study of the United Kingdom." *Research Policy* 44(3), 748–764.

Guerrero, M. and D. S. Siegel (2024). "Schumpeter meets Teece: Proposed metrics for assessing entrepreneurial innovation and dynamic capabilities in entrepreneurial ecosystems in an emerging economy." *Research Policy* 53(5), 104984.

Guerrero, M., D. Urbano, J. Cunningham, and D. Organ (2014). "Entrepreneurial universities in two European regions: a case study comparison." *The Journal of Technology Transfer* 39(3), 415–434.

Guzman, J. and S. Stern (2020). "The State of American Entrepreneurship: New Estimates of the Quantity and Quality of Entrepreneurship for 32 US States, 1988–2014." *American Economic Journal: Economic Policy* 12(4), 212–243.

Haltiwanger, J. (2022). "Entrepreneurship in the twenty-first century." *Small Business Economics* 58(1), 27–40.

Hamermesh, D. S. (2007). "Viewpoint: Replication in Economics." *Canadian Journal of Economics/Revue canadienne d'économique* 40(3), 715–733.

Hébert, R. F. and A. N. Link (2006). "The Entrepreneur as Innovator." *The Journal of Technology Transfer*, 589–597.

Herrera, F., M. Guerrero, and D. Urbano (2018). "Entrepreneurship and Innovation Ecosystem's Drivers: The Role of Higher Education Organizations." In: *Entrepreneurial, Innovative and Sustainable Ecosystems: Best Practices and Implications for Quality of Life*. Ed. by J. Leitão, H. Alves, N. Krueger, and J. Park. Cham: Springer International Publishing, 109–128. DOI: 10.1007/978-3-319-71014-3_6. URL: https://doi.org/10.1007/978-3-319-71014-3_6.

Kortum, S. and J. Lerner (2000). "Assessing the Contribution of Venture Capital to Innovation." *The RAND Journal of Economics* 31(4), 674–692.

Lerner, J. (2009). "The Empirical Impact of Intellectual Property Rights on Innovation: Puzzles and Clues." *American Economic Review* 99(2), 343–348.

Link, A. N. and D. S. Siegel (2005). "Generating science-based growth: an econometric analysis of the impact of organizational incentives on university–industry technology transfer." *The European Journal of Finance* 11(3), 169–181.

Neuhäusler, P., R. Frietsch, and O. Rothengatter (2021). *Documentation of RISIS Datasets: ISI-Trademark Data Collection (ISI-TM)*. Tech. rep. Zenodo. DOI: 10.5281/zenodo.4633061.

Qian, H., Z. J. Acs, and R. R. Stough (2013). "Regional Systems of Entrepreneurship: The Nexus of Human Capital, Knowledge and New Firm Formation." *Journal of Economic Geography* 13(4), 559–587.

Rocha, A., R. Brown, and S. Mawson (2021). "Capturing Conversations in Entrepreneurial Ecosystems." *Research Policy* 50(9), 104317.

Samila, S. and O. Sorenson (2011). "Venture Capital, Entrepreneurship, and Economic Growth." *The Review of Economics and Statistics* 93(1), 338–349.

Stam, E. (2015). "Entrepreneurial Ecosystems and Regional Policy: A Sympathetic Critique." *European Planning Studies* 23(9), 1759–1769.

Stam, E. and A. Van De Ven (2021). "Entrepreneurial ecosystem elements." *Small Business Economics* 56(2), 809–832.

Tartari, V. and S. Stern (2021). *More than an Ivory Tower: The Impact of Research Institutions on the Quantity and Quality of Entrepreneurship*. w28846. Cambridge, MA: National Bureau of Economic Research, w28846. DOI: 10.3386/w28846. URL: http://www.nber.org/papers/w28846.pdf.

Table 1: Legal Forms of Companies in the France, UK and Germany

	Number of Firms	Share
Panel A: France		
Société par actions simplifiée	1,514,405	50.35%
Société à responsabilité limitée (SARL)	1,489,526	49.52%
Société anonyme à conseil d'administration	3,379	0.11%
Société anonyme à directoire	238	<0.01%
Société européenne	96	<0.01%
Total	3,007,644	100%
Panel B: UK		
Private Limited Company	4,954,260	99.92%
Public Limited Company	3,952	0.08%
European Public Limited-Liability Company (SE)	42	<0.01%
United Kingdom Societas	9	<0.01%
Private limited company (Sect. 30)	1	<0.01%
Total	4,958,264	100%
Panel C: Germany		
GmbH	1,331,085	84.44%
GmbH & Co. KG	215,982	13.70%
AG	11,569	0.73%
Firma (Ausland)	10,918	0.69%
freie Berufe	6,430	0.41%
Gewerbebetrieb	173	0.01%
KG	157	<0.01%
BGB-Gesellschaft	74	<0.01%
OHG	37	<0.01%
BGB-Gesellschaft - Arbeitsgemeinschaft	4	<0.01%
Einzelfirma	4	<0.01%
eV	3	<0.01%
Total	1,576,436	100%

Notes: This table reports the distribution of firms by legal form in France, the United Kingdom, and Germany.

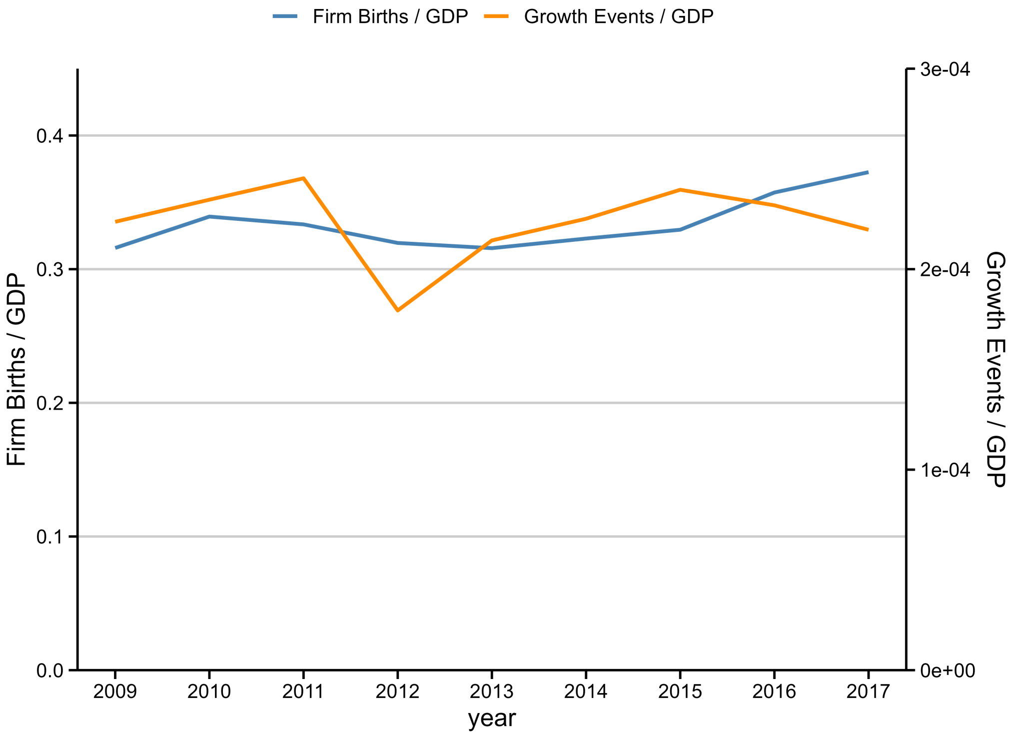
Table 2: Summary Statistics

Variable	France		UK		Germany	
	Mean	St. Dev.	Mean	St. Dev.	Mean	St. Dev.
Outcome variable						
Growth	0.0004	0.0210	0.0009	0.0303	0.0013	0.0362
Corporate form observables						
Public Company	0.0012	0.0351	0.0008	0.0284	0.0073	0.0853
Popular Business Structure	0.4928	0.4999	0.9992	0.0284	0.6149	0.4866
Name-based observables						
Short Name	0.8987	0.3017	0.5715	0.4949	0.8010	0.3992
Eponymous	0.1042	0.3055	0.0923	0.2894	0.1532	0.3602
Intelectual property observables						
Patent	0.0007	0.0272	0.0015	0.0382	0.0098	0.0987
Trademark	0.0009	0.0298	0.0015	0.0387	0.0041	0.0642
Industry Measures						
Mining and quarrying	0.0001	0.0105	0.0019	0.0435	0.0005	0.0218
Manufacturing	0.0365	0.1876	0.0345	0.1826	0.0511	0.2201
Electricity, gas, steam and air conditioning supply	0.0100	0.0996	0.0036	0.0603	0.0217	0.1456
Water supply; sewerage, waste management	0.0020	0.0445	0.0028	0.0529	0.0029	0.0534
Construction	0.1369	0.3437	0.0851	0.2790	0.0681	0.2520
Wholesale and retail trade; repair of vehicles	0.1830	0.3867	0.0970	0.2960	0.1342	0.3408
Transportation and storage	0.0365	0.1876	0.0327	0.1777	0.0274	0.1632
Accommodation and food service activities	0.0990	0.2987	0.0492	0.2164	0.0325	0.1772
Information and communication	0.0550	0.2280	0.0824	0.2750	0.0601	0.2377
Financial and insurance activities	0.0707	0.2563	0.0272	0.1627	0.1028	0.3037
Real estate activities	0.0540	0.2260	0.0550	0.2280	0.0877	0.2829
Professional, scientific and technical activities	0.1441	0.3511	0.1124	0.3158	0.1547	0.3617
Administrative and support service activities	0.0579	0.2336	0.0905	0.2868	0.1315	0.3379
Public administration and defence	0.0000	0.0023	0.0016	0.0403	0.0004	0.0207
Education	0.0127	0.1120	0.0137	0.1162	0.0101	0.1000
Human health and social work activities	0.0146	0.1201	0.0400	0.1959	0.0195	0.1382
Arts, entertainment and recreation	0.0125	0.1111	0.0192	0.1373	0.0117	0.1077
Other service activities	0.0338	0.1808	0.0474	0.2126	0.0404	0.1970
Missing	0.0405	0.1972	0.2037	0.4028	0.0428	0.2024
Observations	3,007,644		4,958,264		1,576,436	

Notes: This table reports summary statistics from the national business registry data of France, the United Kingdom, and Germany. The unit of observation is a limited liability company incorporated between 2009 and 2023 for France and Germany, and between 2010 and 2020 for the United Kingdom. Industry indicators are based on 1-digit NACE codes.

Figure 1: Firm Births versus Number of Growth Events per Cohort

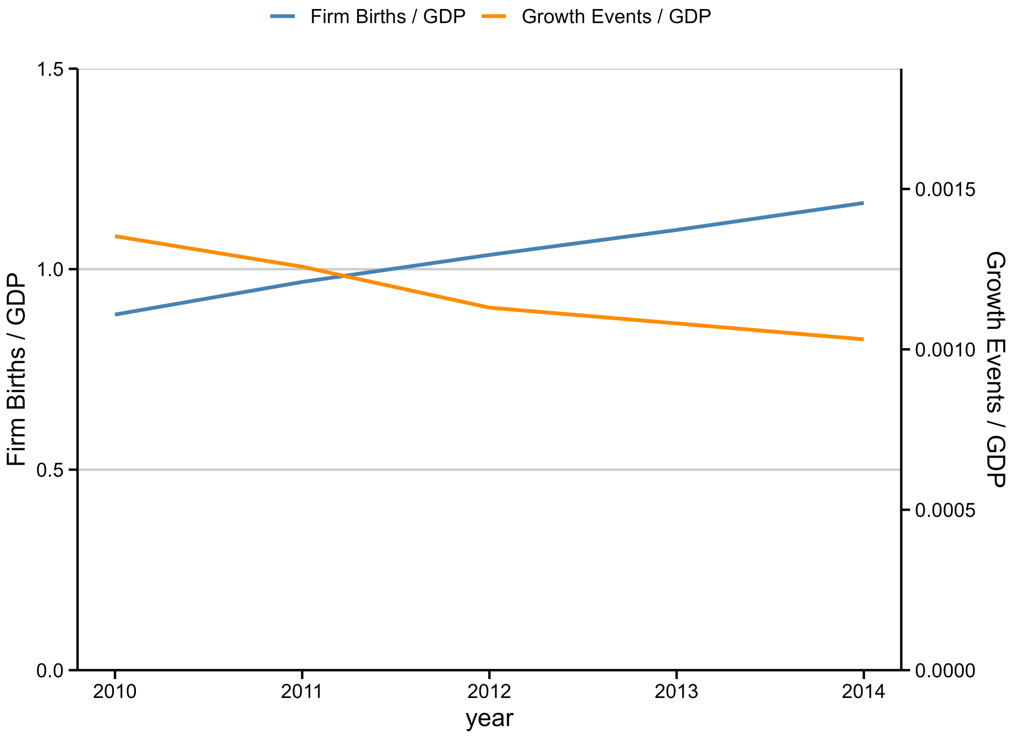
Panel A: France (2009-2017)



Notes: This figure plots the ratio of firm births to GDP (left axis) and the ratio of growth events to GDP (right axis) for France over the period 2009–2017. Growth events are defined as firms reaching IPO or acquisition within six years of founding.

Figure 1: Firm Births versus Number of Growth Events per Cohort (*continued*)

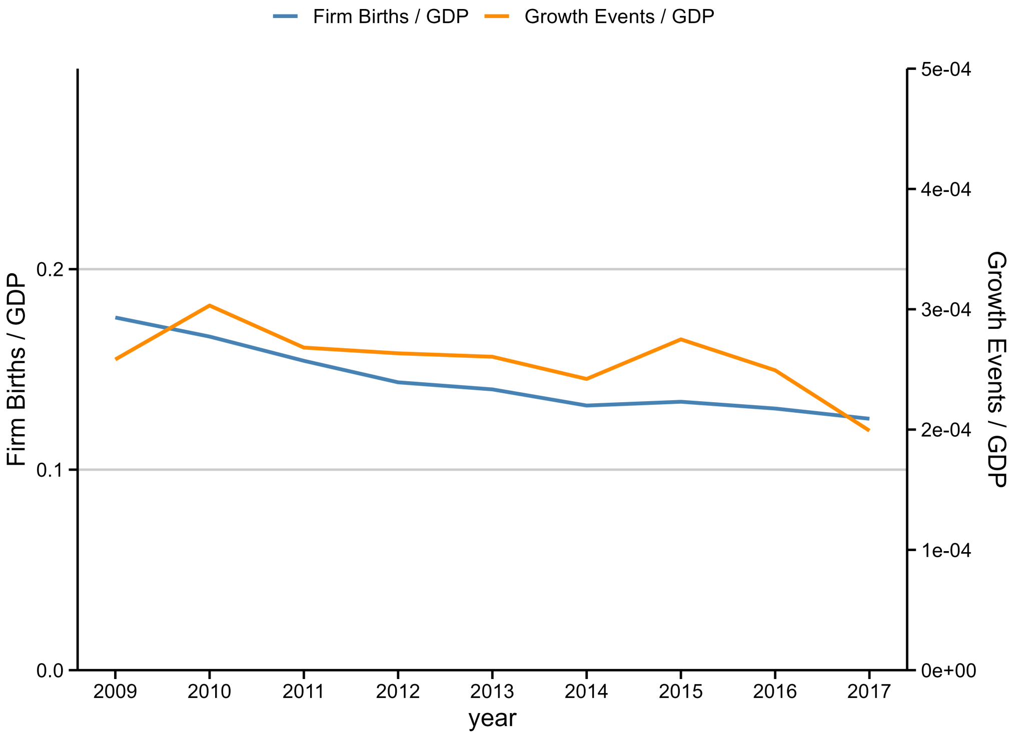
Panel B: UK (2010-2014)



Notes: This figure plots the ratio of firm births to GDP (left axis) and the ratio of growth events to GDP (right axis) for the United Kingdom over the period 2010–2014. Growth events are defined as firms reaching IPO or acquisition within six years of founding.

Figure 1: Firm Births versus Number of Growth Events per Cohort (*continued*)

Panel C: Germany (2009-2017)



Notes: This figure plots the ratio of firm births to GDP (left axis) and the ratio of growth events to GDP (right axis) for Germany over the period 2009–2017. Growth events are defined as firms reaching IPO or acquisition within six years of founding.

Table 3: Logit Univariate Regressions

	France		UK		Germany	
Variable	Coefficient	Pseudo R^2	Coefficient	Pseudo R^2	Coefficient	Pseudo R^2
Name-based measures						
Short Name	1.604***	0.001	1.111**	0.000	1.525***	0.002
	(0.191)		(0.046)		(0.103)	
Eponymous	0.332***	0.004	0.666***	0.001	0.138***	0.011
	(0.053)		(0.054)		(0.024)	
Corporate form measures						
Public Company	15.974***	0.005	151.423***	0.062	9.102***	0.014
	(3.316)		(10.068)		(0.818)	
Popular Business Structure	4.450***	0.032	0.007***	0.061	2.240***	0.009
	(0.299)		(0.000)		(0.131)	
Intellectual property measures						
Patent	32.678***	0.012	26.221***	0.012	11.470***	0.025
	(5.323)		(2.703)		(0.859)	
Trademark	12.324***	0.003	12.151***	0.003	11.417***	0.009
	(2.865)		(1.951)		(1.362)	
Industry measures						
Mining and quarrying	13.854***	0.000	9.682***	0.003	0.992	0.000
	(9.851)		(1.419)		(0.993)	
Manufacturing	1.656***	0.001	2.884***	0.004	2.560***	0.005
	(0.209)		(0.213)		(0.187)	
Electricity, gas, steam and air conditioning supply	0.865	0.000	4.501***	0.002	1.302*	0.000
	(0.275)		(0.672)		(0.180)	
Water supply; sewerage, waste management	0.486	0.000	2.344***	0.000	0.979	0.000
	(0.486)		(0.589)		(0.439)	
Construction	0.142***	0.011	0.295***	0.003	0.113***	0.006
	(0.029)		(0.042)		(0.032)	
Wholesale and retail trade; repair of vehicles	0.511***	0.003	0.716***	0.000	0.700***	0.001
	(0.050)		(0.063)		(0.056)	
Transportation and storage	0.252***	0.002	0.478***	0.000	0.790	0.000
	(0.084)		(0.098)		(0.133)	
Accommodation and food service activities	0.139***	0.008	0.550***	0.001	0.237***	0.002
	(0.034)		(0.078)		(0.066)	
Information and communication	8.382***	0.041	2.314***	0.005	4.157***	0.017
	(0.562)		(0.125)		(0.253)	
Financial and insurance activities	2.316***	0.004	6.624***	0.015	1.162*	0.000
	(0.214)		(0.420)		(0.095)	
Real estate activities	0.484***	0.001	0.775*	0.000	0.297***	0.004
	(0.104)		(0.104)		(0.047)	
Professional, scientific and technical activities	1.684***	0.002	1.836***	0.003	0.807***	0.000
	(0.129)		(0.096)		(0.058)	
Administrative and support service activities	0.793	0.000	1.271***	0.000	1.453***	0.001
	(0.121)		(0.086)		(0.100)	
Public administration and defence	0.005***	0.000	0.847	0.000	1.247	0.000
	(0.002)		(0.600)		(1.249)	
Education	0.517	0.000	0.481***	0.000	0.192***	0.001
	(0.212)		(0.129)		(0.111)	
Human health and social work activities	0.481*	0.000	0.467***	0.001	1.017	0.000
	(0.197)		(0.073)		(0.187)	
Arts, entertainment and recreation	0.443**	0.000	0.818	0.000	0.449**	0.000
	(0.182)		(0.139)		(0.142)	
Other service activities	0.162***	0.002	0.486***	0.001	0.958	0.000
	(0.066)		(0.074)		(0.118)	
Missing	0.455***	0.001	0.235***	0.018	0.009***	0.008
	(0.103)		(0.016)		(0.009)	

Notes: This table reports results from univariate logit regressions estimating the probability that a firm experiences growth, defined as achieving an IPO or acquisition within six years of incorporation. Each column corresponds to a specification including a single predictor. Industry indicators are based on 1-digit NACE codes. Reported coefficients are incidence rate ratios. Robust standard errors are in parentheses. Significance levels: *p<0.1; **p<0.05; ***p<0.01.

Table 4: Growth Predictive Model - Logit Regression on IPO or Acquisition within Six Years - France

	Preliminary Models			Nowcasting	Full Model
	(1)	(2)	(3)	(4)	(5)
Corporate form measures					
Public Company	35.043***			19.793***	16.541***
	(7.471)			(4.364)	(3.612)
Popular Business Structure	4.808***			3.346***	3.270***
	(0.332)			(0.239)	(0.235)
Name-based measures					
Short Name		1.505***		1.266*	1.239*
		(0.180)		(0.153)	(0.149)
Eponymous		0.343***		0.514***	0.515***
		(0.055)		(0.084)	(0.085)
Intellectual property measures					
Patent			28.301***		10.162***
			(4.971)		(1.934)
Trademark			8.526***		4.811***
			(2.209)		(1.234)
Industry Fixed Effects	NO	NO	NO	YES	YES
Observations	1,548,392	1,548,392	1,548,392	1,548,392	1,548,392
Pseudo R^2	0.039	0.005	0.014	0.121	0.129

Notes: This table reports estimates from logit regressions where the dependent variable, *Growth*, equals 1 if a firm undergoes an IPO or acquisition within six years of incorporation and 0 otherwise. Columns (1)–(3) report models including only one class of predictors (name-based, corporate form, or intellectual property measures, respectively). Column (4) includes only business registry–based information, while Column (5) augments the model with intellectual property information. Reported coefficients are incidence rate ratios. Industry fixed effects, based on 2-digit NACE codes, are included only in Columns (4) and (5). Robust standard errors are shown in parentheses. Significance levels: *p<0.1; **p<0.05; ***p<0.01.

Table 5: Growth Predictive Model - Logit Regression on IPO or Acquisition within Six Years - UK

	Preliminary Models			Nowcasting	Full Model
	(1)	(2)	(3)	(4)	(5)
Corporate form measures					
Public Company	151.215***			94.850***	90.333***
	(9.970)			(8.104)	(7.815)
Name-based measures					
Short Name		1.074*		1.031	1.007
		(0.045)		(0.044)	(0.043)
Eponymous		0.673***		0.745***	0.757***
		(0.055)		(0.062)	(0.063)
Intellectual property measures					
Patent			22.759***		10.286***
			(2.524)		(1.327)
Trademark			7.178***		5.077***
			(1.309)		(0.978)
Industry Fixed Effects	NO	NO	NO	YES	YES
Observations	2,127,382	2,127,382	2,127,382	2,127,382	2,127,382
Pseudo R^2	0.062	0.001	0.015	0.117	0.125

Notes: This table reports estimates from logit regressions where the dependent variable, *Growth*, equals 1 if a firm undergoes an IPO or acquisition within six years of incorporation and 0 otherwise. Columns (1)–(3) report models including only one class of predictors (name-based, corporate form, or intellectual property measures, respectively). Column (4) includes only business registry–based information, while Column (5) augments the model with intellectual property information. Reported coefficients are incidence rate ratios. Industry fixed effects, based on 2-digit NACE codes, are included only in Columns (4) and (5). Robust standard errors are shown in parentheses. Significance levels: *p<0.1; **p<0.05; ***p<0.01.

