Artificial Intelligence Adoption and the Direction of Innovation

Xin Deng
Alliance Manchester Business School
xin.deng@manchester.ac.uk

Maria Garcia-Vega
School of Economics, University of Nottingham
maria.garcia-vega@nottingham.ac.uk

Cher Li
Aston Business School
c.li@aston.ac.uk

Richard Upward
School of Economics, University of Nottingham
richard.upward@nottingham.ac.uk

June 9, 2026
DRAFT

Abstract

We study how the adoption of predictive artificial intelligence (AI) - the foundational technology in our sample and still the primary driver of industrial impact - reshapes firms' innovation strategies by shifting the balance between exploitation and exploration. Using UK firm-level patent data and the near-universe of online job vacancies from 20172021, we map these strategies to patenting activity by defining exploitation as innovation within familiar technological domains and exploration as entry into entirely new technology classes. We find that AI adoption increases firms overall patenting but shifts innovation portfolios towards familiar domains (exploitation), with little evidence of systematic entry into new technology classes (exploration). We interpret this as suggesting predictive AI reducing the cost of optimisation and local search in data-rich areas while raising the opportunity cost of exploration in data-scarce domains. These effects are strongest for small firms, for adoption through managerial occupations, and in sectors with high demand volatility. By contrast, adoption through scientific and technical occupations is more strongly associated with patenting in new domains. Overall, our research suggests that in the pre-generative era, AI primarily deepened incumbents' established technological trajectories rather than broadening the scope of inventive search, with important implications for innovation strategy and policy.

JEL classification: L25; L20; O31; O32; O33

Keywords: Artificial Intelligence; Predictive AI; Innovation Direction; Exploitation and Exploration

1 Introduction

Artificial intelligence (AI) is increasingly integrated into firms' core decision systems, from demand forecasting and pricing to hiring, resource allocation and R&D activities. A large and fast-growing empirical literature documents that AI adoption is associated with higher productivity, faster firm growth and greater innovative output, often measured through product-related outcomes such as trademarks, product launches and product-oriented patents (Alekseeva et al., 2020, 2024; Babina et al., 2024; **?**; Brynjolfsson et al., 2024; Czarnitzki et al., 2023; **?**; Engberg et al., 2025; Igna and Venturini, 2023; Kanazawa et al., 2025; Rammer et al., 2022). While these studies establish the performance benefits of AI, **?** propose a deeper theoretical framing, characterising neural network-based technologies as an emerging general method of invention that reshapes the knowledge production function by identifying valuable new combinations. This raises a central question for innovation management and firm growth: whether AI fundamentally alters the *direction* of technological change within firms, either by promoting exploration into unfamiliar technological domains or by deepening innovation along established research trajectories.

Understanding whether AI adoption shifts the technological direction of firms inventive activity matters because innovation is cumulative. Firms' inventions build on prior knowledge and organisational capabilities, and the balance between exploitation and exploration has important implications for long-run learning, competitive dynamics and the rate at which new technological pathways emerge (Abernathy and Clark, 1985; Manso, 2011; Manso et al., 2023; March, 1991).

In this paper, we provide new evidence on how the adoption of AI technologies by firms affects both the extent and the direction of innovation. We do this by combining UK firm-level patent data with the near-universe of online job vacancies from 20172021, allowing us to connect job postings requiring AI-related skills with subsequent patenting activity. Our analysis focuses on predictive AI (Agrawal et al., 2018; Fontanelli, 2025), a set of technologies already widely adopted by firms, including applications such as demand forecasting, inventory optimisation, predictive maintenance, dynamic pricing, fraud detection, credit scoring and customer personalisation. In operations and R&D, these systems are also used for monitoring, experimentation, quality control and improvements in products and production processes. While we do not formally distinguish predictive from generative AI, focusing on the pre-generative AI period allows us to provide a clean test of the effects of predictive AI adoption on innovation outcomes, without confounding influences from the subsequent diffusion of generative AI technologies. Note that predictive AI remains a core component of the broader AI landscape.

The theoretical arguments for the effects of AI on innovation offer contrasting predictions. On the one hand, AIs predictive and optimisation capabilities make these technologies especially effective at improving established routines, enabling firms to extract value from existing knowledge bases (Agrawal et al., 2018; Johnson et al., 2022; Liu et al., 2020). For instance, machine-learning tools are widely deployed to forecast demand, coordinate supply chains and production, optimise operations and tailor products and prices using customer analytics. For example, Amazon applies predictive models to anticipate purchases; Toyota uses AI to improve production scheduling; Bosch employs AI to enhance operational efficiency; and Nike analyses customer behaviour to personalise product offerings. These applications thrive in data-rich environments and they naturally support incremental improvements to existing products and processes (**?**).

An implication is that AI might induce a *data-visibility bias* in inventive search. If AI makes the expected payoffs of research projects close to a firms existing knowledge base more predictable and less risky, the opportunity cost of exploring data-scarce domains might increase. This mechanism is closely related to the "streetlight" effect in data-driven exploration: search tends to concentrate in areas where data are most readily available (Hoelzemann et al., 2025). Consequently, AI may reinforce existing search trajectories by shifting innovation efforts toward domains that are already known to the firm.

On the other hand, AI may also facilitate exploration. A growing body of work argues that AI can be understood as a general-purpose technology that reshapes the innovation process itself (Agrawal et al., 2019; Brynjolfsson et al., 2019; Cockburn et al., 2019). By lowering the cost of analysing adjacent knowledge and screening large combinatorial spaces, AI can reduce barriers to experimentation and help firms operate beyond their core technological expertise. For example, Merck uses AI to analyse genomic and real-world data in drug development, and BASF applies AI-enabled methods to accelerate materials discovery. Breakthrough systems such as DeepMinds AlphaFold similarly illustrate how predictive AI can expand scientific capabilities by forecasting protein structures at scale, enabling advances in biologics. These cases highlight AIs potential to support innovation at the intersection of technological domains, a key source of new invention (Fleming, 2001; Makri et al., 2006; Nesta and Saviotti, 2005). Recent theory suggests that AIs effect on innovation depends on how it reshapes research incentives (Gans, 2025): when AI primarily improves interpolation within dense regions of existing knowledge, it may reinforce incremental search; but when AI expands the feasible set of experiments and recombinations, it may shift effort toward more exploratory directions, with ambiguous implications for long-run growth.[1] Thus, whether AI primarily supports innovation in familiar domains or enables exploration into new ones remains an open empirical question.

Our study provides new evidence on how AI shapes the direction of firm-level innovation. We combine firm-level patent data with the near-universe of online job vacancies from 2017-2021 in the UK, which we use to measure organisational adoption of AI-related skill. Job postings provide a timely signal of adoption and indicate *where* inside the firm AI capabilities are being built (Acemoglu et al., 2022; Alekseeva et al., 2021; Babina et al., 2024). We classify vacancies as AI-related when they require skills

[1] For a broad overview of how AI tools are being deployed across the research and innovation pipeline, see Chen et al. (2025).

associated with AI techniques (e.g., machine learning, natural language processing) and construct firm-year measures of AI adoption based on entry (first AI vacancy) and intensity (share of AI vacancies). This new dataset provides a link from firms' adoption of AI, as measured by relevant job postings, to firm-level measures of innovation.

To measure innovation direction, we distinguish between patents in technological domains where the firm has prior experience and patents in domains that are new to the firm (Balsmeier et al., 2017, 2024; Gao et al., 2018). Specifically, a patent is known if it falls in a 3-digit IPC class where the firm patented in the previous five years, and new otherwise. This mapping closely corresponds to exploitation versus exploration in firms' inventive search.

Our empirical strategy combines entropy balancing (Hainmueller, 2012) with difference-in-differences designs, including staggered adoption estimators that account for continuous treatment adoption as well as potential lagged effects (De Chaisemartin and d'Haultfoeuille, 2024), to compare changes in patenting outcomes between AI-adopting and non-adopting firms.

As a robustness check, we employ alternative measures of technological novelty and search direction. First, we construct a technology proximity measure following Jaffe (1986), which captures the cosine similarity between a firm's current patent portfolio and its accumulated portfolio. Second, we use citation-based exploration and exploitation measures following (Custódio et al., 2019; Levine et al., 2020; He and Hirshleifer, 2022). These measures classify patents based on the extent to which their backward citations draw on the firm's recent knowledge base. Furthermore, we examine whether results differ across the intensive and extensive margins of innovation. In addition, we perform a series of robustness checks on our empirical strategy. We verify that our results are not sensitive to the choice of sample period by considering both shorter and longer time windows. To address potential differences between adopters and non-adopters, we re-estimate our models using only not-yet-treated firms as controls. We also employ an alternative balancing approach based on propensity score reweighting combined with a difference-in-differences framework.

Several findings stand out: For the average firm, AI adoption increases firms overall patenting, but the additional patents are overwhelmingly concentrated in known technological domains. In contrast, we find little evidence that Predictive AI induces systematic entry into new technology classes. We interpret this "null" on new-domain patenting not as a failure of invention, but as a rational portfolio response to the incentives created by predictive technologies: AI reduces uncertainty and lowers the cost of local search and optimisation in data-rich domains, thereby raising the opportunity cost of exploration in data-sparse regions of the knowledge space. Our evidence supports the view that, for the average firm, Predictive AI operates primarily as connecting and refining near knowledge. This is line with Gans (2025) who considers that AI works as a local interpolation technology when AIs coverage is limited.

The average effect masks important sources of heterogeneity that clarify how AI interacts with organisational structure, complementary assets and the external environment. First, the organisational locus of AI adoption shapes innovation direction. When AI-related skills enter through managerial vacancies, our findings show that patenting increases mainly in familiar technological domains. This pattern suggests that AI in managerial roles predominantly improves project evaluation, prioritisation and coordination, core decision-making functions that rely on prediction and information processing as suggested by Babina et al. (2024). By mitigating managerial myopia and information-processing constraints (Lou et al., 2025), AI may sharpen resource allocation toward projects that are closer to the firm's existing knowledge base, thereby reinforcing exploitation. In contrast, when AI skills enter through scientific and technical vacancies, we observe increases in both known- and new-domain patenting. Embedding AI within R&D functions is more consistent with the "method of invention" perspective, expanding the space of feasible experimentation and recombination (Cockburn et al., 2019). In this configuration, AI directly augments research tasks and might mitigate bottlenecks in the knowledge production process (Jones, 2025), enabling broader search across adjacent and moderately distant domains and supporting both exploitation and exploration.

