WHEN AI MEETS ENTREPRENEURSHIP: EVIDENCE FROM THE COMMERCIALIZATION OF CHATGPT

EDOARDO MARCHESI[†] AND SIXUN TANG[⋆]

March 12, 2026

Click here for the latest version of this paper

Abstract. This paper examines how the commercialization of generative artificial intelligence influences new firm creation, leveraging the release of ChatGPT as an exogenous shock to AI accessibility. We develop a novel industry-level measure of AI compatibility that captures the extent to which AI tools can support early-stage entrepreneurial tasks. Using a difference-in-differences design, we find that industries more compatible with AI experienced an increase of over 10% in firm formation following ChatGPT's release. To investigate why, we examine three mechanisms: reduced experimentation costs, capital reallocation, and human capital complementation. While we find no evidence supporting the latter two, we show that the increase in firm formation is concentrated in industries where AI can effectively assist and experimentation is critical—evidence consistent with a reduction in the cost of experimentation as the key driver. Moreover, new firms created after ChatGPT's release in AI-compatible industries are more likely to survive, grow faster, and attract more educated workers. Taken together, the mechanism and quality results suggest that AI lowers the cost of experimentation in a way that disproportionately benefits high-ability entrepreneurs, leading to both greater and higher-quality firm formation.

JEL Classification Codes: L26, O33

We thank Liudmila Alekseeva (discussant), Gregory Brown, Emanuele Colonnelli, Andrea Contigiani, Bo Cowgill (discussant), Julie Fu, Deivy Houeix, Anders Humlum, Steven Kaplan, Young Soo Jang, Mario Leccese, Ralf Meisenzahl, Pascal Noel; and participants at the 3rd Chicago Entrepreneurial Finance Workshop; Revelio Labs Seminar Series; Inter-Finance PhD Seminar; Chicago Booth Finance Brownbag Seminar; Chicago FED Brownbag Seminar; REER 2025; HEC Paris Entrepreneurship Workshop 2025 for helpful comments and suggestions. All errors remain our own.

[†] Booth School of Business, University of Chicago. email: emarche0@chicagobooth.edu.

[⋆] Booth School of Business, University of Chicago. email: sixun-tang@chicagobooth.edu.

1. Introduction

The recent commercialization of artificial intelligence (AI) has unleashed a wave of transformative possibilities for businesses. While public debate has largely focused on AI's implications for productivity and labor markets, much less is known about its impact on one of the economy's key engines of dynamism: entrepreneurship. A central question is whether the arrival of advanced AI tools like ChatGPT has encouraged more individuals to start firms, and through what mechanisms. This paper provides the first evidence that the release of ChatGPT in late 2022 led to a disproportionate surge in firm formation in industries most compatible with AI, shedding light on how technological breakthroughs reshape entrepreneurial activity.

AI is a general-purpose technology with the potential to shape a wide range of economic outcomes. The rapid diffusion of large language models (LLMs) such as ChatGPT has made advanced AI capabilities available to the general public at near-zero marginal cost, enabling individuals and small teams to perform tasks—such as idea generation, coding, market research, and content creation—that previously required substantial time, expertise, or capital. Although AI has been studied extensively in the context of labor markets and established firms, its impact on entrepreneurship remains largely unexplored. Yet the direction of AI's effect on entrepreneurship is far from obvious. On one hand, AI may reduce barriers to entry by lowering the need for specialized skills, cutting upfront costs, accelerating early-stage processes, and providing guidance that was previously out of reach. These forces suggest that AI could make starting a business more accessible to a broader set of individuals. On the other hand, AI may entrench incumbents: large firms are better positioned to integrate new technologies at scale, attract scarce talent, and leverage proprietary data, thereby raising hurdles for new entrants. Which of these opposing forces dominates—and how AI ultimately reshapes the entrepreneurial landscape—remains an open question.

This paper asks whether the commercialization and widespread accessibility of AI tools, exemplified by the release of ChatGPT in late 2022, have influenced the occurrence of new firm creation. We present novel evidence that the availability of AI-powered virtual assistants stimulates entrepreneurial activity, particularly in industries most compatible with such technologies. Exploiting the release of ChatGPT as a quasi-exogenous shock to AI accessibility, we implement a difference-in-differences framework with varying treatment intensity. Our results show that, following the release of ChatGPT, industries with greater AI compatibility experienced a significant and persistent increase in new firm creation, averaging more than 10%.

A central contribution of this paper is the construction of two novel industry-level measures: the *AI Assistance Score* and the *Experimentation Intensity Score*. These measures are constructed using a bespoke multi-step pipeline that leverages large language models applied to a rich corpus of textual data, including business guides, entrepreneurship handbooks, and case studies. We identify the universe of tasks that entrepreneurs must undertake when starting a firm, and then evaluate each task in terms of how useful AI tools can be in supporting its completion (*AI Assistance Score*), as well as how reliant the task is on trial-and-error experimentation (*Experimentation Intensity Score*). This approach draws inspiration from previous works such as Eloundou et al. (2023) and Eisfeldt et al. (2025), which assess task-level AI exposure in labor market settings using LLMs. However, to the best of our knowledge, our measures are the first to focus explicitly on entrepreneurial activity, providing a new framework for quantifying how AI supports the early-stage business formation process across industries.

To explore the mechanisms driving the observed increase in firm creation, we investigate three potential channels, each motivated by distinct theoretical rationales and accompanied by empirical tests. First, we examine whether AI reduces the cost of entrepreneurial experimentation. Starting a new firm often involves iterative trial-and-error, including idea validation, product testing, and business model refinement. AI tools may lower the cognitive and time costs associated with such activities, particularly in industries where experimentation plays a more central role. If this mechanism is operative, we expect the increase in firm creation to be concentrated in industries that are both highly AI-compatible and experimentation-intensive. We test this hypothesis using a triple-difference (DDD) design: specifically, we interact the post-ChatGPT indicator with both our AI Assistance measure and our Experimentation Intensity measure. The resulting triple interaction isolates whether the post-ChatGPT increase in firm creation among AI-compatible industries is disproportionately larger in industries where experimentation is most important.

Second, we assess whether the effect of AI on entrepreneurship operates through capital reallocation. In this view, AI-compatible industries may attract more funding post-ChatGPT, as investors anticipate future productivity gains. Increased investment could in turn stimulate new firm formation, not because AI directly enables entrepreneurs, but because entrepreneurs follow capital into AI-compatible sectors. We test this hypothesis using mediation analysis, assessing whether investment flows into AI-compatible industries increased following the release of ChatGPT and whether these flows explain the observed rise in firm entry.

Third, we explore whether AI tools help individuals overcome human capital barriers that would otherwise deter them from entering certain industries. Entrepreneurs often refrain from launching businesses in sectors where they lack the requisite skills or domain-specific knowledge. If AI-powered tools like ChatGPT can substitute for some of these missing capabilities, they may allow individuals to pursue self-employment despite skill mismatches. To test this hypothesis, we construct a novel, time-varying measure of individual-level skill gaps vis-à-vis any given industry. This measure compares each individual's accumulated occupational skills—derived from employment histories linked to O*NET data—with the skill requirements of each industry, which is based on the skill portfolio displayed by individuals who successfully became entrepreneurs in the period prior to our study. We then examine whether individuals with larger skill gaps became relatively more likely to enter self-employment after the release of ChatGPT, compared to before.

Our findings provide empirical support for the first mechanism. The increase in firm formation is most pronounced in industries that are both highly AI-compatible and experimentation-intensive, consistent with the view that AI reduces the cost of entrepreneurial experimentation. In contrast, we find little evidence that capital supply or skill-bridging channels are primary drivers of the observed effect. While both mechanisms are theoretically plausible, our data suggest these mechanisms played a limited role during the study period.

Finally, we examine whether newly established firms differ in quality relative to their pre-AI counterparts. Canonical models predict that lower entry barriers may encourage the formation of lower-quality firms (Lucas Jr, 1978; Hopenhayn, 1992). In contrast, we find the opposite: firms founded after ChatGPT in highly AI-compatible industries exhibit higher short-term survival rates, expand faster, and attract more educated workers. This pattern suggests that the productivity gains from AI are tilted toward high-ability entrepreneurs, who benefit disproportionately from the technology, leading to both greater firm entry and higher average quality among new ventures.

This paper contributes to two distinct strands of research. First, it advances the literature on the economic consequences of artificial intelligence. A growing body of work has examined how AI technologies affect worker productivity and firm performance. Field experiments and case studies show that access to AI tools can raise worker-level productivity in a variety of settings (Dell'Acqua et al., 2023; Noy and Zhang, 2023; Brynjolfsson et al., 2025b). However, these effects are heterogeneous across both worker types and occupations. Dell'Acqua et al. (2023); Noy and Zhang (2023); Brynjolfsson et al. (2025b) find that performance improvements primarily accrue to low-skilled workers. The heterogeneous effects of AI are also evident at the occupational level: the threat of AI-driven replacement seems to fall disproportionately on high-paying occupations (Eloundou et al., 2023; Hampole et al., 2025). A recent study further highlights this heterogeneity, showing that workers whose skills rely heavily on textbook and codified knowledge (typically younger workers) are especially vulnerable to job displacement (Brynjolfsson et al., 2025a).

At the firm level, both AI adoption and AI compatibility have been linked to changes in performance, employment, and valuations. Firms more engaged in AI development exhibit higher productivity (Yang, 2022), while firms investing more heavily in AI display higher sales growth and profitability (Babina et al., 2024). AI is also associated with aggregate increases in employment (Yang, 2022; Babina et al., 2024) and higher market valuations (Babina et al., 2024; Eisfeldt et al., 2025).

Much of this research has focused on how AI displaces or complements existing workers and how incumbent firms integrate AI into their operations[1]. By contrast, research on the economic consequences of AI for entrepreneurs is still limited. We contribute to this literature by introducing a novel measure of how much AI can assist entrepreneurs in carrying out early-stage tasks (*AI Assistance*) and by documenting that industries where AI provides greater assistance experienced an increase of over 10% in firm formation. Closest to our work, Otis et al. (2024) examine AI's impact on post-entry outcomes for a small group of entrepreneurs, while Bao et al. (2025) study how AI may enable entrepreneurship in STEM sectors. Our study differs from Otis et al. (2024) by focusing on firm formation rather than post-entry performance, and from Bao et al. (2025) by extending the analysis beyond STEM sectors and by employing an exposure measure tailored to capture the nuances of entrepreneurial activity.

Second, this paper contributes to the literature on the determinants of entrepreneurship. One line of this work studies how entrepreneurial activity responds to exogenous shocks, including economic, environmental, and policy disruptions. Studies show that trade liberalization reshapes local firm entry (Cui and Li, 2023), natural disasters alter entrepreneurial dynamics (Boudreaux et al., 2019), input-price movements reallocate business formation (Davis et al., 1996), and commodity price shocks can lead to firm creation in areas highly dependent on local demand conditions (Bernstein et al., 2022). A subset of this literature focuses on technological shocks, showing that improvements in digital and communication infrastructure can lower entry costs and spur firm creation. Broadband expansion has been linked to increases in new firm formation, particularly in service-oriented sectors (Hjort and Poulsen, 2019; Hasbi, 2020; Stephens et al., 2022), while mobile technology diffusion has generated similar effects in developing economies (Aker and Mbiti, 2010; Hjort and Tian, 2021). Other work examines how information frictions shape trade (Steinwender, 2018). We contribute to this literature by studying the entrepreneurial response to a frontier AI tool (ChatGPT) that directly augments or substitutes for entrepreneurial tasks, rather than to broad improvements in digital infrastructure. This approach allows us to explain heterogeneous effects across industries by linking AI compatibility to post-shock firm creation.

A separate line of inquiry on the determinants of entrepreneurship emphasizes the role of frictions such as access to capital (Evans and Jovanovic, 1989; Gompers et al., 2005; Kerr and Nanda, 2011), skill acquisition and human capital (Lazear, 2005; Levine and Rubinstein, 2017), and the costs of experimentation (Kerr et al., 2014; Manso, 2016). We contribute to

[1]A more extensive review of the current findings on the economic impact of AI can be found in Subsection E.1 of the Online Appendix.

this literature by providing causal evidence on the role of experimentation costs—an important but difficult-to-measure friction—using an exogenous technological shock that plausibly relaxes these constraints in a subset of industries. Our results align with experimental studies showing that AI is particularly effective at supporting idea generation, problem solving, and enhancing human creative work (Boussioux et al., 2024; Doshi and Hauser, 2024; Meincke et al., 2024).

The rest of the paper is structured as follows. Section 2 describes data sources and Section 3 the construction of our AI compatibility, experimentation intensity measures, and other key variables. Section 4 presents the empirical strategy and main findings regarding AI compatibility and firm creation. Section 5 explores three potential mechanisms driving the results and evaluates their validity. Section 6 looks at firm level outcomes and 7 concludes.

2. Data Sources

The objective of this paper is to study firm creation following the arrival of ChatGPT, an event that is very recent. In principle, the ideal data would be administrative records, such as tax filings, which capture the full universe of new businesses. However, these sources are either not available at the frequency needed for our analysis or are released only with long delays, making them unsuitable for examining such a recent event. To circumvent this limitation, we use granular data from Revelio Labs, which provide near–real-time information on new firm creation and workforce histories.

This section presents the various data sources employed to construct our analysis dataset.

2.1. **Revelio Labs Data.** To measure entrepreneurial activity and workforce dynamics, we use granular data from Revelio Labs (RL), a dataset that captures firm creation and employment histories sourced from professional profiles on platforms such as LinkedIn.

We draw on detailed workforce data from Revelio Labs (RL), which provides broad coverage of U.S. firm creation and employment histories from 2020Q4 to 2024Q1. RL aggregates publicly available professional profile information, primarily sourced from platforms such as LinkedIn. Firms appear in the dataset either when they establish a LinkedIn company page or when workers list them as their employer. This dataset has been widely used in recent research to study human-capital flows, workforce composition, and organizational dynamics (Kwan et al., 2025; Labaschin et al., 2025; Wang et al., 2025).

We use RL's firm-level panel to measure new firm creation. Firms in RL data are tagged with North American Industry Classification System (NAICS) codes, allowing us to define new firms as those appearing in the data for the first time in a given quarter. For the firms identified, we know which industry they operate in, their location, their date of registration and date of shutdown (if any).

At the individual-level, we leverage RL's employment history data. RL data allows us to construct comprehensive resumes of workers' education and employment histories, including degree earned, jobs held, job titles, and employment durations. Moreover, each position is classified into occupational categories consistent with O*NET Standard Occupational Classification (SOC) codes. We extract data for all U.S. individuals who were active between 2020Q4 and 2024Q1.

It is important to note that RL data are limited to firms and individuals present on LinkedIn and similar platforms, introducing coverage bias. Consequently, our sample may overrepresent some portions of the economy or employment types.