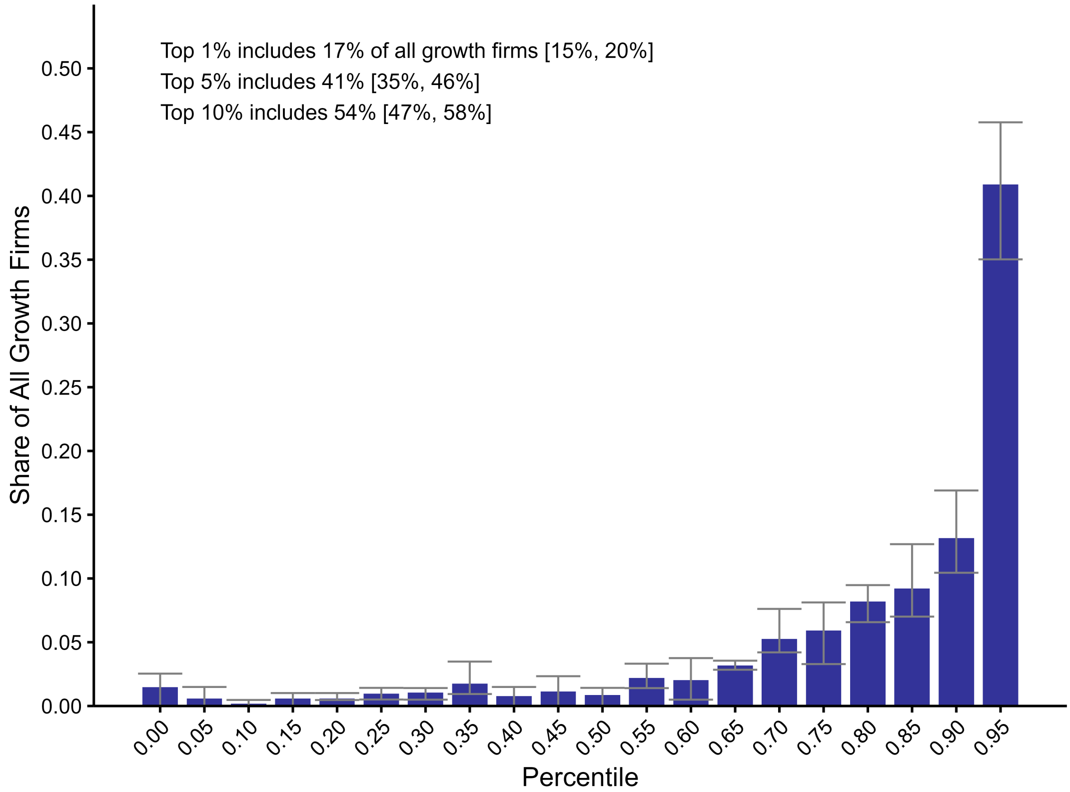
Table 6: Growth Predictive Model - Logit Regression on IPO or Acquisition within Six Years - Germany

	Preliminary Models			Nowcasting	Full Model
	(1)	(2)	(3)	(4)	(5)
Corporate form measures					
Public Company	22.222***			24.598***	21.106***
	(2.379)			(2.677)	(2.342)
Popular Business Structure	3.398***			3.400***	3.187***
	(0.236)			(0.240)	(0.225)
Name-based measures					
Short Name		1.493***		1.268***	1.198***
		(0.101)		(0.087)	(0.083)
Eponymous		0.139***		0.173***	0.172***
		(0.024)		(0.031)	(0.031)
Intellectual property measures					
Patent			10.026***		6.047***
			(0.811)		(0.507)
Trademark			7.316***		3.991***
			(0.971)		(0.532)
Industry Fixed Effects	NO	NO	NO	YES	YES
Observations	922,702	922,702	922,702	922,702	922,702
Pseudo R^2	0.031	0.012	0.032	0.099	0.119

Notes: This table reports estimates from logit regressions where the dependent variable, *Growth*, equals 1 if a firm undergoes an IPO or acquisition within six years of incorporation and 0 otherwise. Columns (1)–(3) report models including only one class of predictors (name-based, corporate form, or intellectual property measures, respectively). Column (4) includes only business registry–based information, while Column (5) augments the model with intellectual property information. Reported coefficients are incidence rate ratios. Industry fixed effects, based on 2-digit NACE codes, are included only in Columns (4) and (5). Robust standard errors are shown in parentheses. Significance levels: *p<0.1; **p<0.05; ***p<0.01.

Figure 2: Entrepreneurial Quality Predictions Validation: 5-Fold Out of Sample Test

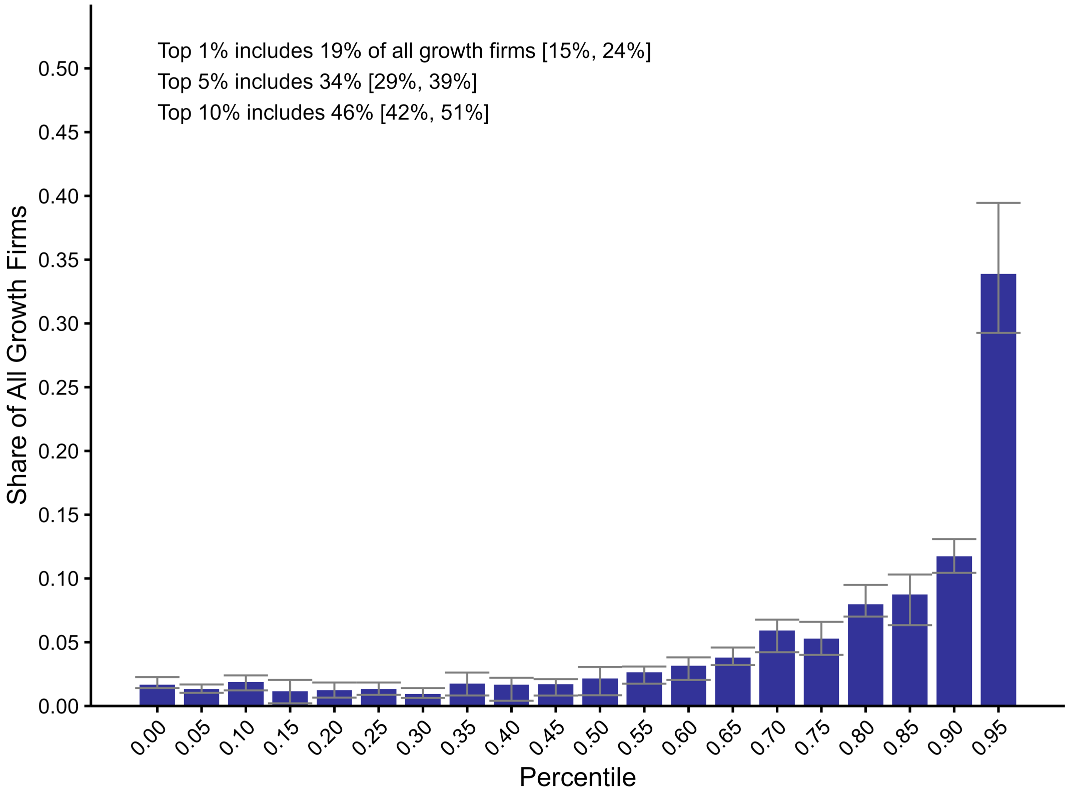
Panel A: France



Notes: This figure presents the results of an out-of-sample cross-validation procedure performed on all firms. We use a fivefold cross-validation and plot the incidence of growth across each 5 percent bin.

Figure 2: Entrepreneurial Quality Predictions Validation: 5-Fold Out of Sample Test (*continued*)

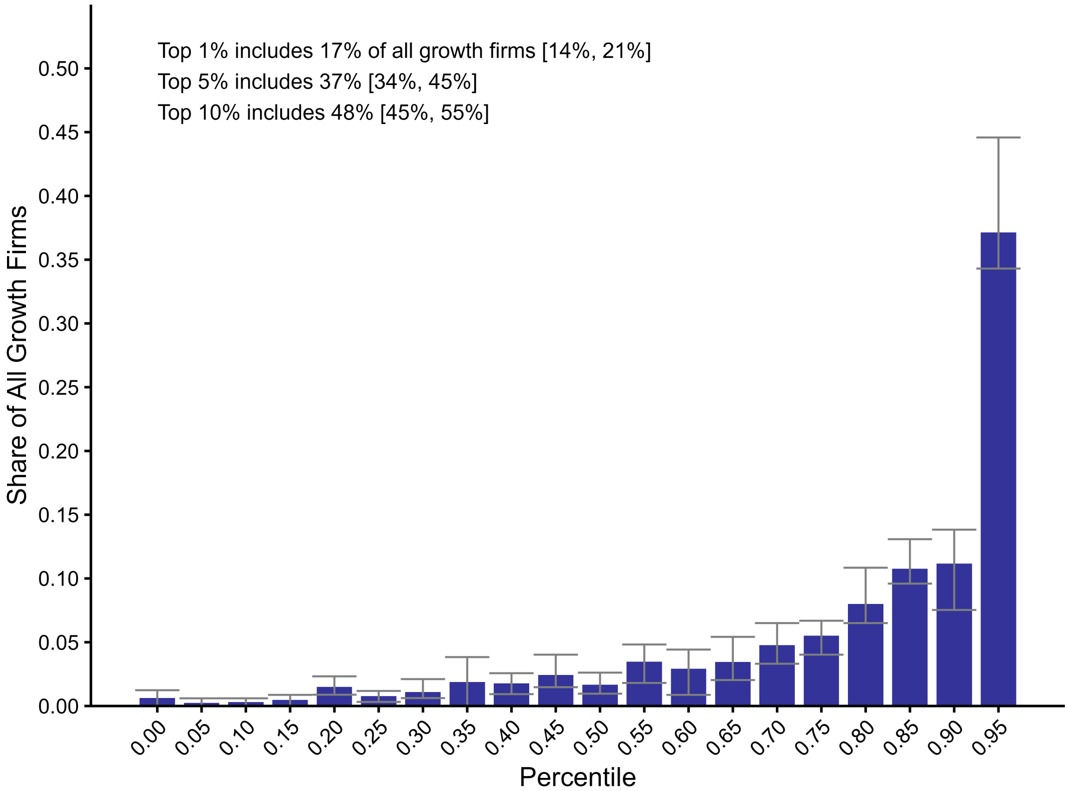
Panel B: UK



Notes: This figure presents the results of an out-of-sample cross-validation procedure performed on all firms. We use a fivefold cross-validation and plot the incidence of growth across each 5 percent bin.

Figure 2: Entrepreneurial Quality Predictions Validation: 5-Fold Out of Sample Test (*continued*)

Panel C: Germany



Notes: This figure presents the results of an out-of-sample cross-validation procedure performed on all firms. We use a fivefold cross-validation and plot the incidence of growth across each 5 percent bin.

Table 7: Entrepreneurial Quality Models with Firm Size Growth Outcomes - France

	Full Model	Employment > 500	Employment > 1000
	(1)	(2)	(3)
Corporate form measures			
Public Company	16.541***	18.738***	8.388***
	(3.612)	(4.478)	(3.474)
Popular Business Structure	3.270***	3.904***	3.386***
	(0.235)	(0.373)	(0.461)
Name-based measures			
Short Name	1.239*	0.724***	0.842
	(0.149)	(0.088)	(0.166)
Eponymous	0.515***	0.657**	0.525**
	(0.085)	(0.111)	(0.146)
Intellectual property measures			
Patent	10.162***	25.138***	24.817***
	(1.934)	(4.553)	(6.003)
Trademark	4.811***	5.951***	7.209***
	(1.234)	(1.892)	(2.892)
Industry Fixed Effects	YES	YES	YES
Observations	1,548,392	1,548,392	1,548,392
Pseudo R^2	0.129	0.197	0.215

Notes: This table reports estimates from logit regressions using the same set of predictors as the full information entrepreneurial quality model for France (Table 4, column 5), but replacing equity growth outcomes with high employment growth outcomes. Our outcome variable is 1 if a firm has high employment withing six years after founding and 0 otherwise, at different thresholds. Column (1) presents results for the baseline specification; Column (2) defines high employment growth as reaching at least 500 employees; Column (3) defines it as reaching at least 1,000 employees. Reported coefficients are incidence rate ratios. All models include industry fixed effects based on 2-digit NACE codes. Robust standard errors are in parentheses. Significance levels: *p<0.1; **p<0.05; ***p<0.01.

Table 8: Entrepreneurial Quality Models with Firm Size Growth Outcomes - UK

	Full Model	Employment > 500	Employment > 1000
	(1)	(2)	(3)
Corporate form measures			
Public Company	90.333***	13.541***	14.799***
	(7.815)	(2.132)	(2.890)
Name-based measures			
Short Name	1.007	0.813***	0.739***
	(0.043)	(0.046)	(0.058)
Eponymous	0.757***	0.489***	0.438***
	(0.063)	(0.060)	(0.078)
Intellectual property measures			
Patent	10.286***	19.770***	21.325***
	(1.327)	(2.861)	(4.027)
Trademark	5.077***	9.263***	12.070***
	(0.978)	(1.888)	(3.002)
Industry Fixed Effects	YES	YES	YES
Observations	2,127,382	2,127,382	2,127,382
Pseudo R^2	0.125	0.151	0.163

Notes: This table reports estimates from logit regressions using the same set of predictors as the full information entrepreneurial quality model for the United Kingdom (Table 5, column 5), but replacing equity growth outcomes with high employment growth outcomes. Our outcome variable is 1 if a firm has high employment withing six years after founding and 0 otherwise, at different thresholds. Column (1) presents results for the baseline specification; Column (2) defines high employment growth as reaching at least 500 employees; Column (3) defines it as reaching at least 1,000 employees. Reported coefficients are incidence rate ratios. All models include industry fixed effects based on 2-digit NACE codes. Robust standard errors are in parentheses. Significance levels: *p<0.1; **p<0.05; ***p<0.01.

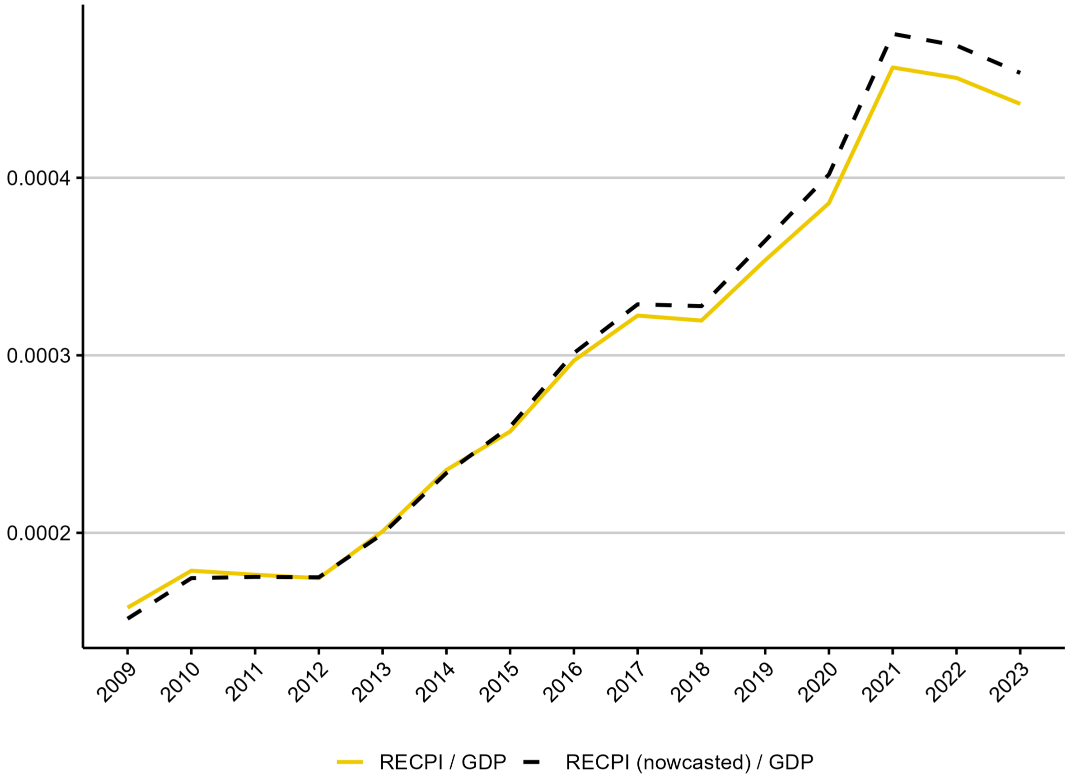
Table 9: Entrepreneurial Quality Models with Firm Size Growth Outcomes - Germany

	Full Model (1)	Employment > 500 (2)	Employment > 1000 (3)
Corporate form measures			
Public Company	21.106***	10.592***	13.167***
	(2.342)	(1.362)	(2.030)
Popular Business Structure	3.187***	2.008***	1.830***
	(0.225)	(0.137)	(0.166)
Name-based measures			
Short Name	1.198***	0.892*	0.943
	(0.083)	(0.057)	(0.081)
Eponymous	0.172***	0.635***	0.733***
	(0.031)	(0.059)	(0.086)
Intellectual property measures			
Patent	6.047***	17.503***	17.274***
	(0.507)	(1.266)	(1.633)
Trademark	3.991***	1.270	1.578
	(0.532)	(0.315)	(0.470)
Industry Fixed Effects	YES	YES	YES
Observations	922,702	922,702	922,702
Pseudo R^2	0.119	0.149	0.149

Notes: This table reports estimates from logit regressions using the same set of predictors as the full information entrepreneurial quality model for Germany (Table 6, column 5), but replacing equity growth outcomes with high employment growth outcomes. Our outcome variable is 1 if a firm has high employment withing six years after founding and 0 otherwise, at different thresholds. Column (1) presents results for the baseline specification; Column (2) defines high employment growth as reaching at least 500 employees; Column (3) defines it as reaching at least 1,000 employees. Reported coefficients are incidence rate ratios. All models include industry fixed effects based on 2-digit NACE codes. Robust standard errors are in parentheses. Significance levels: *p<0.1; **p<0.05; ***p<0.01.

Figure 3: Aggregate Entrepreneurship Regional Entrepreneurship Cohort Potential Index (RECPI) by Year

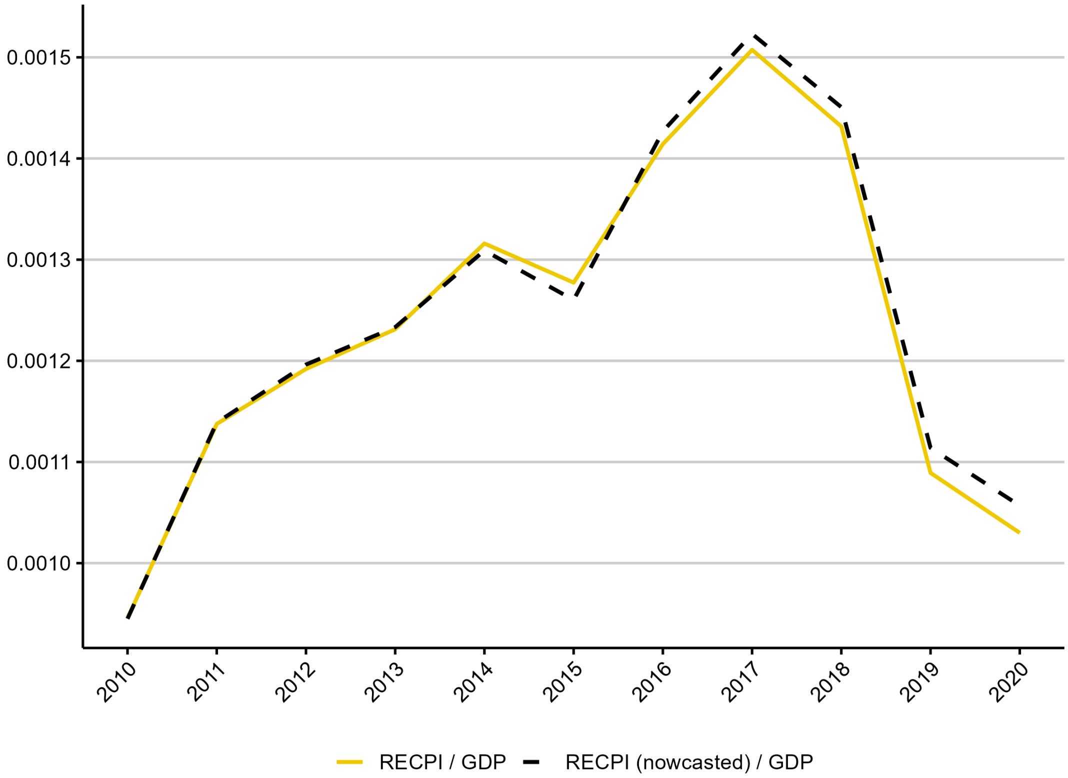
Panel A: France (2009-2023)



Notes: This figure shows the RECPI normalized by GDP for France from 2009 to 2023. The RECPI/GDP captures the total quality-adjusted entrepreneurial output in a region, accounting for the size of the economy in each year.

Figure 3: Aggregate Entrepreneurship Regional Entreprenuership Cohort Potential Index (RECPI) by Year (*continued*)

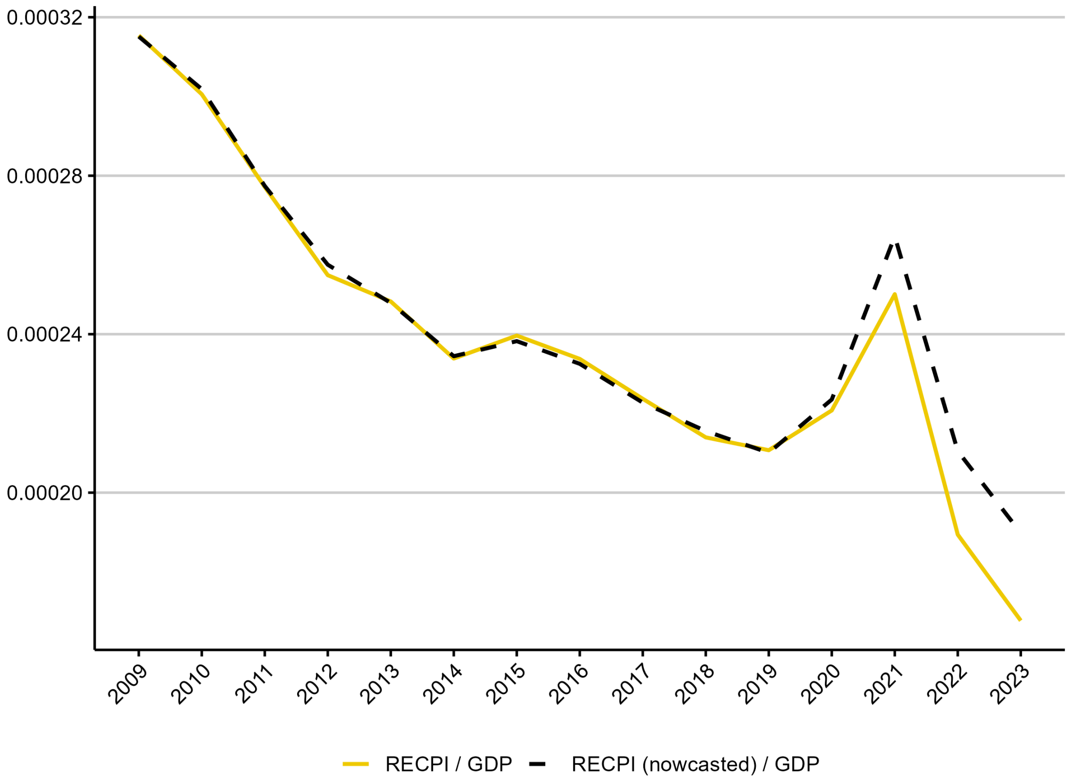
Panel B: UK (2010-2020)



Notes: This figure shows the RECPI normalized by GDP for the United Kingdom from 2010 to 2020. The RECPI/GDP captures the total quality-adjusted entrepreneurial output in a region, accounting for the size of the economy in each year.

Figure 3: Aggregate Entrepreneurship Regional Entreprenuership Cohort Potential Index (RECPI) by Year (*continued*)

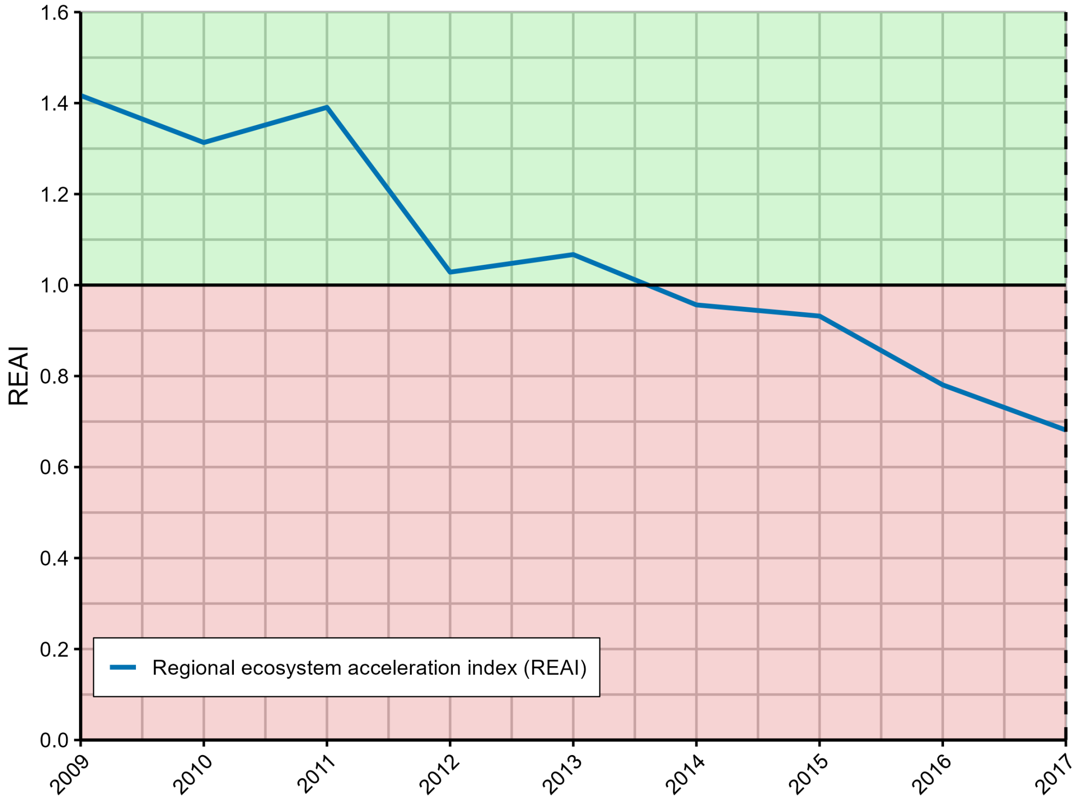
Panel C: Germany (2009-2023)



Notes: This figure shows the RECPI normalized by GDP for Germany from 2009 to 2023. The RECPI/GDP captures the total quality-adjusted entrepreneurial output in a region, accounting for the size of the economy in each year.