Second, firm size conditions the extent to which AI can be leveraged for both exploitation and exploration. Larger firms are more likely to translate AI adoption into both exploitation and exploration. This finding aligns with evidence that AI-powered growth is concentrated among firms with substantial scale and complementary assets (Babina et al., 2024; Igna and Venturini, 2023). From the perspective of AI as a general-purpose research tool, its effectiveness depends on access to proprietary data, computational infrastructure and organisational coordination (Cockburn et al., 2019). Larger firms are better positioned to internalise these complementarities and to deploy AI across multiple units and projects, thereby broadening the effective range over which AI can support both incremental refinement and entry into new domains.

Finally, market environments shape how firms allocate AI across exploration and exploitation. We find that, in sectors with high demand volatility, AI adoption is associated with stronger shifts towards exploitation. When uncertainty about short-run demand is high, improvements in forecasting and rapid feedback generate particularly high marginal returns, as accurate predictions of consumer behaviour, competitor pricing and demand fluctuations become more valuable. As a prediction technology, AI enables firms to learn rapidly from large datasets and update beliefs in real time, reducing uncertainty surrounding near-term decisions. Under such conditions, firms may rationally deploy AI to refine and optimise existing products, adjust production and pricing strategies, and respond swiftly to market fluctuations rather than commit resources to distant exploratory projects with longer and more uncertain payoffs (Lavie et al., 2010).

Our research contributes to the economics of AI and innovation management. Empirically, we provide new evidence that AI adoption reshapes not only the level of inventive activity but also its composition, tilting innovation portfolios towards exploitation in the Predictive AI era. Importantly, this directional effect is conditional. Its magnitude and sign depend on organisational embedding, complementary assets and market context. Specifically, while we find that, for the average firm, AI strengthens incremental innovation in familiar domains, its ability to stimulate exploratory innovation can emerge when firms possess strong complementary capabilities, most notably when AI is adopted within R&D-intensive technical occupations, among large firms with extensive data and organisational resources, and in less volatile market environments.

Conceptually, we interpret these patterns as reflecting that by lowering the cost and uncertainty of prediction in data-rich environments, AI makes innovation in nearby technological domains more attractive. This can generate a data-visibility bias in inventive search and change the direction of innovation. Finally, methodologically, we show how text-based measures of AI skill adoption from vacancy data can be integrated with patent portfolios to analyse organisational AI adoption and innovation strategy at scale, offering a scalable approach to studying how emerging technologies shape innovation composition.

The rest of the paper is organised as follows. Section 2 develops the conceptual framework and relates our mechanisms and measures to the literature. Section 3 presents the data and Section 4 outlines our empirical strategy. Section 5 presents the main results and heterogeneity analyses, and Section 6 concludes.

2 Conceptual framework and related literature

Technological changes can reshape the direction of innovation by altering the opportunity cost of exploration versus exploitation (Acemoglu, 2002; Levinthal, 1997; Manso, 2011). We organise the literature into two streams that jointly motivate our research question and empirical analysis. We begin with the economics of data-driven search and data-visibility bias. We then connect this logic to organisational adoption measured through AI skill demand in vacancies, emphasising how human capital and organisational characteristics influence the impact of AI on innovation.

2.1 The economics of data-driven search and the direction of inventive search

Innovation is often conceptualised as a search problem in which firms balance the marginal cost of experimentation against the expected marginal value of the discovery (March, 1991; Levinthal, 1997; Stuart and Podolny, 1996). AI can affect the direction of innovation by changing the way firms evaluate projects and allocate search efforts across technological opportunities. For example, evaluating a project before physically conducting an experiment depends on the presence of relevant training data and well-specified prediction targets.[2] One mechanism points toward exploitation. Because firms are repositories of historical data generated by their prior activities, AI might be more accurate and valuable in domains where a firm already has dense data exhaust (Gans, 2025). This might create a systematic data-visibility bias in inventive search. Therefore, AI can redirect attention toward projects that are sufficiently observable and predictable and incentives might shift towards exploitation. By contrast, potentially high-value breakthroughs in data-scarce domains might attract less exploratory effort because they lack the informative signals that guide search. This mechanism closely resembles the "streetlight effect" in data-driven exploration modelled by Hoelzemann et al. (2025), where data highlights attractive but ultimately suboptimal projects, narrowing exploration and suppressing breakthrough discovery. This is also consistent with the model of Agrawal et al. (2024), which conceptualises AI as a tool for hypothesis generation that uses predictive models to prioritise candidates in vast combinatorial spaces, potentially focusing search on high-probability local areas.

A complementary perspective emphasises a basic feature of statistical learning: AI is most reliable when applied to problems that resemble the data on which it has been

[2] For instance, in silico screening in drug discovery (computational methods used to evaluate large sets of candidate molecules), or digital twins in manufacturing (virtual models that replicate and simulate real production systems)

trained. As a result, AI might deliver the greatest gains in data-rich parts of a firms existing knowledge base, creating a density dividend: incremental, gap-filling R&D generates additional data and improves predictive fidelity, further raising returns to nearby innovation (Agrawal et al., 2019). In contrast, in unfamiliar technological domains where data are thinner, AI might become noisier and the risk-adjusted payoff to exploration falls.[3]

A countervailing mechanism points toward exploration. AI can reduce the cost of screening alternatives, analysing high-dimensional data and identifying promising combinations of knowledge (Cockburn et al., 2019; Brynjolfsson et al., 2019). By improving the ability to evaluate candidate solutions before costly experimentation, AI may expand the set of feasible experiments and facilitate the recombination of knowledge across technological domains. In this view, AI functions as a research tool that supports invention beyond a firm's core technological base.

These two mechanisms imply that AI may increase total innovative output while having ambiguous effects on innovation direction. If prediction and optimisation effects dominate, firms may concentrate inventive effort in familiar domains. If improvements in screening and recombination dominate, AI may instead support entry into new technological areas. As recent theoretical work emphasises, the net effect depends on how AI reshapes the incentives and feasibility of inventive search (Gans, 2025). This tension motivates our empirical analysis.

2.2 AI skills, organisational adoption and business innovation

A growing empirical literature measures AI adoption using firm investments documents effects on productivity, growth and innovation. A common empirical approach identifies adoption via firms' demand for AI-related skills in job postings, which provides a timely indicator of organisational investment and reveals the tasks for which AI

[3] For a technical discussion of why generalisation in high-dimensional settings can be interpreted as extrapolation relative to observed data, see Balestriero et al. (2021).

capabilities are sought (Acemoglu et al., 2022; Alekseeva et al., 2021; Babina et al., 2024; Lou et al., 2025). This approach aligns with evidence that technology adoption is mediated by complementary human capital, task redesign and organisational change (Fügener et al., 2025). Recent evidence from Engberg et al. (2025), in particular, demonstrates that AI reconfigures the task composition of occupations rather than simply replacing them, specifically complementing activities that involve decision-making, coordination and responsibility. Empirically, AI-investing firms tend to grow faster and exhibit stronger innovation performance, with gains often concentrated among firms possessing richer data, stronger complementary assets and greater organisational capacity (Alekseeva et al., 2020; Babina et al., 2024; Rammer et al., 2022).

For innovation direction, a key implication is that adoption is filtered through *who* inside the firm builds and uses these capabilities. Evidence from human-AI collaboration shows that the benefits of AI depend on users experience, beliefs and calibration and that behavioural frictions can shape the willingness to delegate tasks to AI (Bockstedt and Buckman, 2025; Caplin et al., 2025; Wang et al., 2024). This implies that AI is not a frictionless "automation shock": its impact is mediated by skill complementarity and by the incentives attached to different organisational roles. Additionally, the search-based patterns discussed above further elucidate managerial incentives and organisational control. If AI is deployed primarily to improve forecasting and reduce variance in operational outcomes, managers may rationally allocate attention toward projects with more predictable returns, especially under short-run performance pressures (Lou et al., 2025). Moreover, if data and AI strengthen incumbents' advantages, they may reinforce concentration and further shift private incentives toward exploiting established trajectories (Babina et al., 2024; Krakowski et al., 2023).

In our setting, these considerations motivate distinguishing adoption through *managerial* versus *scientific/technical* vacancies. Managers are typically evaluated on efficiency, predictability and short-run performance, making AI particularly attractive for optimizing known products, markets and processes and thus potentially reinforcing data-visibility bias. By contrast, scientists and engineers may deploy AI to augment specialised R&D tasks (e.g., prioritising experiments and recombining knowledge), which can broaden the effective search set even when prediction performance is imperfect. AI can also automate routine laboratory taskssuch as data processing or experimental screeningfreeing researchers to concentrate on more creative, newer and frontier research activities.

The same reasoning also points to heterogeneity across firms and markets. Firms with larger data resources, stronger computational capacity and broader organisational capabilities may be better positioned to use AI not only for local optimisation but also for more exploratory innovation (Babina et al., 2024; Igna and Venturini, 2023). Conversely, in environments with high demand volatility or stronger short-run performance pressures, the forecasting value of AI may be particularly high, shifting attention toward projects with more predictable returns (Lou et al., 2025). Our empirical analysis examines these sources of heterogeneity directly.

3 Data and descriptive statistics

3.1 Data construction

We start with a comprehensive list of all UK patenting firms obtained from Orbis Intellectual Property (Orbis IP) that filed at least one patent in the UK between 2009 and 2021. Orbis IP provides matched data linking patents to UK-registered firms. We further merge this patent-firm-year-level dataset with FAME, which offers longitudinal firm-level data, including detailed financial and firmographic information for UK-registered firms. This initial dataset identifies 22,893 unique UK-registered patenting firms.