To address these concerns, we benchmark our data against official statistics from the U.S. Census Bureau. Specifically, we benchmark RL against the Business Dynamics Statistics (BDS), the gold-standard source for firm creation data in the United States. In Appendix C.2, we document a series of validation exercises that assess how well RL replicates the key features of new firm formation captured in BDS. These tests compare RL and BDS along several dimensions, including the distribution of new firms across industries, the time series of firm creation activity, and the relationship between firm creation levels in the two datasets. We find that, to a good extent, RL closely mirrors BDS in terms of both the relative share and the rank-ordering of industries by new firm births. Furthermore, the number of new firms recorded in RL tracks that in BDS with high correlation across industries and years, despite RL's known undercoverage of certain traditional or low-tech sectors. These findings suggest that, for the purpose of analyzing patterns of firm creation across industries and over time—as we do in this paper—RL provides a reliable and representative proxy for official data. While RL may not be representative for all economic measures or subpopulations, its coverage is sufficiently broad and consistent to support the main analyses in this paper. In Subsection C.1 of the Online Appendix we provide further information regarding the data generating process underlying RL, while Subsection C.3 we detail the data manipulation steps to create the variables from the various RL datasets.

2.2. **Preqin Data.** To capture shifts in private capital availability that could influence firm creation, we draw on investment data from Preqin, a dataset that collects detailed information on venture capital and private equity funding. We extract data on the investments made between 2020Q4 and 2024Q1 for which we know: (1) the industry of the company receiving the funding and (2) the amount of funding. We subset the dataset to include only funding coming from venture capital and private equity investors that is directed to ventures operating in the U.S. Finally, from this subset of investment we retain only the first instance of equity funding a given company received[2].

[2]The object is to create a comprehensive data set of early stage equity investments injected into U.S. entrepreneurial firms during our period of interest.

Because Preqin primarily sources its data from investors with disclosure obligations, some transactions may be missing or reported incompletely, introducing potential measurement error. However, as long as investors with and without disclosure requirements do not substantially differ in their investment style—in terms of industry targets and timing—, measurement error should be of little concern.

2.3. **O*NET Data.** To examine whether AI facilitates entrepreneurship by bridging human capital gaps, we rely on occupational skill data from the Occupational Information Network (O*NET). O*NET provides standardized measures of skill requirements across occupations, enabling us to quantify individual-level skill gaps.

We utilize occupational skill level data from the Occupational Information Network (O*NET), the primary occupational information database maintained by the U.S. Department of Labor's Employment and Training Administration (USDOL/ETA). The O*NET database provides standardized measures of skills required across nearly 1,000 U.S. occupations. These data are collected through structured surveys administered to job incumbents and occupational experts, ensuring relevance and accuracy in reflecting contemporary occupational requirements.

In the O*NET framework, each skill associated with an occupation is rated along two key dimensions: *importance* and *level*. For our analysis, we focus on the skill *level* ratings, which capture the degree of proficiency or complexity required for effective performance within an occupation. This measure can be interpreted as the intensity of a skill that an individual is expected to possess to perform successfully in a given occupation. Level ratings are determined through survey responses using an anchored numeric scale from 0 (no proficiency required) to 7 (highest proficiency). Behavioral examples provided within O*NET documentation anchor each numeric rating, offering practical interpretations of skill proficiency at different scale points. For example, lower-level anchors reflect basic capabilities suitable for simple tasks, whereas higher-level anchors correspond to advanced capabilities necessary for more complex or specialized occupational tasks.

Each occupation, identified by its O*NET Standard Occupational Classification (SOC) code, is linked systematically to its relevant skill-level ratings, creating a comprehensive occupation-skill dataset.

2.4. **Industry Financials.** We use firm-level financial data from Compustat to construct several control variables that account for pre-existing economic conditions potentially confounding the relationship between AI and firm entry. First, we include Return on Assets (ROA), defined as income before extraordinary items divided by total assets, to control for profitability cycles that could drive entry independently of AI developments. Second, to capture the innovation environment and purge effects confounded with existing technology intensity, we construct R&D intensity as R&D expenses over revenue. Third, CapEx intensity, defined as capital expenditures over revenue, is included to control for investment cycles unrelated to AI-driven entry booms. Fourth, cash holdings over total assets, serve as a proxy for liquidity and internal financing capacity, controlling for entry driven by abundant cash availability. Finally, we incorporate market leverage ratio, calculated as the sum of long-term and current debt divided by the market value of equity, to account for financial risk and capital structure effects that may influence firm formation decisions. Collectively, these controls help isolate the effect of AI on firm entry from broader economic and financial conditions.

2.5. **Textual Data Sources.** To construct our novel industry-level measures of AI compatibility and experimentation intensity, we leverage detailed textual data from entrepreneurship handbooks, business guides, and case studies. This part of the subsection describes these textual sources and motivates their suitability for extracting the universe of tasks that entrepreneurs must perform when establishing new firms.

Since our objective is to establish to which extent—if any—can LLMs facilitate entrepreneurial entry and whether this is conditional on the type of tasks for which they provide assistance, we first generate the universe of tasks that entrepreneurs have to successfully carry out in order to get their business established and starting to operate. One major limitation we face is that, while there are repositories of tasks entailed by different jobs and occupations, to date, there is no source of data listing all the activities a new entrepreneur needs to carry out. To overcome the above mentioned challenge, we take advantage of textual data from various sources such as: (a) books focused on entrepreneurship or becoming an entrepreneur, (b) business guides from official sources[3] and, (c) business school case studies. The unstructured sources of data are then used as raw data for an LLM to extract tasks entrepreneurs have to carry out.

In Subsection D.1 we offer additional information on the textual data sources we used.

3. Construction of Key Measures

Since our objective is to study how the expansion of access to AI influences firm creation, a central requirement of our analysis is a cross-sectional measure of exposure. To this end, we construct two novel industry-level indices: the *AI Assistance Score*, which captures the extent to which large language models can aid entrepreneurs in completing early-stage tasks, and the *Experimentation Intensity Score*, which quantifies how reliant those tasks are on trial-and-error processes. Together, these measures allow us to capture systematic variation across industries in how AI technologies can support entrepreneurial activity.

[3]Examples of these sources are the Small Business Administration and Chambers of Commerce

Beyond these task-based indices, this section also introduces a complementary measure designed to evaluate whether AI helps potential entrepreneurs overcome human capital barriers. Specifically, we develop an individual-by-industry *Skill Gap* measure that compares an individual's accumulated skills to the requirements of different industries, enabling us to test whether AI reduces frictions that would otherwise deter entry.

The remainder of this section details the construction of these key measures (*AI Assistance*, *Experimentation Intensity*, and *Skill Gap*) and explains how they provide the foundation for our empirical analysis of AI and firm creation.

3.1. **AI Assistance and Experimentation Intensity Scores.** To measure cross-sectional exposure, we construct two novel indices: *AI Assistance* and *Experimentation Intensity*. A key challenge is that, unlike occupations, there is no existing repository that systematically lists the set of tasks entrepreneurs must undertake when starting a firm. We address this limitation by leveraging textual sources such as entrepreneurship handbooks, business guides, and case studies, which we process with large language models to extract the universe of entrepreneurial tasks. In what follows, we detail the workflow used to derive the two measures. In the interest of transparency, we report the prompts and system messages used in Subsection D.2 of the Online Appendix.

In a preliminary step, we take each source of textual data listed in Subsection D.1 of the Online Appendix and divide it into smaller chunks to ensure that the LLM can process the content effectively[4]. Each chunk is then submitted to ChatGPT-4o via the API for processing.

In the first layer of the workflow, we instruct the model to impersonate an experienced entrepreneur and focus on relevant sections of its knowledge base when processing each request. Each text chunk is presented individually to the LLM, which is then asked to identify any task related to establishing and first running a firm, and to produce a detailed description of the task. The resulting list of tasks includes activities such as *"Complete and submit necessary paperwork for firm registration"*, *"Design a minimum viable product (MVP)"*, *"Acquire early customer feedback on product/service"*, *"Establish and negotiate initial relations with suppliers"*.

Once the model in the first layer has produced the universe of identified tasks from the textual data, we run into the presence of duplicates both across and within text sources, as the same task may be present multiple times within the same text source or recur in different ones. We use a second LLM model to identify duplicates based on semantic similarities and

[4]This pre-processing step is standard practice. Operationally, LLMs have a maximum context window (e.g., 4k, 32k, even 200k tokens depending on the model); if the text to be analyzed exceeds this limit, it cannot be processed in a single pass. Moreover, even when a text technically fits within the context window, very long inputs can reduce the model's ability to attend to all parts equally. Chunking allows the model to focus on coherent segments, thereby improving accuracy and consistency.

produce condensed versions of the task description. More specifically, we first embed task descriptions to identify groups of similar tasks. Text embedding converts textual content (e.g., words, sentences, or documents) into numerical vectors for use in machine learning models. The key idea is to represent text in a way that captures its semantic meaning—so that similar texts are close to each other in vector space. Once we have grouped the tasks, we ask an LLM to come up with a task name for the group and produce a "collapsed" version of the task description which distills the main elements of the task descriptions collapsed. Through this process we are able to produce the universe of unique entrepreneurial tasks.

At the current stage, we have the universe of tasks entrepreneurs have to carry out, however, a given task may not apply to all industries and even if it does, it may not be specific enough in its description to convey the nuances linked to carrying out that task in that specific industry. We overcome this issue by bringing a third LLM layer into play. For this layer, we ask ChatGPT to impersonate an entrepreneur with extensive experience in running a firm in a specific industry. We then create all the possible combination of tasks and a 3-digit NAICS industry and present these combinations to the model. Namely, for each pair, the model will be fed the task name, task description, the industry name and industry description and be asked whether or not the task is relevant to the industry at hand. Then, if the task is deemed relevant the model will be asked to use the information in the industry description to tailor the task description to the industry. This process results in the list of relevant entrepreneurial tasks for each industry, with their description being crafted to better incorporate specifics about the sector. We extract a total of 15,592 tasks across all industries. There is a total of 847 unique tasks and each industry has on average 197 distinct tasks. Figure 1 presents the distribution of number of tasks across industries.

At present, we have the list of tasks that are relevant to each 3-digit NAICS industry. We introduce the fourth and last LLM layer to score these tasks across 3 dimensions. The first dimension represents how "important" the task is—where importance is defined as how recurrent, time-consuming or prevalent the task is for entrepreneurs starting a company in a specific industry. The second dimension is the degree of AI compatibility of the task, which we call AI Assistance Score. The AI Assistance Score reports the degree to which an LLM can provide support in completing the task, which higher scores representing higher levels of support. The third dimension, which we term Experimentation Intensity, summarizes the degree of experimentation required by an entrepreneur in order to successfully carry out a task; in a nutshell, this dimension measures "linearity" versus trial-and-error based task execution.

We ask the fourth LLM to impersonate an entrepreneur with extensive industry knowledge—based on the industry at hand—and provide it with scoring rubrics to produce values for task Importance, AI Assistance and Experimentation Intensity. To score the tasks across the three dimensions, we implement Chain-of-thought (CoT), which means that instead of asking the model to provide a numerical score, we first ask the model to reason and produce a rationale as to why a task may be "Important," "AI Assistable," or "Experimentation Intense," and then, based on the rationale, produce a numerical score. Chain-of-thought helps models reason more effectively and accurately, especially on complex or multi-step problems. Instead of jumping directly to an answer, CoT encourages the model to generate intermediate reasoning steps, mimicking how humans solve problems by thinking out loud. Once the model has produced a rationale concerning the dimension at hand, the LLM is then asked to produce a numerical score based on it. We run the score estimation procedure three times and average the results in order to smooth out possible instances of hallucinations.

Since our analysis is conducted at the industry level, we collapse the task-level scores into industry-level measures by weighting each task by its *Importance score*, ensuring that marginal or rarely performed tasks do not receive disproportionate weight. Figure 2 presents the distributions of the AI Assistance (Panel (a)) and Experimentation Intensity (Panel (b)) scores.

In Table 1 we report the top 20 and bottom 20 industries by AI Assistance Score. Industries at the top of the AI Assistance distribution—Web Search Portals & Information Services (NAICS 519), Data-Processing and Hosting Services (518), and the broader family of Professional, Scientific & Technical Services (541)—all revolve around the creation, manipulation, or dissemination of digital information. Their core tasks (e.g. software engineering, data curation, text generation, legal drafting, and graphic design) are language– or pattern–based and can be executed, accelerated, or even autonomously handled by large language models and other machine-learning systems. Seemingly "old-economy" leaders such as Electronics Manufacturing (334) and several retail subsectors (Electronics & Appliance Stores, General Merchandise, Clothing & Accessories) also score highly because their competitiveness now hinges on AI-driven recommendation engines, demand forecasting, inventory optimisation, and personalised marketing—all areas where contemporary AI excels. By contrast, industries at the bottom—Crop Production, Forestry and Logging, Oil & Gas Extraction, Primary Metal Manufacturing, and Pipeline Transportation—are dominated by highly physical, location-specific production processes governed by engineering constraints, heavy regulation, and harsh operating environments. While narrow AI can aid predictive maintenance or drilling optimisation, the share of tasks that can be materially substituted or augmented by foundation models remains small, keeping their AI Assistance scores low.

Table 2 reports the top 20 and bottom 20 industries by Experimentation Intensity. Topping this distribution are creative and digital content sectors such as Motion Picture & Sound Recording (512), Broadcasting & Content Providers (516), and again Web Content & Information Services (519). These industries live on rapid A/B testing, audience feedback, and continual format tweaks, making experimentation intrinsic to value creation. Food Services & Drinking Places (722) also score highly because menu design, pricing, and service concepts are constantly refined in response to local consumer tastes. Even the "plumbing" of the digital economy, Data-Hosting & Web-Infrastructure (518), ranks near the top: cloud providers roll out new features, security patches, and pricing models continuously, learning from usage telemetry in real time. At the other extreme lie capital-intensive, heavily regulated sectors such as Pipeline Transportation (486), Utilities (221), Heavy & Civil Engineering Construction (237), and Primary Metal Manufacturing (331). Their production functions entail large irreversible investments and stringent safety or environmental standards that severely limit scope for mid-course experimentation. Similarly, Hospitals (622) and Nursing & Residential Care Facilities (623) face ethical and regulatory constraints that suppress rapid trial-and-error, anchoring them to the lower tail despite growing analytics adoption[5].

3.2. **Skill Gap Measure.** To empirically evaluate the hypothesis that AI tools reduce entry barriers by bridging human capital deficits, we construct a novel, time-varying measure of individual skill gaps relative to industry requirements. In what follows, we detail the construction of this measure and clarify how it enables rigorous tests of the human-capital bridging mechanism.