Figure 4: Regional Entrepreneurial Acceleration Index (REAI)

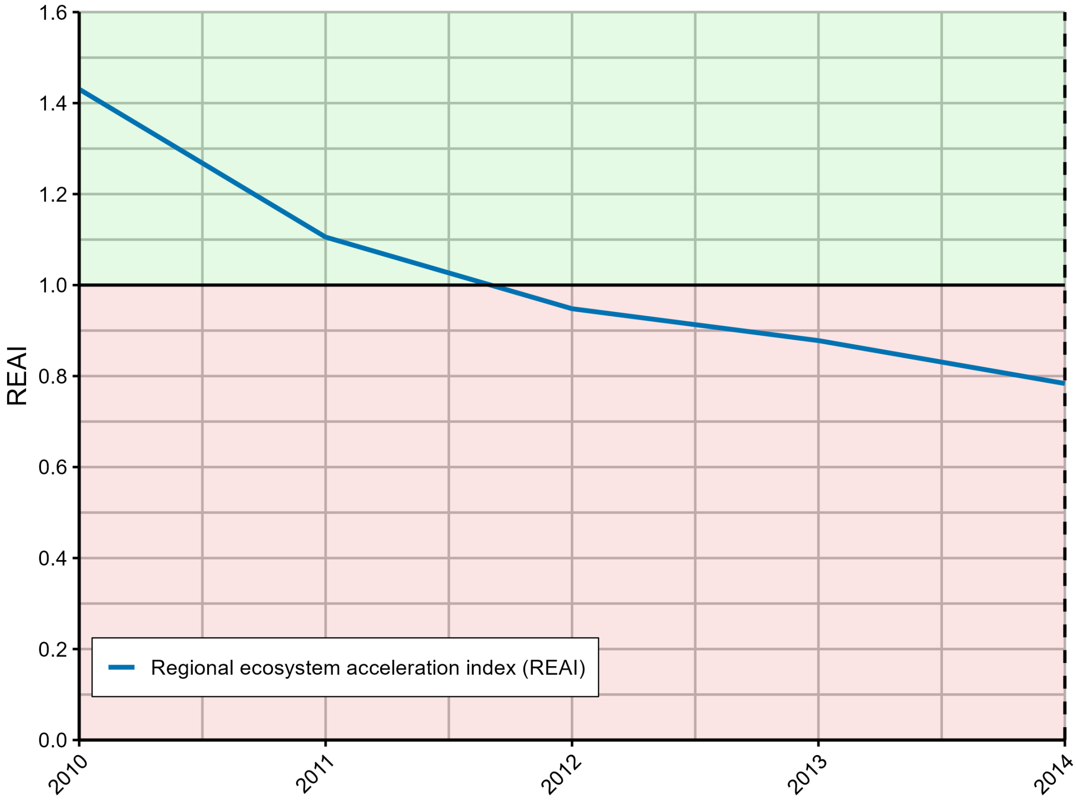
Panel A: France (2009-2019)



Notes: This figure shows the Regional Entrepreneurial Acceleration Index (REAI) for France from 2009 to 2019. The REAI measures how effectively an entrepreneurial ecosystem translates its potential into realized performance. It is calculated as the ratio of the number of growth events (IPOs or acquisitions within six years of founding) in a given cohort to that cohort's RECPI value. Projected performance represents preliminary estimates based on the growth events observed to date for each cohort.

Figure 4: Regional Entrepreneurial Acceleration Index (REAI) (*continued*)

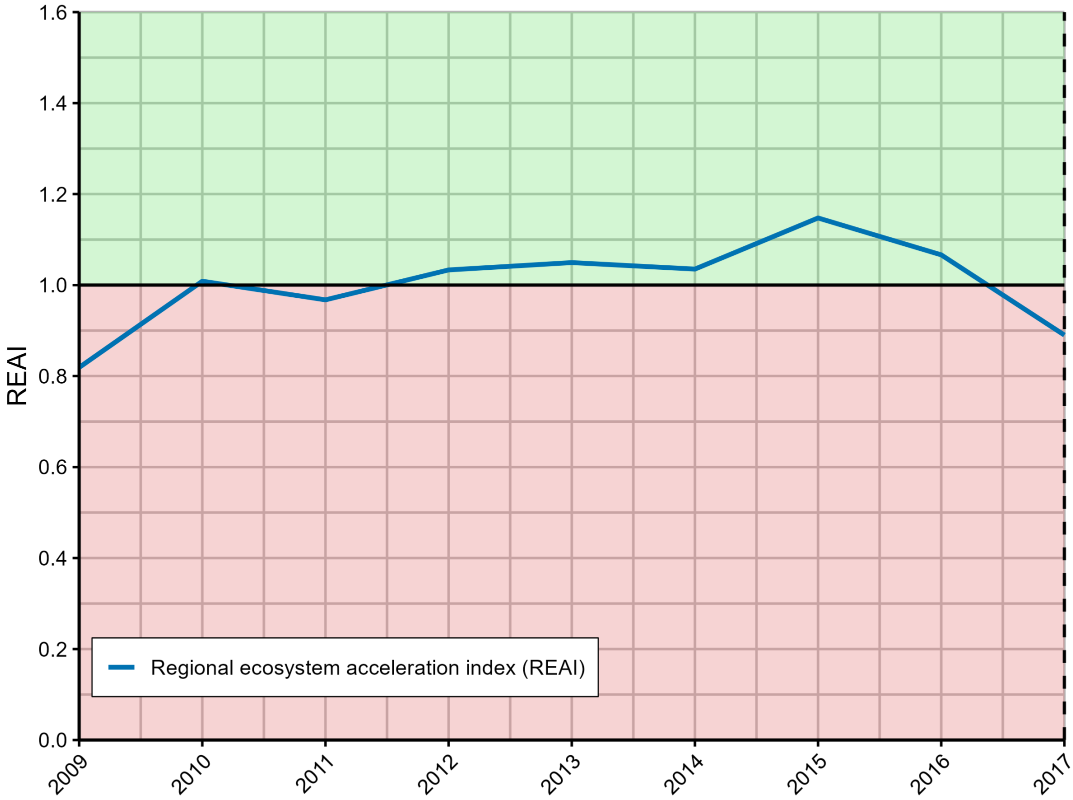
Panel B: UK (2010-2016)



Notes: This figure shows the Regional Entrepreneurial Acceleration Index (REAI) for the United Kingdom from 2010 to 2016. The REAI measures how effectively an entrepreneurial ecosystem translates its potential into realized performance. It is calculated as the ratio of the number of growth events (IPOs or acquisitions within six years of founding) in a given cohort to that cohort's RECPI value. Projected performance represents preliminary estimates based on the growth events observed to date for each cohort.

Figure 4: Regional Entrepreneurial Acceleration Index (REAI) (*continued*)

Panel C: Germany (2009-2019)



Notes: This figure shows the Regional Entrepreneurial Acceleration Index (REAI) for Germany from 2009 to 2019. The REAI measures how effectively an entrepreneurial ecosystem translates its potential into realized performance. It is calculated as the ratio of the number of growth events (IPOs or acquisitions within six years of founding) in a given cohort to that cohort's RECPI value. Projected performance represents preliminary estimates based on the growth events observed to date for each cohort.

Table 10: Summary Statistics by Regions

Panel A: France (N = 2,990,872)

	NUTS 1		NUTS 2		NUTS 3		LAU		FUA	
Variable	Mean	St. Dev.	Mean	St. Dev.	Mean	St. Dev.	Mean	St. Dev.	Mean	St. Dev.
Quantity	15,337.81	16,734.11	9,063.25	14,017.61	2,076.99	3,224.37	10.99	204.353	3,021.08	10,132.60
Average Quality (EQI) x 1,000	0.66	0.18	0.64	0.17	0.62	0.19	0.60	0.92	0.64	0.18
Quality-adjusted Quantity (RECPI)	12.14	18.92	7.17	15.31	1.64	3.90	0.01	0.26	2.39	9.77
Equity Growth (IPO or M&A)	6.77	14.27	4.00	11.36	0.92	3.79	0.00	0.27	1.33	6.83
Number of regions	13		22		96		33,221		66	
Observations	195		330		1,440		272,194		990	

Panel B: UK (N = 4,913,041)

	NUTS 1		NUTS 2		NUTS 3		LAU		FUA	
Variable	Mean	St. Dev.	Mean	St. Dev.	Mean	St. Dev.	Mean	St. Dev.	Mean	St. Dev.
Quantity	37,220.01	34,175.17	10,893.66	8,874.97	2,495.20	2,561.73	1,116.60	1,544.42	9,502.98	27,404.80
Average Quality (EQI) x 1,000	1.14	0.12	1.13	0.18	1.12	0.28	1.14	0.35	1.14	0.19
Quality-adjusted Quantity (RECPI)	44.63	47.00	13.06	13.32	2.99	4.05	1.34	2.330	11.40	35.03
Equity Growth (IPO or M&A)	34.22	40.61	10.02	16.99	2.29	5.79	1.03	3.39	8.74	27.71
Number of regions	12		41		179		400		47	
Observations	132		451		1,969		4,400		517	

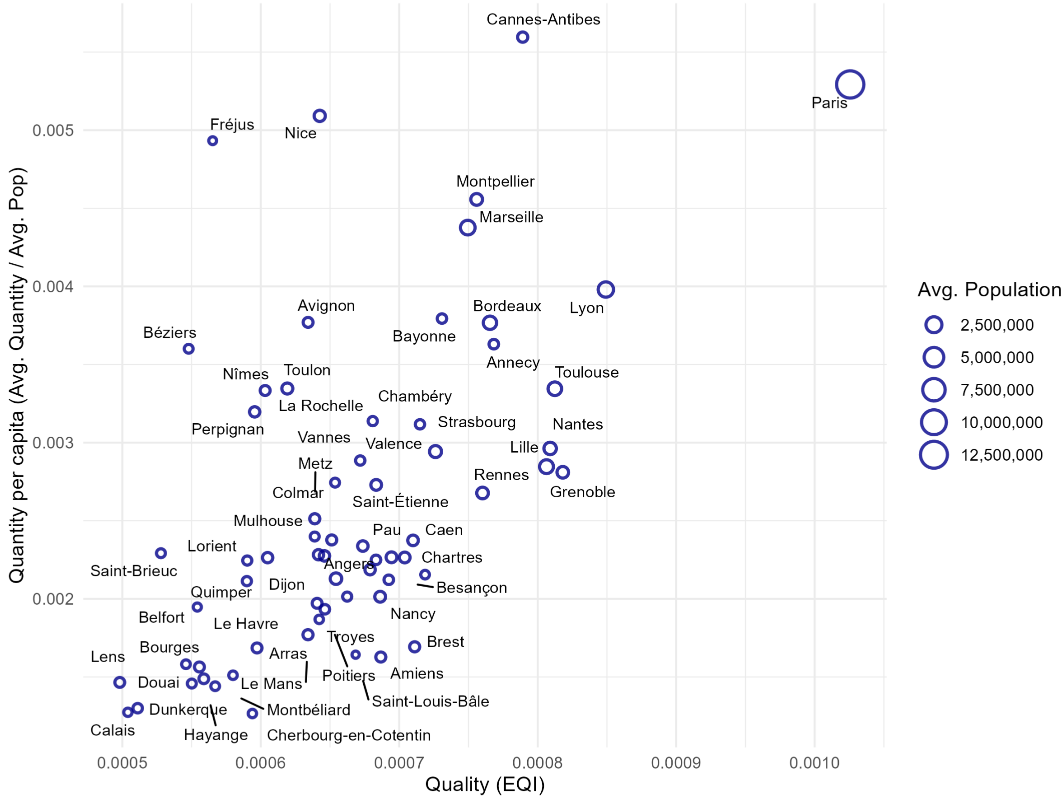
Panel C: Germany (N = 1,537,544)

	NUTS 1		NUTS 2		NUTS 3		LAU		FUA	
Variable	Mean	St. Dev.	Mean	St. Dev.	Mean	St. Dev.	Mean	St. Dev.	Mean	St. Dev.
Quantity	6,406.43	6,245.31	2,697.45	2,377.88	256.26	630.47	15.61	159.56	1,035.38	2,802.68
Average Quality (EQI) x 1,000	1.70	0.220	1.67	0.25	1.59	0.43	1.49	2.11	1.69	0.38
Quality-adjusted Quantity (RECPI)	11.49	11.67	4.84	5.09	0.46	1.36	0.03	0.34	1.86	4.81
Equity Growth (IPO or M&A)	8.58	11.34	3.61	6.33	0.34	1.86	0.02	0.46	1.39	4.48
Number of regions	16		38		400		10,190		99	
Observations	240		570		6,000		98,469		1,485	

Notes: The table reports summary statistics for France (Panel A), the United Kingdom (Panel B), and Germany (Panel C) at different levels of regional aggregation: Nomenclature of Territorial Units for Statistics (NUTS) levels 1-3, Local Administrative Units (LAU), and Functional Urban Areas (FUA). The unit of observation is a limited liability company incorporated in a given region-year between 2009 and 2023 for France and Germany, and between 2010 and 2020 for the United Kingdom Firm observations without valid address information are excluded. For each regional level, we report the mean and standard deviation of four key variables: Quantity, Quality (EQI, rescaled by 1,000), Quality-adjusted Quantity (RECPI), and Equity Growth events (share of firms undergoing an IPO or acquisition within 6 years). The bottom rows indicate the number of regions and total region-year observations at each level of aggregation.

Figure 5: Entrepreneurial Quality and Firm Quantity per Capita across FUAs

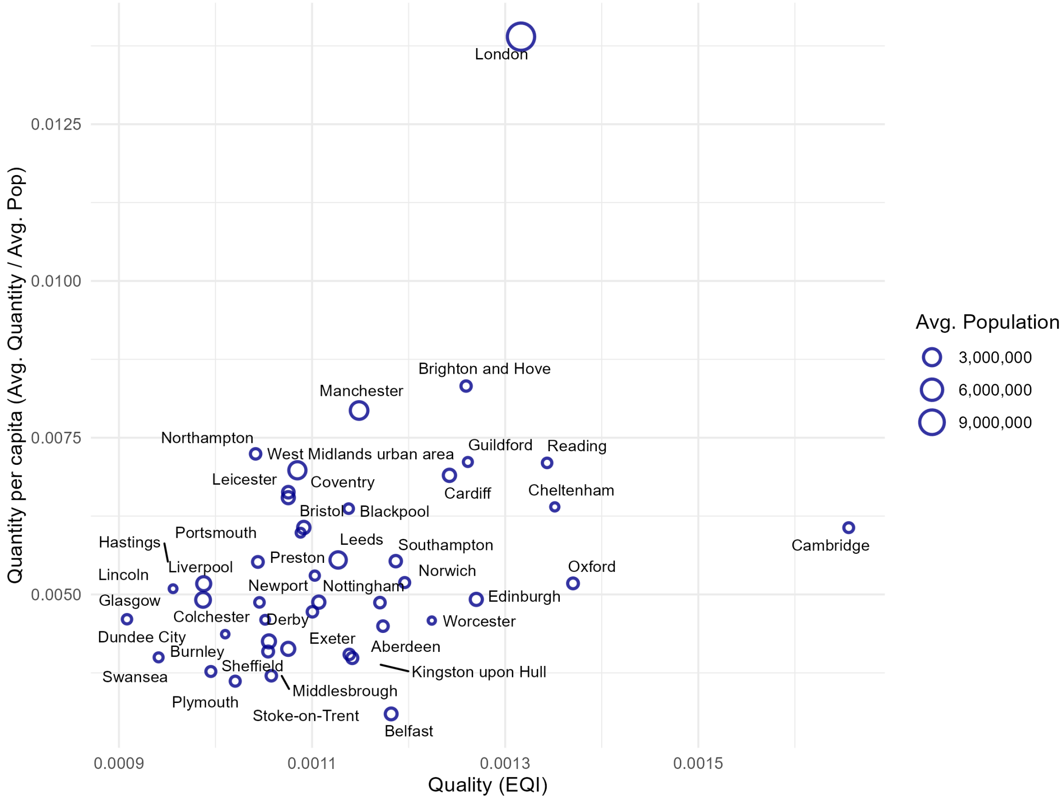
Panel A: France



Notes: This figure shows the relationship between entrepreneurial quality and firm quantity per capita across functional urban areas (FUAs) for France. The x-axis represents the quality index (EQI) and the y-axis represents the average firm quantity per capita, computed as the average number of firms in the FUA divided by the average population. The size of each marker corresponds to the average population of the FUA.

Figure 5: Entrepreneurial Quality and Firm Quantity per Capita across FUAs (*continued*)

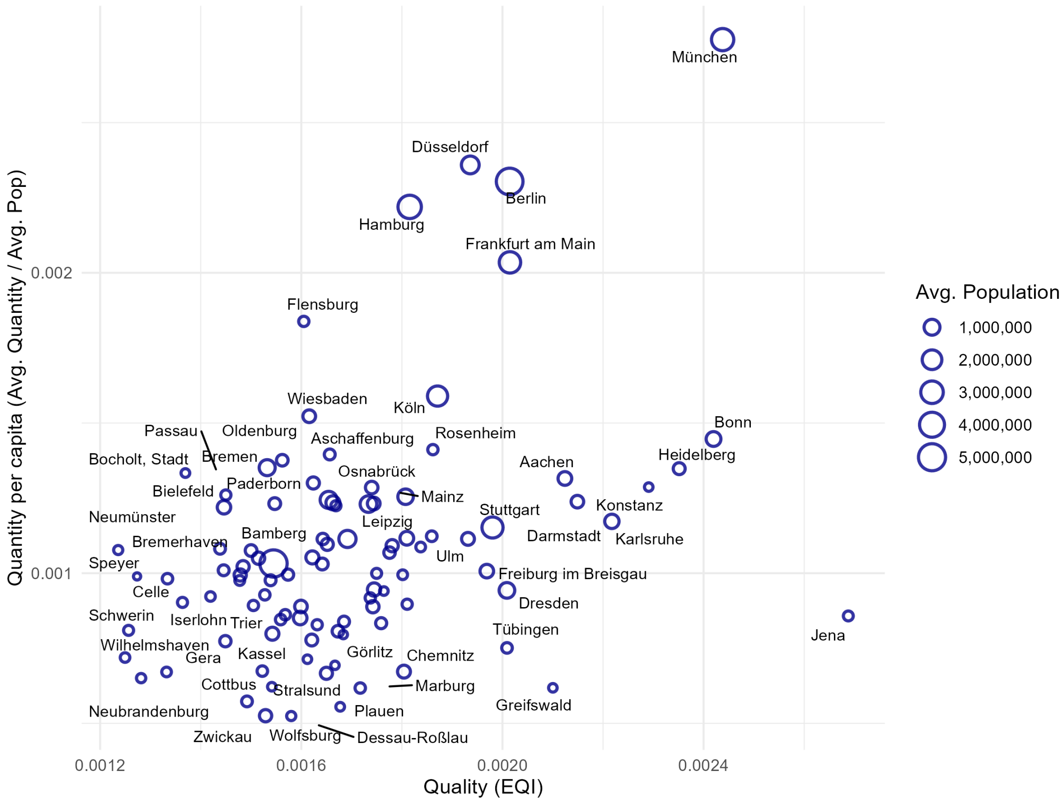
Panel B: UK



Notes: This figure shows the relationship between entrepreneurial quality and firm quantity per capita across functional urban areas (FUAs) for the UK. The x-axis represents the quality index (EQI) and the y-axis represents the average firm quantity per capita, computed as the average number of firms in the FUA divided by the average population. The size of each marker corresponds to the average population of the FUA. Bournemouth and Ipswich are excluded due to missing population data for the corresponding years.

Figure 5: Entrepreneurial Quality and Firm Quantity per Capita across FUAs (*continued*)

Panel C: Germany



Notes: This figure shows the relationship between entrepreneurial quality and firm quantity per capita across functional urban areas (FUAs) for Germany. The x-axis represents the quality index (EQI) and the y-axis represents the average firm quantity per capita, computed as the average number of firms in the FUA divided by the average population. The size of each marker corresponds to the average population of the FUA.

Table 11: Top 15 French FUAs by Entrepreneurial Quality (EQI), 2009–2012 vs. 2020–2023

	2009–2012			2020–2023		
	Top FUAs Avg. EQI	Average Quantity	Avg. EQI	Top FUAs Avg. EQI	Average Quantity	Avg. EQI
1	Paris	52,918.75	0.6998	Paris	88,877.50	1.1968
2	Annemasse-Geneva	615.25	0.6243	Lyon	11,506.50	0.9779
3	Lyon	6,930.00	0.6077	Lille	5,438.25	0.9766
4	Grenoble	1,780.50	0.5804	Toulouse	6,036.25	0.9494
5	Annecy	761.00	0.5795	Bordeaux	6,370.25	0.9422
6	Chambéry	605.25	0.5553	Cannes-Antibes	2,659.75	0.9353
7	Lille	3,595.75	0.5471	Nantes	3,790.25	0.9203
8	Brest	525.75	0.5373	Grenoble	2,353.00	0.9168
9	Nantes	2,275.25	0.5358	Montpellier	4,469.75	0.9100
10	Caen	870.25	0.5280	Rennes	2,609.25	0.9045
11	Toulouse	3,786.50	0.5224	Marseille	10,437.00	0.8862
12	Limoges	576.00	0.5178	Chartres	453.50	0.8825
13	Rennes	1,566.25	0.5167	Bayonne	1,390.25	0.8716
14	Besançon	511.00	0.5064	Strasbourg	2,947.50	0.8644
15	Montpellier	2,731.00	0.4885	Pau	764.25	0.8576

Notes: This table represents the ranking of the top 15 French Functional Urban Areas (FUAs) by average entrepreneurial quality (EQI) in the periods 2009–2012 and 2020–2023. The EQI is computed as the ratio of the total number of quality-adjusted firms (RECPI) to the total number of new limited liability firms in a given FUA-year and is rescaled by 1,000 for ease of interpretation. For each period, we report the FUAs with the highest EQI values, alongside their average number of new limited liability companies registered (Average Quantity).

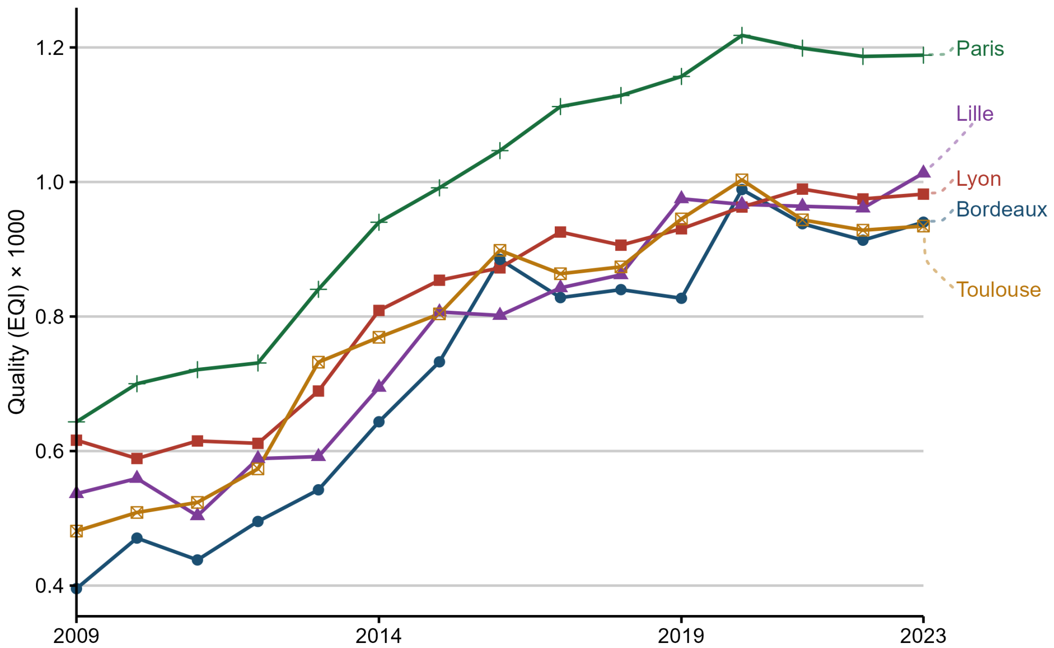
Table 12: Top 15 British FUAs by Entrepreneurial Quality (EQI), 2010–2013 vs. 2017–2020

	2010–2013			2017–2020		
	Top FUAs Avg. EQI	Average Quantity	Avg. EQI	Top FUAs Avg. EQI	Average Quantity	Avg. EQI
1	Cambridge	2,017.75	1.6047	Cambridge	2,150.25	1.8361
2	Reading	2,157.75	1.3987	Cheltenham	1,192.25	1.5209
3	Cheltenham	1,332.00	1.3661	London	159,704.75	1.4981
4	Brighton and Hove	3,682.25	1.2910	Oxford	2,719.00	1.4870
5	Norwich	1,749.50	1.2317	Edinburgh	4,191.00	1.4005
6	Oxford	2,591.75	1.2265	Cardiff	7,224.00	1.3955
7	London	152,110.00	1.2208	Reading	2,150.00	1.3906
8	Southampton	3,506.00	1.2196	Bournemouth	4,141.25	1.3799
9	Ipswich	1,426.00	1.1887	Guildford	1,635.50	1.3548
10	Guildford	2,007.00	1.1829	Manchester	26,800.00	1.3155
11	Worcester	785.50	1.1811	Kingston upon Hull	2,287.50	1.3154
12	Edinburgh	4,123.50	1.1579	Belfast	2,425.75	1.3127
13	Bournemouth	3,468.25	1.1517	Cheshire West and Chester	2,378.75	1.3032
14	Aberdeen	2,350.25	1.1215	Brighton and Hove	3,167.00	1.2795
15	Exeter	1,793.00	1.1182	Southampton	3,499.50	1.2745

Notes: This table represents the ranking of the top 15 British Functional Urban Areas (FUAs) by average entrepreneurial quality (EQI) in the periods 2010–2013 and 2017–2020. The EQI is computed as the ratio of the total number of quality-adjusted firms (RECPI) to the total number of new limited liability firms in a given FUA-year and is rescaled by 1,000 for ease of interpretation. For each period, we report the FUAs with the highest EQI values, alongside their average number of new limited liability companies registered (Average Quantity).