We combine the patenting data with job vacancy data from Adzuna. This database provides the most comprehensive source of online job vacancies in the UK, aggregated from over 1,000 different sources.[4] The dataset has been cleaned to remove duplicate

[4]The Adzuna data is now used by the UK Office for National Statistics to construct its official vacancy

postings and includes 272 weekly snapshots of open vacancies from January 2017 to March 2022. Each snapshot contains between 0.4 and 1.3 million vacancies (see Figure A1 in Appendix A). Each record includes the full text of the advertisement, the job title and the company name. Using job titles, vacancy text and industry classifications, we compute 3-digit Standard Occupational Classification (SOC) codes for each vacancy.[5]

We clean company names in both the Orbis IP and Adzuna data and then match the two datasets together using exact matches on company name. The matching process is manually validated using firms locations and industries to ensure accuracy. Our final sample comprises 4,747 unique UK patenting firms who post vacancies which appear in the Adzuna data.[6] Matched firms are associated with their job vacancies across all Adzuna snapshots, and duplicate vacancies are removed by comparing shared attributes such as job title, location and posting date. This process yields a consolidated dataset of 282,048 unique vacancies linked to the 4,747 patenting firms.

3.2 Measuring innovation activity

The patent data is used to measure firms' innovation search strategies. While not all innovations are patented or patentable, patents are widely recognised as an objective indicator of innovation performance, as they are less susceptible to individual or subjective biases (Bronzini and Piselli, 2016). Moreover, patents offer valuable insights into the quality of innovation, undergoing rigorous evaluation by experts to assess both novelty and practical applicability.[7] For each UK-registered patenting firm, we identify

indicators.

[5]We employ the Python package `occupationcoder` developed by Turrell et al. (2022), available at https://github.com/aeturrell/occupationcoder.

[6]An important characteristic of UK online vacancies is the disproportionately high share accounted for by recruitment agencies. It seems likely that many of the patenting firms do not appear in the Adzuna data because they recruit via agencies, and therefore their company names do not appear in the Adzuna data.

[7]As Griliches (1998) notes, patent activity serves as a proxy for the expansion of economically valuable knowledge, providing a reliable metric for inventive efforts, even though only a subset of all inventions is patented.

the earliest filing date for each granted patent family, retaining the earliest filing date across three key patent offices: the UK Intellectual Property Office (UK IPO), the European Patent Office (EPO) and the United States Patent and Trademark Office (USPTO). According to IPO (2021), the vast majority of UK applicants file their inventions with at least one of these offices.[8] We consider the total number of patents of a given firm as the total number of eventually granted patents that a firm has applied for in a given year.

We follow the approach used by Balsmeier et al. (2017), Gao et al. (2018) and Tzabbar and Kehoe (2014) to classify patents as either *new* or *known* to the firm, based on their technology class. A new patent is defined as one filed in a 3-digit International patent classification (IPC) technology class in which the firm has not filed any patents in the previous five years. Conversely, a known patent is filed in a 3-digit IPC class where the firm has already filed at least one patent during that same five-year period.

3.3 Measuring AI adoption

Following Alekseeva et al. (2020), we use the skill requirements listed in job postings as a proxy for firms adoption of AI technologies. The rationale is that AI adoption, unlike many other forms of technological change, requires employees with highly specialised expertise.[9] To identify AI-related job postings, we apply a set of keywords drawn from the existing literature (e.g., Acemoglu et al., 2022). A job vacancy is classified as AI-related if at least one of these keywords appears in its title or description. As robustness, we also extend this keyword list and use an extended AI-dictionary, shown in Appendix B.

Using this method, we identify 7,612 AI-related job postings, accounting for 3.03% of all vacancy postings made by our sample of 4,747 UK patenting firms between 2017 and 2021. We then aggregate these postings to the firm-year level and merge them with

[8]Specifically, approximately 55% of UK applicants file patents with the UK IPO, while around 31% file with the EPO, either alone or in conjunction with the USPTO. Therefore, we can reasonably assume that these three offices provide comprehensive coverage of the patenting activities of UK firms.

[9]Recall that our sample period predates the widespread use of ChatGPT and other user-friendly large language models.

the patent and accounting datasets used in our analysis. Figure A2 illustrates the share of job postings containing AI keywords in each vacancy snapshot. In the full Adzuna dataset (prior to matching with patenting firms), this share rises from under 0.5% in 2017 to over 1% by 2021.[10]

Our dataset is an unbalanced panel comprising 4,747 UK patenting firms. Within this sample, 408 firms post at least one vacancy requiring AI-related skills during the period, and we classify these firms as AI adopters from the first year in which such a vacancy appears. After removing firms which have missing covariates, our regression sample comprises 220 AI adopters and 2,643 non non-adopters. Because the first instances of AI-related hiring in our data occur in 2017, we assume that AI hiring activity was negligible prior to this date. This assumption is consistent with Schmidt et al. (2024), who document that AI hiring intensity in the UK remained very low before 2017.

Table 1 presents descriptive statistics distinguishing between adopters and non-adopters. On average, AI adopters are larger than non-adopters across several dimensions, including employment, assets and turnover. They also hold more patents and post more job vacancies online. These patterns suggest potential selection into AI adoption at the firm level - an issue we address in our empirical strategy in the following section.

4 Empirical strategy

To study the effect of AI adoption on the direction of innovation at the firm level, we estimate a model that combines matching techniques with difference-in-differences (DID) estimators. The simplest model we estimate can be written as follows:

$$y_{it} = \theta(AI_i \times Post_t) + \alpha_i + \delta_t + \varepsilon_{it} \tag{1}$$

[10] As expected, the share of AI-related vacancies is considerably higher in our sample of patenting firms.

In equation (1), the variable y_{it} represents innovation output. The term AI_i is a time-invariant treatment indicator, which takes the value one if the firm has adopted AI during our sample period, as measured by posting one or more vacancies which require AI skills, and zero otherwise. In our robustness section, we also consider the intensity of the AI posts. We include firm-fixed effects represented by α_i, which means that the effect of AI adoption on innovation is identified from within-firm variation and controls for firm-specific unobserved characteristics which might be correlated with AI adoption and innovation. $Post_t$ is a dummy variable that takes the value of one after the first time we observe in our data that a firm is posting a job with AI skills, δ_t includes a full set of year dummies and ε_{it} is an error term. The coefficient of interest is θ, which measures the difference-in-differences effect of AI adoption on innovation for firms that adopt AI as compared to the control group of non-adopters. In equation (1) standard errors are adjusted for clustering at the firm level.

We restrict our main estimation sample to the years 2010 to 2021, to focus on a time window around the first time we observed AI in our sample (the year 2017). The dependent variable y_{it} represents several innovation outcomes: the total number of patents, the number of patents in technological fields already known to the firm, and the number of patents in new technological domains. Because many firms do not patent in every year and entry into new technological classes is relatively rare, these variables contain a large number of zeros. To accommodate this distributional feature, and following recent recommendations (Agrawal et al., 2020; Balsmeier et al., 2024; Chen and Roth, 2024), we estimate equation (1) using Poisson quasi-maximum likelihood (PQML). In additional robustness checks, we explore alternative estimators and use different time windows around the first treatment event, and in some specifications we include further control variables.

Our identification assumption is that in the absence of AI adoption, the trend in the treated groups outcomes would have been comparable to those of the control group. Although equation (1) allows for differences in the level of innovation activity between

AI adopters and non-adopters, a possible concern is that θ might capture differences in innovation trends and selection into treatment. To address this issue, we implement a balancing method prior to estimating equation (1). Specifically, we employ the entropy balancing algorithm introduced by Hainmueller (2012). Following this methodology, we balance the treated and control samples on pre-treatment values of both the mean and variance of the following variables: the total number of patents (up to two periods prior the treatment), the number of patents known to the firm (up to two periods prior the treatment), employment, assets and number of vacancies. The final DiD analysis is conducted on firms within the common support region, with each firm assigned a weight based on the entropy balancing procedure. The key advantage of entropy balancing, compared to other balancing methods, such as propensity score matching, lies in its ability to reweight the control group to precisely match multiple distributional moments of the covariates in the treatment group. This ensures that the treatment and control groups are similar not only in terms of average characteristics, but also in the variance of their distributions.[11]

Equation (1) imposes a constant treatment effect after treatment. To relax this assumption, we also estimate an event-study specification which allows for heterogeneous treatment effects over time, as in the following specification:

$$y_{it} = \sum_{-11}^{4} \theta_\tau AI_{i\tau} + \alpha_i + \delta_t + \varepsilon_{it} \tag{2}$$

where τ indexes years relative to a firms first AI adoption, with negative values indicating pre-treatment periods and positive values indicating post-treatment years. $AI_{i\tau}$ equals one if firm i is τ years from adoption and zero otherwise and $AI_{i\tau}$ equals zero for non-adopters. The coefficients θ_τ capture differences in innovation trends before adoption and dynamic treatment effects thereafter. We omit θ_{-1} so that effects are measured relative to the year immediately preceding adoption. The parallel-trend assumption is

[11] In our robustness exercises we compare our results using propensity score matching.

assessed by testing whether $\theta_{-2} = \theta_{-3} = \ldots = \theta_{-11} = 0$, while the post-treatment coefficients $(\theta_0, \theta_1, \ldots, \theta_4)$ represent the dynamic effects of AI adoption on innovation.

In addition, we complement these binary treatment specifications with an analysis that captures the intensity of AI adoption. Specifically, we measure AI intensity as the ratio of AI-related vacancies to total vacancies and estimate its effect using the multi-valued treatment difference-in-differences estimator developed by De Chaisemartin and d'Haultfoeuille (2024). This approach generalises the standard DiD framework to settings in which treatment varies continuously over time and across firms. Conceptually, the estimator compares changes in outcomes between periods before and after the treatment level first changes (i.e. when firms first post an AI vacancy), to a control group of firms whose treatment level is the same as the treated firms in the "before" period, but then remains constant until the "after" period. The estimator averages across time of first treatment to yield an average effect for each length of treatment. This "non-normalised" estimate is an estimate of the average effect of being treated rather than untreated for each length of treatment. The estimate can be normalised by the size of the treatment received by each treated unit during the interval between the before and after periods, allowing for an interpretation as the average effect of an incremental rise in treatment intensity. Importantly, it provides consistent estimates in the presence of staggered adoption and heterogeneous treatment timing, addressing the potential biases highlighted in recent critiques of traditional DiD estimators (Goodman-Bacon, 2021; De Chaisemartin and dHaultfoeuille, 2020; Callaway and SantAnna, 2021). This refinement is particularly relevant in our context, as previous AI hiring may continue to affect subsequent innovation outcomes.[12]

The characteristics of firms that adopt AI: Before turning to the main focus of the paper, namely, the effect of AI on firm innovation, we first present estimates for the probability model that we use to obtain the entropy balance for our matching procedure.