One of the potential mechanisms linking AI to firm creation we envisioned revolves around LLMs bridging potential entrepreneurs' human capital gaps. To test this mechanism we estimate the skill gap for individual j, in quarter t vis-a-vis industry i by using the skills s connected to each occupation o reported in O*NET and their relative importance score for the profession.

We first create a measure of skill portfolio, which is the collection of skills retained by an individual j and their respective level of proficiency. The skill portfolio is time varying, as it depends on the occupations covered by the individual over time. Implicitly, we assume that skills do not depreciate. First, let O_j^t be the set of occupations held by individual j at or before time t. Let Skill Intensity$_{so}$ be the intensity required for skill s when an individual is employed in occupation o[6]. Then the skill profile—the skill strength retained by the individual—for individual j at time t with respect to skill s is:

$$\text{Skill Profile}_{jts} = \max_{o \in O_j^t} \{\text{Skill Level}_{so}\} \tag{1}$$

[5]To further clarify what the AI Assistance measure captures, we correlate it with various industry characteristics in Table OA1 (Online Appendix, Subsection A.1). AI Assistance does not exhibit meaningful correlations with these characteristics, underscoring that the measure reflects a distinct dimension of the entrepreneurial process—namely, the extent to which AI can support early-stage entrepreneurial tasks.

[6]The same skill may be used in multiple occupations, but its intensity score is likely different

Since we assume that skills do not depreciate over time, and multiple occupations may require the same skill, when there's overlapping of skills between one or more occupations held by the same individual, for that particular skill, we report the highest level of skill intensity reported among all occupations ever held, hence the $\max_{o \in O_j^t}$.

Next, we create a measure of *industry requirements* using the skill profiles of individuals who successfully became entrepreneurs prior to the period of analysis to identify the skills historically associated with entry in each industry. The intuition is that these profiles represent the empirical combination of capabilities that allowed entrepreneurs to overcome the barriers to entry in a given sector before the widespread availability of generative AI.

Formally, let $\mathcal{E}_i$ denote the set of individuals who became entrepreneurs in industry i between 2010Q1 and 2022Q3, and let t_j^* be the quarter in which individual $j \in \mathcal{E}_i$ first appears as an entrepreneur. For each of these individuals, we extract their skill portfolio in the quarter immediately preceding entry, $\text{Skill Profile}_{j,t_j^*-1,s}$, where each element s corresponds to one of the standardized O*NET skills described above.

We then define the *industry requirement* for industry i and skill s as the mean pre-entry skill level across all historical entrepreneurs in that industry:

$$\text{Industry Requirement}_{is} = \frac{1}{N_i} \sum_{j \in \mathcal{E}_i} \text{Skill Profile}_{j,t_j^*-1,s}, \tag{2}$$

where $N_i = |\mathcal{E}_i|$ is the number of entrepreneurs used to construct the benchmark for industry i.

The resulting vector $\text{Industry Requirement}_i = (\text{Industry Requirement}_{is})_{s=1}^{S}$ provides an empirically grounded representation of the skill bundle that historically characterized founders in each industry prior to the diffusion of AI. By anchoring the benchmark to pre-ChatGPT entrepreneurs, the measure captures the human-capital profile that was *sufficient* for firm entry in the pre-AI regime and remains unaffected by post-treatment changes in the composition of skills among entrepreneurs.

Next we calculate the skill deficit for a given individual j for each skills s in quarter t with respect to industry i. When calculating the skill deficit, we ensure that the deficit captures only instances of an individual's skills' insufficiency with respect to an industry benchmark. The skill deficit for a given individual j for each skills s in quarter t with respect to industry i is:

$$\text{Deficit}_{jits} = \max\{0, \text{Industry Requirements}_{is} - \text{Skill Profile}_{jts}\} \tag{3}$$

To calculate the skill gap for a given individual j in quarter t with respect to industry i we aggregate over the skill deficit values, using a Euclidean Norm to penalize large shortfalls more heavily:

$$\text{Skill Gap}_{jit} = \sqrt{\sum_{s} (\text{Deficit}_{jits})^2} \tag{4}$$

For the analysis, we standardize the skill gap measure using the pre-ChatGPT mean and standard deviation. Anchoring the transformation to the pre-treatment distribution ensures that the scale of the variable is fixed prior to the arrival of ChatGPT, avoiding the risk that post-treatment dynamics mechanically influence the normalization itself. This preserves the comparability of the measure across time and is essential for any attempt at a causal interpretation of pre/post differences. In addition, standardization enhances statistical stability in our regressions: it prevents coefficients from being sensitive to the arbitrary scale of the raw Euclidean norm and facilitates estimation in specifications with interaction terms (e.g., Skill Gap $\times$ Post), where unstandardized variables could otherwise generate unstable or difficult-to-interpret results.

4. The Effects of AI on Firm Creation

In this section, we turn to the central empirical question of the paper: does the commercialization and widespread accessibility of AI technologies influence entrepreneurial activity? To answer this, we exploit variation in industry-level AI compatibility around the release of ChatGPT in late 2022, which we treat as a quasi-exogenous shock to the availability of generative AI. Our empirical design leverages cross-sectional differences in how much entrepreneurs in different industries can benefit from AI tools when starting and operating new ventures.

Our dataset is structured at the industry-by-quarter level from 2020Q4 through 2024Q1. It combines measures of firm creation, financial and economic controls from the sources outlined in Section 2, and our LLM-generated measures of AI Assistance and Experimentation Intensity. Industries are defined at the 3-digit NAICS level, and we restrict the sample to sectors where private firm formation is economically meaningful and comparable. Table 3 reports descriptive statistics for the key variables in our empirical analysis. Panel (a) presents industry-by-quarter variables, including our primary outcome measure—the number of newly established firms—as well as investment flows and financial controls. Firm creation exhibits substantial variation across industries and over time, with a mean of 14,064 new firms per industry-quarter and a standard deviation of over 32,000. Investment flows are similarly heterogeneous, with a mean of \$1,730, 000 and a maximum of nearly \$95 million. Panel (b) summarizes industry-level variables that remain constant over time, including our standardized AI Assistance and Experimentation Intensity measures.

The remainder of this section introduces our empirical strategy, outlines the identification assumptions, and presents baseline estimates of the impact of AI Assistance on new firm creation. Robustness tests and validation exercises supporting the interpretation of the AI Assistance measure appear at the end of the section.

4.1. **Empirical Strategy.** Our goal is to estimate how the increased accessibility of generative AI tools, exemplified by the public release of ChatGPT, affects entrepreneurial activity.

To answer our research question, we exploit the fact that potential entrepreneurs are differentially exposed to the benefits and challenges of using AI and LLMs when starting and operating a business, depending on the industry in which they plan to operate[7].

The release of ChatGPT in November of 2022 was an unprecedented and unanticipated shock to the availability of LLM technologies for the public. While LLMs had been available as early as 2018, their use was restricted due to the proprietary nature of the technology. The release of ChatGPT marked the first instance of true democratization of the technology. We treat the arrival of ChatGPT as an unexpected shock to the availability of the technology and use it to study its effects on firm formation. The choice of this event as a shock is guided by the fact that the timing of the release was driven by ethical and safety considerations and model readiness—factors unrelated to business dynamism. Figure 3 provides additional evidence of the absence of anticipation using Google Trends search data for the term "ChatGPT." Search intensity is flat at or near zero throughout 2022 until the week beginning November 27, 2022 (which contains the November 30 release date), when it rises sharply. The lack of any prior trend in search activity confirms that the public debut of ChatGPT did not follow a period of increasing attention, further supporting our interpretation of the release as an unanticipated shock to AI tool accessibility.

In our DiD setting, cross-sectional exposure arises from the ex-ante benefit entrepreneurs derive from using AI virtual assistants to start a firm in a given industry.

We use a difference-in-differences design with varying treatment intensity to understand how the arrival of ChatGPT impacts business formation, conditional on a given industry's compatibility with AI.

At the Industry-by-Quarter level, the main specification is:

$$\text{IHS(N. New Firms)}_{it} = \alpha_0 + \alpha_1 \text{AI Assistance}_i \times \text{Post}_t + \text{Controls}_{it} + \mu_i + \mu_t + \epsilon_{it} \tag{5}$$

where the dependent variable $\text{IHS(N. New Firms)}_{it}$ is the inverse hyperbolic sine of the total number of new firms. AI Assistance_i is the industry level proxy measuring LLMs' ability to support entrepreneurs through the completion of various tasks, which we standardize for ease of interpretation. Post_t is an indicator equal to 1 from 2022Q4 onward, the periods during which ChatGPT became publicly available. μ_i are industry fixed effects used to control for industries' intrinsically higher levels of firm creation, which may not be related to the arrival

[7]This is the case, as different industries represent distinct collections of tasks that are differentially exposed to the benefits of using AI in the process of task completion.

of ChatGPT. μ_t are quarter fixed effects, which allow us to net out common and aggregate time trends unrelated to the dissemination of the new technology, which may still shape business formation. We also include a series of industry level, time-varying controls which collectively help isolate the effect of AI on firm entry from broader economic and financial conditions.

In our setting, recovering the Average Treatment on the Treated (ATT) effect requires (conditional) comparability across industries prior to the shock in terms of outcomes. To lend credibility to this assumption we inspect the coefficients from the dynamic model

(6)
$$\text{IHS(N. New Firms)}_{it} = \beta_0 + \sum_{t \neq 2022q3} \beta_t \text{AI Assistance}_i \times 1\{D = t\}_t + \text{Controls}_{it} + \mu_i + \mu_t + \epsilon_{it}$$

The loadings on the interaction between the AI Assistance measure and a quarter dummy identify the differential change in outcome due to a change in the independent variable (holding within industry and quarter characteristics fixed). For industries to be comparable, the coefficients on the leads (β_t) need to be insignificant.

4.2. **Results and Interpretation.** The first question we are trying to answer is whether there is an association between levels of firm formation and the assistance provided by LLMs to entrepreneurs.

The main analysis uses as level of industry aggregation 3-digit NAICS industries, over the period 2020Q4 to 2024Q1. In this analysis, we exclude certain industries that are either related to public administration or sectors where firm formation by private entrepreneurs is not applicable[8]. These sectors are excluded because they primarily consist of government functions or nonprofit organizations where traditional firm entry is either not possible or not relevant to entrepreneurial dynamics studied in this paper.

To understand industries' comparability across different AI Assistance levels, we begin by examining the coefficients from the dynamic specification presented in Equation 6. As shown by the insignificant coefficients on the leads, we do not observe meaningful pre-trends in firm creation. Moreover, the coefficients on the lags are positive and fairly similar across quarters, suggesting the presence of an effect with a rather persistent impact (Figure 4).

[8]Specifically, we drop the following 3-digit NAICS industries: 491 (Postal Service), 521 (Monetary Authorities-Central Bank), 525 (Funds, Trusts, and Other Financial Vehicles), 813 (Religious, Grantmaking, Civic, Professional, and Similar Organizations), 711 (Performing Arts, Spectator Sports, and Related Industries), 712 (Museums, Historical Sites, and Similar Institutions), 713 (Amusement, Gambling, and Recreation Industries), and the series 921 to 928 which cover various segments of Public Administration (921 – Executive, Legislative, and Other General Government Support; 922 – Justice, Public Order, and Safety Activities; 923 – Administration of Human Resource Programs; 924 – Administration of Environmental Quality Programs; 925 – Administration of Housing Programs, Urban Planning, and Community Development; 926 – Administration of Economic Programs; 927 – Space Research and Technology; 928 – National Security and International Affairs).

These results remain robust after controlling for broader economic and financial conditions (Figure 4, Panel (b)).

Next, we test for existence of ATT using the compact regression in Equation 5. Industries that are more compatible with the use of LLMs in entrepreneurial tasks—measured by AI's ability to assist entrepreneurs in the process of venture formation—experience a positive and significant increase in firm creation (Table 4, Column (1)). The effects stemming from a widening in LLM access on firm formation do not appear to be driven by contemporaneous economic and financial conditions, the coefficient on the main variable hardly changes (Table 4, Column (2)). When evaluated at the pre-treatment mean number of firms (13,673), this 0.105 coefficient corresponds to an increase of roughly 11 percent.[9] To gauge economic significance, a coefficient of 0.105 implies that a one–standard deviation increase in AI Assistance raises new firm creation by roughly 10 percent. Evaluated at the pre-treatment mean number of new firms (13,673), this corresponds to about 1,513 additional firms, not a new total. This calculation uses the mean of the outcome in levels, since interpreting magnitudes requires returning to the original scale rather than the IHS-transformed mean.

Taken together, these findings suggest that the increased accessibility of AI tools coincided with a notable rise in new firm creation in AI-compatible industries.

4.3. **Validation of Interpretation and Robustness.** We begin by evaluating the robustness of our main findings along several dimensions. To bolster confidence in our causal interpretation and address potential threats to validity, we conducted a battery of robustness tests, detailed in Subsection A.1 of the Online Appendix. Specifically, we implemented richer fixed-effects structures (sector-by-quarter fixed effects) to mitigate omitted variable bias arising from broader sectoral shocks. We also tested alternative binary definitions of treatment (industries above the mean and median of AI compatibility), finding qualitatively similar results with even stronger magnitudes, thereby alleviating concerns about measurement error, bias, and sensitivity to outliers. A falsification exercise, using a placebo treatment date significantly preceding the actual event, confirmed no pre-existing differential trends, strongly supporting the parallel trends assumption. Moreover, we conducted a heterogeneity analysis by deciles of AI compatibility, revealing potential nonlinearities and threshold effects, further enhancing the robustness, interpretability, and policy relevance of our findings. Another concern is that the observed post-ChatGPT increase in firm creation could simply reflect

[9] With $\beta = 0.105$ and pre-treatment mean $y_0 = 13{,}673$, the implied post-period change for a one–standard-deviation increase in AI Assistance is

$$\Delta y = y_0(\cosh\beta - 1) + \sqrt{1 + y_0^2}\,\sinh\beta \approx 1{,}513 \text{ firms},$$

which corresponds to

$$\frac{\Delta y}{y_0} \approx e^{\beta} - 1 \approx 11.07\%.$$

broader industry dynamics rather than an AI-specific response. We examine this possibility in the context of entry surges during the COVID-19 period and long-term growth trajectories in value added (1970–2010) in Subsubsection A.1.6 of the Online Appendix, and confirm that our results are not driven by these confounding factors but instead reflect the effects of AI.

Overall, these robustness checks substantially strengthen our main results, reinforcing the validity of our difference-in-differences design with continuous treatment exposure.