Figure 6: Evolution of Entrepreneurial Quality for selected Functional Urban Areas (FUA)

Panel A: France (2009-2023)



Notes: The figure plots the evolution of the Entrepreneurial Quality Index (EQI) for selected French Functional Urban Areas (FUAs) between 2009 and 2023. The EQI is computed as the ratio of the total number of quality-adjusted firms (RECPI) to the total number of newly established limited liability companies in each FUA-year.

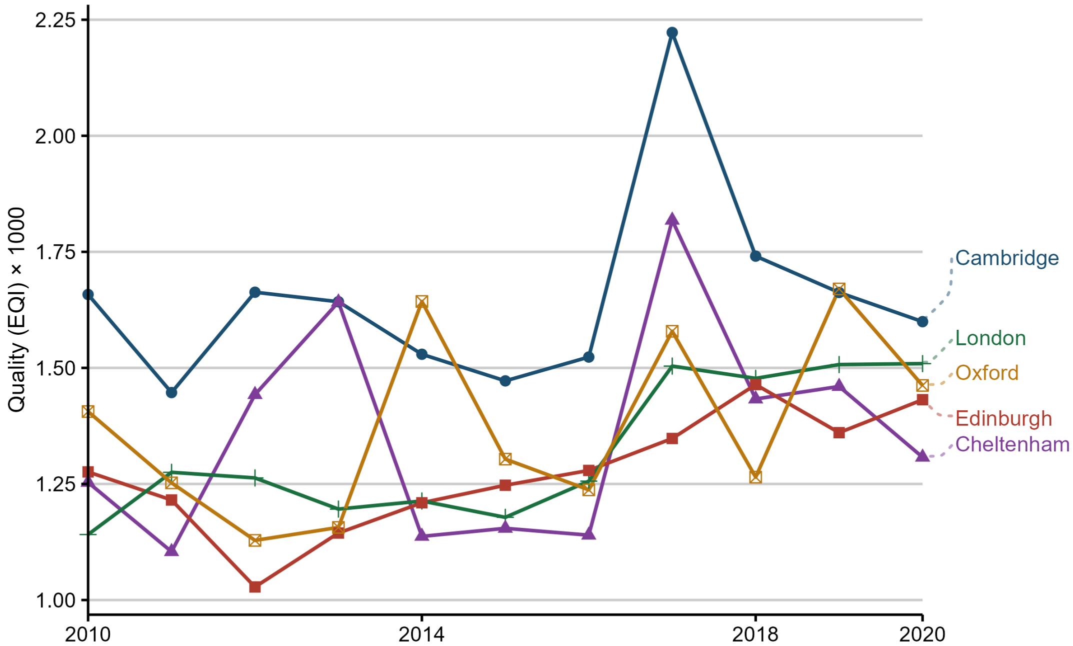
Table 13: Top 15 German FUAs by Entrepreneurial Quality (EQI), 2009–2012 vs. 2020–2023

	2009–2012			2020–2023		
	Top FUAs Avg. EQI	Average Quantity	Avg. EQI	Top FUAs Avg. EQI	Average Quantity	Avg. EQI
1	Greifswald	68.75	2.8182	Bonn	1,244.25	2.3818
2	Jena	153.25	2.5391	München	8,358.50	2.2988
3	München	7,710.50	2.4738	Heidelberg	510.25	2.1645
4	Heidelberg	428.00	2.4600	Jena	127.50	2.1599
5	Bonn	1,006.25	2.3738	Dresden	900.75	2.0493
6	Konstanz	123.00	2.3199	Karlsruhe	936.75	2.0300
7	Karlsruhe	772.75	2.3100	Aachen	889.50	2.0127
8	Frankfurt am Main	5,135.25	2.1604	Greifswald	64.75	1.9649
9	Ulm	516.75	2.1465	Berlin	12,533.00	1.9640
10	Darmstadt	542.00	2.1172	Darmstadt	550.00	1.9453
11	Aachen	690.00	2.1011	Konstanz	140.50	1.9387
12	Rosenheim	273.00	2.0495	Frankfurt am Main	5,456.00	1.9162
13	Düsseldorf	3,188.75	2.0434	Düsseldorf	3,647.00	1.9093
14	Tübingen	146.75	2.0374	Görlitz	50.50	1.9017
15	Regensburg	547.50	2.0364	Chemnitz	315.50	1.8808

Notes: This table represents the ranking of the top 15 German Functional Urban Areas (FUAs) by average entrepreneurial quality (EQI) in the periods 2009–2012 and 2020–2023. The EQI is computed as the ratio of the total number of quality-adjusted firms (RECPI) to the total number of new limited liability firms in a given FUA-year and is rescaled by 1,000 for ease of interpretation. For each period, we report the FUAs with the highest EQI values, alongside their average number of new limited liability companies registered (Average Quantity).

Figure 6: Evolution of Entrepreneurial Quality for selected Functional Urban Areas (FUA) (*continued*)

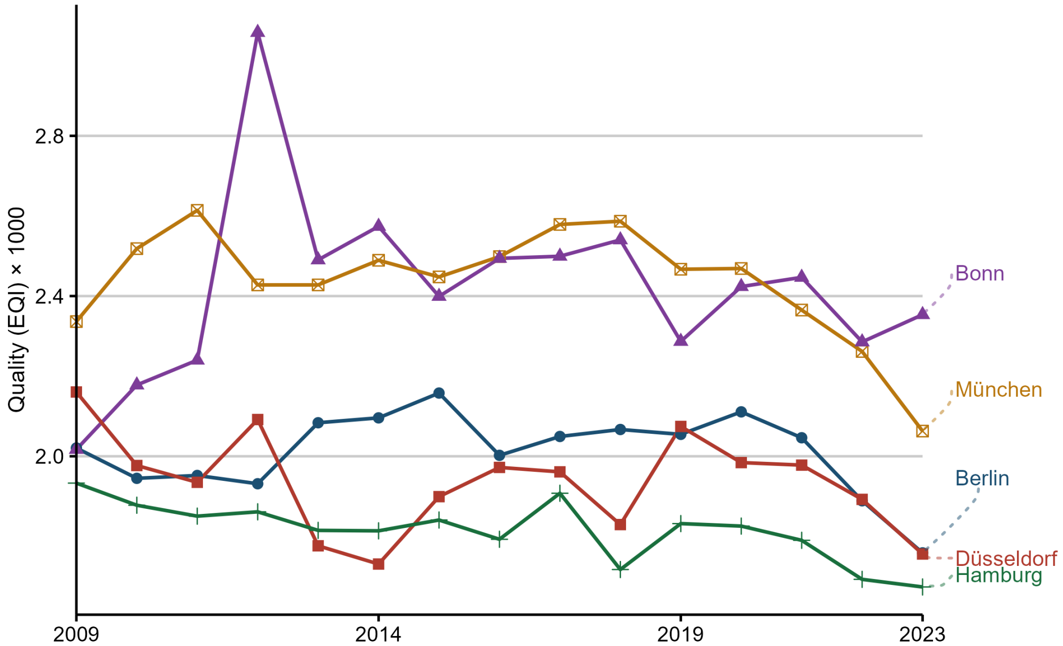
Panel B: UK (2010-2020)



Notes: The figure plots the evolution of the Entrepreneurial Quality Index (EQI) for selected British Functional Urban Areas (FUAs) between 2010 and 2020. The EQI is computed as the ratio of the total number of quality-adjusted firms (RECPI) to the total number of newly established limited liability companies in each FUA-year.

Figure 6: Evolution of Entrepreneurial Quality for selected Functional Urban Areas (FUA) (*continued*)

Panel C: Germany (2009-2023)



Notes: The figure plots the evolution of the Entrepreneurial Quality Index (EQI) for selected German Functional Urban Areas (FUAs) between 2009 and 2023. The EQI is computed as the ratio of the total number of quality-adjusted firms (RECPI) to the total number of newly established limited liability companies in each FUA-year.

Supplementary Material to: "The State of European Entrepreneurship: Trends in Quantity and Quality in France, Germany, and the UK (2009–2023)"

Massimo G. Colombo[1], Lena Füner[2,3,4], Massimiliano Guerini[1], Hanna Hottenrott[2,3], and Daniel Souza[1]

[1]Polytechnic University of Milan
[2]Leibniz Centre for European Economic Research (ZEW)
[3]Technical University Munich
[4]IAB Institute for Employment Research

Appendix A Additional Data Descriptives

In this section, we provide supplementary descriptive statistics related to our datasets. Table A1 reports summary statistics for our industry classification aggregated at the two-digit NACE level. Table A2 presents corresponding statistics based on the OECD High-Technology Industry and Knowledge-Intensive Services classification.

Figure A1 illustrates the ratio of firm births to GDP (left axis) alongside the five-year log change in GDP relative to contemporaneous GDP, $\ln(GDP_{t+5}/GDP_t)$ (right axis), for France (Panel A), the United Kingdom (Panel B), and Germany (Panel C). These series highlight the relationship between firm entry intensity and subsequent macroeconomic growth dynamics.

Table A1: Summary Statistics - Aggregated NACE 2-digit codes

	France		UK		Germany	
Variable	Mean	St. Dev.	Mean	St. Dev.	Mean	St. Dev.
Industry Measures						
Aerospace	0.0001	0.0088	0.0001	0.0121	0.0001	0.0102
Agriculture, forestry and fishing	0.0001	0.0105	0.0016	0.0402	-	-
Biotechnology	0.0004	0.0203	0.0006	0.0254	0.0007	0.0266
Chemicals	0.0010	0.0320	0.0009	0.0303	0.0027	0.0523
Communications and Media	0.0380	0.1912	0.0483	0.2144	0.0293	0.1686
Computers and electronic components	0.0057	0.0752	0.0035	0.0593	0.0070	0.0835
Construction	0.1373	0.3441	0.0827	0.2754	0.0616	0.2405
Consultancy	0.0991	0.2988	0.0703	0.2557	0.1121	0.3155
Cultural and creative industries	0.0240	0.1530	0.0282	0.1655	0.0117	0.1074
Energy & environment	0.0120	0.1089	0.0062	0.0787	0.0239	0.1528
Financial & Insurance	0.0707	0.2563	0.0271	0.1625	0.1019	0.3025
Machinery and equipment	0.0060	0.0771	0.0036	0.0600	0.0151	0.1220
Medical/Health/Life Science	0.0292	0.1683	0.0728	0.2598	0.0510	0.2200
Motor vehicles	0.0344	0.1822	0.0163	0.1264	0.0114	0.1061
Non high-tech manufacturing	0.0303	0.1715	0.0221	0.1469	0.0239	0.1528
Other non high-tech services	0.1261	0.3319	0.0658	0.2480	0.0410	0.1983
R&D and Engineering	0.0352	0.1842	0.0329	0.1784	0.0374	0.1898
Real estate	0.0540	0.2260	0.0550	0.2280	0.0858	0.2800
Software	0.0367	0.1881	0.0523	0.2226	0.0440	0.2051
Sport	0.0093	0.0961	0.0101	0.0998	0.0098	0.0987
Support services	0.0583	0.2343	0.0909	0.2875	0.1284	0.3346
Trasportation	0.0365	0.1876	0.0311	0.1735	0.0256	0.1578
Wholesale and retail trade	0.1151	0.3191	0.0574	0.2325	0.0782	0.2685
Other	-	-	0.0164	0.1269	0.0544	0.2268
Missing	0.0405	0.1972	0.2037	0.4028	0.0428	0.2024

Notes: This table reports summary statistics from the national business registry data of France, the United Kingdom, and Germany. The unit of observation is a limited liability company incorporated between 2009 and 2023 for France and Germany, and between 2010 and 2020 for the United Kingdom. Industry measures are based on an alternative industry classification aggregated from NACE 2-digit codes.

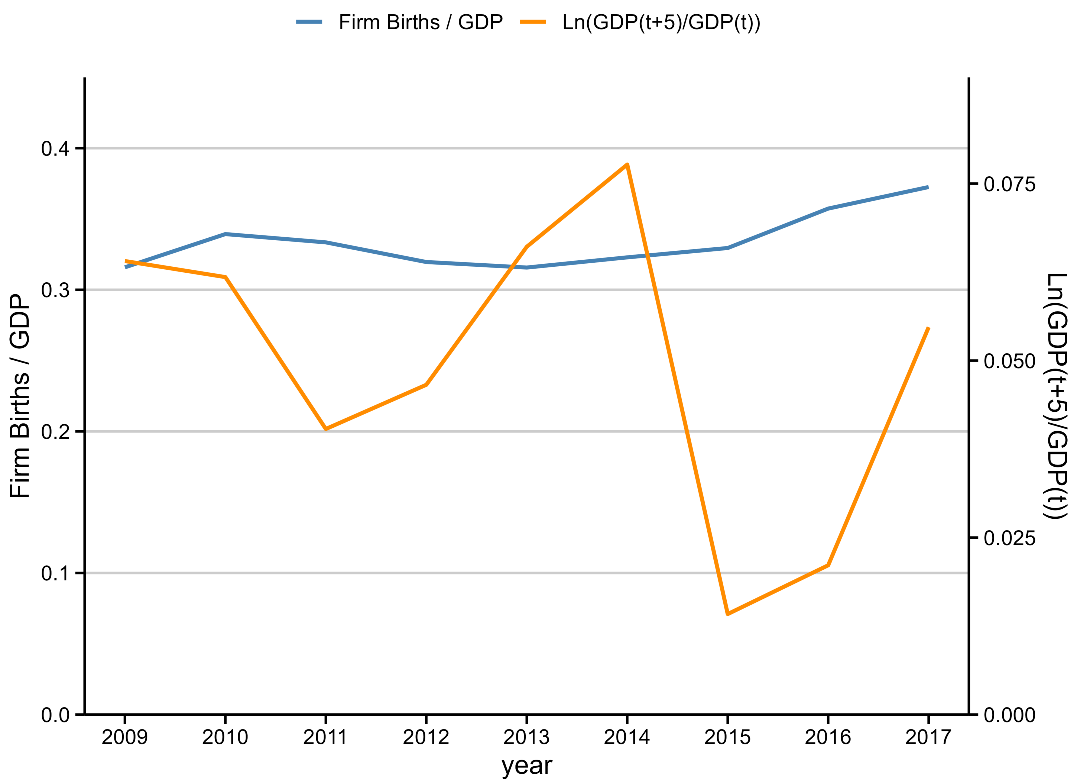
Table A2: Summary Statistics - OECD High-Tech and Knowledge-Intensive Industry Classification

	France		UK		Germany	
Variable	Mean	St. Dev.	Mean	St. Dev.	Mean	St. Dev.
Industry Measures						
High Technology	0.0007	0.0259	0.0013	0.0364	0.0037	0.0603
Medium-high-technology	0.0027	0.0519	0.0041	0.0640	0.0127	0.1120
Medium-low-technology	0.0082	0.0904	0.0106	0.1025	0.0149	0.1213
Low technology	0.0249	0.1558	0.0185	0.1347	0.0198	0.1392
High tech knowledge-intensive services	0.0511	0.2202	0.0781	0.2683	0.0616	0.2404
Knowledge-intensive financial services	0.0707	0.2563	0.0272	0.1627	0.1028	0.3037
Knowledge-intensive market services	0.1495	0.3566	0.1254	0.3312	0.1629	0.3693
Other knowledge-intensive services	0.0461	0.2098	0.0815	0.2737	0.0455	0.2084
Less knowledge-intensive market services	0.4236	0.4941	0.3071	0.4613	0.3987	0.4896
Other less knowledge-intensive services	0.0330	0.1785	0.0490	0.2158	0.0420	0.2006
Non-tech, not knowledge-intense	0.1490	0.3561	0.0934	0.2910	0.0926	0.2899
Missing	0.0405	0.1972	0.2037	0.4028	0.0428	0.2024

Notes: This table reports summary statistics from the national business registry data of France, the United Kingdom, and Germany. The unit of observation is a limited liability company incorporated between 2009 and 2023 for France and Germany, and between 2010 and 2020 for the United Kingdom. Industry measures are based on the OECD High-Tech and Knowledge-Intensive Industry Classification.

Figure A1: Firm Births versus GDP Growth Over Next 5 Years

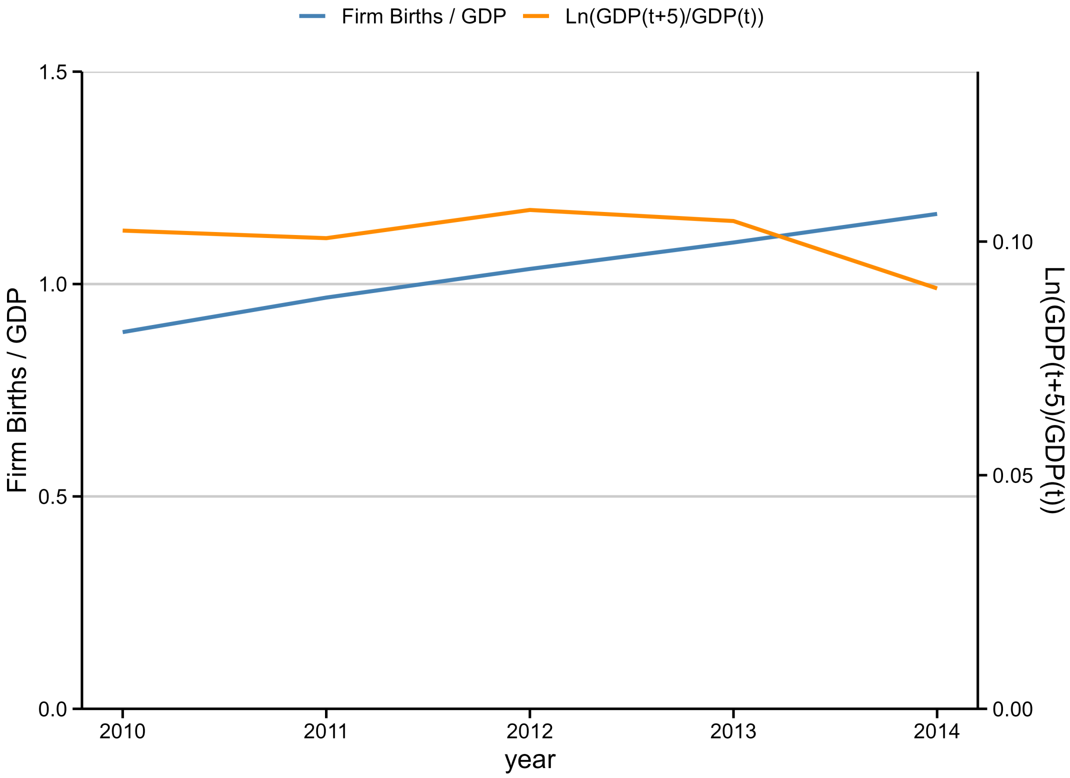
Panel A: France



Notes: This figure plots the ratio of firm births to GDP (left axis) and the log change in GDP five years ahead relative to contemporaneous GDP, $\ln(GDP_{t+5}/GDP_t)$ (right axis), for France over the period 2009–2017. The series illustrate the relationship between entry intensity and subsequent macroeconomic growth dynamics.

Figure A1: Firm Births versus GDP Growth Over Next 5 Years (*continued*)

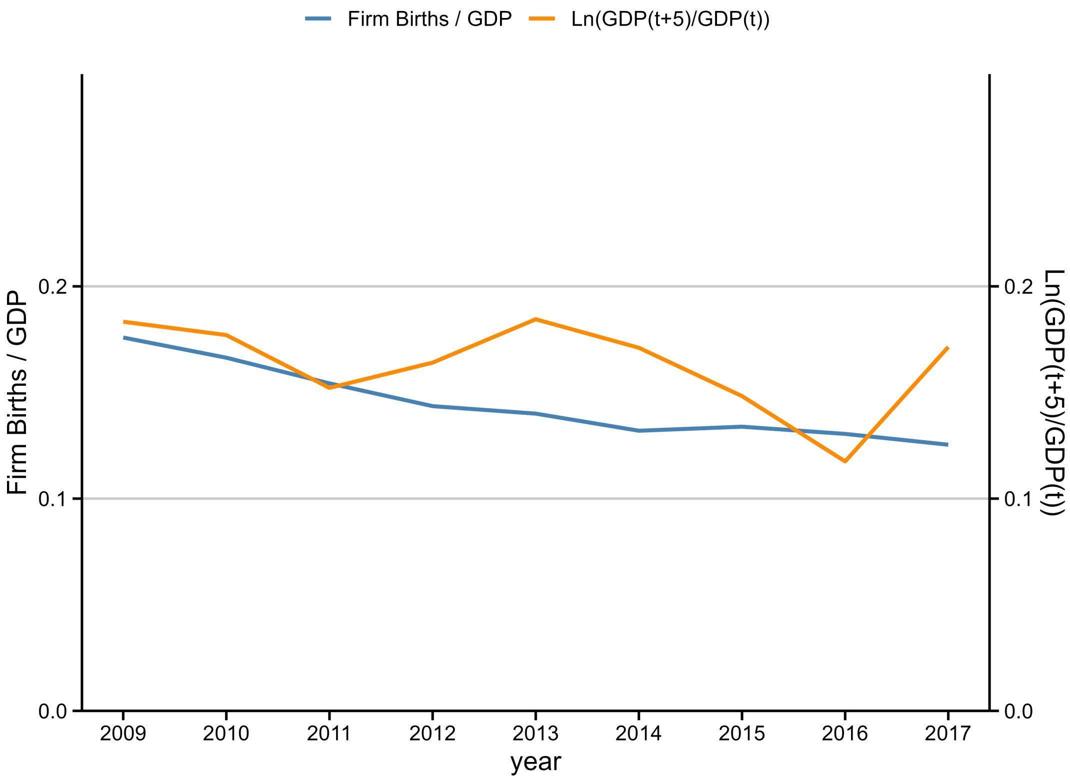
Panel B: UK



Notes: This figure plots the ratio of firm births to GDP (left axis) and the log change in GDP five years ahead relative to contemporaneous GDP, $\ln(GDP_{t+5}/GDP_t)$ (right axis), for the United Kingdom over the period 2010–2014.

Figure A1: Firm Births versus GDP Growth Over Next 5 Years (*continued*)

Panel C: Germany



Notes: This figure plots the ratio of firm births to GDP (left axis) and the log change in GDP five years ahead relative to contemporaneous GDP, $\ln(GDP_{t+5}/GDP_t)$ (right axis), for Germany over the period 2009–2017. The series illustrate the relationship between entry intensity and subsequent macroeconomic growth dynamics.

Appendix B Additional Results

In this section, we provide additional results stemming from our predictive model. In Table B1 we present the coefficients of a series of univariate regressions of our main Growth variable (IPO or M&A within 6 years of fouding) on each industry based on NACE 2 code. Table B2 and Table B3 do the same, but for our alternative aggregation of NACE 2-digit code and the OECD High-Technology Industry and Knowledge-Intensive Services classification.