[12] We estimate this model using the `did_multiplegt_dyn` command in Stata, which implements the dynamic version of this estimator.

This also allows us to describe the characteristics of AI adopters in comparison to firms who do not adopt. In Table 2, we show the estimates from a linear probability model with firm and year-fixed effects, where we regress different firm lagged characteristics on AI adoption. The results indicate that larger firms, both in terms of employment and total assets, are more likely to adopt AI. Additionally, while overall patenting activity appears to reduce the likelihood of AI adoption, patenting in established technologies seems to increase the probability of patenting.

In Table 3, we present a comparison of sample means and variances before and after the entropy balancing procedure. Before balancing (top panel) there are large differences in the distributions of characteristics of adopters and non-adopters. After balancing (bottom panel), these differences are almost entirely eliminated. The matched estimation sample consists of 211 AI adopters and 1,635 non-adopters, achieving close equivalence in their pre-treatment observable characteristics, particularly in terms of innovation outcomes.[13]

5 AI adoption and the direction of innovation

5.1 Baseline results

In this section, we present evidence of the effect of AI adoption on innovation at the firm level. Table 4 presents our baseline estimates of equation (1). In column (1), the dependent variable is the total number of patents, in column (2), the dependent variable is the number of patents known to the firm, and in column (3), the dependent variable is the number of patents that are new to the firm. In the table, we present different combinations of fixed effects and controls. In Panel A, we include firm and year fixed-effects. In Panel B, we add industry-year fixed effects to account for sectoral shocks that

[13] Moreover, the estimated coefficients of the differences in innovation trends between treated and control before treatment for the weighted sample are -0.0004 with standard error 0.106 for the number of patents, -0.0006 with standard error 0.114 for known patents and -0.053 with standard error 0.244 for the new patents

might not be accounted for the firm fixed effects, and in Panel C, we include as firm-level controls size (measured as the natural logarithm of the number of employees), the natural logarithm of the firm turnover and the natural logarithm of the firm assets.

Through the table, the results show that AI adoption increases the total number of patents. The estimated coefficients range from 47.1% (column 1 in Panel A) to 18.4% (in column 1 Panel C). The effect for patents that are known to the firm is positive and significant at standard levels, with values between 47.3% (column 2 in Panel A) to 18.5% (column 2 in Panel C). The effects for patents that are novel to the firm are positive, although of a smaller magnitude than for the other variables in Panel A and less precisely estimated in Panels B and C. This suggests that the effects of AI adoption on innovation are positive, but they are most clearly concentrated on technologies known to the firm. These are proportionately large effects, but note that the mean number of patents in our sample is 0.18 per year, so our estimates in panel (c) indicate an increase from 0.18 to 0.21.

Next, in Figure 1 we present the results for the time heterogeneous-treatment effect using the event-study specification as in equation (2). Panels (a) and (b) plot the estimates for total and known patents. After balancing treated and control groups, all estimated coefficients for the pre-adoption period are equal to zero except for the sixth lead. Moreover, there is no clear trend before adoption. The post-treatment effect is positive and long-lasting. It increases over time during the first three years and shows a small decline in the last year. Panel (c) plots the estimates for new patents. The result is quite different from the first two panels. Pre-adoption trends are again small and insignificant, but now the post-adoption effects are small and not significant at standard levels. This suggests that AI adoption has a negligible effect on the technologies that are new to the firm.

Next, we present the results based on the $DID_{+\ell}$ estimator of De Chaisemartin and d'Haultfoeuille (2024), which accounts for staggered and continuous treatment adoption as well as potential lagged effects on the outcome variable. In this framework, ℓ represents the number of periods since the initial increase in AI intensity. The results, reported in Table 5, show that greater AI intensity is associated with a significant increase in total patenting, with this effect again concentrated in the firms established technological domains.

5.2 Robustness checks

We next perform several robustness checks on our main findings. In these robustness checks, we consider different pre-treatment windows, not-yet-treated firms as controls, an alternative balancing matching technique, estimations using linear models distinguishing between the intensive and extensive margins, alternative measures of innovation and an alternative measure of AI adoption using an extended dictionary.

Our baseline results span the years from 2010 to 2021, but the first time we observe AI vacancies is in 2017. To check the sensitivity of the time window, in Panel A of Table 6, we show the results for a shorter time window (from 2012 to 2021), and in Panel B, we present results for a longer time window (from 2005 to 2021). The estimated coefficients are similar in terms of magnitude and significance to our baseline specification.

Our weighting procedure and event study show adopters and non-adopters follow a similar innovation trajectory prior to adoption. To further strengthen the validity of our approach, we conduct a weighted regression using only not-yet-treated firms as the control group. This might reduce the potential bias arising from structural differences between adopters and firms that do not adopt within our sample period. The results of these estimations are presented in Panel A of Table 7. The estimated coefficients are smaller than those in the baseline case and they are consistent with the conclusion that AI adoption primarily enhances the technologies already familiar to the firm.

Next, we use p-score as an alternative balancing method, which we present in Panel

B of Table 7. In these regressions, we combine the p-score reweighting estimator with the DiD approach. We first conduct a logit regression for AI adoption with the same covariates that we have used for the entropy balance. The DiD analysis is performed on firms within the common support that satisfy the balancing property for average covariate values. The results align with previous estimates regarding the impact of AI adoption on total, known and new patents.

In Table 8, we perform linear estimations for the intensive (Panel A) and extensive margin (Panel B). The estimated effects on the intensive margin of innovation, measured by the natural logarithms of innovation variables for firms with positive patent counts in the previous year, indicate that AI adoption increases the intensity of the total and known patents. For new patents, the estimated coefficient is positive, but not statistically significant at standard levels. The extensive margin is represented by a dummy variable that equals one if a firm reports positive innovation in the current period after having no registered innovation in the previous year. This measure captures the likelihood of starting to innovate. The results from this extensive margin analysis indicate that AI adoption does not increase the likelihood of starting any type of innovation. This suggests that initiating innovation entails significant costs that AI adoption does not alleviate.

Next, we consider alternative measures of known technologies, which we present in Table 9. First, in column (1), we consider the technology proximity distance of Jaffe (1986). This measure is constructed as the cosine similarity of patent portfolios. We compare a firm's patent portfolio at a given time t with its portfolio from the preceding five years in the following way: $Pmt = \frac{F_t F'_{t-5}}{(F_t F'_t)^(1/2)(F_{t-5}F'_{t-5})^(1/2)}$, where $Ft = (N_{1t}, .., N_{kt})$ is the technological space of a given firm at t, F_{t-5} is the technological space of a given firm between $t-6$ to $t-1$, and N_{kt} is the number of patents that the firm holds in the technological category k. This vector is constructed using the distribution of the firms patents in the different technological areas (the same 135 areas, as in the previous section).

In columns (2) and (3) of Table 9, we present estimations of exploitation and exploration indexes, which consider patents filed within the past five years and the previous citations of these patents. In particular, a patent is classified as exploitative if at least 60% of its backward citations - references to earlier patents - relate to the firms prior patents or to those cited by its other patents over the past five years. A patent is classified as exploratory if at least 60% of its backward citations are unrelated to the firms previous patents or those cited by the firms other patents within the last five years (Custódio et al., 2019; Levine et al., 2020; He and Hirshleifer, 2022). Exploratory patents reflect the extent to which a firm engages in innovative efforts that diverge from its previous inventions and existing search trajectories.

As in Balsmeier et al. (2017) and Gao et al. (2018), we then construct an exploitation (exploration) index as the ratio of a firms number of exploitation (exploration) patents to its total number of patent applications in a given year. The percentage of exploitative or exploratory patents reflects whether a firms innovation strategy is oriented toward exploring new or known technologies (Brav et al., 2018). In columns (1) and (2), we find a positive and statistically significant effect of AI adoption on technology close patents and the exploitation index. In column (3), the estimated effect on the exploration index is negative, the coefficient is small and non-significant at standard statistical levels. This is in line with our previous findings and corroborates the idea that AI adoption is positive for innovation, but it seems that AI particularly increases technologies that are known to the firm.

Table 10 reports results using an extended dictionary of AI-related terms.[14] This expanded list builds on the AI skill classifications from Acemoglu et al. (2022) and Babina et al. (2024), incorporating additional terms and grouping related concepts under broader, unified categories.[15] The results using this extended dictionary remain consistent with our baseline findings. We continue to observe a positive and statistically

[14]This dictionary was developed by Matias Golman as part of his PhD dissertation (Golman, 2026). We are grateful to him for generously sharing it with us.
[15]The full dictionary is provided in Appendix B.

significant effect on total patent output and on patents in technological domains already familiar to the firm. In contrast, the effect on entirely new technological areas remains small and statistically insignificant at standard levels.

5.3 Heterogeneous Effects of AI on the Direction of Innovation

The preceding analysis shows that, on average, AI adoption stimulates firm innovation, with effects concentrated in technological areas already known to the firm. In this section, we investigate how this relationship varies across firms and occupations to uncover the mechanisms driving these aggregate results. Consistent with the conceptual framework outlined in Section 2, we focus on three dimensions of heterogeneity: the type of occupation in which AI-related skills are adopted, firm size and the market environment in which firms operate. This analysis allows us to examine whether AI primarily strengthens incremental innovation within established domains or also enables more exploratory activity when supported by specific organisational and environmental conditions.

Types of occupations: The first dimension of heterogeneity we consider is the nature of the job in which AI skills are deployed. We focus on two occupational groups where AI adoption is likely to have the greatest impact on firm strategy: "managers or senior officials" and "professional occupations". Managers identify key problems, data needs and organisational processes required for effective AI implementation activities shown to raise sales and market valuation (Alekseeva et al., 2020). Professional occupations such as natural scientists, engineers, IT developers, or R&D specialists are the most frequent roles requiring AI knowledge (Alekseeva et al., 2021), and they are directly involved in research, technical experimentation and innovation. In our data, professional and associate professional occupations constitute by far the largest group of AI-related vacancies, whereas managerial roles account for a smaller share.[16] This

[16] We follow the SOC classification, which defines professional occupations as: natural and social science, engineering, IT and telecommunication, conservation and environment and R&D positions. Managers include managerial, chief executive and senior official roles. In our data, professional and associate

concentration reflects the fact that, during our sample period, AI implementation relies heavily on technical and analytical expertise, making professional roles the primary locus of AI investment, while managerial adoption remains more limited.