A key concern with any new measure is whether it captures the underlying concept it is designed to reflect. To validate our AI Assistance Score, we compare it to two widely used measures in the literature. The first one is the O*NET-based measure of LLM exposure (AI Task Exposure) developed by Eloundou et al. (2023) and employed by several papers such as Gmyrek et al. (2023); Auer et al. (2024); Lindebaum and Fleming (2024); Chen et al. (2025); Eisfeldt et al. (2025); Ghosh et al. (2025) among others. This benchmark evaluates the extent to which the tasks performed in each occupation can be automated or augmented by large language models, and is widely accepted as a proxy for how much LLMs can support human work. We aggregate it to the industry level using employment shares, so that it is directly comparable to our AI Assistance score. The second measure is inspired by Acemoglu et al. (2022) and capture the share of industry level job postings in the Revelio Labs' data requiring AI-related skills using keywords in job descriptions; we term this measure AI Demand. This benchmark evaluates the industry-level demand for AI skills and, to the extent that demand for skills in higher in industries where AI usage is post to be more beneficial, it can proxy for some sort of complementation between an industry's operations and AI tools.

For the analysis, we fix both measure at pre-ChatGPT levels (2022Q3) and standardize them. We begin with simple correlation analysis between each alternative measure and our AI Assistance measure. The Pearson correlation coefficient between AI Assistance and the AI Task Exposure measure is two standardized measures is 0.53 ($p = 0.000$), while the estimate is 0.32 ($p = 0.004$) with the AI Demand. The Spearman rank correlation between AI Assistance and AI Task Exposure is 0.43 ($p = 0.000$), and 0.27 ($p = 0.015$) between the AI Assistance and AI Demand. These results indicate that at the industry level, neither alternative measures perfectly correlates with the AI Assistance variable, highlighting that the our variable conveys information that is not capture by broad industry-level AI compatibility measure.

To systematically determine whether the information conveyed by the AI Assistance measure can be superseded by other measures that broadly describe industry level AI compatibility, we use a horse race design. The idea of this approach is straightforward: we place AI Assistance and an alternative measure of AI–industry relevance side-by-side in the same difference-in-differences specification and evaluate which variable retains explanatory power when both compete to explain the same outcome. The specification is

$$
\begin{aligned}
\text{IHS(N. New Firms)}_{it} = \alpha'_0 + \alpha'_1 \text{AI Assistance}_i \times \text{Post}_t \\
+ \alpha'_2 \text{Alternative Measure}_i \times \text{Post}_t + \mu_i + \mu_t + \epsilon_{it}
\end{aligned} \tag{7}
$$

where the benchmark is either AI Task Exposure or AI Demand. The key test in this design is whether the coefficient on the competing benchmark measure remains significant once AI Assistance is included.

We report the results of the horse race test in Table 5, in Column (1) we present the baseline results from the main analysis and in Columns (2) and (3), we show the results of the tests. Across all specifications, we find that the interaction between AI Assistance and the post-ChatGPT period remains large and statistically significant, whereas the coefficients on both AI Task Exposure (Table 5, Column (2)) and AI Demand (Table 5, Column (3)) are small and statistically insignificant. This pattern indicates that the variation in AI Assistance that predicts firm creation is not subsumed by existing measures of AI exposure or AI adoption.

The horse race design therefore provides evidence that AI Assistance captures an aspect of AI complementarity that is specific to entrepreneurship—namely, the extent to which AI helps founders perform the broad, cross-functional tasks involved in starting and running a new business—rather than characteristics of AI that are relevant primarily for incumbent firms.

5. Potential Mechanisms Linking AI and Firm Creation

Thus far, we have shown that AI-compatible industries experienced relatively higher levels of firm formation following the release of ChatGPT, suggesting that AI technology has the potential to stimulate entrepreneurial activity. In this section we present three potential mechanisms linking AI availability and firm formation.

5.1. **Does AI Reduce the Cost of Experimentation for New Entrepreneurs?** A central feature of entrepreneurship is iterative trial-and-error, encompassing idea validation, product testing, and business model refinement. If AI tools lower the costs associated with these activities, then the positive effect of AI on firm creation should be especially pronounced in industries where experimentation plays a central role. To test this mechanism, we build directly on the dataset introduced in Section 4, which combines measures of firm creation with industry-level indicators of AI Assistance and Experimentation Intensity. This ensures that our analysis of the experimentation channel is based on the same empirical foundation as the main results.

5.1.1. *Entrepreneurship as an Experimental Process.* Scholars increasingly conceptualize entrepreneurship as a process of experimentation, emphasizing the iterative and exploratory nature of creating and developing new ventures. This perspective views entrepreneurship not merely as the pursuit of opportunities, but as a systematic approach to testing hypotheses, learning from outcomes, and adapting strategies accordingly. The lean startup methodology, championed by Ries (2011), advocates for building minimal viable products, conducting experiments to validate business models, and iterating rapidly in response to market reactions. This experimental approach aligns with the notion that entrepreneurial success often hinges on the ability to pivot and refine ideas through continuous testing and learning (Blank, 2020).

AI has the potential to increase entrepreneurs' productivity in completing experimental tasks related to setting up a firm. This, in turn, reduces the cost of experimentation, making it relatively more appealing to start a firm and leading to higher levels of firm formation.

Our hypothesis is that industries in which AI significantly reduces experimentation costs should experience a greater increase in business formation after ChatGPT's release. To proxy for the benefit accrued from reducing the cost of experimentation, we interact our measure of AI Assistance with the measure of Experimentation Intensity we devised.

5.1.2. *Empirical Strategy for Testing the Experimentation Channel.* We study the relation between the cost of experimentation and firm formation by first sorting industries into groups based on their levels of Experimentation Intensity. Specifically, we define a binary variable equal to one if an industry falls within the top quartile of the distribution of Experimentation Intensity scores, and zero otherwise. This coarse specification serves two purposes. First, it mitigates concerns about noise in the original continuous measure, which stems from the subjective nature of task scoring and aggregation, and which may lead to instability in more complex regression specifications such as our triple interaction model. Second, using a top-quartile indicator allows us to focus on industries where experimentation is most salient, without relying on arbitrary thresholds or assuming linearity in treatment intensity. Relative to alternative splits (e.g., above/below median or quintiles), the top quartile definition isolates industries where experimentation is plausibly a first-order constraint in the firm formation process. This focus sharpens our test of the mechanism that AI reduces the cost of experimentation, by contrasting the most experimentation-intensive industries with all others.

We start our mechanism analysis by revisiting Equation 5, now interacting AI Assistance with the Experimentation Intense indicator presented above in a DDD design. At the

Industry-by-Quarter level, the main specification is:

$$
\begin{aligned}
\text{IHS(N. New Firms)}_{it} = \gamma_0 &+ \gamma_1 \text{AI Assistance}_i \times \text{Experimentation Intense}_i \times \text{Post}_t \\
&+ \gamma_2 \text{AI Assistance}_i \times \text{Post}_t + \gamma_3 \text{Experimentation Intense}_i \times \text{Post}_t \\
&+ \text{Controls}_{it} + \mu_i + \mu_{It} + \varepsilon_{it}
\end{aligned} \tag{8}
$$

where the dependent variable IHS(N. New Firms)$_{it}$ is the inverse hyperbolic sine of the total number of new firms. AI Assistance$_i$ is the industry level proxy measuring LLMs' ability to support entrepreneurs through the completion of various tasks, which we standardize for ease of interpretation. Experimentation Intense is a binary variable equal to one if an industry falls within the top quartile of the distribution of Experimentation Intensity scores, and zero otherwise. Post$_t$ is an indicator switching to 1 from 2022Q4 forward—periods during which ChatGPT is publicly available. μ_i are industry fixed effects used to control for industries' intrinsically higher levels of firm creation, which may not be related to the arrival of ChatGPT. In this specification we replace quarter fixed effects (μ_t) with sector-by-quarter fixed effects (μ_{It}), where sector is defined at the 2-digits NAICS level. We introduce a richer fixed effects structure to ensure that our results are not driven by sector-wide trends that may correlate with Experimentation Intensity (e.g., sectoral adoption of AI tools, recovery dynamics post COVID or broad sector policy changes). This allows us to isolate the variation across 3-digit industries within a given sector and quarter, thereby ensuring that our estimates are not confounded by broader sectoral dynamics and reflect genuine heterogeneity in how AI interacts with the cost of experimentation. In particular, we are concerned about sector-specific trends that could disproportionately affect experimentation-intensive industries. For example, some sectors—such as Information (NAICS 51) and Professional Services (NAICS 54)—may have seen a wave of AI-related innovation, investment, and entrepreneurial interest after the release of ChatGPT. These sectors also tend to include experimentation-heavy sub-industries (e.g., digital media, marketing, consulting), where trial-and-error is critical to new firm development. If these sector-wide surges in firm creation are not controlled for, we risk mistakenly attributing them to the interaction between AI Assistance and Experimentation Intensity, when in fact they reflect a general sectoral response to the AI shock.

In our specification, the triple interaction AI Assistance × Experimentation Intense × Post implements a triple-difference design. Intuitively, the estimator compares the post-ChatGPT increase in firm creation for AI-compatible industries relative to less AI-compatible industries, and then asks whether this differential effect is itself larger in industries where experimentation plays a more central role. In other words, the DDD isolates the *additional* post-ChatGPT increase in entry that arises only when an industry is both (i) highly AI-assistable and (ii) highly experimentation-intensive. By netting out the DiD effects associated with AI compatibility alone and experimentation intensity alone, the triple interaction coefficient γ_1 captures the incremental effect attributable specifically to a reduction in experimentation costs. This allows us to test directly whether industries that rely more heavily on trial-and-error benefit disproportionately from the arrival of AI tools.

We also revise the dynamic specification to tailor it to the mechanism we are considering so to visually inspect parallel trends. The event study specification is

(9)
$$
\begin{aligned}
\text{IHS(N. New Firms)}_{it} = \delta_0 &+ \sum_{t \neq 2022q3} \delta_t^1 \text{AI Assistance}_i \times \text{Experimentation Intense}_i \times 1\{D=t\}_t \\
&+ \sum_{t \neq 2022q3} \delta_t^2 \text{AI Assistance}_i \times 1\{D=t\}_t \\
&+ \sum_{t \neq 2022q3} \delta_t^3 \text{Experimentation Intense}_i \times 1\{D=t\}_t \\
&+ \text{Controls}_{it} + \mu_i + \mu_{It} + \varepsilon_{it}
\end{aligned}
$$

Here, we are interested in studying the sign and significance of the loadings on the triple interaction (δ_t^1), as they inform us as to whether different industries differ in the joint effect of AI Assistance and Experimentation Intensity prior to the shock.

5.1.3. *Results: AI Assistance in Experimentation Intense Industries.* As a first step in investigating the mechanism, we examine the coefficients from the dynamic regression in Equation 9. Analyzing the significance of the event study coefficients allows us to assess whether industries systematically differed in business formation based on their AI compatibility and reliance on experimentation prior to the arrival of ChatGPT. As shown in the event study plots (Figure 5), the coefficients for the lead periods are statistically insignificant and relatively stable in magnitude, suggesting that pre-existing cross-industry differences are unlikely to drive the results, and that there is no evidence of anticipatory effects. The effects on the lags are positive and incremental in size over quarters, probably driven by differential rate of take up over time, which are to be expected for complex technologies (Figure 5). The findings are unaffected when including the vector controls to account for varying financial and economic conditions (Figure 5, Panel (b)).

Next, we test for existence of an ATT using the compact regression in Equation 8. Industries where the benefits from reducing the cost of experimentation are higher—industries where experimentation is prominent and AI can strongly assist with task completion—experience a positive and significant increase in firm creation (Table 6, Column (1)). Moreover, the fact that neither the coefficient on AI Assistance$\times$Post and Experimentation Intense$\times$ Post are significant, suggest that the differences in firm creation levels after the arrival of

ChatGPT are fully explained by a reduction in the cost of experimentation. It is important to acknowledge that AI may still influence firm creation through mechanisms that operate uniformly within broad sectors. However, because the specification includes sector-by-quarter fixed effects, such common sector-wide effects are absorbed and cannot be separately identified in our analysis.

The results in the analysis display economically significant effects: one standard deviation increase in the AI Assistance measure leads to about 3,500 new firms being created in high Experimentation Intensity industries, which equates to roughly a 25%[10] increase from the pre-treatment mean.

To the extend to which our battery of industry-by-quarter varying financial and economic controls can proxy for time varying within industry economic and financial conditions, the results are not driven by that. On the contrary, the inclusion of the control set sightly increases the magnitude of the effect of the mechanism (Table 6, Column (2)).

In Section B.2 of the Online Appendix we show that our results are unchanged when using *Log*(N. New Firms) in place of IHS, assuaging concerns regarding the results being an artifact of the IHS specification. We also test the robustness of our results by providing difference binarized version of the Experimentation Intense indicator. In Subsection A.2 of the Online appendix we categorize industries as being Experimentation Intense using as cutoff either the sample mean or median. The results remain qualitatively unchanged.

Overall, the evidence presented in this section supports the interpretation that the rise in firm creation following the release of ChatGPT is concentrated in industries where AI can meaningfully reduce the cost of experimentation. These findings are not explained by pre-trends, sector-wide shocks, or time-varying financial and economic conditions, reinforcing the role of reduced experimentation costs as a key mechanism.

5.2. **AI as a Catalyst of Private Financing.** Beyond reducing the costs of experimentation, AI may also influence firm creation by shaping how capital is allocated. If investors perceive AI as enhancing founders' productivity or mitigating early-stage risks, industries more compatible with AI should attract relatively more private financing. Such reallocation of capital could further amplify entry in AI-compatible sectors. To test this mechanism, we rely on the same dataset described in Section 4, which integrates firm creation and capital allocation data with industry-level measures of AI Assistance.

5.2.1. *Do Entrepreneurs Cater to Investors?* Industries inherently more compatible with AI should experience productivity gains after the deployment of the technology, consequently

[10]This does not mean the mechanism regression contradicts the baseline results of %11; rather, it shows that the baseline average of 1,500 firms aggregates both low- and high-experimentation industries, and once we isolate the channel where AI is most relevant, the impact is amplified. In other words, the mechanism identifies the subset of industries where AI's ability to lower experimentation costs is most potent.

attracting more capital, due to higher expected returns. The heightened flow of capital into highly AI exposed industries entices entrepreneurs, as it increases their perceived probability of success. More specifically, entrepreneurs cater to investor by starting companies in industries where financing is abundant. Therefore, an alternative explanation as to why we see increased firm creation in AI-compatible industries—following the arrival of ChatGPT—is that industries where AI is highly beneficial for entrepreneurs experience an increase in funding by investors, which encourages prospective entrepreneurs to start companies in that same industry.