Table B1: Logit Univariate Regressions - NACE Code 2-digits

	France		UK		Germany	
2-digit NACE	Coefficient	Pseudo R^2	Coefficient	Pseudo R^2	Coefficient	Pseudo R^2
Industry measures						
05	-	-	26.019***	0.000	-	-
			(18.674)			
06	21.666***	0.000	13.590***	0.002	-	-
	(21.833)		(3.333)			
07	0.005***	0.000	30.805***	0.002	-	-
	(0.002)		(9.068)			
08	13.590***	0.000	24.430***	0.002	-	-
	(13.658)		(6.413)			
09	0.002***	0.000	0.740	0.000	-	-
	(0.000)		(0.524)			
10	0.252***	0.001	4.601***	0.001	1.877**	0.000
	(0.126)		(0.862)		(0.472)	
11	4.523***	0.000	2.549**	0.000	0.539	0.000
	(2.619)		(1.143)		(0.539)	
12	-	-	0.001***	0.000	0.001***	0.000
			(0.000)		(0.000)	
13	1.813	0.000	1.555	0.000	3.595***	0.000
	(1.815)		(0.779)		(1.205)	
14	0.781	0.000	0.402	0.000	1.639	0.000
	(0.553)		(0.284)		(0.949)	
15	4.881**	0.000	0.000***	0.000	2.318	0.000
	(3.460)		(0.000)		(2.324)	
16	2.040	0.000	0.535	0.000	0.000***	0.000
	(1.180)		(0.535)		(0.000)	
17	12.788***	0.000	13.619***	0.001	3.986**	0.000
	(9.090)		(4.147)		(2.311)	
18	0.450	0.000	1.898*	0.000	1.017	0.000
	(0.450)		(0.673)		(0.456)	
19	-	-	7.895**	0.000	0.000***	0.000
			(7.932)		(0.000)	
20	17.378***	0.002	7.557***	0.001	3.816***	0.001
	(5.298)		(2.405)		(1.028)	
21	16.075***	0.000	7.893***	0.000	6.199***	0.001
	(16.169)		(3.549)		(2.361)	
22	3.365*	0.000	6.289***	0.001	2.662***	0.000
	(2.384)		(1.907)		(0.946)	
23	0.000***	0.000	5.054***	0.000	1.446	0.000
	(0.000)		(2.269)		(0.725)	
24	15.191***	0.000	6.545***	0.000	3.808***	0.000
	(10.806)		(3.287)		(1.447)	
25	1.401	0.000	2.919***	0.000	1.769***	0.000
	(0.628)		(0.712)		(0.380)	
26	7.425***	0.001	7.664***	0.001	3.804***	0.001
	(3.337)		(1.931)		(0.819)	
27	6.398***	0.000	0.522	0.000	3.680***	0.001
	(3.707)		(0.522)		(0.900)	
28	4.971***	0.000	5.974***	0.001	3.801***	0.002
	(2.232)		(1.553)		(0.570)	
29	5.765**	0.000	1.434	0.000	1.399	0.000
	(4.088)		(1.015)		(0.991)	
30	13.739***	0.001	2.733**	0.000	3.141**	0.000
	(6.187)		(1.370)		(1.411)	
31	0.000***	0.000	0.560	0.000	1.752	0.000
	(0.000)		(0.396)		(1.014)	
32	1.014	0.000	2.724***	0.001	1.850**	0.000
	(0.586)		(0.501)		(0.516)	

	France		UK		Germany	
2-digit NACE	Coefficient	Pseudo R^2	Coefficient	Pseudo R^2	Coefficient	Pseudo R^2
33	0.963	0.000	1.212	0.000	1.100	0.000
	(0.432)		(0.287)		(0.390)	
35	0.865	0.000	4.501***	0.002	1.302*	0.000
	(0.275)		(0.672)		(0.180)	
36	0.000***	0.000	4.783***	0.000	0.000***	0.000
	(0.000)		(2.770)		(0.000)	
37	0.000***	0.000	0.000***	0.000	0.000***	0.000
	(0.000)		(0.000)		(0.000)	
38	0.755	0.000	2.661***	0.000	1.148	0.000
	(0.755)		(0.771)		(0.576)	
39	0.000***	0.000	0.724	0.000	1.228	0.000
	(0.000)		(0.725)		(1.229)	
41	0.255***	0.001	0.242***	0.001	0.077***	0.002
	(0.104)		(0.067)		(0.054)	
42	0.825	0.000	0.286**	0.000	0.734	0.000
	(0.826)		(0.165)		(0.425)	
43	0.124***	0.010	0.337***	0.002	0.098***	0.004
	(0.029)		(0.058)		(0.035)	
45	0.054***	0.003	0.505***	0.000	0.442***	0.000
	(0.039)		(0.131)		(0.123)	
46	1.002	0.000	1.317**	0.000	0.774**	0.000
	(0.135)		(0.170)		(0.084)	
47	0.474***	0.002	0.528***	0.001	0.742**	0.000
	(0.065)		(0.070)		(0.089)	
49	0.063***	0.003	0.226***	0.001	0.418**	0.000
	(0.045)		(0.080)		(0.171)	
50	0.000***	0.000	0.000***	0.000	0.615	0.000
	(0.000)		(0.000)		(0.435)	
51	0.000***	0.000	2.604*	0.000	10.366***	0.000
	(0.000)		(1.305)		(5.237)	
52	1.786	0.000	1.407	0.000	0.949	0.000
	(0.732)		(0.446)		(0.200)	
53	1.931	0.000	0.431	0.000	0.316	0.000
	(1.933)		(0.305)		(0.316)	
55	0.635	0.000	2.697***	0.000	0.656	0.000
	(0.260)		(0.543)		(0.268)	
56	0.099***	0.008	0.310***	0.001	0.154***	0.002
	(0.030)		(0.061)		(0.058)	
58	9.365***	0.008	2.447***	0.001	1.658*	0.000
	(1.349)		(0.436)		(0.502)	
59	1.309	0.000	1.036	0.000	0.693	0.000
	(0.380)		(0.194)		(0.311)	
60	19.201***	0.001	1.568	0.000	3.114	0.000
	(11.173)		(0.907)		(2.210)	
61	5.425***	0.001	3.231***	0.001	3.100***	0.001
	(1.578)		(0.576)		(0.834)	
62	7.325***	0.024	2.322***	0.004	4.717***	0.016
	(0.584)		(0.150)		(0.310)	
63	11.829***	0.009	2.326***	0.001	3.082***	0.001
	(1.674)		(0.347)		(0.543)	
64	2.387***	0.004	7.778***	0.014	1.307***	0.000
	(0.239)		(0.538)		(0.118)	
65	0.000***	0.000	5.284***	0.001	0.571	0.000
	(0.000)		(1.421)		(0.571)	
66	1.822***	0.000	3.009***	0.001	0.792	0.000
	(0.384)		(0.506)		(0.146)	
68	0.484***	0.001	0.775*	0.000	0.297***	0.004
	(0.104)		(0.104)		(0.047)	
69	0.116**	0.001	1.757***	0.000	0.092**	0.001
	(0.116)		(0.226)		(0.093)	
70	1.322***	0.000	2.007***	0.002	0.459***	0.003
	(0.142)		(0.142)		(0.052)	
71	1.549***	0.000	0.695*	0.000	1.163	0.000
	(0.244)		(0.146)		(0.182)	
72	18.928***	0.007	9.685***	0.003	4.470***	0.002
	(3.528)		(1.552)		(0.778)	
73	3.094***	0.001	2.867***	0.001	1.703***	0.000
	(0.604)		(0.428)		(0.320)	
74	1.103	0.000	0.845	0.000	1.173	0.000
	(0.243)		(0.116)		(0.220)	

	France		UK		Germany	
2-digit NACE	Coefficient	Pseudo R^2	Coefficient	Pseudo R^2	Coefficient	Pseudo R^2
75	0.000***	0.000	3.775***	0.000	12.283***	0.001
	(0.000)		(1.340)		(5.561)	
77	0.721	0.000	0.845	0.000	0.499	0.000
	(0.274)		(0.346)		(0.224)	
78	0.395	0.000	1.494**	0.000	0.231**	0.000
	(0.395)		(0.279)		(0.133)	
79	3.782***	0.001	1.002	0.000	1.537	0.000
	(1.268)		(0.355)		(0.466)	
80	0.000***	0.001	0.648	0.000	0.198	0.000
	(0.000)		(0.245)		(0.198)	
81	0.214***	0.001	0.290***	0.000	0.170***	0.001
	(0.107)		(0.130)		(0.085)	
82	1.426*	0.000	1.450***	0.001	1.979***	0.003
	(0.295)		(0.110)		(0.142)	
84	0.005***	0.000	0.847	0.000	1.247	0.000
	(0.002)		(0.600)		(1.249)	
85	0.517	0.000	0.481***	0.000	0.192***	0.001
	(0.212)		(0.129)		(0.111)	
86	0.358*	0.000	0.386***	0.001	1.734***	0.000
	(0.207)		(0.073)		(0.365)	
87	0.000***	0.000	1.546	0.000	1.283	0.000
	(0.000)		(0.548)		(0.575)	
88	0.805	0.000	0.510*	0.000	0.162**	0.001
	(0.466)		(0.208)		(0.115)	
89	-	-	-	-	0.000***	0.000
					(0.000)	
90	0.570	0.000	0.542**	0.000	0.000***	0.000
	(0.404)		(0.164)		(0.000)	
91	0.002***	0.000	0.000***	0.000	1.526	0.000
	(0.000)		(0.000)		(1.529)	
92	11.325**	0.000	5.292***	0.000	0.697	0.000
	(11.373)		(2.656)		(0.349)	
93	0.303**	0.000	0.938	0.000	0.372**	0.000
	(0.175)		(0.211)		(0.167)	
94	0.001***	0.000	2.577**	0.000	0.000***	0.000
	(0.000)		(1.156)		(0.000)	
95	1.109	0.000	0.588	0.000	0.000***	0.000
	(0.785)		(0.416)		(0.000)	
96	0.114***	0.003	0.434***	0.001	1.028	0.000
	(0.057)		(0.072)		(0.127)	
99	-	-	-	-	0.000***	0.000
					(0.000)	
Missing	0.455***	0.001	0.235***	0.018	0.009***	0.008
	(0.103)		(0.016)		(0.009)	

Notes: This table reports results from univariate logit regressions estimating the probability that a firm experiences growth, defined as achieving an IPO or M&A within six years of incorporation. Industry indicators are based on NACE 2-digit code. Reported coefficients are incidence rate ratios, with robust standard errors in parentheses. Significance levels: *p<0.1; **p<0.05; ***p<0.01.

Table B2: Logit Univariate Regressions - Aggregate NACE 2-digit Codes

	France		UK		Germany	
Variable	Coefficient	Pseudo R^2	Coefficient	Pseudo R^2	Coefficient	Pseudo R^2
Industry measures						
Aerospace	35.947***	0.001	2.843	0.000	0.000***	0.000
	(21.032)		(2.848)		(0.000)	
Agriculture, forestry and fishing	13.854***	0.000	11.061***	0.004	-	0.000
	(9.851)		(1.639)			
Biotechnology	26.306***	0.003	19.415***	0.003	5.236***	0.000
	(7.704)		(3.869)		(2.151)	
Chemicals	12.249***	0.002	4.820***	0.000	2.409***	0.000
	(3.571)		(1.712)		(0.730)	
Communications and Media	3.664***	0.007	1.586***	0.001	1.906***	0.001
	(0.360)		(0.132)		(0.207)	
Computers and electronic components	2.015***	0.000	2.837***	0.000	2.387***	0.001
	(0.543)		(0.623)		(0.427)	
Construction	0.141***	0.011	0.302***	0.003	0.138***	0.005
	(0.029)		(0.043)		(0.037)	
Consultancy	1.399***	0.001	2.119***	0.004	0.545***	0.002
	(0.134)		(0.126)		(0.053)	
Cultural and creative industries	0.511**	0.000	0.644***	0.000	0.555*	0.000
	(0.148)		(0.103)		(0.176)	
Energy & environment	0.806	0.000	3.662***	0.002	1.230	0.000
	(0.245)		(0.480)		(0.167)	
Financial & Insurance	2.316***	0.004	6.641***	0.015	1.166*	0.000
	(0.214)		(0.422)		(0.096)	
Machinery and equipment	2.180***	0.000	2.604***	0.000	2.764***	0.002
	(0.550)		(0.546)		(0.332)	
Medical/Health/Life Science	0.618**	0.000	0.616***	0.001	1.386***	0.000
	(0.151)		(0.065)		(0.131)	
Motor vehicles	0.108***	0.003	0.547**	0.000	0.585*	0.000
	(0.054)		(0.133)		(0.170)	
Non high-tech manufacturing	0.992	0.000	2.387***	0.002	1.926***	0.001
	(0.173)		(0.232)		(0.225)	
Other	-	0.000	1.263	0.000	0.617***	0.001
	-		(0.207)		(0.079)	
Other non high-tech services	0.145***	0.010	0.541***	0.001	0.180***	0.003
	(0.031)		(0.067)		(0.052)	
R&D and Engineering	2.018***	0.002	0.907	0.000	1.430***	0.000
	(0.241)		(0.106)		(0.157)	
Real estate	0.484***	0.001	0.775*	0.000	0.294***	0.004
	(0.104)		(0.104)		(0.048)	
Software	8.731***	0.032	2.363***	0.004	4.764***	0.017
	(0.648)		(0.150)		(0.310)	
Sport	0.401*	0.000	1.089	0.000	0.371***	0.000
	(0.201)		(0.224)		(0.141)	
Support services	0.804	0.000	1.263***	0.000	1.516***	0.001
	(0.121)		(0.086)		(0.104)	
Trasportation	0.252***	0.002	0.442***	0.000	0.807	0.000
	(0.084)		(0.097)		(0.140)	
Wholesale and retail trade	0.409***	0.004	0.752***	0.000	0.592***	0.001
	(0.054)		(0.083)		(0.065)	
Missing	0.455***	0.001	0.235***	0.018	0.009***	0.008
	(0.103)		(0.016)		(0.009)	

Notes: This table reports results from univariate logit regressions estimating the probability that a firm experiences growth, defined as achieving an IPO or M&A within six years of incorporation. Industry indicators are based on aggregated NACE 2-digit codes. Reported coefficients are incidence rate ratios, with robust standard errors in parentheses. Significance levels: $^{*}p<0.1$; $^{**}p<0.05$; $^{***}p<0.01$.

Table B3: Logit Univariate Regressions - OECD High-Tech and Knowledge-Intensive Industry Classification

	France		UK		Germany	
Variable	Coefficient	Pseudo R^2	Coefficient	Pseudo R^2	Coefficient	Pseudo R^2
Industry measures						
High Technology	8.166***	0.001	7.739***	0.001	4.215***	0.002
	(3.353)		(1.703)		(0.793)	
Low technology	0.744	0.000	2.468***	0.001	1.635***	0.000
	(0.164)		(0.263)		(0.227)	
Medium-high-technology	9.391***	0.004	3.745***	0.001	3.677***	0.004
	(1.871)		(0.668)		(0.413)	
Medium-low-technology	1.416	0.000	2.344***	0.001	1.802***	0.001
	(0.381)		(0.317)		(0.262)	
High tech knowledge-intensive services	8.275***	0.039	2.491***	0.006	4.473***	0.020
	(0.565)		(0.134)		(0.265)	
Knowledge-intensive market services	1.338***	0.001	1.613***	0.002	0.648***	0.001
	(0.108)		(0.085)		(0.050)	
Knowledge-intensive financial services	2.316***	0.004	6.624***	0.015	1.162*	0.000
	(0.214)		(0.420)		(0.095)	
Less knowledge-intensive market services	0.337***	0.014	0.801***	0.001	0.719***	0.002
	(0.026)		(0.041)		(0.039)	
Other less knowledge-intensive services	0.140***	0.002	0.479***	0.001	0.943	0.000
	(0.063)		(0.073)		(0.115)	
Other knowledge-intensive services	1.609***	0.001	0.735***	0.000	0.790*	0.000
	(0.201)		(0.066)		(0.104)	
Non-tech, not knowledge-intense	0.197***	0.010	0.862*	0.000	0.440***	0.002
	(0.033)		(0.071)		(0.053)	
Missing	0.455***	0.001	0.235***	0.018	0.009***	0.008
	(0.103)		(0.016)		(0.009)	

Notes: This table reports results from univariate logit regressions estimating the probability that a firm experiences growth, defined as achieving an IPO or M&A within six years of incorporation. Industry indicators are based on OECD classification of high-tech industries and knowledge-intensive services. Reported coefficients are incidence rate ratios, with robust standard errors in parentheses. Significance levels: *p<0.1; **p<0.05; ***p<0.01.

We also present in Tables B4 - B6 the ranking of Functional Urban Areas in terms of RECPI / population and in Tables B7 - B9 the average REAI for France, UK and Germany respectively.

Table B4: Top 15 French FUAs by Quality-Adjusted Entrepreneurship, 2009–2012 vs. 2020–2023

	2009–2012			2020–2023		
	Top FUAs Avg. RECPI / Avg. Pop.	Average Quantity	Avg. RECPI / Avg. Pop.	Top FUAs Avg. RECPI / Avg. Pop.	Average Quantity	Avg. RECPI / Avg. Pop.
1	Paris	52,918.75	0.0029	Paris	88,877.50	0.0081
2	Cannes-Antibes	1,951.00	0.0024	Cannes-Antibes	2,659.75	0.0063
3	Lyon	6,930.00	0.0020	Montpellier	4,469.75	0.0050
4	Montpellier	2,731.00	0.0019	Marseille	10,437.00	0.0049
5	Marseille	6,574.75	0.0017	Lyon	11,506.50	0.0049
6	Annecy	761.00	0.0017	Nice	3,736.00	0.0048
7	Nice	2,698.75	0.0016	Bordeaux	6,370.25	0.0043
8	Fréjus	482.75	0.0016	Bayonne	1,390.25	0.0043
9	Toulouse	3,786.50	0.0015	Fréjus	759.00	0.0042
10	Grenoble	1,780.50	0.0015	Annecy	1,387.50	0.0039
11	Chambéry	605.25	0.0014	Toulouse	6,036.25	0.0039
12	Bordeaux	3,702.75	0.0014	Lille	5,438.25	0.0035
13	Avignon	1,042.50	0.0014	Avignon	1,546.50	0.0035
14	Nantes	2,275.25	0.0014	Nantes	3,790.25	0.0034
15	Lille	3,595.75	0.0013	Chambéry	1,053.25	0.0033

Notes: This table represents the rank the top 15 French Functional Urban Areas (FUAs) by average quality-adjusted entrepreneurial activity, measured as RECPI per capita (Avg. RECPI / Avg. Pop.), in the periods 2009–2012 and 2020–2023. For each period, we report the FUAs with the highest RECPI per capita, alongside their average number of new limited liability companies (Average Quantity).

Table B4 ranks the 15 leading French FUAs by quality-adjusted entrepreneurial activity (RECPI per capita) in 2009–2012 and 2020–2023. Paris dominates both periods, with a marked rise in RECPI per capita, while Cannes–Antibes also remains consistently second. Lyon, Marseille, and Montpellier retain prominent positions, though with some reordering. FUAs such as Annecy, Fréjus, and Chambéry also persist in the top 15. New entrants in 2020–2023 include Bayonne and Lille, while Grenoble drops out.

Table B5 ranks the 15 leading British FUAs by quality-adjusted entrepreneurial activity (RECPI per capita) in 2010–2013 and 2017–2020. London dominates both periods, with RECPI per capita rising from 0.0162 to 0.0194. In the second tier, notable changes occur: Brighton and Hove, Reading, and Cambridge all feature in the top ranks across both periods, while Cardiff, Manchester, and Leicester gain prominence in 2017–2020. Smaller FUAs such as Cheltenham, Guildford, and Preston remain present, suggesting persistent localized entrepreneurial activity. Meanwhile, new entrants like West Midlands and Leicester replace cities such as Bristol in the later ranking.

Table B6 reports the 15 leading German FUAs by quality-adjusted entrepreneurial activity (RECPI per capita) in 2009–2012 and 2020–2023. München consistently ranks first, although its RECPI per capita declines slightly in the later period. Berlin rises from fifth to second place, reflecting its growing entrepreneurial prominence, while Düsseldorf, Hamburg, and Frankfurt am Main remain among the top five. Several medium-sized FUAs such as Heidelberg, Konstanz, and Darmstadt appear in both periods, pointing to persistent localized activity. At the same time, mobility is evident: Bonn, Aachen, Karlsruhe, and Osnabrück enter the top 15 in 2020–2023,

Table B5: Top 15 British FUAs by Quality-Adjusted Entrepreneurship, 2010–2013 vs. 2017–2020

	2010–2013			2017–2020		
	Top FUAs Avg. RECPI / Avg. Pop.	Average Quantity	Avg. RECPI / Avg. Pop.	Top FUAs Avg. RECPI / Avg. Pop.	Average Quantity	Avg. RECPI / Avg. Pop.
1	London	152,110.00	0.0162	London	159,704.75	0.0194
2	Brighton and Hove	3,682.25	0.0110	Cardiff	7,224.00	0.0110
3	Reading	2,157.75	0.0097	Cambridge	2,150.25	0.0106
4	Cheltenham	1,332.00	0.0092	Manchester	26,800.00	0.0105
5	Guildford	2,007.00	0.0091	Reading	2,150.00	0.0090
6	Cambridge	2,017.75	0.0091	Brighton and Hove	3,167.00	0.0089
7	Manchester	22,496.50	0.0072	Cheltenham	1,192.25	0.0086
8	Coventry	4,326.75	0.0067	Guildford	1,635.50	0.0081
9	Blackpool	1,959.00	0.0065	West Midlands	22,564.75	0.0080
10	Southampton	3,506.00	0.0065	Leicester	6,148.25	0.0080
11	Bristol	5,435.00	0.0063	Northampton	3,455.75	0.0078
12	Oxford	2,591.75	0.0061	Oxford	2,719.00	0.0074
13	West Midlands	17,087.00	0.0060	Blackpool	1,969.00	0.0073
14	Preston	1,391.50	0.0058	Preston	1,507.00	0.0072
15	Northampton	2,820.25	0.0058	Coventry	4,580.00	0.0070

Notes: This table represents the rank the top 15 British Functional Urban Areas (FUAs) by average quality-adjusted entrepreneurial activity, measured as RECPI per capita (Avg. RECPI / Avg. Pop.), in the periods 2010–2013 and 2017–2020. For each period, we report the FUAs with the highest RECPI per capita, alongside their average number of new limited liability companies (Average Quantity). Bournemouth and Ipswich are excluded due to missing population data for the corresponding years.

while Wiesbaden, Regensburg, and Aschaffenburg drop out.

Table B7 reports the 15 leading French FUAs ranked by their average Regional Ecosystem Acceleration Index (REAI) in 2009–2012 and 2020–2023. The REAI measures the ratio of equity growth outcomes—IPOs and M&A events within six years of incorporation—to the number of quality-adjusted firms (RECPI). In 2009–2012, smaller FUAs such as Cherbourg-en-Cotentin, Pau, and Bourges lead the ranking, while Paris, despite its large entrepreneurial population, ranks 14th. By 2020–2023, the top positions shift to Montbéliard, Annemasse-Geneva, and Quimper, reflecting changes in the efficiency of local entrepreneurial ecosystems. Larger metropolitan areas like Paris, Bordeaux, and Lille appear lower in the ranking, highlighting the distinction between ecosystem acceleration and sheer entrepreneurial scale. Overall, the table illustrates substantial mobility in REAI rankings over time.

Table B8 reports the 15 leading British FUAs ranked by their average Regional Ecosystem Acceleration Index (REAI) in 2010–2013 and 2017–2020. In 2010–2013, smaller FUAs such as Preston, Exeter, and Aberdeen top the ranking, while larger metropolitan areas like Manchester and Leeds appear lower. By 2017–2020, the leading positions shift to Burnley, Cheshire West and Chester, and Aberdeen, reflecting changes in local ecosystem efficiency. Cities such as Bristol, Nottingham, and Southampton rise in the ranking, whereas Preston and Exeter move down. Overall, we also observe substantial variation in the REAI rankings over time in the UK.

Table B9 reports the 15 leading German FUAs ranked by their average Regional Ecosystem Acceleration Index (REAI) in 2009–2012 and 2020–2023. In 2009–2012, smaller FUAs such as Plauen, Remscheid, and Neubrandenburg top the ranking, while major metropolitan areas like Berlin and Munich appear lower, similar to the patterns observed in France and the UK. By 2020–2023, the leading positions shift to Greifswald, Konstanz, and Rostock, with Munich moving into 12th place, reflecting increased activity in larger hubs but persistent prominence of smaller, highly efficient ecosystems. Again, we observe a pattern of variation in the rankings

Table B6: Top 15 German FUAs by Quality-Adjusted Entrepreneurship, 2009–2012 vs. 2020–2023

	2009–2012			2020–2023		
	Top FUAs Avg. RECPI / Avg. Pop.	Average Quantity	Avg. RECPI / Avg. Pop.	Top FUAs Avg. RECPI / Avg. Pop.	Average Quantity	Avg. RECPI / Avg. Pop.
1	München	7,710.50	0.0069	München	8,358.50	0.0063
2	Düsseldorf	3,188.75	0.0046	Berlin	12,533.00	0.0049
3	Frankfurt am Main	5,135.25	0.0045	Düsseldorf	3,647.00	0.0047
4	Hamburg	7,408.75	0.0043	Hamburg	7,727.75	0.0039
5	Berlin	9,250.00	0.0039	Frankfurt am Main	5,456.00	0.0039
6	Flensburg	390.75	0.0033	Bonn	1,244.25	0.0037
7	Bonn	1,006.25	0.0031	Köln	3,753.75	0.0031
8	Heidelberg	428.00	0.0031	Heidelberg	510.25	0.0031
9	Konstanz	123.00	0.0030	Aachen	889.50	0.0030
10	Rosenheim	273.00	0.0030	Konstanz	140.50	0.0028
11	Köln	3,361.00	0.0030	Flensburg	316.75	0.0028
12	Wiesbaden	785.25	0.0029	Karlsruhe	936.75	0.0027
13	Darmstadt	542.00	0.0028	Darmstadt	550.00	0.0024
14	Regensburg	547.50	0.0026	Rosenheim	281.25	0.0024
15	Aschaffenburg	327.50	0.0026	Osnabrück	519.75	0.0023

Notes: This table represents the rank the top 15 German Functional Urban Areas (FUAs) by average quality-adjusted entrepreneurial activity, measured as RECPI per capita (Avg. RECPI / Avg. Pop.), in the periods 2009–2012 and 2020–2023. For each period, we report the FUAs with the highest RECPI per capita, alongside their average number of new limited liability companies (Average Quantity).

over time consistent with the French and UK rankings.