To examine the effect of AI adoption by managers and professional occupations, we use estimate two different variations of equation (1) each of them with a different treatment. In the first estimation, the treatment is adoption of AI in managerial occupations, as measured by vacancies posted which are identified as being in managerial occupations and which require AI skills. In the second estimation, the treatment is adoption of AI in professional occupations, measured in the same way. In each of the estimations, we reweight the sample using entropy balancing, where we balance each treatment group against a control group using the pre-treatment values of the mean and variance of the same variables as in the previous estimations.

We present the results in Table 11. In Panel A, we show results for managers and in panel B for professional occupations. Focusing on panel A, the results indicate that AI skills in managers lead to an increase in the number of patents, with this effect being concentrated on known patents. The estimations in panel B suggest that AI in professional occupations increase both known and new patents. This is consistent with the idea that for scientists, knowledge of AI decreases the opportunity costs of the known technologies, thereby freeing resources to expand the firm innovation capacity beyond its knowledge base. An alternative explanation is that AI skills are directly used in knowledge experimentation.

On the other hand, AI skills in managers might be more effective when AI is used on the managers' existing knowledge (Kim et al., 2024). Our results suggest that AI enables managers to improve the efficiency of their known processes, possibly because it provides accurate data on market conditions, but adoption of AI skills in managers is

professional occupations account for 9.8% of all AI vacancies. Within this group, natural and social sciences represent 7.8%, engineering 15.5%, IT and telecommunication 4%, conservation and environment 12% and R&D roles 0.9%, with the remainder in other specialised professional positions. Managerial occupations account for 2% of AI vacancies.

less likely to drive breakthroughs outside the company's established fields.

Large and small firms: The second dimension of heterogeneity we consider is firm size. A fundamental feature of AI is that it provides companies with real-time data on market conditions, customers or competitors. This can be particularly relevant for large firms if they have better data availability to forecast economic conditions. The ability of larger firms to leverage AI for exploration may also stem from their deeper digital foundations. As Igna and Venturini (2023) show, a firms probability of innovating in AI is path-dependent and significantly higher for those with existing engagement in Information and Communication Technologies. To account for potential differences between large and small firms, we include a triple interaction term in equation (1). We interact the term $AI_i \times Post_t$ times a dummy variable that measures whether a company is small. Firms are classified as large if their total assets pre-sample (between 2005 and 2009) fall at or above the 50th percentile of the total assets distribution, and small otherwise.[17] The dummy variable measuring firm size is time-invariant to avoid potential reverse causality issues.

The results, presented in Table 12, suggest that AI adoption has a positive effect on the total and known patents for both small and large firms, as shown by the positive and significant double interaction coefficient. The triple interaction term indicates that these positive effects are greater for large firms than for small ones. This suggests that AI adoption is important for all firms to generate innovations in fields known to the firm, but it seems to benefit especially large firms that might be more established in a market. For patents new to the firm, the estimated coefficients for the double and triple interaction are positive and significant at standard levels. This suggests that there are differences by firm size and that AI might lead to the agglomeration of new innovations by large firms.

Uncertainty in the market: Finally, we consider how the market environment in-

[17] The results are robust for different time windows such as between 2007 to 2010

teracts with the effect of AI on innovation. AI adoption helps companies to gather and analyse market data such as consumer behaviour, competitor pricing, or demand fluctuations. This information can be especially valuable in highly uncertain market environments, as it can reduce the risks associated with the returns of the innovation, for example, due to changes in consumers' preferences or value chain disturbances.[18] Since companies are more likely to obtain these data in their existing markets, AI may be more relevant for driving innovation in fields where the company already has expertise. We explore this possibility by splitting the sample between firms that are in sectors with high and low volatility in the spirit of Balsmeier et al. (2024). To avoid potential reverse causality issues, we calculate the firm average volatility (as the mean over the standard deviation of the firm turnover) for the pre-sample period (2005 to 2009). Then, we calculate the average sectoral value by aggregating volatility across firms. We construct a dummy variable that takes the value one if a firm is in a sector with volatility above the median. We add a triple interaction term to equation (1), where we interact the term $AI_i \times Post_t$ times the dummy variable measuring high market volatility.

We present the results in Table 13. The results from these estimations indicate that the effect of AI on total and known patents is more important in sectors with high volatility. We do not find a significant effect for technologies unknown to the company. Overall, the results suggest that AI encourages firms to innovate by offering information into areas where firms have already expertise.

6 Conclusion

This paper examines how the adoption of predictive AI technologies reshapes the *direction* of firm innovation. Linking UK firm-level patent data to the near-universe of online job vacancies, we show that AI adoption is associated with higher inventive output, but this increase is concentrated in technological domains already familiar to the firm. On average, we find little evidence that AI adoption induces entry into new

[18]See Czarnitzki and Toole (2011) for detailed study of market uncertainty on innovation.

technology classes. We interpret this pattern as consistent with a *data-visibility bias* in inventive search: prediction technologies lower the cost and uncertainty of optimisation and local search where data are abundant, thereby increasing the opportunity cost of exploration in data-scarce parts of the opportunity set.

A central implication of our research is that the effect on the direction of innovation depends on organisational and market conditions. Results from our heterogeneity analysis reveal that adoption entering through managerial occupations is associated with deeper inventive activity in established domains, consistent with predictive tools being deployed to improve coordination, monitoring and operational decision-making. By contrast, adoption through scientific and technical occupations is more strongly linked to new-domain patenting, indicating that when AI is integrated into R&D workflows it can support experimentation and extension into less familiar areas. Additionally, firm capabilities and the external environment matter as well: larger firms, with deeper complementary assets and broader data footprints, are better positioned to translate AI adoption into both exploitation and (to a limited extent) exploration, while high-demand-volatility settings strengthen the tilt towards local, exploitative invention.

A natural next step is to examine how these patterns differ across research domains and types of AI, particularly generative AI. Our results provide clean evidence on the impacts of predictive AI technologies in use before the diffusion of generative AI technologies. Key strengths of our paper are its full coverage of firm research areas and the identification of predictive AI models through job ads. However, important differences may arise between domains with abundant data and those where data are more limited. While generative models can expand the feasible design space and reduce the cost of early-stage experimentation (Agrawal et al., 2024; Hoffmann et al., 2025; Wang and Wu, 2025), emerging evidence suggests they may also encourage homogenisation by directing search toward similar outputs, concentrating activity in high-probability regions of the design space (**??**) and reducing the diversity of creative exploration (**????**), which suggests that the strength of the core mechanisms we identify may depend on the specific research field. Future work could use experimental designs to examine how AI affects innovation processes in environments with differing levels of data availability and technological uncertainty.

Our findings have implications for strategy and policy that go beyond the binary question of whether AI raises innovation. For managers, they underscore that the same technology can reinforce existing trajectories or broaden search depending on adoption pathways and governance. Embedding AI primarily in managerial control systems is likely to deliver reliable gains through refinement; embedding it in technical and scientific roles, and pairing it with resources for experimentation, is more conducive to exploration. For policymakers, the results imply that diffusion of AI may raise inventive output while simultaneously reinforcing path dependence and widening gaps between firms with strong complementary assets and those without. This points to a role for policies that not only facilitate adoption, but also support technical skills, experimentation capacity and data generation in under-observed domains.

Bibliography

Abernathy, W. J. and Clark, K. B. (1985). Innovation: Mapping the winds of creative destruction. *Research Policy*, 14(1):3–22.

Acemoglu, D. (2002). Directed technical change. *The review of economic studies*, 69(4):781–809.

Acemoglu, D., Hazell, J., and Restrepo, P. (2022). AI and jobs: Evidence from online vacancies. *Journal of Labor Economics*, A40(S1):S293–S340.

Agrawal, A., Gans, J., and Goldfarb, A. (2018). *Prediction Machines: The Simple Economics of Artificial Intelligence*. Harvard Business Review Press, Boston, MA.

Agrawal, A., Gans, J., and Goldfarb, A., editors (2019). *The Economics of Artificial Intelligence: An Agenda*. University of Chicago Press, Chicago and London.

Agrawal, A., McHale, J., and Oettl, A. (2024). Artificial intelligence and scientific discovery: A model of prioritized search. *Research Policy*, 53(5):104989.

Agrawal, A., Rosell, C., and Simcoe, T. (2020). Tax credits and small firm R&D spending. *American Economic Journal: Economic Policy*, 12(2):1–21.

Alekseeva, L., Azar, J., Giné, M., and Samila, S. (2024). AI adoption and the demand for managerial expertise. Technical Report 4986379, SSRN.

Alekseeva, L., Azar, J., Giné, M., Samila, S., and Taska, B. (2021). The demand for AI skills in the labor market. *Labour Economics*, 71:102002.

Alekseeva, L., Gine, M., Samila, S., and Taska, B. (2020). AI Adoption and Firm Performance: Management versus IT. *SSRN Working Paper*, 3677237.

Babina, T., Fedyk, A., He, A. X., and Hodson, J. (2024). Artificial intelligence, firm growth, and product innovation. *Journal of Financial Economics*, 151:103745.

Balestriero, R., Pesenti, J., and LeCun, Y. (2021). Learning in high dimension always amounts to extrapolation. *arXiv preprint arXiv:2110.09485*.

Balsmeier, B., Fleming, L., and Manso, G. (2017). Independent boards and innovation. *Journal of Financial Economics*, 123:536–557.

Balsmeier, B., Fleming, L., Stiebale, J., and Veihl, M. (2024). The unintended impact of R&D tax credits on innovative search. *Review of Economics and Statistics*, pages 1–47.

Bockstedt, J. C. and Buckman, J. R. (2025). Humans use of AI assistance: The effect of loss aversion on willingness to delegate decisions. *Management Science*.

Brav, A., Jiang, W., Ma, S., and Tian, X. (2018). How does hedge fund activism reshape corporate innovation? *Journal of Financial Economics*, 130:237–264.

Bronzini, R. and Piselli, P. (2016). The impact of R&D subsidies on firm innovation. *Research Policy*, 45(2):442–457.

Brynjolfsson, E., Li, X., and Raymond, L. R. (2024). Generative AI at work. NBER Working Paper No. 31161.