5.2.2. *Mediation Strategy to Test the Capital Allocation Channel.* We investigate the validity of this potential mechanism using mediation analysis. First, we estimate the Total Impact of AI-compatibility on firm creation:

$$\text{IHS(N. New Firms)}_{it} = \zeta_0 + \zeta_1 \text{AI Assistance}_i \times \text{Post}_t + \mu_i + \mu_{It} + \epsilon_{it} \tag{10}$$

where the dependent variable $\text{IHS(N. New Firms)}_{it}$ is the inverse hyperbolic sine of the total number of new firms. AI Assistance_i is the industry level proxy measuring LLMs' ability to support entrepreneurs through the completion of the tasks of various tasks, which we standardize for ease of interpretation. Post_t is an indicator switching to 1 from 2022Q4 forward—periods during which ChatGPT is publicly available. μ_i are industry fixed effects used to control for industries' intrinsically higher levels of firm creation, which may not be related to the arrival of ChatGPT. μ_{It} are sector-by-quarter[11] fixed effects, which allow us to net out time trends that are common to macro industry groups and which may shape business formation. If ζ_1 is significant, it supports the idea that firm creation increases more in AI-compatible industries after the release ChatGPT.

Next, we check whether AI compatibility results in higher levels of capital allocated post-ChatGPT:

$$\text{IHS(Investments)}_{it} = \eta_0 + \eta_1 \text{AI Assistance}_i \times \text{Post}_t + \mu_i + \mu_{It} + \epsilon_{it} \tag{11}$$

where the dependent variable $\text{IHS(Investments)}_{it}$ is the inverse hyperbolic sine of the total amount of VCPE investments in million of dollars[12]. AI Assistance_i is the industry level proxy measuring LLMs' ability to support entrepreneurs through the completion of the tasks of various tasks, which we standardize for ease of interpretation. Post_t is an indicator

[11]This specification differs from Equation 5 in Section 4, as it is more stringent in terms of fixed effects structure. We include sector-by-quarter fixed effects to keep the regression specification consistent with the test presented in Subsection 5.1.

[12]It is important to notice that while the total capital allocated to a given industry is indeed an equilibrium outcome, which could be driven by both demand and supply factors, the tests are not aimed at understanding which side is responding, rather their objective is to elicit the presence of an overall capital shift to the arrival of AI that might coincide with changes in the level of firm creation and potentially act as the main driver of increased firm creation.

switching to 1 from 2022Q4 forward—periods during which ChatGPT is publicly available. μ_i are industry fixed effects used to control for industries' intrinsically different levels of allocated capital, which may not be related to the arrival of ChatGPT. μ_{It} are sector by quarter fixed effects, which allow us to net out time trends that are common to macro industry groups and which may shape business formation. The sign and significance of η_1 will determine whether funding availability acts as a conduit for the effect of the popularization of AI technologies on firm formation.

Finally, we estimate the Mediated Effect of the treatment after controlling for investment flows. We include Investment in the firm creation regression as a mediator:

(12)

$$\text{IHS(N. New Firms)}_{it} = \theta_0 + \theta_1 \text{AI Assistance}_i \times \text{Post}_t + \theta_2 \text{IHS(Investments)}_{it} + \mu_i + \mu_{It} + \epsilon_{it}$$

where all the variables are measured as usual. θ_1 measures how much AI compatibility directly affects firm creation, controlling for investment (Direct Effect). θ_2 is the residual explanatory power of capital supply after conditioning on the treatment (AI Assistance $\times$ Post).

5.2.3. *Results: Investment Flows and Firm Creation.* As a preliminary step for the mediation analysis, we examine the event study for the total amount of investments (Figure 6). The coefficients on the leads are all insignificant with no clear pattern arising, suggesting two things. First, more AI compatible industries where not experiencing differential levels of investments prior to the arrival of ChatGPT. Second, investors did not anticipate the productivity boosting effect of the technology and rushed to fund AI compatible industries.

We present the results of the mediation analysis in Table 7. The treatment displays a positive and significant total impact: industries that are more compatible with LLMs display relatively higher firm creation following the arrival of ChatGPT (Column (1)). On the other hand, when we look at the association between the treatment and the total amount invested by VCPE funds—the mediating variable—, we find that the coefficient is negative and significant, suggesting that the arrival of AI coincides with lower levels of capital deployed in AI compatible industries (7, Column (2)). This finding is not consistent with the hypothesis that the effect of an increase in access to AI impacts firm creation via a capital allocation channel, as we actually observe reallocation away from AI compatible industries. In 7, Column (3), we examine the direct impact of the treatment, now mediating for investment levels. We find that the effect of the release of ChatGPT remains unchanged after including the mediator, which itself is statistically insignificant, suggesting that the effect of the treatment is not flowing through a change in capital allocation. These findings also help address the concern that the main effects (Table 4) might be driven by changes in the investment opportunity set rather than by AI lowering the cost of entrepreneurial experimentation. If the release of ChatGPT had primarily expanded profitable investment opportunities—either directly, by creating new AI-related markets, or indirectly, by increasing incumbents' productivity and stimulating downstream or upstream demand—we would expect to observe a corresponding rise in aggregate investment following the diffusion of AI. The results in this section show no evidence that investment rose more in high-AI-assistance industries relative to less AI-compatible industries in the post-ChatGPT period. The absence of an investment surge suggests that the overall pool of available capital remains stable, and that the observed increase in firm formation arises from entrepreneurs' enhanced ability to operate and experiment with existing resources rather than from a broad expansion in profitable opportunities. In this sense, the evidence is more consistent with AI reducing the fixed costs of entry and facilitating experimentation, rather than shifting the investment opportunity set.

All in all, given the current evidence at hand, it is unlikely that firm formation is driven by entrepreneurs catering to investors, as we do not find evidence that firm creation increases in AI compatible industries following the arrival of ChatGPT because investors anticipate productivity growth and flood the industries with funds. It is important to highlight that our findings do not suggest that relatively less capital was allocated to AI industries following the arrival of ChatGPT. Rather, our treatment group is defined as the set of industries where entrepreneurs would benefit the most from using LLMs to carry out the tasks involved in establishing and running a firm. Indeed, private investment in AI-related companies has increased substantially in recent years (Maslej et al., 2025). Hence, these findings further confirm that our measure of AI compatibility does not perfectly overlap with the traditional group of AI and AI-related industries. This distinction ensures that our design, which exploits the arrival of ChatGPT as an exogenous shock, does not merely capture a technological shock confined to technology firms.

While the use of IHS is innocuous for firm creation, since the dependent variable is strictly positive, the critique of Chen and Roth (2024) has sharper relevance in the context of investment flows. In some industry–quarter cells, total VCPE investment is zero, which means that conventional log specifications are undefined and IHS implicitly combines extensive and intensive margins. To address this concern, in Subsection B.3 we implement a battery of robustness tests. In particular, we (i) replicate the specifications using logs whenever feasible, (ii) re-estimate the investment regressions using Poisson pseudo-maximum likelihood (PPML), which is transformation-free and well defined in the presence of zeros, and (iii) explicitly separate investments into extensive and intensive margins in a two-part specification. Across all approaches, the coefficient on *AI Assistance* $\times$ *Post* remains stable and the mediator continues to be insignificant. This demonstrates that our conclusion—that capital reallocation does not explain the rise in firm creation in AI-compatible industries—is not an artifact of the IHS functional form.

Lastly, we acknowledge that the relation between capital allocation and AI compatibility is not ex-ante obvious. On the one hand, higher expected productivity may attract greater investment into AI-compatible industries. On the other hand, firms in these industries may require less external financing if AI reduces the cost of key inputs, in which case average deal sizes would decline rather than increase. In Subsection A.3 of the Online Appendix we test this alternative channel by replacing total capital with the average size of investment rounds as a mediator. We find little evidence in support of this mechanism: while average deal sizes are somewhat smaller in AI-compatible industries after ChatGPT, these changes do not predict entry, and the treatment effect on firm creation remains unchanged. Overall, capital does not appear to be reallocated in AI-compatible industries, and it is therefore unlikely that financing dynamics are the primary driver of the increase in firm creation.

5.3. **Does AI Bridge Human Capital Gaps?** A further channel through which AI may affect entrepreneurship is by alleviating human capital constraints. Many potential founders possess incomplete skill sets relative to the demands of operating a business in their chosen industry. If AI tools complement individual abilities and mitigate these deficiencies, then industries where potential founders have larger skill gaps should see a disproportionate boost in entry following the arrival of ChatGPT. To test this hypothesis, we construct a novel individual-level dataset from Revelio Labs, which tracks employment histories and job titles of U.S. workers. By linking these histories to standardized task and skill data from O*NET, we build dynamic skill portfolios for each worker and compare them against industry-specific requirements. This allows us to quantify skill gaps and examine whether AI reduces the barriers to firm creation posed by human capital limitations.

5.3.1. *Missing Skills as a Barrier to Entrepreneurial Entry.* A lack of skills is a legitimate barrier that prevents potential entrepreneurs from entering certain industries. AI—particularly LLM-powered virtual assistants—can supplement the skill sets of entrepreneurs by allowing them to complete tasks they previously could not perform. Individuals whose skill set is better complemented by AI see a reduction in barriers to entry, and should be more likely to sort into self-employment. Therefore, a potential channel linking AI availability to firm creation is LLMs' ability to bridge human capital gaps. This channel is distinct from a reduction in the cost of experimentation: the latter applies to tasks that individuals can already perform but can now complete more efficiently thanks to the presence of LLMs. Here, AI supplements the entrepreneur's abilities, enabling them to perform tasks for which they previously lacked the skills. AI, in this sense, fills skill gaps, letting entrepreneurs operate businesses they otherwise wouldn't be able to run.

An implication of this mechanism is that individuals whose skill sets are better complemented by AI should sort themselves into unfamiliar industries at higher rates following the arrival of ChatGPT.

If following the arrival of ChatGPT, entrepreneurs become more likely to start firms in industries for which they do not have the appropriate skill set, but these are also industries where ChatGPT can help bridge the skill gap to start a venture, then firm formation is likely fueled by LLMs' ability to bridge skill gaps.

5.3.2. *Empirical Design for Measuring Skill Gaps and Entry.* We estimate the regression at the Individual-by-Industry-by-Quarter level (jit)

$$\text{P(Entrepreneur)}_{jit} = \iota_0 + \iota_1 \text{Skill Gap}_{jit} \times \text{Post}_t + \iota_2 \text{Skill Gap}_{jit} + \mu_{ji} + \mu_{jt} + \mu_{it} + \epsilon_{jit} \tag{13}$$

$\text{P(Entrepreneur)}_{jit}$ is an indicator reporting 10000 if individual j started a firm in industry i in quarter t 0 otherwise[13]. Skill Gap_{jit} is the skill gap value for individual j with respect to industry i at quarter t. Post_t is an indicator for the post-ChatGPT period (2022Q4 onward). μ_{ji} are industry by individual fixed effects, which control for any time-invariant preference or sorting between an individual and an industry; for example, if an individual is more likely to start a business in a given industry due to prior experience or interest, this FE absorbs that. μ_{jt} are individual by quarter fixed effects, which control for individual-level shocks or trends over time, such as job loss, career shifts, or general changes in propensity to become an entrepreneur. Finally, μ_{it} are industry by quarter fixed effects, which control for all industry-specific shocks over time, such as business cycle exposure, regulatory changes, or industry-specific AI trends.

The coefficient ι_2 reflects the importance of a skill gap for an individual's ability to penetrate a given industry, while ι_1 represents the differential effect of the existence of a skill gap in penetrating a given industry pre and post the arrival of ChatGPT. When ι_1 is positive, entrepreneurs sort themselves into harder to penetrate industries at higher rates after the arrival of ChatGPT. Moreover, when ι_2 is negative, we know that a given industry is unconditionally harder to penetrate given the individual's skill set. Therefore, if ι_1 is positive and ι_2 is negative, we can conclude that entrepreneurs sort themselves into industries where ChatGPT is able to provide support to bridge the skill gap and said industries are generally harder to break into for entrepreneurs without the right set of skills, implying that ChatGPT eases the barrier to entry due to skill gaps.

5.3.3. *Results: Skill Gaps and Entry into Entrepreneurship.* In Table 8 we presents the results of the skill gap mechanism test. The sample used includes individuals present on RL who were actively working between 2020q4 and 2024q1; for computational easy we draw a random sample of 200,000 individuals to use for the analysis. First, we assess the reliability of our skill gap measure (Table 8, Column (1)) and find that our proxy for skill mismatch predicts whether an individual will become an entrepreneur: the coefficient is negative and significant, one standard deviation increase in the skill gap measure—i.e. moving away from

[13]We multiply the indicator by a constant for ease of exposition of the results.

the optimal skill benchmark—leads to a decline of over 0.002265%[14] in the likelihood of starting a firm in a specific industry-quarter, which equates to a 3% reduction with respect to the pre-treatment entrepreneurial flow (0.755%).

Next, we estimate Equation 13 to test the mechanism (Table 8, Column (2)). The skill-gap coefficient is negative and significant, whereas the interaction with the post-ChatGPT indicator is positive but statistically insignificant. Although the signs align with the hypothesized mechanism, only the main effect (ι_2) is statistically significant, reiterating that entrepreneurs tend to enter industries that better match their skills. On the other hand, the coefficient on the interaction between the skill gap and the post variable is insignificant, although it is positive, consistent with the mechanism's prediction.

We next restrict the sample to individuals who eventually become entrepreneurs during the sample period and re-estimate Equation 13. This exercise isolates variation in where entrepreneurs choose to enter—holding fixed the decision to become a founder—and allows us to test whether the availability of AI tools shifts entrepreneurs toward industries that are farther from their prior skill set. First, the skill mismatch variable accurately predicts the industry in which the individual will seek to start a firm: a higher level of skill misalignment with respect to a given industry's requirement correlates negatively with the likelihood of becoming an entrepreneur in the same industry (Table 9, Column (1)). On the other hand, the interaction captures whether post-ChatGPT entrants are relatively more likely to start firms in industries in which they initially lacked traditional skills. The results, reported in Table 9, Column (2), show that the interaction coefficient remains positive but statistically insignificant. This suggests that, conditional on deciding to become an entrepreneur, the introduction of generative AI has not yet led to a meaningful reallocation of founders across industries according to their prior skill gaps.

Overall, the combined evidence in Table 8 and 9 suggests that while AI might modestly ease entry into industries requiring unfamiliar skills, the effect is not strong enough to be statistically detected. At the same time, the strong negative relationship between the skill-gap measure and entry prior to the AI shock supports the validity of the measure as a proxy for human-capital mismatch across industries. Hence, AI's ability to bridge such mismatches appears limited during our sample period and is unlikely to be a primary driver of the post-ChatGPT rise in firm creation.

6. Who are the New Entrepreneurs? Evidence on Post-AI Firm Quality

As a final step in our analysis, we examine whether firms created after the release of ChatGPT differ intrinsically from comparable firms established before its arrival.

[14]This is because the coefficient is result from multiplying the binary outcome by 1,000 to enhance readability.