Table B7: Top 15 French FUAs by Average Regional Ecosystem Acceleration Index (REAI), 2009–2012 vs. 2020–2023

	2009–2012			2020–2023		
	Top FUAs Avg. REAI	Average Quantity	Avg. REAI	Top FUAs Avg. REAI	Average Quantity	Avg. REAI
1	Cherbourg-en-Cotentin	158.00	3.7274	Montbéliard	311.25	1.1720
2	Pau	551.25	3.0351	Annemasse-Geneva	1,278.50	0.5402
3	Bourges	267.25	2.8367	Quimper	642.75	0.5132
4	La Rochelle	646.75	2.7635	Avignon	1,546.50	0.4393
5	Hayange	397.75	2.6168	Reims	1,120.25	0.3620
6	Grenoble	1,780.50	2.4522	Mulhouse	1,094.00	0.3582
7	Valence	616.25	2.2227	Dijon	1,190.50	0.2964
8	Brest	525.75	2.1329	Nîmes	1,464.00	0.2932
9	Dunkerque	328.00	2.0161	Metz	1,120.50	0.2632
10	Saint-Brieuc	375.25	1.9940	Nancy	1,328.75	0.2601
11	Rennes	1,566.25	1.8262	Perpignan	1,569.00	0.2488
12	Angers	803.50	1.8010	Cannes-Antibes	2,659.75	0.2060
13	Toulon	1,551.50	1.7740	Tours	1,538.50	0.1910
14	Paris	52,918.75	1.6959	Bordeaux	6,370.25	0.1769
15	Troyes	387.25	1.6121	Lille	5,438.25	0.1467

Notes: The table reports the top 15 French Functional Urban Areas (FUAs) ranked by their average Regional Ecosystem Acceleration Index (REAI) during the periods 2009–2012 and 2020–2023. The REAI is defined as the ratio between equity growth outcomes (number of IPOs and M&A events within six years of incorporation) and the total number of quality-adjusted firms (RECPI) in a given FUA-year. For each period, we report the FUAs with the highest average REAI, alongside their average number of new limited liability companies registered (Average Quantity).

Table B8: Top 15 British FUAs by Average Regional Ecosystem Acceleration Index (REAI), 2010–2013 vs. 2017–2020

	2010–2013			2017–2020		
	Top FUAs Avg. REAI	Average Quantity	Avg. REAI	Top FUAs Avg. REAI	Average Quantity	Avg. REAI
1	Preston	1,391.50	2.3875	Burnley	811.75	1.2292
2	Exeter	1,793.00	1.9893	Cheshire West and Chester	2,378.75	1.0365
3	Aberdeen	2,350.25	1.9823	Aberdeen	1,902.25	0.9922
4	Cardiff	4,422.75	1.7082	Belfast	2,425.75	0.9272
5	Lincoln	1,510.75	1.6594	Bristol	5,165.50	0.9190
6	Norwich	1,749.50	1.5891	Nottingham	4,275.75	0.8851
7	Plymouth	1,314.75	1.5343	Southampton	3,499.50	0.8469
8	Sheffield	4,606.50	1.5043	Edinburgh	4,191.00	0.8447
9	Dundee City	972.75	1.4257	Swansea	1,428.75	0.7985
10	Belfast	2,155.50	1.4071	Guildford	1,635.50	0.7902
11	Newcastle upon Tyne	4,322.50	1.3973	Oxford	2,719.00	0.7839
12	Burnley	674.50	1.3819	Cambridge	2,150.25	0.7807
13	Manchester	22,496.50	1.3784	Leeds	14,156.00	0.7248
14	Leeds	12,679.25	1.3560	Norwich	2,067.25	0.7208
15	Edinburgh	4,123.50	1.3237	Coventry	4,580.00	0.7066

Notes: The table reports the top 15 British Functional Urban Areas (FUAs) ranked by their average Regional Ecosystem Acceleration Index (REAI) during the periods 2010–2013 and 2017–2020. The REAI is defined as the ratio between equity growth outcomes (number of IPOs and M&A events within six years of incorporation) and the total number of quality-adjusted firms (RECPI) in a given FUA-year. For each period, we report the FUAs with the highest average REAI, alongside their average number of new limited liability companies registered (Average Quantity).

Table B9: Top 15 German FUAs by Average Regional Ecosystem Acceleration Index (REAI), 2009–2012 vs. 2020–2023

	2009–2012			2020–2023		
	Top FUAs Avg. REAI	Average Quantity	Avg. REAI	Top FUAs Avg. REAI	Average Quantity	Avg. REAI
1	Plauen	60.00	4.5181	Greifswald	64.75	2.7486
2	Remscheid	176.75	3.7464	Konstanz	140.50	1.3186
3	Neubrandenburg	98.75	3.4835	Rostock	356.50	1.2342
4	Solingen	160.00	2.5104	Iserlohn	163.75	0.9417
5	Jena	153.25	2.4800	Kempten	184.25	0.7280
6	Erlangen	295.50	2.0961	Bamberg	242.50	0.5897
7	Wilhelmshaven	110.25	1.8834	Koblenz	363.50	0.4949
8	Kassel	393.25	1.8368	Rosenheim	281.25	0.4942
9	Lübeck	401.50	1.8094	Erlangen	343.00	0.4648
10	Offenburg	158.75	1.7818	Kassel	411.25	0.4143
11	Freiburg im Breisgau	547.25	1.7513	Freiburg im Breisgau	660.00	0.3765
12	Erfurt	306.75	1.6333	München	8,358.50	0.3526
13	Berlin	9,250.00	1.6321	Paderborn	485.00	0.3257
14	Bayreuth	197.75	1.5720	Köln	3,753.75	0.3028
15	Bamberg	244.75	1.5235	Braunschweig-Salzgitter	504.25	0.2963

Notes: The table reports the top 15 German Functional Urban Areas (FUAs) ranked by their average Regional Ecosystem Acceleration Index (REAI) during the periods 2009–2012 and 2020–2023. The REAI is defined as the ratio between equity growth outcomes (number of IPOs and M&A events within six years of incorporation) and the total number of quality-adjusted firms (RECPI) in a given FUA-year. For each period, we report the FUAs with the highest average REAI, alongside their average number of new limited liability companies registered (Average Quantity).

Appendix C Data Appendix

This section provides a detailed account of the business registry data used in this study and the steps taken to construct the measures used in the analysis. We document cross-country differences in data coverage and structure, outline how we adapted and implemented the measures proposed by Guzman and Stern (2020), and explain the choices made to ensure comparability across contexts. The section also describes procedures for data cleaning, integration with external datasets, and harmonization between countries. Particular attention is paid to institutional and reporting differences between European registries and the strategies we adopted to mitigate these discrepancies to produce results that are as comparable as possible.

C.1 Business Registration Records in France, the United Kingdom, and Germany

Our analysis centers on three European countries of major economic significance: France, the United Kingdom, and Germany. These countries constitute the largest national economies in Europe and, taken together, represent approximately 50% of the combined output of the European Union (27 member states) and the United Kingdom in 2020, according to OECD data. Despite their economic weight, the institutional arrangements that govern business registration differ markedly among the three. The registries vary in the procedures required to establish a firm, the format and content of the information disclosed, and the degree of accessibility offered to external users, such as researchers. Although the main text outlines the conceptual rationale for focusing on these economies, this appendix turns to the practical aspects of registry access, data availability, and the institutional characteristics that condition the type and quality of information recorded.

Business registration systems vary across European jurisdictions not only in terms of the legal formats available for incorporation, but also in the precise definitions of registry fields, the level of detail recorded, and the usability of the raw data. Despite these differences, registration is a legal requirement at the point of firm creation, which ensures that the registries capture businesses at a comparable stage of their life cycle across countries. However, heterogeneity in coverage, formatting, and access procedures required careful harmonization to construct consistent cross-country measures. In what follows, we provide country-specific details on these aspects, beginning with France.

C.1.1 France

In France, the business registration process has historically involved multiple administrative steps and institutions. For most of the period under analysis, entrepreneurs were required to submit their registration applications to the local chamber of commerce (*Chambre de Commerce et d'Industrie*). These applications were then reviewed and validated by a court clerk at the local commercial court (*Greffe du Tribunal de Commerce*). Since January 1, 2023, this decentralized system has been replaced by a fully online, centralized portal - the *Guichet des Formalités des Entreprises* - which is managed by the *Institut National de la Propriété Industrielle* (INPI). This reform abolished the former *Centres de Formalités des Entreprises* (CFE) and standardized the procedures for the creation, modification, and cessation of the activity of the company in all types of business, regardless of legal form or sector. Once an application is approved, the firm is officially entered into the national registry, the *Registre du Commerce et des Sociétés* (RCS). Its incorporation is then published in the official bulletin (*Bulletin Officiel des Annonces Civiles et Commerciales*, BODACC) and transmitted to the national statistical institute, INSEE (*Institut National de la Statistique et des Études Économiques*). INSEE assigns each enterprise a unique identification number and records it in the national business directory (*Système d'Identification* *du Répertoire des Entreprises*, SIRENE), which serves as the central repository for the French business data.

Business Registry Data. The French business registry data used in this study were obtained from the SIRENE database, which has been freely available as open data since January 2017.[1] We draw on two complementary files that were released in May 2024: the current stock of legal entities and their associated historical records. From the present-day registry, we extracted a set of variables describing firm identity, legal structure, and incorporation timing. Specifically, we retained the enterprise identifier (SIREN), the date of incorporation, the legal status, enterprise category, last update date, the registered name (denomination), the primary activity code and classification, and the location identifier of the registered headquarters (NIC).

The historical registry was merged with the present-day dataset using the unique firm identifier (SIREN). Observations lacking a valid incorporation date were excluded from the analysis. To harmonize information on incorporation and activity start dates, two cases were distinguished. When the incorporation date and the activity start date coincided, the observation was flagged accordingly. In cases where the two dates diverged, the earliest start date was retained and the discrepancy explicitly flagged. This procedure ensured that each firm was represented by a unique and consistent pair of dates, while preserving the ability to identify mismatches between legal incorporation and reported commencement of activity.

To focus on firms with the potential for growth, the analysis is restricted to limited liability companies. For the relatively small number of legal unities where historical records did not include a legal structure, this information was imputed from the present-day dataset. The study restricted attention to limited liability companies, defined as entities with 2-digit legal status codes between 54 and 58, as defined by INSEE's official classification. This category includes the main forms of French limited liability corporations: Société à responsabilité limitée (SARL), Société anonyme à conseil d'administration, Société anonyme à directoire, Société par actions simplifiée (SAS), and Société européenne. Finally, the sample was limited to firms incorporated between 2009 and 2023, ensuring comparability with the data available for other countries. The resulting dataset, therefore, consists of French limited liability companies incorporated during 2009–2023, with harmonized legal status information, consistently defined incorporation and activity start dates, and the variables necessary to construct indicators of startup quality. In a second step, establishment-level information was integrated into the dataset. These data were drawn from both the current and historical SIRENE files. To ensure consistency with the firm-level analysis, establishments were first restricted to those associated with the limited liability companies already included in the sample. From the current registry, we extracted a broad set of variables, including firm and establishment identifiers (SIREN and SIRET, respectively), indicators of whether the unit functioned as a headquarters, the dates of establishment creation and activity start, postal codes and address identifiers, geographic coordinates, administrative status, denomination, and detailed activity codes.

The historical establishment records were then merged with the current dataset using both firm and establishment identifiers. Observations lacking a valid creation date were excluded. Finally, the merged establishment data were linked to the firm-level dataset, enabling each establishment to be associated with the incorporation date of its parent company. To facilitate temporal analyses, the year and month were systematically extracted from both the creation date and the activity start date of each establishment.

The dataset was subsequently restricted to headquarters establishments, identified by matching the establishment's identifying code (NIC) to the headquarters code associated with the legal entity. This ensured that each firm was represented by its primary registered location. Additional consistency checks were performed by comparing the firm-level incorporation date with

[1] The dataset and its documentation are accessible at: https://www.sirene.fr/sirene/public/static/acces-donnees and https://www.sirene.fr/sirene/public/static/documentation.

the establishment creation date. Records where the two dates coincided were flagged, while establishments with discrepancies were excluded. In the final step, the sample was restricted to headquarters establishments for which the reported start date of activities matched the incorporation date of the corresponding legal entity.

The resulting dataset therefore contains the headquarters establishments of limited liability companies incorporated in France between 2009 and 2023, with harmonized and consistent creation and activity start dates, as well as geocoded location and address information.

Industry Indicators. The next step was to harmonize and classify firm-level activity codes to construct measures of industry and technological intensity. The activity codes provided in the SIRENE registry follow the French Nomenclature of Activities (NAF Rev. 2), which is directly linked to the European NACE Rev. 2 classification. The processing began by retaining for each firm the primary activity codes recorded in both the present and historical registers. To harmonize these with international standards, the official correspondence tables published by INSEE were used, mapping NAF codes across five levels of aggregation (from section to five-digit codes). For each firm, both two-digit and four-digit NACE codes were extracted. Several adjustments were made to correct inconsistencies in the raw data. Specifically, invalid activity codes (e.g., “00.00Z”) were replaced with the present-day code of the corresponding firm.

In addition to standard industry classification, firms were categorized according to technological and knowledge intensity using the OECD taxonomy of manufacturing and service sectors. For manufacturing, this taxonomy distinguishes between high-, medium-high-, medium-low-, and low-technology industries. For services, it differentiates between knowledge-intensive and less knowledge-intensive activities, with further disaggregation into market, financial, and high-technology services when a more granular classification is applicable. To ensure comparability across countries and provide a basis for robustness checks, this OECD-based taxonomy was supplemented with an alternative classification scheme developed by the authors at the four-digit NACE level. In cases where industry codes could not be reliably assigned, despite the presence of valid NACE sections, firms were placed in a residual “Other” category in order to maintain coverage of the registry.

At the same time, a small set of sectors was flagged for exclusion, specifically NACE sections A, T, and U. Section A covers agriculture, forestry, and fishing; Section T encompasses households as employers of domestic personnel and the production of goods and services for own use; and Section U includes activities of extraterritorial organizations and bodies such as embassies or international institutions. These activities were excluded because they fall outside the scope of entrepreneurial dynamics that are relevant for scaling up startups. Agricultural production (section A) is shaped by land use, seasonality, and natural resource constraints rather than by the innovation-driven growth models that characterize high-potential firms. Household activities (section T) are not market-oriented businesses in the conventional sense and therefore cannot be meaningfully compared with incorporated enterprises. Extraterritorial organizations (section U) represent administrative or diplomatic entities that are not subject to the same entrepreneurial or competitive pressures as private firms.

The final dataset thus provides, for each firm in the cleaned registry, harmonized NACE codes, OECD-based classifications of technological and knowledge intensity, and excludes activities outside the scope of analysis.

Extraction of Founders and Shareholders from BODACC. An essential component of our analysis involves extracting information on the individuals associated with the creation of new firms in France, with particular attention to founders and shareholders. This step is central to the construction of our eponymous variable, which relies on identifying whether firm names are derived from those of their founders. To this end, we made use of raw announcements published in the *Bulletin Officiel des Annonces Civiles et Commerciales* (BO-

DACC), the official source for legally binding notices concerning firm creations, dissolutions, and modifications. BODACC is freely accessible in open-data format via its official platform (https://www.bodacc.fr/pages/donnees-ouvertes-et-api/).

Our focus was specifically on the dataset of commercial creation notices (*annonces commerciales – création*), which provides detailed information on the legal constitution of firms at the moment of their registration. Each record includes a range of structured variables, among which we extracted the announcement identifier, the date of publication, the type and family of the notice, the codes of the department and region, the tribunal of registration, firm identifiers (SIREN and SIRET), as well as the registered name and location of the enterprise.

Of particular importance are two JSON-formatted fields embedded within the BODACC files. The first encodes structured information on establishments, including addresses and activity details. The second provides information on the individuals formally associated with the company, such as shareholders, administrators, and directors. This latter field is especially critical for our purposes, as it enables the systematic identification of persons linked to firm ownership and governance, and thus forms the basis for constructing indicators of founder identity and related quality signals.

The main challenge in using BODACC data arises from the unstructured encoding of individual-level roles. The contains text strings where roles and names are concatenated, often in inconsistent formats. For example, one might encounter entries such as "gérant: Jean Dupont; associé gérant: Marie Martin", where names and positions are concatenated in free text.

The text data required substantial cleaning prior to extraction. We normalized all entries by removing accents, converting to lowercase, stripping punctuation, and eliminating commas. To bring structure to this information, we constructed a controlled vocabulary that standardized the observed roles into a consistent typology.

In total, we identified 40 distinct role designations appearing in the BODACC notices. These covered a wide spectrum of company functions, ranging from common executive and managerial roles—such as *gérant*, *associé gérant*, *co-gérant*, *président*, *directeur général*, *directeur général exécutif*, *membre du directoire*, and *membre du conseil d'administration*—to more specialized functions, including *membre du conseil de surveillance*, *commissaire aux comptes titulaire*, *commissaire aux comptes suppléant*, *contrôleur des comptes*, *contrôleur de gestion*, and *responsable à l'étranger*. This comprehensive classification ensured that all possible designations of individuals associated with newly created firms were systematically captured.

For the purposes of our analysis, however, not all of these roles were relevant. Since the central objective is to identify founders and shareholders, we restricted attention to positions that imply ownership stakes or founding responsibilities. A set of roles was therefore excluded, corresponding to functions unrelated to ownership. Specifically, the following categories were filtered out: *contrôleur de gestion*, *contrôleur des comptes*, *membre du conseil de surveillance*, *responsable à l'étranger*, *commissaire aux comptes*, *commissaire aux comptes titulaire*, *commissaire aux comptes suppléant*, *responsable en France*, *personne ayant le pouvoir d'engager la société*, and *personne ayant le pouvoir d'engager à titre habituel la société*. These correspond primarily to statutory auditors, controllers, supervisory board members, and external representatives. From the full BODACC vocabulary, the following roles are kept:

- *Gérant* / *Gerante* (manager, often equivalent to a founder in small firms)
- *Président* (chairman or statutory head, often a shareholder-founder)
- *Président de la société* (presidency role with ownership implications in certain corporate structures)
- *Directeur général* (CEO / Managing Director role, relevant when linked to ownership)
- *Directeur général délégué* (delegated managing director, tied to executive authority)

- *Directeur général exécutif* (executive managing director, combining management and ownership functions)
- *Administrateur* (board member, may imply ownership depending on corporate structure)
- *Personne ayant le pouvoir de diriger et gérer* (individual formally empowered to direct and manage the company)

Ultimately, we constructed a cleaned and structured mapping linking each firm identifier (SIREN) to the individuals associated with it and their corresponding roles. Duplicate records were removed, and missing values were systematically flagged. By restricting the dataset to roles directly related to ownership or founding functions, we ensured that the retained information accurately reflects the individuals responsible for the creation and control of the firm. This mapping forms the basis for our analysis of startup quality through founder identity and shareholder structures.

One limitation of the data is important to note: for reasons that remain unclear, only about half of the firm registrations recorded in SIRENE for the period 2009–2023 are observed in the BODACC notices of newly registered firms. This discrepancy implies that some founders and firms may be missing from our extraction. Ongoing efforts aim to address these gaps and improve the completeness of the eponymous measures.

Identifying Eponymous and Business Name Length. The second stage of the analysis is devoted to detecting eponymous firms - those whose legal name incorporates the personal name of one of their founders or top managers. Achieving this requires linking two complementary sources of information: (i) the official firm names recorded in the SIRENE register, and (ii) the names of individuals associated with firms, extracted from BODACC announcements in the previous step, focusing specifically on those occupying managerial or founding roles. The identification procedure involves several steps, including systematic cleaning, standardization, and pattern matching of both firm and personal names.

Firm names from SIRENE are first standardized through a cleaning process that removes extraneous characters, accents, and formatting inconsistencies. Two versions of each firm name are then generated. The first version applies only generic cleaning, preserving most of the original content. The second incorporates the list of over 1,000 business-related terms and legal-entity endings compiled by Belenzon et al. (2017), removing legal entity markers (e.g., "SARL", "SAS") and common business words (e.g., "services", "industries"), producing a more neutral string better suited for detecting personal names embedded in the company name. For both versions, we calculated name length measures, including string length and word count, and constructed indicators for short names based on absolute thresholds and distributional percentiles. We used in the main paper the indicator for short name as *word count* less than or equal to 3. Furthermore, these preprocessing steps create a structured representation of firm names that can be systematically compared to the cleaned and standardized names of founders and managers, enabling robust identification of eponymous relationships.

Given that the BODACC dataset captures only a subset of newly registered firms, we complement the identification of eponymous firms using an alternative approach based on the presence of personal names within the firm's legal name. The underlying assumption is that it is uncommon for businesses to be named after individuals other than their founders, so detecting a personal name in the firm name provides a strong signal of founder eponymy. Even in the uncommon cases where the name refers to a non-founder, this measure still aligns with the conceptual intent of the eponymous variable as defined by Stern and Guzman. To implement this approach, we relied on an auxiliary dataset of all given names and surnames of natural persons that registered a business in France and have their name recorded in Sirene[2]. This

[2]Dataset publicly available at https://www.data.gouv.fr/datasets/liste-de-prenoms-et-patronymes/

dataset includes frequency information for each first name and surname, allowing us to identify the most common personal names among entrepreneurs in France. While we acknowledge that this method is not perfect, extensive manual checks indicate that it performs well in practice. We believe that the errors introduced by this strategy are likely to only add noise rather than systematic bias.

We begin by cleaning the BODACC dataset to ensure that only actual individuals are considered. Any shareholder or manager whose name contains business words (e.g., "boulangerie") or legal entity markers (e.g., "SARL") compiled by Belenzon et al. (2017) is flagged and excluded. Additional parsing extracts usage names (*nom d'usage*) and maiden names (*nom de célibataire*), standardizes them, and separates surnames from given names. The resulting first name and surname datasets are then further processed. Names are cleaned, restricted to those with more than three characters, and filtered to retain the most common thousand entries. This produces two reference lists of high-frequency French first names and surnames, which are subsequently combined into a single list of "popular names" used for generic eponym detection in firm names.

The key identification step consists of two complementary procedures. First, a regular expression is built from the list of popular names, allowing the detection of generic eponymy, i.e., cases where a company name contains a common French personal name. Second, the names of *dirigeants* extracted from BODACC are merged with the SIRENE firm name dataset. For each firm, a regular expression is compiled from the *dirigeants'* names, and the cleaned legal denomination of the firm is searched for a match. This produces a firm-level indicator of specific eponymy, identifying companies explicitly named after one of their founders or top managers.

The final output provides three distinct measures: (i) *eponymous common name*, flagging firms whose legal name contains any frequently occurring French personal name; (ii) *eponymous from BODACC*, flagging firms whose name matches directly with the names of registered *dirigeants*; and (iii) *eponymous combined*, which integrates the two approaches by prioritizing direct matches while using the generic name method as a fallback when BODACC data are incomplete. This procedure ensures that the identification of eponymous firms is both comprehensive, by considering common personal names, and precise, by leveraging the administrative linkage between firm names and their registered founders or managers in BODACC.

C.1.2 United Kingdom

In the United Kingdom, company registration is fully centralized through Companies House, the national registrar. Applications are reviewed by registry officials, and once approved, a unique company number is assigned. A notice of incorporation is published in the official public record (The Gazette), and the information is recorded in the publicly accessible Companies House registry. The registry data are available as monthly bulk downloads at https://download.companieshouse.gov.uk/.

A key limitation is that Companies House excludes dissolved companies from the main public dataset. To address this, we supplement the registry with data from the StatBooks® CHS platform (https://statbooks.co.uk/), which provides Companies House–derived data products covering both active and dissolved entities. Specifically, we obtained data on all dissolved firms between 2010 and 2020. This constraint explains why the UK time window in our analysis differs from that used for France and Germany. To assess coverage and reliability, we validated the StatBooks dataset against official business demography statistics from the UK Office for National Statistics (ONS). The combined information, i.e., active firms from Companies House and dissolved firms from StatBooks, closely matches official ONS figures, confirming its reliability.

Business Registry Data For this study, we use Companies House June 2024 snapshot and StatBooks dissolved companies from January 2025. The first stage of data processing involves constructing a comprehensive dataset of dissolved firms. This requires harmonizing company identifiers, incorporation and dissolution dates, addresses, and industry codes (SIC). Company category labels are further standardized to ensure internal consistency across the dataset. To align with the scope of analysis, we restrict the combined dataset (active and dissolved firms) to companies incorporated between 2010 and 2020. Observations with missing incorporation dates or company category information are excluded. The analysis is then restricted to limited liability companies, including: private limited companies, public limited companies, European public limited-liability companies (SEs), United Kingdom Societas, and companies incorporated under Section 30 of the Companies Act. This restriction ensures a homogeneous and comparable set of firms with respect to governance structure, liability regime, and corporate obligations.

Historical company names are then incorporated. Companies House provides up to 10 previous names for each firm. These columns are identified and sorted from most recent to oldest. For each firm, the most recent non-missing historical name is selected as the business name at founding; if no previous names are available, the current name is retained. This step allows the dataset to track name changes over time, which is critical for creating the name-based indicators.

We also extracted and organized the registered address information of companies, retaining all fields relevant to firm location. These include the care-of field, P.O. Box, multiple address lines, postal town, county, country, and primary postal code. Historical address data are not available in the basic Companies House records, making it infeasible to identify the original business address at incorporation, as is possible in the French business registry. Consequently, our address dataset more closely resembles the original sample in Guzman and Stern (2020), where for most U.S. states, only current address information was available. Their robustness checks using the Massachusetts sample - where historical addresses could be observed - demonstrated that predictive model results were not sensitive to firm relocations. This provides reassurance that using current registered addresses is a valid approach for our subsequent spatial analysis, including geocoding and regional mapping.