Brynjolfsson, E., Rock, D., and Syverson, C. (2019). Artificial intelligence and the modern productivity paradox: A clash of expectations and statistics. In Agrawal, A., Gans, J., and Goldfarb, A., editors, *The Economics of Artificial Intelligence: An Agenda*, pages 23–57. University of Chicago Press, Chicago and London.

Callaway, B. and SantAnna, P. (2021). Difference-in-differences with multiple time periods. *Journal of Econometrics*, 225:200–230.

Caplin, A., Deming, D., Li, S., Martin, D., Marx, P., Weidmann, B., and Ye, K. J. (2025). The abcs of who benefits from working with AI: Ability, beliefs, and calibration. *Management Science*.

Chen, J. et al. (2025). Ai4research: A survey of artificial intelligence for scientific research. *arXiv preprint*.

Chen, J. and Roth, J. (2024). Logs with zeros? Some problems and solutions. *The Quarterly Journal of Economics*, 139(2):891–936.

Cockburn, I. M., Henderson, R., and Stern, S. (2019). The impact of artificial intelligence on innovation: An exploratory analysis. In Agrawal, A., Gans, J., and Goldfarb, A., editors, *The Economics of Artificial Intelligence: An Agenda*, pages 115–148. University of Chicago Press, Chicago and London.

Custódio, C., Ferreira, M. A., and Matos, P. (2019). Do general managerial skills spur innovation? *Management Science*, 65:459–476.

Czarnitzki, D., Fernández, V., and Rammer, C. (2023). Artificial intelligence and firm-level productivity. *Journal of Economic Behavior & Organization*, 211:188–205.

Czarnitzki, D. and Toole, A. (2011). Patent protection, market uncertainty, and R&D investment. *Review of Economics and Statistics*, 1(93):147159.

De Chaisemartin, C. and d'Haultfoeuille, X. (2024). Difference-in-differences estimators of intertemporal treatment effects. *Review of Economics and Statistics*, Feb 9:1–45.

De Chaisemartin, C. and dHaultfoeuille, X. (2020). Two-way fixed effects estimators with heterogeneous treatment effects. *American Economic Review*, 110:2964–2996.

Engberg, E., Koch, M., Lodefalk, M., and Schroeder, S. (2025). Artificial intelligence, tasks, skills, and wages: Worker-level evidence from Germany. *Research Policy*, 54(8):105285.

Fleming, L. (2001). Recombinant uncertainty in technological search. *Management Science*, 47(1):117–132.

Fontanelli, L. (2025). Predictive ai and productivity growth dynamics. *Journal of Economic Behavior & Organization*.

Fügener, A., Walzner, D. D., and Gupta, A. (2025). Roles of artificial intelligence in collaboration with humans: Automation, augmentation, and the future of work. *Management Science*.

Gans, J. S. (2025). Growth in AI knowledge. NBER Working Paper No. 33907.

Gao, H., Hsu, P.-H., and Li, K. (2018). Innovation strategy of private firms. *Journal of Financial and Quantitative Analysis*, 53(1):1–32.

Golman, M. (2026). *Artificial Intelligence and Personality Traits*. PhD thesis, Nottingham University School of Economics.

Goodman-Bacon, A. (2021). Difference-in-differences with variation in treatment timing. *Journal of Econometrics*, 225:254–277.

Griliches, Z. (1998). Patent statistics as economic indicators: A survey. In Griliches, Z., editor, *R&D and Productivity: The Econometric Evidence*, pages 287–343. University of Chicago Press.

Hainmueller, J. (2012). Entropy balancing for causal effects: A multivariate reweighting method to produce balanced samples in observational studies. *Political Analysis*, 20(1):25–46.

He, Z. and Hirshleifer, D. (2022). The exploratory mindset and corporate innovation. *Journal of Financial and Quantitative Analysis*, 57:127–169.

Hoelzemann, J., Manso, G., Nagaraj, A., and Tranchero, M. (2025). The streetlight effect in data-driven exploration. NBER Working Paper No. 32401.

Hoffmann, M., Boysel, S., Nagle, F., Peng, S., and Xu, K. (2025). Generative ai and the nature of work. *Harvard Business School Strategy Unit Working Paper*, (25-021):25–021.

Igna, I. and Venturini, F. (2023). The determinants of AI innovation across European firms. *Research Policy*, 52(2):104661.

IPO (2021). Analysing the global filing activities of UK patent applicants supporting the innovation ecosystem: Building the evidence base on the drivers of ip. *Intellectual Property Office Research Paper*.

Jaffe, A. (1986). Technological opportunity and spillovers of R&D: Evidence from firms' patents, profits, and market value. *American Economic Review*, 76(5):984–1001.

Johnson, P. C., Laurell, C., Ots, M., and Sandström, C. (2022). Digital innovation and the effects of artificial intelligence on firms' research and development Automation or augmentation, exploration or exploitation? *Technological Forecasting and Social Change*, 179(April 2021).

Jones, B. F. (2025). Artificial intelligence in research and development. NBER Working Paper No. 34312.

Kanazawa, K., Kawaguchi, D., Shigeoka, H., and Watanabe, Y. (2025). Ai, skill, and productivity: The case of taxi drivers. *Management Science*.

Kim, A., Kim, D., Muhn, M., and Nikolaev, V. (2024). AI, investment decisions, and inequality. *Working Paper of the Becker Friedman Institute*, (2025-02).

Krakowski, S., Luger, J., and Raisch, S. (2023). Artificial intelligence and the changing sources of competitive advantage. *Strategic Management Journal*, 44(6):1425–1452.

Lavie, D., Stettner, U., and Tushman, M. L. (2010). Exploration and exploitation within and across organizations. *Academy of Management annals*, 4(1):109–155.

Levine, R., Lin, C., Wei, L., and Xie, W. (2020). Competition laws and corporate innovation. *National Bureau of Economic Research*, No. w27253.

Levinthal, D. A. (1997). Adaptation on rugged landscapes. *Management Science*, 43(7):934–950.

Liu, J., Chang, H., Forrest, J. Y.-L., and Yang, B. (2020). Influence of artificial intelligence on technological innovation: Evidence from the panel data of china's manufacturing sectors. *Technological Forecasting and Social Change*, 158:120142.

Lou, B., Ma, X., and Wu, L. (2025). Artificial intelligence, ceo turnover, and exploration orientation in firm innovation. *Information Systems Research*.

Makri, M., Hitt, M. A., and Lane, P. J. (2006). Complementary technologies, knowledge relatedness, and invention outcomes. *Research Policy*, 35(4):532–551.

Manso, G. (2011). Motivating innovation. *The journal of finance*, 66(5):1823–1860.

Manso, G., Balsmeier, B., and Fleming, L. (2023). Heterogeneous innovation over the business cycle. *Review of Economics and Statistics*, 105:1224–1236.

March, J. G. (1991). Exploration and exploitation in organizational learning. *Organization Science*, 2(1):71–87.

Nesta, L. and Saviotti, P. P. (2005). Coherence of the knowledge base and the firm's innovative performance. *Research Policy*, 34(2):212–232.

Rammer, C., Fernández, G. P., and Czarnitzki, D. (2022). Artificial intelligence and industrial innovation: Evidence from German firm-level data. *Research Policy*, 51(7):104555.

Schmidt, J., Pilgrim, G., and Mourougane, A. (2024). Measuring the demand for AI skills in the United Kingdom. *OECD Publishing*, (25).

Stuart, T. E. and Podolny, J. M. (1996). Local search and the evolution of technological capabilities. *Strategic Management Journal*, 17(S1):21–38.

Turrell, A., Speigner, B., Djumalieva, J., Copple, D., and Thurgood, J. (2022). *Transforming Naturally Occurring Text Data into Economic Statistics: The Case of Online Job Vacancy Postings*, pages 173–208. University of Chicago Press, Chicago.

Tzabbar, D. and Kehoe, R. R. (2014). Can opportunity emerge from disarray? An examination of exploration and exploitation following star scientist turnover. *Journal of Management*, 40(2):449–482.

Wang, W., Gao, G. G., and Agarwal, R. (2024). Friend or foe? Teaming between artificial intelligence and workers with variation in experience. *Management Science*, 70(9):5753–5775.

Wang, X. and Wu, L. (2025). Artificial intelligence, lean startup method, and product innovations. *Management Science*, 70(9).

Tables

Table 1: Descriptive statistics: Full sample, AI adopters, and non-adopters (2010–2021)

	Full sample		**AI adopters**		**Non-adopters**	
Variable	**Mean**	**SD**	**Mean**	**SD**	**Mean**	**SD**
No. AI posts	0.073	2.585	1.248	10.663	0.000	0.000
No. total posts	5.277	96.641	36.343	345.023	3.361	50.180
Treated (firm-year)	0.057	0.233	1.000	0.000	0.000	0.000
No. patents	0.905	7.600	5.331	25.457	0.636	4.536
Known patents	0.749	7.393	4.972	25.046	0.492	4.319
New patents	0.156	0.635	0.359	1.009	0.144	0.603
Total employment (log)	4.673	1.843	5.709	2.610	4.609	1.766
Total assets (log)	2.818	2.296	4.175	3.387	2.735	2.185
Turnover (log)	3.540	1.962	4.709	3.096	3.464	1.839
Return on turnover	1.319	2.782	1.031	0.841	1.337	2.862
Observations (firm-years)	19,840		1,139		18,701	

Notes: The table reports means and standard deviations (SD) for the full firm-year sample, AI adopters, and non-adopters over 2010–2021. Adopters are firms that post at least one AI-related vacancy at any point and remain coded as adopters thereafter. *No. AI posts* counts vacancies requiring AI-related skills; *No. total posts* counts all vacancies; *No. patents* refers to eventually granted patent applications. *Known* and *New* patents classify applications based on whether the firm patented in the 3-digit IPC class within the previous five years. *Total employment* is the logarithm of total personnel employed at the firm on December 31st. *Total assets* is the logarithm of the total assets of the firm on December 31st. *Turnover* is the logarithm of the total turnover of the firm on December 31st.