A central implication of our mechanism is that AI lowers the cost of experimentation, thereby reducing barriers to entry for potential entrepreneurs. Canonical economic models predict that while entry increases, following a reduction in the barriers to entry, this influx of new entrants dilutes average firm quality (Lucas Jr, 1978; Hopenhayn, 1992). If lower entry costs enable the formation of ventures that previously would not have been viable, the post-AI increase in firm formation might be accompanied by a composition shift toward lower-quality firms: a "lemon market" for new ventures. However, this need not always be the case. In an alternative framework, the effect of lower entry barriers on average firm quality depends on how these barriers are reduced. If AI primarily lowers fixed entry costs—a uniform reduction that benefits all potential founders equally—then canonical predictions hold and average firm quality declines as marginal, low-ability entrepreneurs enter. By contrast, if AI disproportionately reduces the convex costs of experimentation and scaling, the gains from AI are *ability-complementary*: higher-ability individuals benefit more because they engage in more experimentation and operate at larger scale. In this case, the post-AI payoff gain is increasing and convex in entrepreneurial ability, inducing both higher entry rates and a shift in the composition of entrants toward higher-ability types. Consequently, firm formation rises while average firm quality improves.

Guided by this distinction, we turn to the data to test whether the new firms spawned in the wake of ChatGPT differ in quality from their predecessors.To do so, we construct a firm-level dataset from Revelio Labs that tracks incorporation dates, industry affiliations, survival, employment growth, and average worker education for firms established between 2020Q4 and 2024Q1, with outcomes observed through 2025Q1[15]. By linking each firm to its industry's AI compatibility score, we can directly compare proxies of firm quality for post-AI entrants with those of earlier cohorts.

6.1. **Empirical Strategy.** We assess whether AI-driven firm creation results in lower-quality entrants or in the formation of higher-quality ventures by examining three firm-level post-entry outcomes that capture complementary dimensions of venture quality.

First, *firm survival* reflects basic viability: higher-quality firms are more likely to endure beyond the initial quarters of operation, whereas lower-quality entrants tend to exit quickly. Second, *employment growth* proxies for a firm's ability to expand and scale, which depends on managerial capability, resource acquisition, and market success—key manifestations of operational quality. Third, the *average education level of employees* captures the quality of the workforce that new firms are able to attract. Because firm and worker quality tend to be positively assortatively matched, the education level of employees provides an informative

[15]We refresh the workforce data up to 2025Q1, so that we are able to calculate proxies of quality for firms registered in most recent years.

proxy for underlying firm quality: if newly founded firms in highly AI-compatible industries attract more educated workers, they are likely to be of higher quality.

Because firm quality is not directly observable, we examine multiple complementary dimensions—viability, growth capacity, and human capital composition—to provide a broader picture of how the post—AI surge in firm creation relates to the quality of new entrants.

Our empirical strategy uses a difference-in-differences framework that leverages cross-sectional variation in industry AI Assistance Scores and temporal variation around the release of ChatGPT in 2022Q44. Specifically, we estimate the following cross-sectional firm-level specification:

$$\mathrm{Y}_f^{(\tau)} = \kappa_0 + \kappa_1 \text{AI Assistance}_{i(f)} \times \text{Post}_{t(f)} + \mu_{i(f)} + \mu_{t(f)} + \mu_{s(f)} + \varepsilon_f \tag{14}$$

$\mathrm{Y}f^{(\tau)}$ denotes one of three outcomes observed τ quarters after registration: an indicator for survival ($\mathrm{P(Survived)}f^{(\tau)}$), the proportional change in employment since establishment (Emp. Growth$f^{(\tau)}$), or the average years of education among the firm's active workforce (Avg. Education$f^{(\tau)}$), with $\tau \in 4, 6, 8$. AI Assistance$_{i(f)}$ is the standardized AI compatibility score of firm f's industry, and Post$_{t(f)}$ is an indicator for registration after the release of ChatGPT. We include industry ($\mu_{i(f)}$), registration quarter ($\mu_{t(f)}$), and state ($\mu_{s(f)}$) fixed effects to account for unobserved heterogeneity.

6.2. **Results and Interpretation.** Table 10 shows that firms in highly AI-compatible industries founded after the release of ChatGPT exhibit significantly higher survival probabilities across all three time horizons. The interaction coefficient κ_1 is positive and statistically significant in all specifications, with point estimates ranging from 0.003 to 0.004. The estimated effects on firm survival are modest in absolute terms but economically meaningful. At one year, a one-standard deviation increase in AI Assistance raises survival by 0.3 percentage points, equivalent to a 0.36 percent relative increase over the pre-ChatGPT baseline survival rate of roughly 84%. At 18 months, the effect rises to 0.4 percentage points (0.52% relative to a 77% baseline), and at two years it remains at 0.3 percentage points (0.41% relative to a 74% baseline).

While the magnitudes appear smaller at first sight, they are well within the range of effects documented in the literature. For instance, Myers et al. (2025) finds that receipt of an SBIR grant raises annual survival probabilities by about 1.4 percentage points on a baseline survival rate of 80%, whereas Schwartz (2011) shows that business incubators often have no effect (and in some cases a negative effect) on long-term survival. It is important to note, however, that the SBIR study examines a highly idiosyncratic sample, as the treated firms are those able to secure a competitive R&D grant, while our estimates reflect the broad effect of AI tools across entire industries.

Turning to employment dynamics, Table 11 reports the effect of AI compatibility on post-entry size. We find that firms founded after the release of ChatGPT in industries with higher AI Assistance scores expand more rapidly in their early years. At one year, a one-standard deviation increase in AI Assistance is associated with 0.7 percentage points higher employment growth, which corresponds to a 4.7% increase relative to the pre-ChatGPT mean growth rate of 15%. At 18 months, the effect is 0.9 percentage points, or 5.6% relative to the baseline mean of 16%. At two years the estimated coefficient smaller: 0.6 percentage points, which implies a 3% increase relative to the baseline mean of 20%. This pattern suggests that the positive impact of AI on employment growth is mostly concentrated in the early stages of firm development.

In addition, we produce results regarding whether post-AI firms differ in terms level of the human capital they attract, conditional on the level of AI compatibility (Table 12). The analysis shows that firms founded after the arrival of ChatGPT in industries with higher AI Assistance scores tend to have a more educated workforce. At one year, a one–standard–deviation increase in AI Assistance is associated with an additional 0.796 years of educational attainment among the active workforce, a 5% increase relative to the pre-ChatGPT baseline mean of 15.914. At 18 months, the effect is 0.812 years (5.1% relative to the baseline mean of 15.954), and at two years the estimated coefficient remains positive though somewhat smaller at 0.719 years (a 4.5% increase relative to the baseline mean of 15.975). This pattern suggests that the positive impact of AI on firm creation is concentrated among entrepreneurs and ventures able to attract higher-quality human capital.

Taken together, these results are consistent with the alternative framework in which AI reduces the convex costs of experimentation and early-stage scaling, rather than merely lowering fixed entry costs. In such a setting, AI acts as an ability-complementary technology: it enhances the effectiveness of skilled founders in refining products, reaching customers, and organizing production, thereby improving the survival and early growth prospects of new firms. The finding that post-ChatGPT entrants in AI-compatible industries both survive longer and expand faster suggests that AI tools do not simply enable more entry, but also help new ventures execute more effectively once they are formed. In other words, generative AI appears to lower the cost of experimentation in a way that raises—not dilutes—the average quality of entrepreneurial outcomes.

7. Conclusion

This paper examines how the commercialization and widespread accessibility of artificial intelligence—exemplified by the release of ChatGPT in late 2022—has influenced entrepreneurial activity in the United States. While prior research has explored how AI reshapes labor markets and firm-level productivity, we focus on a previously underexplored margin: new firm creation. Specifically, we ask whether AI tools can lower barriers to entry for prospective entrepreneurs by supporting early-stage business formation tasks.

To answer this question, we construct novel industry-level measures of AI compatibility and experimentation intensity using a multi-step pipeline built on large language models. These measures are designed to capture (i) how useful AI is in assisting entrepreneurs with task completion in a given industry (AI Assistance), and (ii) the extent to which early-stage firm formation relies on trial-and-error (Experimentation Intensity). We link these measures to newly established firms across industries and quarters between 2020 and 2024.

Using a difference-in-differences framework with continuous treatment intensity, we exploit the arrival of ChatGPT as an exogenous shock to AI availability. Our results show that industries more compatible with AI experienced a significant increase in new firm creation following the release of ChatGPT. A one standard deviation increase in AI Assistance leads to roughly a 10% increase in firm entry, holding other factors constant. This effect is not driven by pre-existing trends and remains robust to controls for funding conditions and industry fixed effects.

The mechanism linking AI availability and firm formation is one entailing reductions in the cost of experimentation: industries that both rely on experimentation and benefit strongly from AI assistance exhibit the largest post-ChatGPT increase in firm creation. In contrast, we find little empirical support for alternative explanations based on capital supply or AI bridging human capital gaps. While these mechanisms are theoretically plausible, their contributions appear limited in practice during the period we study.

Moreover, our analysis of firm-level outcomes shows that the rise in entrepreneurship in AI-compatible industries is not accompanied by a decline in firm quality. On the contrary, new firms created after the release of ChatGPT are more likely to survive, expand faster, and attract more educated workers, indicating that AI lowers frictional barriers to entry without encouraging the formation of lower-quality ventures.

Our findings contribute to a growing literature on the economic implications of AI by offering new evidence that generative AI tools can foster entrepreneurial activity—not merely by enhancing productivity, but by lowering cognitive and operational barriers to firm formation. More broadly, our findings highlights how democratized access to advanced technologies can unlock new entry margins, especially in sectors where innovation and experimentation are central.

References

Acemoglu, D., D. Autor, J. Hazell, and P. Restrepo (2022): "Artificial intelligence and jobs: Evidence from online vacancies," *Journal of Labor Economics*, 40, S293–S340. 4.3

Aker, J. C. and I. M. Mbiti (2010): "Mobile phones and economic development in Africa," *Journal of economic Perspectives*, 24, 207–232. 1

Auer, R., D. Köpfer, and J. Švéda (2024): "The rise of generative AI: modelling exposure, substitution, and inequality effects on the US labour market," Tech. rep., CESifo Working Paper. 4.3

Babina, T., A. Fedyk, A. He, and J. Hodson (2024): "Artificial intelligence, firm growth, and product innovation," *Journal of Financial Economics*, 151, 103745. 1, E.1

Bao, J., B. Lou, and H. Sun (2025): "Generative AI and Entrepreneurial Entry," *Available at SSRN 5167018*. 1, E.1

Bernstein, S., E. Colonnelli, D. Malacrino, and T. McQuade (2022): "Who creates new firms when local opportunities arise?" *Journal of Financial Economics*, 143, 107–130. 1

Bick, A., A. Blandin, and D. J. Deming (2024): "The rapid adoption of generative ai," Tech. rep., National Bureau of Economic Research. E.1

Blank, S. (2020): *The four steps to the epiphany: successful strategies for products that win*, John Wiley & Sons. 5.1.1, E.2

Bonney, K., C. Breaux, C. Buffington, E. Dinlersoz, L. Foster, N. Goldschlag, J. Haltiwanger, Z. Kroff, and K. Savage (2024): "The impact of AI on the workforce: Tasks versus jobs?" *Economics Letters*, 244, 111971. E.1

Boudreaux, C. J., M. P. Escaleras, and M. Skidmore (2019): "Natural disasters and entrepreneurship activity," *Economics Letters*, 182, 82–85. 1

Boussioux, L., J. N. Lane, M. Zhang, V. Jacimovic, and K. R. Lakhani (2024): "The crowdless future? Generative AI and creative problem-solving," *Organization Science*, 35, 1589–1607. 1, E.1

Brynjolfsson, E., B. Chandar, and R. Chen (2025a): "Canaries in the Coal Mine? Six Facts about the Recent Employment Effects of Artificial Intelligence," . 1

Brynjolfsson, E., D. Li, and L. Raymond (2025b): "Generative AI at work," *The Quarterly Journal of Economics*, qjae044. 1, E.1

Chen, J. and J. Roth (2024): "Logs with zeros? Some problems and solutions," *The Quarterly Journal of Economics*, 139, 891–936. 5.2.3, B.1, **??**, **??**, B.2, **??**, **??**, B.3, **??**, B.3.2, **??**, **??**

Chen, Q., J. Ge, H. Xie, X. Xu, and Y. Yang (2025): "Large language models at work in China's labor market," *China Economic Review*, 102413. 4.3

Contigiani, A. (2023): "Experimentation and appropriability in early-stage ventures: Evidence from the US software industry," *Strategic Management Journal*, 44, 2128–2174. E.2

Contigiani, A. and T. Young-Hyman (2022): "Experimentation, planning, and structure in early-stage ventures: Evidence from pitch decks," *Strategic Entrepreneurship Journal*, 16, 425–459. E.2

Cui, C. and L. S.-Z. Li (2023): "Trade policy uncertainty and new firm entry: Evidence from China," *Journal of Development Economics*, 163, 103093. 1

Davis, S. J., J. Haltiwanger, and S. Schuh (1996): "Small business and job creation: Dissecting the myth and reassessing the facts," *Small business economics*, 8, 297–315. 1

Dell'Acqua, F., E. McFowland III, E. R. Mollick, H. Lifshitz-Assaf, K. Kellogg, S. Rajendran, L. Krayer, F. Candelon, and K. R. Lakhani (2023): "Navigating the jagged technological frontier: Field experimental evidence of the effects of AI on knowledge worker productivity and quality," *Harvard Business School Technology & Operations Mgt. Unit Working Paper*. 1, E.1

Doshi, A. R. and O. P. Hauser (2024): "Generative AI enhances individual creativity but reduces the collective diversity of novel content," *Science advances*, 10, eadn5290. 1, E.1

Eisfeldt, A. L., G. Schubert, B. Taska, and M. B. Zhang (2025): "Generative AI and Firm Values," *Available at SSRN 4436627*. 1, 4.3, E.1

Eloundou, T., S. Manning, P. Mishkin, and D. Rock (2023): "Gpts are gpts: An early look at the labor market impact potential of large language models," *arXiv preprint arXiv:2303.10130*. 1, 4.3, E.1

Ericson, R. and A. Pakes (1995): "Markov-perfect industry dynamics: A framework for empirical work," *The Review of economic studies*, 62, 53–82. E.2

Evans, D. S. and B. Jovanovic (1989): "An Estimated Model of Entrepreneurial Choice under Liquidity Constraints," *Journal of Political Economy*, 97, 808–827. 1

Ewens, M., R. Nanda, and M. Rhodes-Kropf (2018): "Cost of experimentation and the evolution of venture capital," *Journal of Financial Economics*, 128, 422–442. E.2

Ghosh, D., R. Ghosh, S. Roy Chowdhury, and B. Ganguly (2025): "AI-exposure and labour market: a systematic literature review on estimations, validations, and perceptions," *Management Review Quarterly*, 75, 677–704. 4.3