The final dataset for the United Kingdom includes, for each limited liability company incorporated between 2010 and 2020, the current company name, historical names, unique company identifier, legal category, incorporation date, and current addresses.

Industry Indicators The United Kingdom uses the SIC 2007 classification, which is identical to NACE Rev. 2 up to the 4-digit level, and still requires mandatory information to be added in the business application after Brexit. A limitation relative to the French registry is the absence of historical SIC codes, which are not systematically preserved in the publicly available Companies House bulk downloads. While these codes can be retrieved from alternative open-access government repositories, their systematic collection remains part of an ongoing long-term project.

To ensure comparability, we apply the same procedure as for the French data. Companies House SIC codes are first mapped to their corresponding NACE equivalents. Firms are then classified according to OECD technological and knowledge-intensity categories and flagged for exclusion if they operate in sectors outside the scope of analysis, specifically NACE sections A (Agriculture), T (Activities of households as employers; undifferentiated goods- and services-producing activities of households for own use), and U (Activities of extraterritorial organizations). After standardizing the SIC codes, the reference table is cleaned and reformatted to reflect the hierarchical NACE structure. The table is then merged with our alternative NACE-based industry dataset at the 4-digit level, ensuring consistent assignment of each SIC code to both an OECD technology/knowledge-intensity category and an alternative NACE industry classification. Finally, firms with missing sectoral information are explicitly flagged.

The resulting dataset records, for each company, its primary SIC code, the corresponding NACE section and division, an alternative industry classification, the assigned OECD technology/knowledge-intensity category, and an explicit flag for cases with missing sectoral information.

Extracting Information from Person with Significant Control Snapshot We then extract information on individuals with significant ownership stakes, which is essential for identifying founders and major shareholders. Our goal is to construct a dataset of individuals holding at least 50% of a company's shares, including their names and relevant attributes for subsequent analysis. The underlying dataset, Persons with Significant Control (PSC), is publicly available from Companies House (https://download.companieshouse.gov.uk/en_pscdata.html); for this study, we use the February 2025 snapshot. This dataset covers only current PSCs associated with active firms. While historical PSC data could theoretically be accessed via the Companies House API or direct data requests, this was not feasible for the present work. Collecting comprehensive historical PSC data for all firms remains part of an ongoing long-term project.

Data processing proceeds in several steps. First, we retain only individual persons, excluding corporate entities and other legal persons, following the approach of Belenzon et al. (2017) by identifying firms based on a list of business words and legal-entity markers. Second, we select a set of relevant variables, including the nature of control, name components (forename, middle name, surname, title), country of residence, nationality, and date of birth. Third, we account for cases where an individual may hold multiple forms of control (e.g., ownership and voting rights) by exploding these entries into distinct records. Fourth, we restrict the dataset to individuals with ownership stakes between 50 and 100%, thereby ensuring that only major shareholders or founders with decisive control are included.

After cleaning, individual names are standardized by concatenating name components into full names. The data are then aggregated at the company level, producing, for each firm, a list of major individual owners (forenames, surnames, and full names). This step enables the detection of eponymous companies. The final dataset provides, for each UK company, a structured record of individuals with significant control, including their full names, ownership share, and nature of control.

Identifying Eponymous Firms and Business Name Length For the UK dataset, we developed a systematic framework to identify eponymous companies—firms whose registered names incorporate the personal names of founders or major shareholders. The procedure integrates cleaned historical company names, population-level name distributions, and information on Persons with Significant Control (PSC) from Companies House, while filtering out legal and generic business terms to reduce false positives.

The first step involves standardizing company names. This process unescapes HTML entities, normalizes accented characters, removes punctuation and extraneous symbols, collapses multiple spaces, and trims whitespace. We then compute features such as string length, word count, and the distribution of words within each name. Indicators for short names are generated based on two thresholds: (i) names with three words or fewer, and (ii) names with character counts falling below the 10th or 25th percentile of the full distribution. To prepare company names for eponym detection, we leverage PSC data to perform founder-specific eponymous identification. For each company, the dataset provides the full names of individuals with significant ownership or control. These names are normalized, converted into regular expression patterns, and matched against cleaned company names to generate a PSC-based eponymous indicator. This approach offers a precise, founder-linked measure of eponymous identity.

Not all companies, however, have an identified PSC—either because they are dissolved (and thus not included) or no individual is listed. To address these gaps, we supplement the PSC-based approach with a population-level name dataset for the UK, NameDataset[3], which includes first names, surnames, gender, and country of origin (Remy, 2021). From NameDataset, we calculate name and surname lengths and generate frequency counts. To focus on common personal names, we retain only the top 0.1% most frequent names and exclude names shorter than three characters

[3]This dataset, compiled by Remy (2021) and accessible via the Python library names-dataset, draws on a large corpus of publicly available Facebook user names, filtered to include only UK-based entries.

to reduce noise from short or ambiguous strings. These selected names are then compiled into a comprehensive regular expression that matches any occurrence of a common first or last name within a company name. Companies matching this pattern are flagged as eponymous based on common names, creating the eponymous popular names indicator, analogous to the procedure applied to the French dataset.

To construct a unified measure, we combine the two approaches: the PSC-based indicator is used whenever available, while the population-based indicator serves as a fallback in cases where PSC data are missing. This strategy ensures comprehensive coverage of firms, preserving verified founder-level matches where possible while still allowing eponymous detection for companies lacking PSC information. The resulting composite indicator provides a robust measure of eponymous entrepreneurship in the UK, linking company naming practices to founder identity, ownership structures, and broader firm characteristics.

C.1.3 Germany

In Germany, the registration process is decentralized at the state level, similar to the United States. Companies are registered with the local commercial register (*Handelsregister*) through the regional chamber of commerce and confirmed by its respective district court (*Amtsgericht*). Upon approval, the company is formally entered into the *Handelsregister* and its information published in the federal public register (*Bundesanzeiger*).

Business Registry Data Germany established a digital Business Registry in 2007[4] , which has been free of charge for most services since 2022. Access to additional information, such as annual financial statements, requires a fee of one euro per firm. However, the registry is not fully harmonized and exhibits variation across federal states.

For our analysis, we therefore rely on the Mannheim Enterprise Panel (MUP), a firm-level panel dataset jointly maintained by Creditreform, Germany's largest credit rating agency, and ZEW, the Leibniz Centre for European Economic Research. Unlike the French and UK sources, the MUP is not publicly available. It builds on the official German Business Registry but extends coverage substantially by incorporating firms that are not legally required to register. Specifically, firms are included if they (i) appear in the official Handelsregister, (ii) are identified through media reports, or (iii) are recorded by Creditreform in response to client requests. Through this procedure, the MUP captures virtually the entire population of economically active firms in Germany.

The dataset covers firms founded between 1995 and 2023, with systematic half-yearly updates since 1992. For comparability across countries, we restrict our analysis to the period 2009–2023. The statistical unit of observation is the legally independent enterprise. The MUP records detailed firm-level information, including exact foundation and closure dates, insolvency proceedings, registered address, legal form, five-digit NACE industry classification, employment and sales, and—in a subset of cases—balance sheet data. Importantly for our study, the dataset also includes historical information on addresses, industry classification, and individuals with ownership or managerial responsibilities (e.g., owners, managing directors, partners, board members, and majority shareholders).

Relative to the French and UK registries, the MUP offers higher-quality information owing to long-standing manual verification and harmonization efforts by Creditreform and ZEW. To ensure comparability, we construct our indicators following the same procedures applied to France and the UK: we define short company names as those with three words or fewer; exclude firms in NACE sections A, T, and U; identify eponymous firms using majority shareholder information; and supplement missing matches with a list of common names observed in the

[4] Available at: https://www.unternehmensregister.de.

registry. Thus, while the German data are of higher quality and include historical coverage, they remain subject to the same conceptual limitations inherent to this approach.

C.2 Additional External Datasets

To replicate the methodology of Guzman and Stern (2020), we complement business registry data with additional external sources in order to construct the full set of variables required. These cover startup growth outcomes, business registration details, and further observables such as geocoding and regional information.

Growth To measure growth, we rely on data from the Moody's Analytics Orbis M&A database (formerly Zephyr, Bureau van Dijk) to identify whether a startup undergoes an initial public offering (IPO) or is acquired within six years of its registration date. Orbis M&A is the largest comparable dataset on private European companies. As with other Moody's Analytics datasets used in this study, coverage is not universal and varies over time (Bajgar et al., 2020). Nevertheless, there is evidence that the database provides relatively good coverage of high-value acquisitions, similar to Thomson Reuters SDC data used in the original Guzman and Stern (2020) study (Bollaert and Delanghe, 2015). Any limitations in coverage are likely to induce attenuation bias, consistent with the original analysis. Over the respective coverage periods for each country, we observe 1,322 growth events in France (82 IPOs), 4,562 in the United Kingdom (671 IPOs), and 2,069 in Germany (132 IPOs).

To test the robustness of our predictive models, we construct alternative growth measures. These include varying the observation window (e.g., nine years or unlimited), considering IPOs alone instead of the combined IPO and M&A outcome, and defining growth in terms of employment thresholds. Using Orbis balance sheet and employment data, we identify firms that reach milestones of 100, 250, 500, or 1,000 employees, or attain the European Commission's classification as a medium or large enterprise[5] within six years of registration. Firms not matched in Orbis are conservatively assumed not to have achieved the growth metric. While this approach has limitations, Bajgar et al. (2020) show that Orbis coverage is particularly strong for larger firms, suggesting that these alternative measures provide a reasonable approximation of high-growth outcomes.

Intellectual property measures To complement registry-based measures, we incorporate indicators of startup quality based on intellectual property. Specifically, we draw on Moody's Analytics ORBIS IP dataset and the Fraunhofer Institute for Systems and Innovation Research (ISI) Trademark Data Collection (ISI-TM).

The ORBIS IP dataset links patent applications and grants from major patent offices worldwide to corporate entities covered in the ORBIS universe. It consolidates information from sources such as the European Patent Office (EPO), the United States Patent and Trademark Office (USPTO), the World Intellectual Property Organization (WIPO), and selected national patent offices. Through its firm-level harmonization, ORBIS IP enables the systematic attribution of patent activity to legally registered companies, making it a widely used resource in research on innovation, intellectual property, and firm dynamics.

The ISI Trademark Data Collection (ISI-TM), developed by the Fraunhofer Institute for Systems and Innovation Research (ISI), compiles firm-level information on trademark applications filed with the European Union Intellectual Property Office (EUIPO) and the United States Patent and Trademark Office (USPTO). The dataset is publicly available via the Zenodo repository (https://zenodo.org/records/4633061), with detailed documentation provided by

[5]European Commission (2003), "Recommendation 2003/361/EC: Definition of micro, small and medium-sized enterprises." An enterprise is classified as medium if it has fewer than 250 employees and either an annual turnover $\leq$ €50 million or a balance sheet total $\leq$ €43 million; it is classified as large if it exceeds these thresholds.

Neuhäusler et al. (2021). ISI-TM provides comprehensive coverage of EU- and US-level trademark filings, but does not include applications made exclusively at national intellectual property offices in Europe. This limitation notwithstanding, European Union trademarks in particular are widely viewed as robust indicators of firm quality, as they signal a broader market orientation beyond the domestic sphere and entail higher application costs relative to national filings.

Although ISI-TM does not capture trademarks filed solely at the national level, European trademarks are widely regarded as stronger signals of firm quality and market orientation. We construct two binary indicators: *Patent*, equal to 1 if a firm holds a patent within the first year of registration, and *Trademark*, equal to 1 if a firm applies for a European trademark within the first year of registration.

Address Geocoding and Regional Mapping To geocode business locations and construct regional identifiers, we combine information from national statistical offices, Eurostat, and the OECD. The procedure differs slightly across countries, but in each case complements the core business registry data.

The spatial data used in this study are derived from official Eurostat GISCO releases. Specifically, we employ the NUTS 2021 shapefiles (20M resolution, EPSG:4326) for NUTS level boundaries[6]. Local Administrative Units (LAU) are obtained from both the 2019 and 2021 shapefiles (1M resolution, EPSG:4326)[7]. The inclusion of the 2019 version is necessary to cover the United Kingdom, which is not included in the LAU 2021 release. The two LAU datasets are concatenated to ensure complete coverage of European territories. We then link LAU units to NUTS 2021 codes and incorporate the DEGURBA classification, based on Eurostat sources. In addition, we use the OECD Data Explorer platform[8] to obtain GDP and population figures for countries and Functional Urban Areas (FUAs), which are merged into our datasets. This harmonized spatial data enables a consistent assignment of establishments to local and regional units across countries.

For France, we implemented a multi-step geocoding process. First, we retrieved official French postcode geometries from the Eurostat Postal Codes 2024 dataset[9], which provides polygon boundaries and centroids of all valid postal codes. Second, we extracted establishment-level address records from the Sirene historical address dataset, including unique identifiers (SIREN, SIRET, address IDs), postal codes, commune names, and Lambert coordinates where available. Finally, we complemented this with the official Sirene geolocation file[10], which reports geographic coordinates (latitude/longitude) and municipality codes for establishments.

The dataset was then split into two subsets. For establishments lacking direct geographic coordinates, postal codes were matched to the Eurostat dataset, and geometries were imputed using postcode centroids. The resulting geocoded addresses were spatially joined to the harmonized LAU/NUTS reference file, assigning each establishment to the appropriate Local Administrative Unit (LAU) and NUTS region. Missing values were systematically addressed, and establishments without valid regional matches were flagged. Higher-level regional identifiers (NUTS-1 and NUTS-2) were derived from the NUTS-3 code hierarchy to enable consistent regional analyses.

For the UK, we geolocate companies and link them to administrative regions using postal code information, providing precise regional identifiers for each firm at ZIP code level. We

[6] Eurostat GISCO (2021), NUTS 2021 shapefiles, 20M resolution, EPSG:4326. Available at: https://gisco-services.ec.europa.eu/distribution/v2/nuts/shp/

[7] Eurostat GISCO (2019), LAU 2019 shapefiles, 1M resolution, EPSG:4326. Available at: https://gisco-services.ec.europa.eu/distribution/v2/lau/shp/; Eurostat GISCO (2021), LAU 2021 shapefiles, 1M resolution, EPSG:4326. Available at: https://gisco-services.ec.europa.eu/distribution/v2/lau/shp/

[8] Available at: https://data-explorer.oecd.org/

[9] Available at: https://ec.europa.eu/eurostat/web/gisco/geodata/administrative-units/postal-codes

[10] Available at: data.gouv.fr (Sirene geolocation file)

use the reference dataset of UK postcodes from the ONS Open Geography Portal[11], which contains latitude, longitude, and other geospatial attributes. These coordinates are converted into point geometries suitable for spatial operations. Next, company addresses from the registry are merged with the postal code reference dataset, matching each company's postal code to its corresponding geolocation. The resulting dataset contains company identifiers, postal codes, and associated latitude/longitude coordinates. Firms lacking coordinates are retained but flagged with a missing-location indicator. The point geometries of company addresses are then spatially joined to the LAU and NUTS polygons, assigning each firm to the relevant administrative boundaries, including LAU codes, NUTS-3 codes, counties, districts, and wards. Higher-level NUTS codes (NUTS-1 and NUTS-2) are derived from the NUTS-3 assignments, as previously.

For Germany, firms in the Mannheim Enterprise Panel (MUP) are already geolocated by Creditreform and ZEW. We supplemented these coordinates with Eurostat LAU/NUTS reference data that built and applied the same spatial harmonization procedure used for France and the UK, assigning each firm to the appropriate LAU and NUTS regions and deriving higher-level NUTS codes from NUTS-3 where necessary.

[11] Available at https://geoportal.statistics.gov.uk/datasets/957329d0bfee40349a03c07652ae64a7/about

References

Bajgar, M., G. Berlingieri, S. Calligaris, C. Criscuolo, and J. Timmis (2020). *Coverage and Representativeness of Orbis Data*. OECD Science, Technology and Industry Working Papers 2020/06. DOI: 10.1787/c7bdaa03-en. URL: https://www.oecd-ilibrary.org/science-and-technology/coverage-and-representativeness-of-orbis-data_c7bdaa03-en.

Belenzon, S., A. K. Chatterji, and B. Daley (2017). "Eponymous Entrepreneurs." *American Economic Review* 107(6), 1638–1655.

Bollaert, H. and M. Delanghe (2015). "Securities Data Company and Zephyr, Data Sources for M&A Research." *Journal of Corporate Finance* 33, 85–100. (Visited on 09/21/2025).

Guzman, J. and S. Stern (2020). "The State of American Entrepreneurship: New Estimates of the Quantity and Quality of Entrepreneurship for 32 US States, 1988–2014." *American Economic Journal: Economic Policy* 12(4), 212–243.

Neuhäusler, P., R. Frietsch, and O. Rothengatter (2021). *Documentation of RISIS Datasets: ISI-Trademark Data Collection (ISI-TM)*. Tech. rep. Zenodo. DOI: 10.5281/zenodo.4633061.

Remy, P. (2021). *Name Dataset*. https://github.com/philipperemy/name-dataset.

Appendix D Robustness Checks

In this section, we report a series of robustness checks conducted to assess potential biases in our measures and to evaluate heterogeneity across regions.

Tables D1–D3 present results obtained using alternative industry fixed effects based on different levels of aggregation of industry codes from the business registries. Across all countries, the estimated coefficients remain largely consistent in both magnitude and significance, indicating that our main results are not driven by the choice of sectoral aggregation.

Tables D4–D6 report results from models that include regional fixed effects in our preliminary, Nowcasting, and Full model specifications. While the inclusion of regional fixed effects slightly increases the Pseudo R^2, the overall impact is minimal, and the substantive conclusions of our analyses remain unchanged.

We also examine the role of dynamic regional effects by incorporating both region fixed effects and region-specific time trends, as shown in Tables D7–D9. Given the large number of dummy variables required in the logistic regression framework, precise estimation of robust standard errors is challenging. To address this, we employ an estimation approach that absorbs the fixed effects, allowing for more efficient coefficient estimation while producing standard errors equivalent to those obtained from models including explicit dummy variables. Some observations are lost due to singletons in the fixed effects, but the estimated coefficients remain remarkably stable across specifications.

Additionally, we explore alternative measures of firm size growth. Specifically, we consider thresholds corresponding to medium- and large-sized firms as defined by the European Commission, as well as employment-based thresholds of more than 100 and 250 employees. The results are broadly consistent with those using the 500- and 1000-employee thresholds, reinforcing the findings reported in the main results: models predicting scaling through equity growth differ markedly from those predicting growth in employment outcomes across France, the UK, and Germany.

We also test the robustness of our results using alternative definitions of growth. Specifically, we re-estimate the models with the same set of predictors as the full-information entrepreneurial quality model, but replace the baseline equity growth outcome with: (i) IPOs within six years of foundation, (ii) IPOs or M&As within nine years of foundation, and (iii) whether a firm ever undergoes an IPO. The results are consistent across countries: relative to the baseline model, specifications relying only on IPOs yield different coefficient magnitudes and significance levels, whereas models including both IPOs and M&As over a longer time horizon produce results that are qualitatively similar to the baseline and, in most cases, show improved explanatory power (with the exception of the UK, where the pseudo R^2 is slightly lower). These findings suggests that IPOs in Europe capture a distinct dimension of entrepreneurial success, where M&As represent a more common channel of equity growth.

Finally, we test whether our predictive model is mostly driven by regions with a disproportionate number of growth outcomes or if it has external validity. Table D16 reports goodness-of-fit measures for the entrepreneurial quality model across regions in France, the UK, and Germany, using out-of-sample 5-fold cross-validation at the NUTS-1 level. The table evaluates model performance along three dimensions: (i) the median share of high-growth firms captured within the top 5% and 10% of predicted probabilities for each region, and (ii) the correlation between region-specific and national models.

France (Panel A): Model performance is strongest in Île de France, with 39% of growth events captured in the top 5% and a high correlation (0.94) with the national model. Smaller regions such as Bourgogne-Franche-Comté and Corsica are excluded due to very few growth events, while median capture rates in other regions range from 22% to 57%, with correlations generally moderate (0.30–0.81). The overall regional average shows 37% of growth events in the top 5%, 49% in the top 10%, and a 0.57 correlation with national predictions.

UK (Panel B): Performance is highest in London, capturing 41% of growth events in the top 5% and 52% in the top 10%, with a very high correlation (0.97) with national predictions. Regional medians range from 18% to 41% in the top 5%, and correlations are generally high (0.51–0.97), with lower values in Wales (0.65) and Northern Ireland (0.16). Overall, the UK shows slightly lower median capture rates than France (29% top 5%, 42% top 10%) but a stronger average correlation (0.76).

Germany (Panel C): The model performs well in both large and small regions, with top regions such as Berlin and Bayern capturing 40% and 33% of growth events in the top 5%, respectively. Certain smaller regions (e.g., DEE – Sachsen-Anhalt; DEG – Thüringen) show very high top-5% capture rates (60–67%) but lower correlations with the national model (0.33–0.55), reflecting idiosyncratic regional patterns. On average, German regions achieve 37% of growth events in the top 5%, 51% in the top 10%, and a correlation of 0.61 with national predictions.

Overall, the model reliably identifies high-growth firms in major regions across all three countries, with the strongest alignment between regional and national predictions observed in large metropolitan areas. Smaller or peripheral regions exhibit greater variability, often achieving very high top-5% capture rates but lower correlations, indicating that regional idiosyncrasies can limit the generalizability of the model, particularly in France and Germany, and to a lesser extent in the UK. Nevertheless, average correlations with the national models remain above 0.5, indicating that the model performs reasonably well across regions.

Table D1: Alternative Industry Fixed Effects in the Predictive Model (France)

	Dependent variable:			
	Growth (IPO/M&A within 6 years)			
	(1)	(2)	(3)	(4)
Public Company	19.211***	16.541***	18.513***	18.373***
	(4.158)	(3.612)	(4.018)	(3.985)
Popular Business Structure	3.527***	3.270***	3.451***	3.575***
	(0.251)	(0.235)	(0.247)	(0.254)
Short Name	1.288**	1.239*	1.241*	1.290**
	(0.154)	(0.149)	(0.149)	(0.156)
Eponymous	0.474***	0.515***	0.494***	0.478***
	(0.077)	(0.085)	(0.081)	(0.078)
Patent	15.148***	10.162***	12.330***	11.666***
	(2.622)	(1.934)	(2.260)	(2.118)
Trademark	5.549***	4.811***	5.372***	5.487***
	(1.400)	(1.234)	(1.353)	(1.414)
NACE 1-digit Fixed Effects	YES	NO	NO	NO
NACE 2-digit Effects	NO	YES	NO	NO
Aggregated NACE 2-digit Fixed Effects	NO	NO	YES	NO
OECD High-tech/KIS Fixed Effects	NO	NO	NO	YES
Observations	1,548,392	1,548,392	1,548,392	1,548,392
Pseudo R^2	0.108	0.129	0.115	0.103

Notes: The table reports estimates from logit regressions using the same set of predictors as the baseline entrepreneurial quality model, but with alternative specifications of industry fixed effects. Column (1) includes NACE 1-digit fixed effects; Column (2) includes NACE 2-digit fixed effects (baseline); Column (3) includes an alternative aggregation of NACE 2-digit fixed effects; and Column (4) includes fixed effects based on the OECD High-tech industry/Knowledge-Intensive Services (KIS) classification. Reported coefficients are incidence rate ratios. Robust standard errors are in parentheses. Significance levels: * p<0.1; ** p<0.05; *** p<0.01.

Table D2: Alternative Industry Fixed Effects in the Predictive Model (UK)

	Dependent variable:			
	Growth (IPO/M&A within 6 years)			
	(1)	(2)	(3)	(4)
Public Company	100.334***	90.333***	98.535***	118.940***
	(8.341)	(7.815)	(8.224)	(9.922)
Short Name	1.019	1.007	1.009	1.014
	(0.044)	(0.043)	(0.043)	(0.043)
Eponymous	0.736***	0.757***	0.744***	0.733***
	(0.061)	(0.063)	(0.061)	(0.060)
Patent	12.368***	10.286***	12.132***	12.017***
	(1.528)	(1.327)	(1.530)	(1.525)
Trademark	5.198***	5.077***	5.440***	5.294***
	(1.004)	(0.978)	(1.040)	(1.037)
NACE 1-digit Fixed Effects	YES	NO	NO	NO
NACE 2-digit Effects	NO	YES	NO	NO
Aggregated NACE 2-digit Fixed Effects	NO	NO	YES	NO
OECD High-tech/KIS Fixed Effects	NO	NO	NO	YES
Observations	2,127,382	2,127,382	2,127,382	2,127,382
Pseudo R^2	0.115	0.125	0.115	0.108

Notes: The table reports estimates from logit regressions using the same set of predictors as the baseline entrepreneurial quality model, but with alternative specifications of industry fixed effects. Column (1) includes NACE 1-digit fixed effects; Column (2) includes NACE 2-digit fixed effects (baseline); Column (3) includes an alternative aggregation of NACE 2-digit fixed effects; and Column (4) includes fixed effects based on the OECD High-tech industry/Knowledge-Intensive Services (KIS) classification. Reported coefficients are incidence rate ratios. Robust standard errors are in parentheses. Significance levels: * p<0.1; ** p<0.05; *** p<0.01.