Table 2: Determinants of AI adoption

	(1)	(2)	(3)	(4)	(5)
Log Employment at $t-1$	0.012***	0.008**	0.008***	0.008***	0.011***
	(0.002)	(0.003)	(0.003)	(0.003)	(0.002)
Log Assets at $t-1$	0.007***	0.003	0.001	0.002	0.007***
	(0.002)	(0.003)	(0.003)	(0.003)	(0.002)
Turnover/Assets		0.000	0.000	0.000	
		(0.000)	(0.000)	(0.000)	
No. posts at $t-1$		0.000*	0.000	0.000	0.000
		(0.000)	(0.000)	(0.000)	(0.000)
No. patents at $t-1$			0.001	−0.005***	−0.004**
			(0.001)	(0.002)	(0.002)
No. known patents at $t-1$				0.006***	0.005**
				(0.002)	(0.002)
N	28029	18406	14493	14493	19737
R^2	0.006	0.006	0.008	0.011	0.011

Notes: The table reports the estimated coefficients from a linear probability model with firm and year FEs. Robust standard errors in parentheses. $^{*}p < 0.10,^{**} p < 0.05,^{***} p < 0.01$.

Table 3: Balancing test: Distributions of Covariates of Treated and Untreated Firms, Before and After Re-Weighting (Entropy Balancing)

	Mean Treated	Var Treated	Mean Control	Var Control
Before matching				
Patents at $t-1$	4.057	599.8	0.9063	46.56
Known patents at $t-1$	3.768	591.1	0.732	43.62
Total patents at $t-2$	3.806	488.2	0.9104	43.88
Known patents at $t-2$	3.488	476.6	0.7281	40.97
Total employment at $t-1$	4.354	7.223	4.606	3.533
Total assets at $t-1$	2.49	13.02	2.706	5.394
Total vacancies at $t-1$	27.01	45134	2.527	1553
After matching				
Patents at $t-1$	4.057	599.8	4.057	599.8
Known patents at $t-1$	3.768	591.1	3.768	591.1
Total patents at $t-2$	3.806	488.2	3.806	488.2
Known patents at $t-2$	3.488	476.6	3.488	476.6
Total employment at $t-1$	4.354	7.223	4.354	7.223
Total assets at $t-1$	2.49	13.02	2.49	13.02
Total vacancies at $t-1$	27.01	45134	27.01	45132

Notes: The table reports the mean and variance of firm characteristics for the treated and control groups, after applying the entropy balance re-weighting algorithm of Hainmueller (2012). The weights assigned to treated and non-treated firms are constructed to equate the mean and variance of all covariates. Patents is the total number of eventually granted patents applied for in a given year. Known is the number of eventually granted patents that are filed in a 3-digit IPC technology class where the firm has previously filed in the preceding five years. New is the number of eventually granted patents that are filed in a 3-digit IPC technology class where the firm has not previously filed in the preceding five years.

Table 4: Poisson Pseudo Maximum Likelihood estimation for the weighted sample

Dependent variable	(1) Total patents	(2) Known patents	(3) New patents
Panel A: Firm and year fixed-effects			
$AI_i \times$ Post	0.471***	0.473***	0.388*
	(0.139)	(0.143)	(0.221)
N	14071	6990	12403
Panel B: Firm, year and industry-year fixed effects			
$AI_i \times$ Post	0.281***	0.233***	0.317
	(0.066)	(0.074)	(0.214)
N	12677	6491	11163
Panel C: Firm, year, industry-year fixed effects and controls			
$AI_i \times Post$	0.184**	0.185**	0.279
	(0.091)	(0.085)	(0.241)
N	9775	5228	8755

Notes: This table shows results from Poisson Pseudo Maximum Likelihood estimation of equation (1) after balancing using entropy balancing weights. The dependent variable in column (1) is the total number of patents eventually granted that a firm applied for in a given year. The dependent variable in column (2) is the number of patents known to the firm, that is, the number of patents that are filed in a 3-digit technology class where the given firm did patent in years t-1 to t-5 and the dependent variable in column (3) is the number of patents new to the firm, that is, that is, the number of patents that are filed in a 3-digit technology class where the given firm did not patent in years t-1 to t-5. In Panel A, we control for firm and year FEs, in Panel B, we control for firm, year and industry-year FEs. In Panel C, we control for firm, year, industry-year FEs and employment, turnover and total assets. We present robust standard errors clustered at the firm-level in parenthesis.* p-value<0.10 ** p-value<0.05 *** p-value<0.01.

Table 5: The effects of AI intensity on the direction of innovation

Dependent variable	(1) Total patents	(2) Known patents	(3) New patents
AI intensity	4.056***	4.113***	-0.057
	(1.091)	(1.089)	(0.120)
N	7110	7110	7110

Notes: AI intensity is measured as the total number of AI posts over the total number of posts. We use the methodology of De Chaisemartin and d'Haultfoeuille (2024). In all regressions, we control for firm and year FEs. We present robust standard errors clustered at the firm-level in parenthesis.* p-value<0.10 ** p-value<0.05 *** p-value<0.01.

Table 6: The effects of AI adoption for different time windows

Dependent variable	(1) Total patents	(2) Known patents	(3) New patents
Panel A: Shorter time window from 2012 to 2021			
$AI_i \times$ Post	0.459***	0.459***	0.436*
	(0.133)	(0.137)	(0.224)
N	11816	5876	10136
Panel B: Longer time window from 2005 to 2021			
$AI_i \times$ Post	0.477***	0.479***	0.398*
	(0.137)	(0.140)	(0.217)
N	14236	7105	12544

Notes: This table shows results from Poisson Pseudo Maximum Likelihood estimation of equation (1) for the weighted sample. In Panel A, we consider the time window from 2012 to 2021. In Panel B, we consider the time window from 2005 to 2021. In all regressions, we control for firm and year FEs. We present robust standard errors clustered at the firm-level in parenthesis.* p-value<0.10 ** p-value<0.05 *** p-value<0.01.

Table 7: The effects of AI adoption for not yet treated firms as control group and alternative balancing algorithm

Dependent variable	(1) Total patents	(2) Known patents	(3) New patents
Panel A: Not yet treated firms as control group			
$AI_i \times$ Post	0.273***	0.280***	0.253
	(0.072)	(0.064)	(0.313)
N	1259	835	1082
Panel B: Balancing algorithm p-score			
$AI_i \times$ Post	0.312***	0.311***	0.169
	(0.146)	(0.154)	(0.331)
N	14095	6989	12422

Notes: This table shows results from Poisson Pseudo Maximum Likelihood estimation of equation (1) for the weighted sample. In Panel A, we consider not yet treated firms as controls instead of never treated. In Panel B, we consider as an alternative balancing algorithm the p-score. In all regressions, we control for firm and year FEs. We present robust standard errors clustered at the firm-level in parenthesis.* p-value<0.10 ** p-value<0.05 *** p-value<0.01.

Table 8: The effects of AI adoption. Intensive and extensive margin

Panel A: Intensive margin			
Dependent variable	(1) $Ln(patent_{it})$	(2) $Ln(known_{it})$	(3) $Ln(new_{it})$
$AI_i \times$ Post	0.394***	0.386**	0.358
	(0.149)	(0.186)	(0.244)
N	2295	1578	856
Panel B: Extensive margin			
Dependent variable	(1) $1\{Patent_{it} > 0\}$	(2) $1\{Known_{it} > 0\}$	(3) $1\{New_{it} > 0\}$
$AI_i \times$ Post	−0.012	−0.047**	0.018
	(0.052)	(0.023)	(0.048)
N	15749	15749	15749

Notes: Panel A uses those observations for which patenting is strictly positive. Panel B uses all observations. In all regressions, we control for firm and year FEs. We present robust standard errors clustered at the firm-level in parenthesis.* p-value<0.10 ** p-value<0.05 *** p-value<0.01.

Table 9: Alternative measures of known and new technologies

Dependent variable	(1) Jaffe proximity	(2) Exploitation index	(3) Exploration index
$AI_i \times$ Post	0.688***	1.107*	−0.029
	(0.151)	(0.610)	(0.048)
N	1825	1052	1052

Notes: In column (1) the measure of known technology is the Jaffe (1986) proximity measure defined as the cosine similarity of the patent portfolio over time of a given firm. We compare a firm's patent portfolio at a given time t with its portfolio from the preceding five years in the following way: $Pmt = \frac{F_t F'_{t-5}}{(F_t F'_t)^{(1/2)}(F_{t-5} F'_{t-5})^{(1/2)}}$, where $Ft = (N_{1t}, .., N_{kt})$ is the technological space of a given firm at t, F_{t-5} is the technological space of a given firm between $t-6$ to $t-1$, and N_{kt} is the number of patents that the firm holds in the technological category k. In columns (2) and (3), we present estimations of exploitation and exploration indexes. A patent is classified as exploitative if at least 60% of its backward citations relate to the firms prior patents or to those cited by its other patents over the past five years. A patent is classified as exploratory if at least 60% of its backward citations are unrelated to the firms previous patents or those cited by the firms other patents within the last five years (Custódio et al., 2019; Levine et al., 2020; He and Hirshleifer, 2022). As in Balsmeier et al. (2017); Gao et al. (2018), we then construct exploitation (exploration) index as the ratio of a firms number of exploitation (exploration) patents to its total number of patent applications in a given year. We present a Poisson Pseudo Maximum Likelihood estimation for the weighted sample. We present robust standard errors clustered at the firm-level in parenthesis.* p-value<0.10 ** p-value<0.05 *** p-value<0.01.

Table 10: Alternative measure of AI using an extended AI dictionary

Dependent variable	(1) Total patents	(2) Known patents	(3) New patents
AI_i *extended* $\times$ Post	0.275***	0.308***	0.046
	(0.103)	(0.114)	(0.161)
N	19001	9070	17554

Notes: We show regressions where the AI adoption variable is calculated using the extended-AI dictionary shown in Appendix B and created by Matias Goldman. We present robust standard errors clustered at the firm-level in parenthesis.* p-value<0.10 ** p-value<0.05 *** p-value<0.01.