Gmyrek, P., J. Berg, and D. Bescond (2023): "Generative AI and jobs: A global analysis of potential effects on job quantity and quality," *ILO Working paper*, 96. 4.3

Gompers, P., J. Lerner, and D. Scharfstein (2005): "Entrepreneurial spawning: Public corporations and the genesis of new ventures, 1986 to 1999," *The journal of Finance*, 60, 577–614. 1

Haltiwanger, J. C. (2022): "Entrepreneurship during the COVID-19 pandemic: Evidence from the business formation statistics," *Entrepreneurship and Innovation Policy and the Economy*, 1, 9–42. A.1.6

Hampole, M., D. Papanikolaou, L. D. Schmidt, and B. Seegmiller (2025): "Artificial intelligence and the labor market," Tech. rep., National Bureau of Economic Research. 1, E.1

Hasbi, M. (2020): "Impact of very high-speed broadband on company creation and entrepreneurship: Empirical Evidence," *Telecommunications Policy*, 44, 101873. 1

Heller, D. and D. Asam (2024): "Generative AI and Firm-level Productivity: Evidence from Startup Funding Dynamics," *Available at SSRN 4877505*. E.1

Hellmann, T. and M. Puri (2000): "The interaction between product market and financing strategy: The role of venture capital," *The review of financial studies*, 13, 959–984. E.2

——— (2002): "Venture capital and the professionalization of start-up firms: Empirical evidence," *The journal of finance*, 57, 169–197. E.2

Hjort, J. and J. Poulsen (2019): "The arrival of fast internet and employment in Africa," *American Economic Review*, 109, 1032–1079. 1

Hjort, J. and L. Tian (2021): "The economic impact of internet connectivity in developing countries," *Annual Review of Economics*, 17. 1

Hopenhayn, H. A. (1992): "Entry, Exit, and Firm Dynamics in Long Run Equilibrium," *Econometrica*, 60, 1127–1150. 1, 6

Hui, X., O. Reshef, and L. Zhou (2024): "The short-term effects of generative artificial intelligence on employment: Evidence from an online labor market," *Organization Science*, 35, 1977–1989. E.1

Humlum, A. and E. Vestergaard (2025): "The unequal adoption of ChatGPT exacerbates existing inequalities among workers," *Proceedings of the National Academy of Sciences*, 122, e2414972121. E.1

Jovanovic, B. (1982): "Selection and the evolution of industry," *Econometrica*, 50, 649–670. E.2

Kerr, W. R. and R. Nanda (2011): "Financing constraints and entrepreneurship," *Handbook of Research on Innovation and Entrepreneurship. Cheltenham: Elgar*, 88–103. 1

Kerr, W. R., R. Nanda, and M. Rhodes-Kropf (2014): "Entrepreneurship as experimentation," *Journal of Economic Perspectives*, 28, 25–48. 1, E.2

Knight, F. H. (1921): *Risk, uncertainty and profit*, Houghton Mifflin Company. E.2

Koning, R., S. Hasan, and A. Chatterji (2022): "Experimentation and start-up performance: Evidence from A/B testing," *Management Science*, 68, 6434–6453. E.2

Kwan, A., B. Matthies, R. R. Townsend, and T. Xu (2025): "Entrepreneurial spawning from remote work," Tech. rep., National Bureau of Economic Research. 2.1

Labaschin, B., T. Eloundou, S. Manning, P. Mishkin, and D. Rock (2025): "Extending "GPTs Are GPTs" to Firms," in *AEA Papers and Proceedings*, American Economic Association 2014 Broadway, Suite 305, Nashville, TN 37203, vol. 115, 51–55. 2.1

Lazear, E. P. (2005): "Entrepreneurship," *Journal of Labor Economics*, 23, 649–680. 1

Levine, R. and Y. Rubinstein (2017): "Smart and illicit: who becomes an entrepreneur and do they earn more?" *The Quarterly journal of economics*, 132, 963–1018. 1

Lindebaum, D. and P. Fleming (2024): "ChatGPT undermines human reflexivity, scientific responsibility and responsible management research," *British Journal of Management*, 35, 566–575. 4.3

Lucas Jr, R. E. (1978): "On the size distribution of business firms," *The Bell Journal of Economics*, 508–523. 1, 6

Manso, G. (2011): "Motivating innovation," *The journal of finance*, 66, 1823–1860. E.2

——— (2016): "Experimentation and the Returns to Entrepreneurship," *Review of Financial Studies*, 29, 2319–2340. 1, E.2

Maslej, N., L. Fattorini, R. Perrault, Y. Gil, V. Parli, N. Kariuki, E. Capstick, A. Reuel, E. Brynjolfsson, J. Etchemendy, et al. (2025): "Artificial intelligence index report 2025," *arXiv preprint arXiv:2504.07139*. 5.2.3

Meincke, L., E. R. Mollick, and C. Terwiesch (2024): "Prompting diverse ideas: Increasing AI idea variance," *arXiv preprint arXiv:2402.01727*. 1, E.1

Myers, K. R., L. Lanahan, and E. Johnson (2025): "Small Business Innovation Applied to National Needs," Tech. rep., National Bureau of Economic Research. 6.2

Noy, S. and W. Zhang (2023): "Experimental evidence on the productivity effects of generative artificial intelligence," *Science*, 381, 187–192. 1, E.1

Otis, N., R. Clarke, S. Delecourt, D. Holtz, and R. Koning (2024): "The uneven impact of generative AI on entrepreneurial performance," *Available at SSRN 4671369*. 1, E.1

Ries, E. (2011): *The lean startup: How today's entrepreneurs use continuous innovation to create radically successful businesses*, Crown Currency. 5.1.1, E.2

Schwartz, M. (2011): "Incubating an Illusion? Long-term incubator firm performance after graduation," *Growth and change*, 42, 491–516. 6.2

Steinwender, C. (2018): "Real effects of information frictions: When the states and the kingdom became united," *American Economic Review*, 108, 657–696. 1

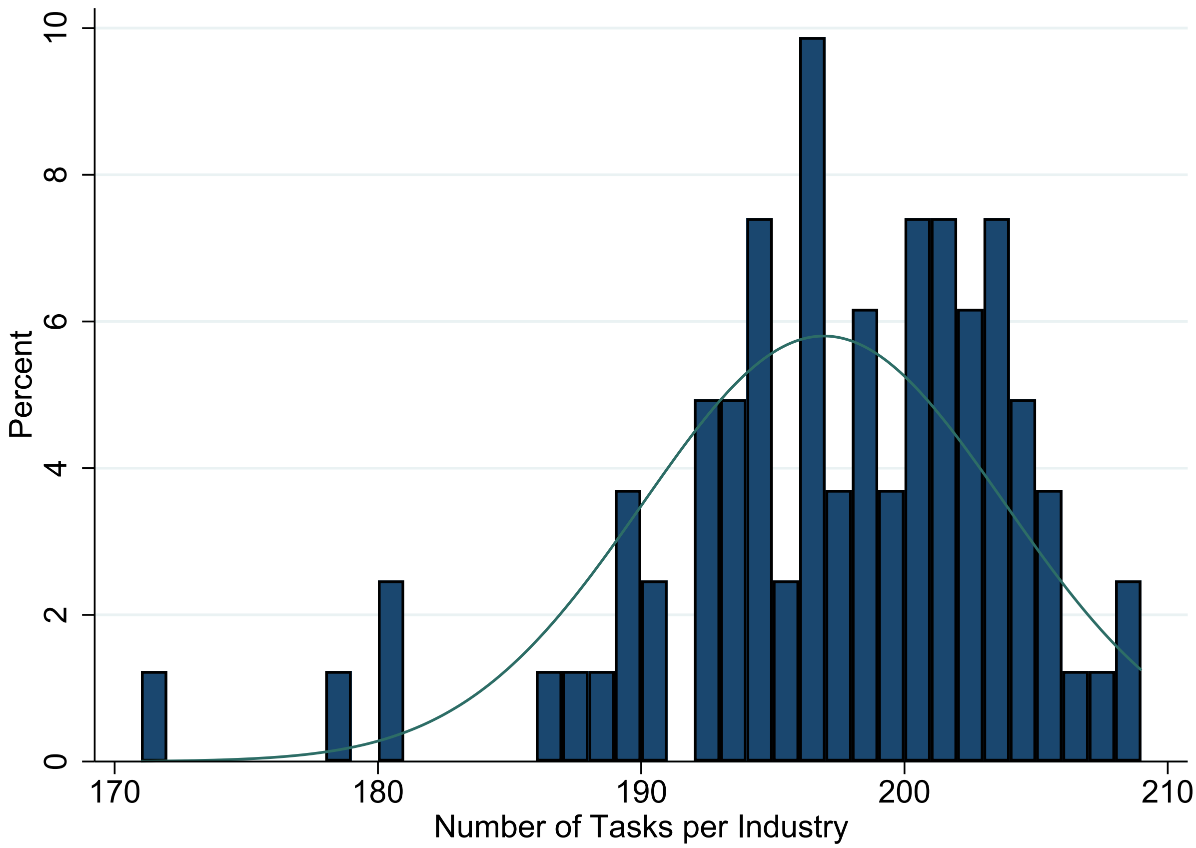
Stephens, H. M., E. A. Mack, and J. Mann (2022): "Broadband and entrepreneurship: An empirical assessment of the connection between broadband availability and new business activity across the United States," *Telematics and Informatics*, 74, 101873. 1

Wang, X., C. Feng, and T. Sun (2025): "AI Spillover is Different: Flat and Lean Firms as Engines of AI Diffusion and Productivity Gain," *arXiv preprint arXiv:2511.02099*. 2.1

Yang, C.-H. (2022): "How artificial intelligence technology affects productivity and employment: Firm-level evidence from Taiwan," *Research Policy*, 51, 104536. 1

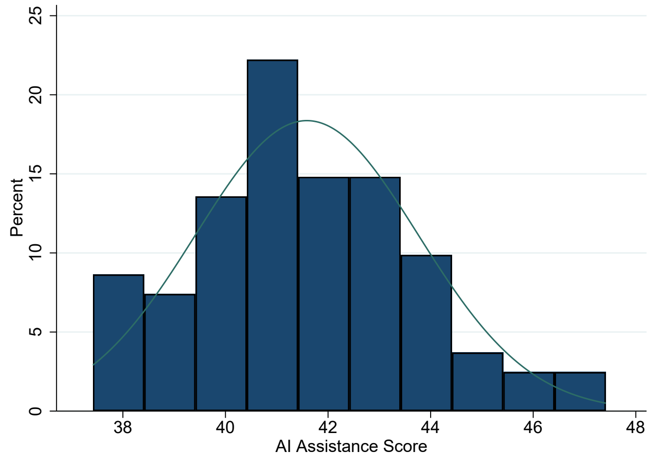
Figures and Tables

Figure 1. Distribution of the Number of Tasks per Industry

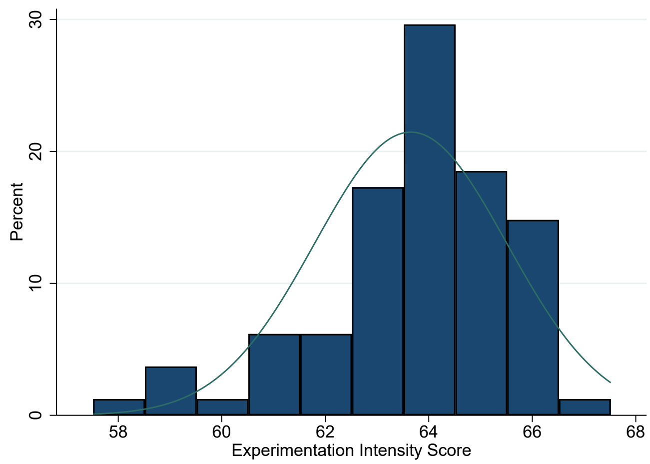


Notes: The figure presents the distribution of the number of task across all 81 three-digit NAICS industries in our sample. Tasks are extracted from the corpus of several texts focusing on entrepreneurship. After all the tasks have been extracted and deduplicated, an LLM assesses the relevance of any given task with respect to a particular industry. Once the universe of industry-specific tasks is created, we calculate the number of tasks per industry and plot it. The horizontal axis reports the number of tasks assigned to an industry; the vertical axis shows the share of industries in each bin. For reference, we overlay a normal density curve fitted to the sample mean and standard deviation.