Table D3: Alternative Industry Fixed Effects in the Predictive Model (Germany)

	Dependent variable:			
	Growth (IPO/M&A within 6 years)			
	(1)	(2)	(3)	(4)
Public Company	22.487***	21.106***	22.202***	23.223***
	(2.484)	(2.342)	(2.466)	(2.598)
Popular Business Structure	3.202***	3.187***	3.218***	3.081***
	(0.225)	(0.225)	(0.226)	(0.215)
Short Name	1.217***	1.198***	1.197***	1.209***
	(0.084)	(0.083)	(0.082)	(0.083)
Eponymous	0.159***	0.172***	0.165***	0.157***
	(0.028)	(0.031)	(0.029)	(0.028)
Patent	6.857***	6.047***	6.613***	6.371***
	(0.559)	(0.507)	(0.541)	(0.528)
Trademark	4.182***	3.991***	4.147***	4.146***
	(0.555)	(0.532)	(0.551)	(0.554)
NACE 1-digit Fixed Effects	YES	NO	NO	NO
NACE 2-digit Effects	NO	YES	NO	NO
Aggregated NACE 2-digit Fixed Effects	NO	NO	YES	NO
OECD High-tech/KIS Fixed Effects	NO	NO	NO	YES
Observations	922,702	922,702	922,702	922,702
Pseudo R^2	0.107	0.119	0.109	0.099

Notes: The table reports estimates from logit regressions using the same set of predictors as the baseline entrepreneurial quality model, but with alternative specifications of industry fixed effects. Column (1) includes NACE 1-digit fixed effects; Column (2) includes NACE 2-digit fixed effects (baseline); Column (3) includes an alternative aggregation of NACE 2-digit fixed effects; and Column (4) includes fixed effects based on the OECD High-tech industry/Knowledge-Intensive Services (KIS) classification. Reported coefficients are incidence rate ratios. Robust standard errors are in parentheses. Significance levels: * $p<0.1$; ** $p<0.05$; *** $p<0.01$.

Table D4: Growth Predictive Model with Regional Fixed Effects - Logit Regression on IPO or Acquisition within Six Years - France

	Preliminary Models			Nowcasting	Full Model
	(1)	(2)	(3)	(4)	(5)
Corporate governance measures					
Public Company	29.348***			17.669***	14.978***
	(6.277)			(3.943)	(3.275)
Popular Business Structure	4.508***			3.206***	3.135***
	(0.315)			(0.232)	(0.228)
Name-based measures					
Short Name		1.420***		1.252*	1.227*
		(0.170)		(0.151)	(0.148)
Eponymous		0.383***		0.527***	0.529***
		(0.062)		(0.087)	(0.087)
Intellectual property measures					
Patent			26.311***		9.707***
			(4.605)		(1.840)
Trademark			8.223***		4.727***
			(2.127)		(1.203)
Industry Fixed Effects	NO	NO	NO	YES	YES
Region Fixed Effects	YES	YES	YES	YES	YES
Observations	1,548,392	1,548,392	1,548,392	1,548,392	1,548,392
Pseudo R^2	0.053	0.022	0.032	0.129	0.136

Notes: Significance levels: *p<0.1; **p<0.05; ***p<0.01. Robust standard errors in parentheses. The predictive growth model is a logit regression where Growth is a binary variable equal to 1 if a firm achieves IPO or acquisition within 6 years, and 0 otherwise.

Table D5: Growth Predictive Model with Regional Fixed Effects - Logit Regression on IPO or Acquisition within Six Years - UK

	Preliminary Models			Nowcasting	Full Model
	(1)	(2)	(3)	(4)	(5)
Corporate governance measures					
Public Company	149.465***			93.135***	88.766***
	(10.064)			(8.076)	(7.789)
Name-based measures					
Short Name		1.068		1.032	1.008
		(0.045)		(0.044)	(0.044)
Eponymous		0.679***		0.747***	0.759***
		(0.055)		(0.062)	(0.063)
Intellectual property measures					
Patent			23.047***		10.302***
			(2.554)		(1.328)
Trademark			7.123***		5.143***
			(1.299)		(0.990)
Industry Fixed Effects	NO	NO	NO	YES	YES
Region Fixed Effects	YES	YES	YES	YES	YES
Observations	2,127,382	2,127,382	2,127,382	2,127,382	2,127,382
Pseudo R^2	0.062	0.002	0.016	0.118	0.127

Notes: Significance levels: $^{*}p<0.1$; $^{**}p<0.05$; $^{***}p<0.01$. Robust standard errors in parentheses. The predictive growth model is a logit regression where Growth is a binary variable equal to 1 if a firm achieves IPO or acquisition within 6 years, and 0 otherwise.

Table D6: Growth Predictive Model with Regional Fixed Effects - Logit Regression on IPO or Acquisition within Six Years - Germany

	Preliminary Models			Nowcasting	Full Model
	(1)	(2)	(3)	(4)	(5)
Corporate governance measures					
Public Company	25.688***			23.974***	20.636***
	(2.813)			(2.616)	(2.299)
Popular Business Structure	3.270***			3.365***	3.152***
	(0.228)			(0.237)	(0.223)
Name-based measures					
Short Name		1.482***		1.284***	1.217***
		(0.101)		(0.089)	(0.084)
Eponymous		0.146***		0.181***	0.181***
		(0.026)		(0.032)	(0.032)
Intellectual property measures					
Patent			10.085***		6.111***
			(0.815)		(0.515)
Trademark			6.640***		3.737***
			(0.878)		(0.499)
Industry Fixed Effects	NO	NO	NO	YES	YES
Region Fixed Effects	YES	YES	YES	YES	YES
Observations	922,702	922,702	922,702	922,702	922,702
Pseudo R^2	0.043	0.023	0.042	0.106	0.125

Notes: Significance levels: *p<0.1; **p<0.05; ***p<0.01. Robust standard errors in parentheses. The predictive growth model is a logit regression where Growth is a binary variable equal to 1 if a firm achieves IPO or acquisition within 6 years, and 0 otherwise.

Table D7: Predictive Model with Region Fixed Effects and Region-Specific Time Trends (France)

	Dependent variable:			
	Growth (IPO/M&A within 6 years)			
	(1)	(2)	(3)	(4)
Corporate governance measures				
Public Company	14.978***	15.536***	15.507***	15.507***
	(3.275)	(3.386)	(3.400)	(3.400)
Popular Business Structure	3.135***	3.775***	3.780***	3.780***
	(0.228)	(0.310)	(0.311)	(0.311)
Name-based measures				
Short Name	1.227*	1.258*	1.257*	1.257*
	(0.148)	(0.151)	(0.151)	(0.151)
Eponymous	0.529***	0.536***	0.536***	0.536***
	(0.087)	(0.088)	(0.088)	(0.088)
Intellectual property measures				
Patent	9.707***	8.834***	8.807***	8.807***
	(1.840)	(1.684)	(1.690)	(1.690)
Trademark	4.727***	4.477***	4.445***	4.445***
	(1.203)	(1.142)	(1.138)	(1.138)
Industry Fixed Effects	YES	YES	YES	YES
Region Fixed Effects	YES	YES	YES	YES
Year Fixed Effects	NO	YES	NO	YES
Region Trends	NO	NO	YES	YES
Observations	1,532,913	1,532,913	1,470,533	1,470,533
Pseudo R^2	0.135	0.139	0.142	0.142

Notes: The table reports estimates from logit regressions using the same set of predictors as the baseline entrepreneurial quality model, with alternative specifications of year and region effects. Column (1) includes industry and region fixed effects (NUTS 1 level). Column (2) additionally includes year fixed effects. Column (3) replaces year fixed effects with region-specific time trends. Column (4) includes both year fixed effects and region-specific time trends. Reported coefficients are incidence rate ratios. Robust standard errors are reported in parentheses. Significance levels: * p<0.1; ** p<0.05; *** p<0.01.

Table D8: Predictive Model with Region Fixed Effects and Region-Specific Time Trends (UK)

	Dependent variable:			
	Growth (IPO/M&A within 6 years)			
	(1)	(2)	(3)	(4)
Corporate governance measures				
Public Company	88.766***	87.547***	88.055***	88.055***
	(7.789)	(7.662)	(7.706)	(7.706)
Name-based measures				
Short name	1.008	1.011	1.012	1.012
	(0.044)	(0.044)	(0.044)	(0.044)
Eponymous	0.759***	0.755***	0.755***	0.755***
	(0.063)	(0.062)	(0.063)	(0.063)
Intellectual property measures				
Patent	10.302***	10.207***	10.202***	10.202***
	(1.328)	(1.321)	(1.321)	(1.321)
Trademark	5.143***	5.318***	5.320***	5.320***
	(0.990)	(1.018)	(1.017)	(1.017)
Industry Fixed Effects	YES	YES	YES	YES
Region Fixed Effects	YES	YES	YES	YES
Year Fixed Effects	NO	YES	NO	YES
Region Trends	NO	NO	YES	YES
Observations	2,124,983	2,124,983	2,124,983	2,124,983
Pseudo R^2	0.126	0.129	0.131	0.131

Notes: The table reports estimates from logit regressions using the same set of predictors as the baseline entrepreneurial quality model, with alternative specifications of year and region effects. Column (1) includes industry and region fixed effects (NUTS 1 level). Column (2) additionally includes year fixed effects. Column (3) replaces year fixed effects with region-specific time trends. Column (4) includes both year fixed effects and region-specific time trends. Reported coefficients are incidence rate ratios. Robust standard errors are reported in parentheses. Significance levels: * p<0.1; ** p<0.05; *** p<0.01.

Table D9: Predictive Model with Region Fixed Effects and Region-Specific Time Trends (Germany)

	Dependent variable:			
	Growth (IPO/M&A within 6 years)			
	(1)	(2)	(3)	(4)
Corporate governance measures				
Public Company	20.636***	20.782***	21.109***	21.109***
	(2.299)	(2.318)	(2.350)	(2.350)
Popular Business Structure	3.152***	3.155***	3.154***	3.154***
	(0.223)	(0.223)	(0.223)	(0.223)
Name-based measures				
Short Name	1.217***	1.211***	1.212***	1.212***
	(0.084)	(0.084)	(0.084)	(0.084)
Eponymous	0.181***	0.181***	0.180***	0.180***
	(0.032)	(0.032)	(0.032)	(0.032)
Intellectual property measures				
Patent	6.111***	6.126***	6.175***	6.175***
	(0.515)	(0.516)	(0.519)	(0.519)
Trademark	3.737***	3.721***	3.764***	3.764***
	(0.499)	(0.498)	(0.503)	(0.503)
Industry Fixed Effects	YES	YES	YES	YES
Region Fixed Effects	YES	YES	YES	YES
Year Fixed Effects	NO	YES	NO	YES
Region Trends	NO	NO	YES	YES
Observations	915,848	915,848	890,614	890,614
Pseudo R^2	0.124	0.125	0.128	0.128

Notes: The table reports estimates from logit regressions using the same set of predictors as the baseline entrepreneurial quality model, with alternative specifications of year and region effects. Column (1) includes industry and region fixed effects (NUTS 1 level). Column (2) additionally includes year fixed effects. Column (3) replaces year fixed effects with region-specific time trends. Column (4) includes both year fixed effects and region-specific time trends. Reported coefficients are incidence rate ratios. Robust standard errors are reported in parentheses. Significance levels: * p<0.1; ** p<0.05; *** p<0.01.

Table D10: Entrepreneurial Quality Models with Alternative Firm Size Growth Outcomes - France

	Full Model	Medium-sized Firm	Large-sized Firm	Employment > 100	Employment > 250
	(1)	(2)	(3)	(4)	(5)
Corporate governance measures					
Public Company	16.541***	21.191***	29.885***	16.660***	16.884***
	(3.612)	(1.317)	(2.786)	(2.306)	(3.154)
Popular Business Structure	3.270***	4.156***	6.187***	4.385***	4.096***
	(0.235)	(0.070)	(0.226)	(0.186)	(0.266)
Name-based measures					
Short Name	1.239*	0.605***	0.558***	0.709***	0.751***
	(0.149)	(0.013)	(0.023)	(0.037)	(0.063)
Eponymous	0.515***	0.912***	0.903*	0.897*	0.874
	(0.085)	(0.024)	(0.048)	(0.059)	(0.090)
Intellectual property measures					
Patent	10.162***	13.888***	20.170***	17.498***	20.385***
	(1.934)	(0.919)	(1.897)	(1.909)	(2.920)
Trademark	4.811***	5.209***	5.876***	4.857***	5.306***
	(1.234)	(0.451)	(0.773)	(0.823)	(1.252)
Industry Fixed Effects	YES	YES	YES	YES	YES
Observations	1,548,392	1,548,392	1,548,392	1,548,392	1,548,392
Pseudo R^2	0.129	0.171	0.223	0.165	0.167

Notes: This table reports logit regression estimates using the same set of predictors as the full-information entrepreneurial quality model for France, but replacing equity growth with alternative measures of firm size growth. Column (1) presents the full model specification. Columns (2)–(3) use thresholds for medium- and large-sized firms, as defined by the European Commission. Columns (4)–(5) consider firm size thresholds of more than 100 and 250 employees, respectively. Reported coefficients are incidence rate ratios. All models include industry fixed effects based on 2-digit NACE codes. Robust standard errors are reported in parentheses. Significance levels: $^{*}p<0.1$; $^{**}p<0.05$; $^{***}p<0.01$.

Table D11: Entrepreneurial Quality Models with Alternative Firm Size Growth Outcomes - UK

	Full Model	Medium-sized Firm	Large-sized Firm	Employment > 100	Employment > 250
	(1)	(2)	(3)	(4)	(5)
Corporate governance measures					
Public Company	90.333***	16.007***	23.167***	7.777***	11.513***
	(7.815)	(1.112)	(1.832)	(0.895)	(1.532)
Name-based measures					
Short Name	1.007	0.695***	0.571***	0.800***	0.789***
	(0.043)	(0.010)	(0.013)	(0.022)	(0.033)
Eponymous	0.757***	0.596***	0.467***	0.545***	0.499***
	(0.063)	(0.016)	(0.023)	(0.031)	(0.045)
Intellectual property measures					
Patent	10.286***	11.608***	16.904***	16.177***	17.771***
	(1.327)	(0.655)	(1.263)	(1.351)	(2.027)
Trademark	5.077***	6.224***	6.997***	7.582***	8.753***
	(0.978)	(0.495)	(0.789)	(0.942)	(1.417)
Industry Fixed Effects	YES	YES	YES	YES	YES
Observations	2,127,382	2,127,382	2,127,382	2,127,382	2,127,382
Pseudo R^2	0.125	0.154	0.184	0.130	0.140

Notes: This table reports logit regression estimates using the same set of predictors as the full-information entrepreneurial quality model for the United Kingdom , but replacing equity growth with alternative measures of firm size growth. Column (1) presents the full model specification. Columns (2)–(3) use thresholds for medium- and large-sized firms, as defined by the European Commission. Columns (4)–(5) consider firm size thresholds of more than 100 and 250 employees, respectively. Reported coefficients are incidence rate ratios. All models include industry fixed effects based on 2-digit NACE codes. Robust standard errors are reported in parentheses. Significance levels: *p<0.1; **p<0.05; ***p<0.01.

Table D12: Entrepreneurial Quality Models with Alternative Firm Size Growth Outcomes - Germany

	Full Model	Medium-sized Firm	Large-sized Firm	Employment > 100	Employment > 250
	(1)	(2)	(3)	(4)	(5)
Corporate governance measures					
Public Company	21.106***	2.886***	3.985***	7.753***	9.478***
	(2.342)	(0.142)	(0.295)	(0.629)	(0.982)
Popular Business Structure	3.187***	1.418***	1.307***	2.385***	2.203***
	(0.225)	(0.018)	(0.034)	(0.076)	(0.110)
Name-based measures					
Short Name	1.198***	0.797***	0.728***	0.877***	0.860***
	(0.083)	(0.010)	(0.020)	(0.026)	(0.039)
Eponymous	0.172***	0.715***	0.450***	0.766***	0.677***
	(0.031)	(0.014)	(0.021)	(0.032)	(0.045)
Intellectual property measures					
Patent	6.047***	5.987***	11.813***	10.442***	15.210***
	(0.507)	(0.163)	(0.467)	(0.442)	(0.846)
Trademark	3.991***	2.111***	1.782***	2.052***	1.760***
	(0.532)	(0.136)	(0.214)	(0.227)	(0.290)
Industry Fixed Effects	YES	YES	YES	YES	YES
Observations	922,702	922,702	922,702	922,702	922,702
Pseudo R^2	0.119	0.085	0.112	0.133	0.141

Notes: This table reports logit regression estimates using the same set of predictors as the full-information entrepreneurial quality model for Germany, but replacing equity growth with alternative measures of firm size growth. Column (1) presents the full model specification. Columns (2)–(3) use thresholds for medium- and large-sized firms, as defined by the European Commission. Columns (4)–(5) consider firm size thresholds of more than 100 and 250 employees, respectively. Reported coefficients are incidence rate ratios. All models include industry fixed effects based on 2-digit NACE codes. Robust standard errors are reported in parentheses. Significance levels: *p<0.1; **p<0.05; ***p<0.01.

Table D13: Alternative implementations of Predictive Model - France

	Full Model	IPO in 6 Years	Growth in 9 Years	IPO (Ever)
	(1)	(2)	(3)	(4)
Public Company	16.541***		17.749***	
	(3.612)		(3.607)	
Popular Business Structure	3.270***		4.640***	
	(0.235)		(0.315)	
Short Name	1.239*	1.189	1.278**	1.367
	(0.149)	(0.566)	(0.145)	(0.553)
Eponymous	0.515***	0.170*	0.553***	0.597
	(0.085)	(0.174)	(0.085)	(0.286)
Patent	10.162***	41.868***	7.830***	23.267***
	(1.934)	(17.390)	(1.445)	(9.061)
Trademark	4.811***	3.823*	4.191***	6.653***
	(1.234)	(3.295)	(1.051)	(4.207)
Industry Fixed Effects	YES	YES	YES	YES
Observations	1,548,392	1,548,392	985,369	1,548,392
Pseudo R^2	0.129	0.216	0.143	0.182

Notes: This table reports estimates from logit regressions using the same set of predictors as the full model. Column (1) reports results for the baseline specification. Column (2) defines the dependent variable as equal to 1 if the firm has an IPO within six years of incorporation. Column (3) defines the dependent variable as equal to 1 if the firm reaches high growth (IPO or M&A) within nine years. Column (4) defines the dependent variable as equal to 1 if the firm ever has an IPO. Reported coefficients are incidence rate ratios. Robust standard errors are in parentheses. Significance levels: *p<0.1; **p<0.05; ***p<0.01.

Table D14: Alternative implementations of Predictive Model - UK

	Full Model	IPO in 6 Years	Growth in 9 Years	IPO (Ever)
	(1)	(2)	(3)	(4)
Public Company	90.333***		82.833***	
	(7.815)		(10.034)	
Short Name	1.007	1.141	0.937	1.112
	(0.043)	(0.124)	(0.052)	(0.118)
Eponymous	0.757***	0.350***	0.814**	0.360***
	(0.063)	(0.103)	(0.083)	(0.102)
Patent	10.286***	15.988***	8.657***	15.845***
	(1.327)	(3.699)	(1.684)	(3.610)
Trademark	5.077***	9.657***	4.189***	9.191***
	(0.978)	(3.047)	(1.374)	(2.885)
Industry Fixed Effects	YES	YES	YES	YES
Observations	2,127,382	2,127,382	744,629	2,127,382
Pseudo R^2	0.125	0.155	0.118	0.154

Notes: This table reports estimates from logit regressions using the same set of predictors as the full model. Column (1) reports results for the baseline specification. Column (2) defines the dependent variable as equal to 1 if the firm has an IPO within six years of incorporation. Column (3) defines the dependent variable as equal to 1 if the firm reaches high growth (IPO or M&A) within nine years. Column (4) defines the dependent variable as equal to 1 if the firm ever has an IPO. Reported coefficients are incidence rate ratios. Robust standard errors are in parentheses. Significance levels: *p<0.1; **p<0.05; ***p<0.01.

Table D15: Alternative implementations of Predictive Model - Germany

	Full Model	IPO in 6 Years	Growth in 9 Years	IPO (Ever)
	(1)	(2)	(3)	(4)
Public Company	21.106***		18.850***	
	(2.342)		(2.203)	
Popular Business Structure	3.187***		3.164***	
	(0.225)		(0.230)	
Short Name	1.198***	0.759***	1.363***	0.813***
	(0.083)	(0.181)	(0.101)	(0.186)
Eponymous	0.172***	0.125	0.185***	0.166*
	(0.031)	(0.091)	(0.034)	(0.098)
Patent	6.047***	9.427***	6.619***	10.057***
	(0.507)	(2.990)	(0.557)	(2.900)
Trademark	3.991***	8.243**	4.404***	8.683**
	(0.532)	(4.183)	(0.613)	(4.048)
Industry Fixed Effects	YES	YES	YES	YES
Observations	922,702	922,702	615,172	922,702
Pseudo R^2	0.119	0.136	0.125	0.138

Notes: This table reports estimates from logit regressions using the same set of predictors as the full model. Column (1) reports results for the baseline specification. Column (2) defines the dependent variable as equal to 1 if the firm has an IPO within six years of incorporation. Column (3) defines the dependent variable as equal to 1 if the firm reaches high growth (IPO or M&A) within nine years. Column (4) defines the dependent variable as equal to 1 if the firm ever has an IPO. Reported coefficients are incidence rate ratios. Robust standard errors are in parentheses. Significance levels: *p<0.1; **p<0.05; ***p<0.01.

Table D16: Goodness of Fit Measures of Entrepreneurial Quality Model Across Regions

Region	Total Growth Events	Share in Top 5%: Median	Share in Top 10%: Median	Correlation Regional and National Models
Panel A: France				
FR1 - Île de France	736	0.39	0.55	0.94
FRB - Centre-Val de Loire	19	0.33	0.33	0.30
FRC - Bourgogne-Franche-Comté	15	-	-	-
FRD - Normandy	18	0.50	0.50	0.32
FRE - Hauts-de-France	61	0.22	0.50	0.48
FRF - Grand Est	38	0.33	0.57	0.43
FRG - Pays de la Loire	45	0.30	0.40	0.53
FRH - Brittany	48	0.57	0.57	0.68
FRI - Nouvelle-Aquitaine	56	0.27	0.55	0.67
FRJ - Occitania	57	0.38	0.38	0.47
FRK - Auvergne-Rhône-Alpes	147	0.33	0.58	0.81
FRL - Provence-Alpes-Côte d'Azur	80	0.40	0.46	0.65
FRM - Corsica	1	-	-	-
Average		0.37	0.49	0.57
Panel B: UK				
UKC - North East (England)	105	0.23	0.36	0.74
UKD - North West (England)	468	0.27	0.41	0.91
UKE - Yorkshire and the Humber	293	0.30	0.40	0.91
UKF - East Midlands (England)	186	0.24	0.43	0.77
UKG - West Midlands (England)	283	0.37	0.46	0.87
UKH - East of England	292	0.38	0.45	0.86
UKI - London	1706	0.41	0.52	0.97
UKJ - South East (England)	566	0.37	0.44	0.96
UKK - South West (England)	242	0.28	0.33	0.84
UKL - Wales	127	0.18	0.27	0.65
UKM - Scotland	209	0.28	0.44	0.51
UKN - Northern Ireland	40	0.20	0.50	0.16
Average		0.29	0.42	0.76
Panel C: Germany				
DE1 - Baden-Württemberg	209	0.32	0.48	0.74
DE2 - Bayern	472	0.33	0.45	0.91
DE3 - Berlin	332	0.40	0.50	0.79
DE4 - Brandenburg	27	0.25	0.67	0.34
DE5 - Bremen	17	0.50	0.50	0.53
DE6 - Hamburg	164	0.25	0.32	0.81
DE7 - Hessen	178	0.24	0.42	0.73
DE8 - Mecklenburg-Vorpommern	22	0.33	0.33	0.23
DE9 - Niedersachsen	98	0.28	0.45	0.54
DEA - Nordrhein-Westfalen	348	0.36	0.50	0.88
DEB - Rheinland-Pfalz	46	0.36	0.50	0.50
DEC - Saarland	6	-	-	-
DED - Sachsen	47	0.38	0.62	0.76
DEE - Sachsen-Anhalt	21	0.60	0.83	0.33
DEF - Schleswig-Holstein	46	0.23	0.40	0.53
DEG - Thüringen	27	0.67	0.71	0.55
Average		0.37	0.51	0.61

Notes: This table presents goodness-of-fit measures for predicted entrepreneurial quality across regions. Predictions are obtained via an out-of-sample 5-fold cross-validation for each NUTS-1 region. "Total Growth Events" counts firms experiencing an IPO or M&A within six years of registration. "Share in Top 5% / Top 10%: Median" reports the median share of growth events captured among firms ranked in the top 5% or 10% of predicted probabilities. "Correlation Regional and National Models" shows the Pearson correlation between region-specific models' predictions and national predictions. Regions with 15 or fewer growth events are omitted.

Download ZEW Discussion Papers:

https://www.zew.de/en/publications/zew-discussion-papers

or see:

https://www.ssrn.com/link/ZEW-Ctr-Euro-Econ-Research.html
https://ideas.repec.org/s/zbw/zewdip.html

//
IMPRINT

ZEW – Leibniz-Zentrum für Europäische Wirtschaftsforschung GmbH Mannheim

ZEW – Leibniz Centre for European Economic Research

L 7,1 · 68161 Mannheim · Germany
Phone +49 621 1235-01
info@zew.de · zew.de