Table 11: The effects of AI adoption by type of AI on firm innovation

Dependent variable	(1) Total patents	(2) Known patents	(3) New patents
Panel A: Managers or senior officials with AI skills			
Managers with $AI_i \times$ Post	0.887***	0.949***	−0.664
	(0.309)	(0.347)	(0.623)
N	14092	6988	12419
Panel B: Professionals with AI skills			
Professionals with $AI_i \times$ Post	0.430***	0.427***	0.360*
	(0.104)	(0.106)	(0.218)
N	14068	6988	12400

Notes: This table shows results from Poisson Pseudo Maximum Likelihood estimation of equation (1) for the weighted after balancing using entropy balancing weights. The dependent variable in column (1) is the total number of patents eventually granted that a firm applied for in a given year. The dependent variable in column (2) is the number of patents known to the firm, that is, the number of patents that are filed in a 3-digit technology class where the given firm did patent in years t-1 to t-5 and the dependent variable in column (3) is the number of patents new to the firm, that is, that is, the number of patents that are filed in a 3-digit technology class where the given firm did not patent in years t-1 to t-5. In Panel A, we consider managerial, chief executive and senior official occupations that require AI skills in the job vacancy. In Panel B, we consider professional occupations that require AI skills in the job vacancy. These occupations include natural and social science, engineering, IT and telecommunication, conservation and environment and R&D positions. We present robust standard errors clustered at the firm-level in parenthesis.* p-value<0.10 ** p-value<0.05 *** p-value<0.01.

Table 12: The effects of AI adoption for large and small firms

Dependent variable	(1) Total patents	(2) Known patents	(3) New patents
$AI_i \times$ Post	0.478***	0.479***	0.420*
	(0.139)	(0.143)	(0.221)
$AI_i \times$ Post $\times$ Small	-1.970^{***}	-2.006^{***}	-1.576^{***}
	(0.461)	(0.511)	(0.573)
N	14071	6990	12403

Notes: We present a Poisson model for the weighted sample. Firms are classified as large if their total assets in 2016 the year preceding the first observed AI adoption fall at or above the 50th percentile of the total assets distribution, and small otherwise. In all regressions, we control for firm and year FEs. We present robust standard errors clustered at the firm-level in parenthesis.* p-value<0.10 ** p-value<0.05 *** p-value<0.01.

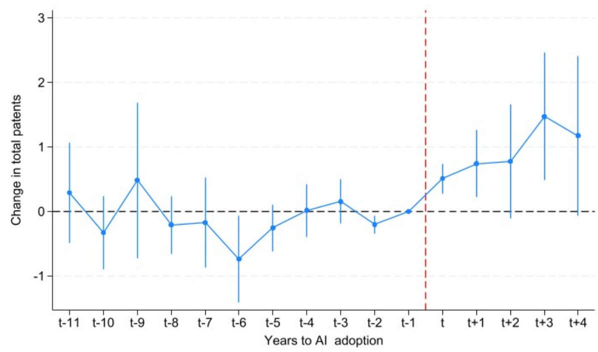
Table 13: The effects of AI adoption for companies in high and low volatile industries

Dependent variable	(1) Total patents	(2) Known patents	(3) New patents
$AI_i \times$ Post	0.340***	0.339***	0.476
	(0.090)	(0.088)	(0.295)
$AI_i \times$ Post $\times$ High volatility	0.357***	0.371***	-0.195
	(0.121)	(0.114)	(0.347)
N	14071	6990	12403

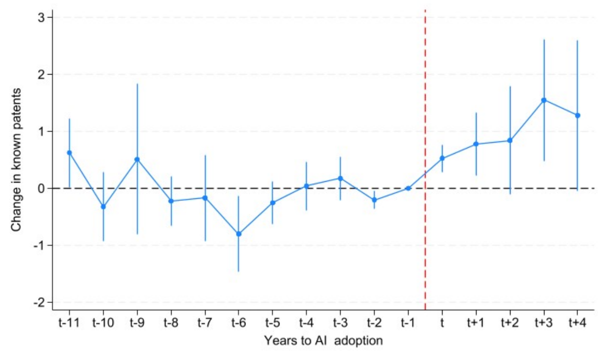
Notes: we present a Poisson model for the weighted sample. Firms are classified in high volatile industries if the industry has a volatility above the median based on pre-sample years. In all regressions, we control for firm and year FEs. We present robust standard errors clustered at the firm-level in parenthesis.* p-value<0.10 ** p-value<0.05 *** p-value<0.01.

Figures

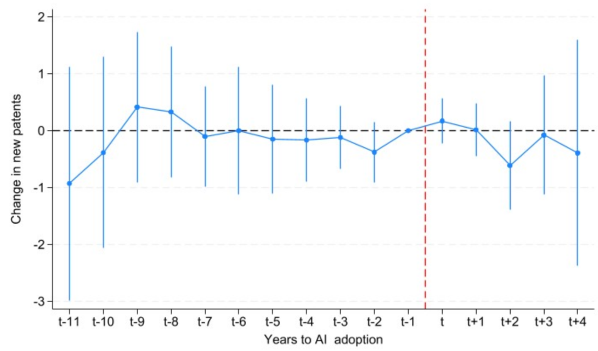
Figure 1: The effect of AI adoption on patenting: heterogeneous time effects



(a) Total patents



(b) Known patents



(c) New patents

Note: Point estimates and corresponding 95% confidence intervals from estimation of equation (2) after balancing using entropy balancing weights. The X-axis refers to the number of years after the first time we observe adopts AI in our sample. The dependent variable in panel (a) is the change in total number of eventually granted patents; in panel (b) the change in the number of known patents; and in panel (c) the change in the number of new patents. All specifications include a full set of year fixed effects to control for varying macroeconomic conditions, and firm fixed effects to control for time-invariant unobserved firm characteristics.

Appendix A

Figure A1: Total number of vacancies in each snapshot of the Adzuna database January 2017 to March 2022.

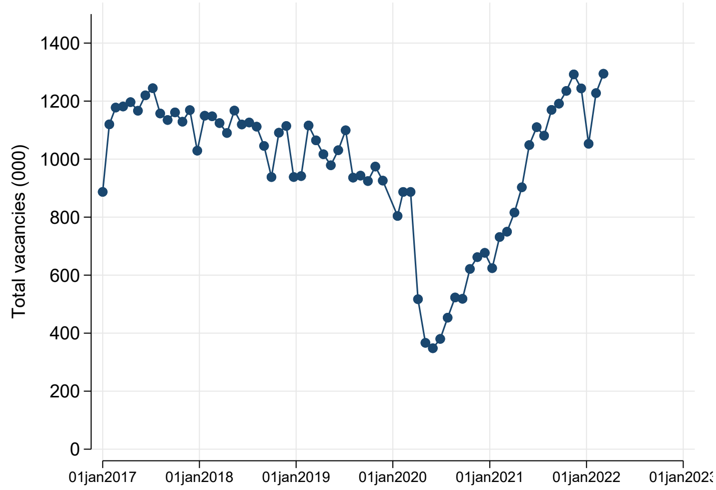
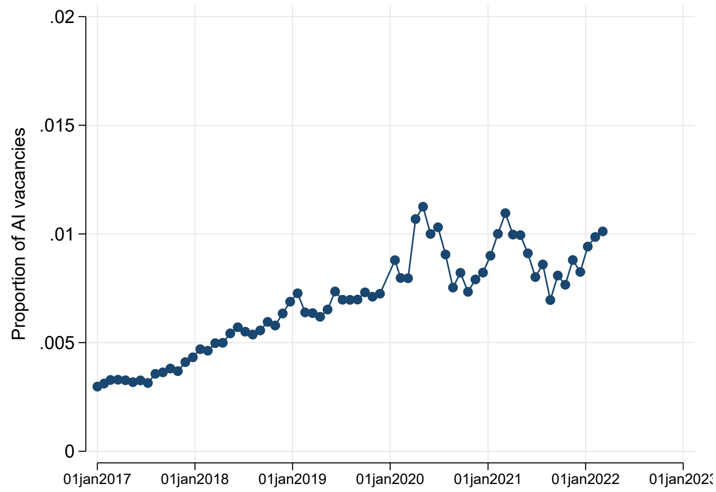


Figure A2: Proportion of vacancies using one of the AI keywords listed in Acemoglu et al. (2022)



Appendix B Dictionaries used to identify AI related skills in job ads

The skills used by Acemoglu et al. (2022) are *machine learning, computer vision, machine vision, deep learning, virtual agents, image recognition, natural language processing, speech recognition, pattern recognition, object recognition, neural networks, ai chatbot, supervised learning, text mining, unsupervised learning, image processing, mahout, recommender systems, support vector machines, random forests, latent semantic analysis (lsa), sentiment analysis, opinion mining, latent dirichlet allocation (lda), predictive models, kernel methods, keras, gradient boosting, openCV, xgboost, libsvm, word2vec, machine translation and sentiment classification.*

The skills used by Babina et al. (2024) that were not in Acemoglu et al. (2022) that were incorporated are *apache uima, artificial intelligence, autoencoders, bayesian networks/methods, caffe deep learning framework, classification algorithms, clustering algorithms, computational linguistics, convolutional neural network (cnn), dbscan (density-based spatial clustering), deeplearning4j, dimensionality reduction, dlib, expectation maximisation algorithm, hidden markov model, information extraction, jung framework, k-means, long short-term memory (lstm), matrix factorisation, maximum entropy classifier, microsoft cognitive toolkit, mlpack, mxnet, naive bayes, natural language toolkit (nltk), nd4j, opennlp, pybrain, recurrent neural network (rnn), scikit-learn, semi-supervised learning, stochastic gradient descent (sgd), theano, unstructured information management architecture, vowpal, wabbit, weka.*

Golman (2026) expanded the list with the following terms: *a b testing, ai, bidirectional encoder (bert), chaid, cloud technology, cntk, cnns, convolutional, convolutional networks, cross domain learning, data augmentation, decision trees, deep embedding, discriminant analysis, dnn techniques, document embeddings, emotion detection, ensemble learning, ensemble methods, entity linking, fastai, feature engineering, federated learning, ffmpeg, gaussian processes, generative adversarial networks (gans), generative models, generative pre-trained transformer (gpt), hierarchical clustering, huggingface, hyperparameter tuning, image fusion, image segmentation, intelligent automation, k-nearest method, kalman filtering, knowledge distillation, kubernetes, mlflow, model compression, multi-task learning, named entity recognition, natural language generation (nlg), natural language understanding (nlu), object tracking, optical character recognition, pattern detection, pcl, predictive churn, predictive modelling, pytorch, quantum computing, recurrent network, reinforcement learning, resnets vaes, robotic process automation, rnns, sequence-to-sequence modelling, shot learning, siamese networks, similarity models, sound recognition, text classification, topic detection, topic modelling, tree based learners, tree-based classifiers, t-sne, umap, uniform manifold, vision processing*.