Figure 2. AI Assistance and Experimentation Intensity Scores Distributions



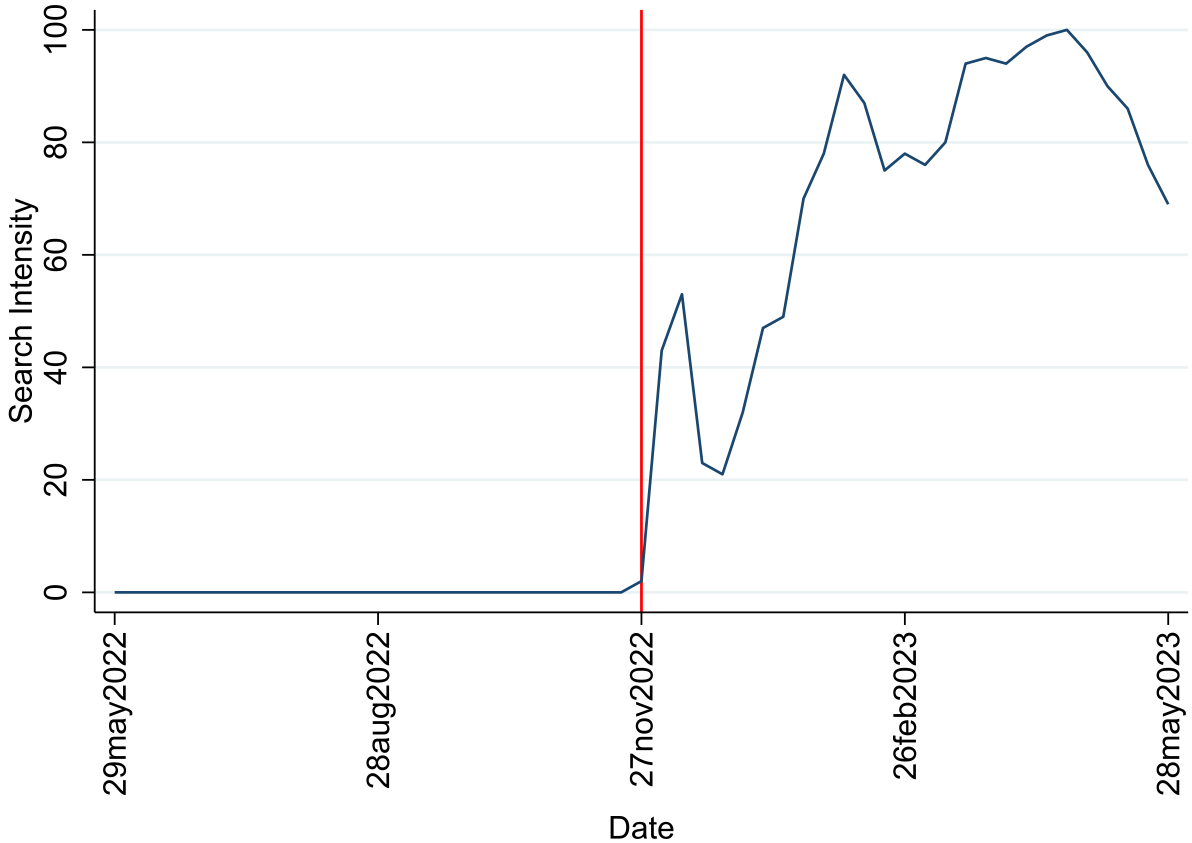
(a) AI Assistance



(b) Experimentation Intensity

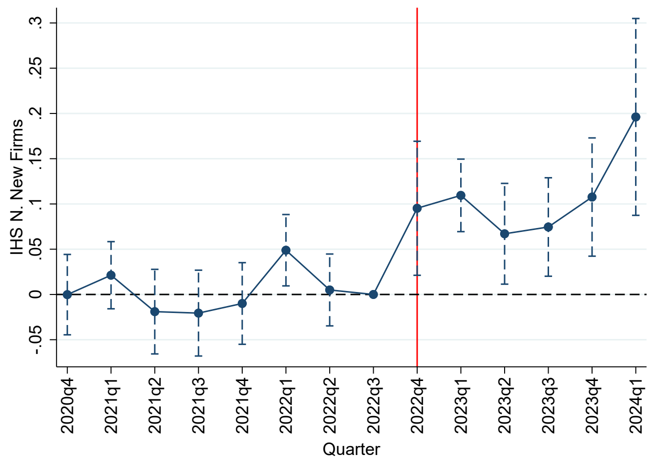
Notes: This panel of plots displays the distributions of the two measures generated through the GenAI workflow across all 81 three-digit NAICS industries in our sample. Tasks are extracted from the corpus of several texts focusing on entrepreneurship. After all the tasks have been extracted and deduplicated, an LLM assesses the relevance of any given task with respect to a particular industry. Once the universe of industry-specific tasks is created, each task is scored on Importance Score, AI Assistance and Experimentation Intensity. We then take the industry level average of the two measures weighting by the Importance Score. Panel (a) shows the distribution of the AI Assistance Score, while panel (b) shows the distribution of the Experimentation Intensity Measure. Both variables are presented at the industry level. For each of the variables, we also overlay a normal distribution curve for comparison.

Figure 3. Google Search Intensity for ChatGPT around Model Release

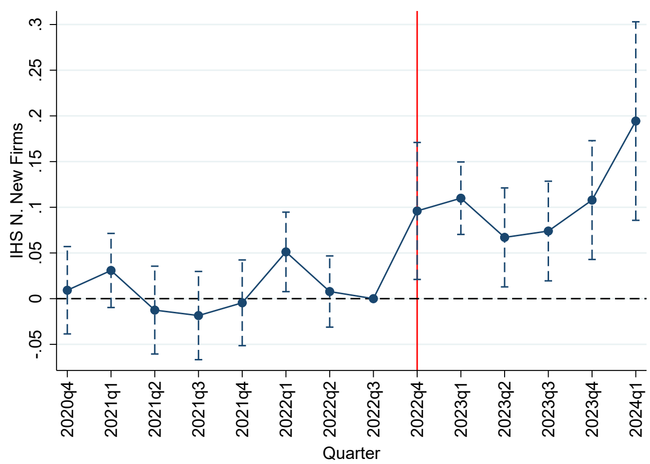


Notes: This figure plots weekly Google Trends search intensity for the term "ChatGPT,". Google Trends normalizes the number of searches within a specific period so to range on a 0–100 scale, where 100 represents the peak search volume during the sample period. The vertical red line marks the week in which ChatGPT was released. ChatGPT launched on November 30, 2022, which falls within the week beginning November 27, 2022 (as shown on the horizontal axis).

Figure 4. Event Study of AI Assistance on Firm Creation Around the Release of ChatGPT



(a) No Controls



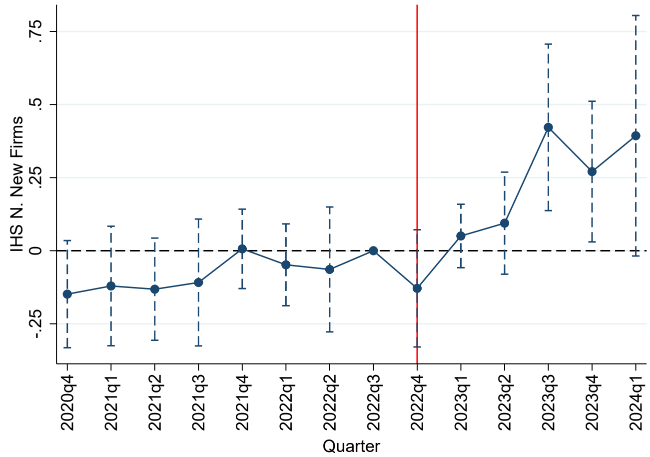
(b) Yes Controls

Notes: The figure presents the event study for the sample of 81 three-digit NAICS industries in our analysis, over the time period 2020Q4–2024Q1. We plot the estimates from the dynamic difference-in-differences model for the interaction of the quarter-year dummies with the AI Assistance$_i$ from the specification

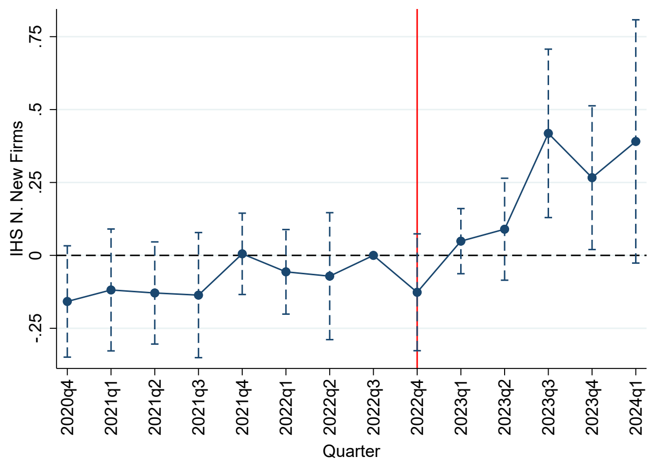
$$\text{IHS(N. New Firms)}_{it} = \beta_0 + \sum_{t \neq 2022q3} \beta_t \text{AI Assistance}_i \times 1\{D = t\}_t + \text{Controls}_{it} + \mu_i + \mu_t + \epsilon_{it}$$

The dependent variable is the inverse hyperbolic sine (IHS) of the number of new firms created. Treatment intensity is proxied by AI Assistance in its standardized form. The specifications include both industry and quarter fixed effects. Standard errors are clustered at the three-digit NAICS industry level. The specification in Panel (a) does not include controls. The specification in Panel (b) includes the following controls: lag of VCPE Investments (IHS), median Return on Assets, median R&D Intensity, median CapEx Intensity, median Liquidity Ratio, and median Market Leverage.

Figure 5. Event Study of Cost of Experimentation on Firm Creation Around the Release of ChatGPT



(a) No Controls



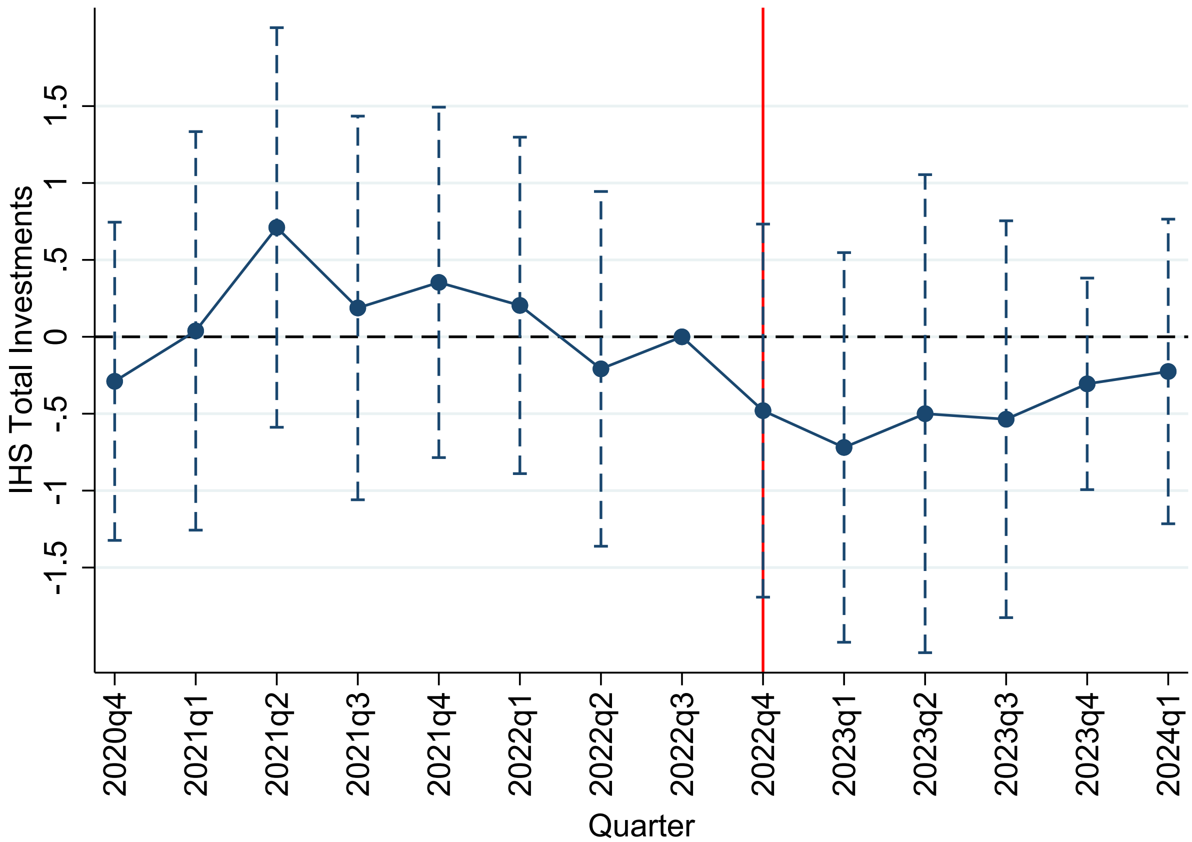
(b) Yes Controls

Notes: The figure presents the event study for the sample of 81 three-digit NAICS industries in our analysis, over the time period 2020Q4–2024Q1. We plot the estimates from the dynamic triple differences model for the interaction of the quarter-year dummies with the AI Assistance$_i$ $\times$ Experimentation Intensity$_i$ from the specification

$$\begin{aligned}\text{IHS(N. New Firms)}_{it} = \delta_0 + &\sum_{t\neq 2022q3} \delta_t^1 \text{AI Assistance}_i \times \text{Experimentation Intense}_i \times 1\{D=t\}_t \\ &+ \sum_{t\neq 2022q3} \delta_t^2 \text{AI Assistance}_i \times 1\{D=t\}_t + \sum_{t\neq 2022q3} \delta_t^3 \text{Experimentation Intense}_i \times 1\{D=t\}_t \\ &+ \text{Controls}_{it} + \mu_i + \mu_{It} + \varepsilon_{it}\end{aligned}$$

The dependent variable is the Inverse Hyperbolic Sine of the number of new firms created. Treatment intensity is proxied by AI Assistance in its standardized form. Treatment heterogeneity comes from the indicator variable Experimentation Intense, which equal 1 if the industry reports a value of Experimentation Intensity within the top quartile and 0 otherwise. The specifications include both industry and sector-by-quarter fixed effects. Standard errors are clustered at the three-digit NAICS industry level. The specification in Panel (a) does not include controls. The specification in Panel (b) includes the following controls: lag of VCPE Investments (IHS), median Return on Assets, median R&D Intensity, median CapEx Intensity, median Liquidity Ratio, and median Market Leverage.

Figure 6. Event Study of AI Assistance on Allocated Capital Around the Release of ChatGPT



Notes: The figure presents the event study for the sample of 81 three-digit NAICS industries in our analysis, over the time period 2020Q4–2024Q1. We plot the estimates from the dynamic difference-in-differences model for the interaction of the quarter-year dummies with the AI Assistance$_i$ from the specification

$$\text{IHS(Investments)}_{it} = \psi_0 + \sum_{t \neq 2022q3} \psi_t \text{AI Assistance}_i \times 1\{D = t\}_t + \text{Controls}_{it} + \mu_i + \mu_{It} + \epsilon_{it}$$

The dependent variable is the Inverse Hyperbolic Sine of the dollar amount of VCPE investments, in thousands. Treatment intensity is proxied by AI Assistance in its standardized form. The specification includes both industry and sector-by-quarter fixed effects. Standard errors are clustered at the three-digit NAICS industry level.

Table 1. Top and Bottom 20 Industries by AI Assistance Score

Spot	Industry Name	NAICS Code	Value
1	Web Search Portals, Libraries, Archives, and Other Information Services	519	47.226
2	Computing Infrastructure Providers, Data Processing, Web Hosting, and Related Services	518	47.021
3	Professional, Scientific, and Technical Services	541	46.140
4	Computer and Electronic Product Manufacturing	334	45.608
5	Broadcasting and Content Providers	516	45.234
6	Furniture, Home Furnishings, Electronics, and Appliance Retailers	449	44.940
7	Telecommunications	517	44.756
8	Administrative and Support Services	561	44.342
9	Merchant Wholesalers, Nondurable Goods	424	44.236
10	Sporting Goods, Hobby, Musical Instrument, Book, and Miscellaneous Retailers	459	44.197
11	Clothing, Clothing Accessories, Shoe, and Jewelry Retailers	458	44.132
12	Merchant Wholesalers, Durable Goods	423	44.070
13	Management of Companies and Enterprises	551	43.771
14	Publishing Industries	513	43.733
15	Health and Personal Care Retailers	456	43.432
16	Educational Services	611	43.248
17	Miscellaneous Manufacturing	339	43.160
18	General Merchandise Retailers	455	43.153
19	Rental and Leasing Services	532	43.113
20	Wholesale Trade Agents and Brokers	425	43.021
⋮	⋮	⋮	⋮
62	Support Activities for Agriculture and Forestry	115	40.001
63	Textile Mills	313	39.914
64	Gasoline Stations and Fuel Dealers	457	39.652
65	Support Activities for Mining	213	39.598
66	Private Households	814	39.546
67	Construction of Buildings	236	39.532
68	Animal Production and Aquaculture	112	39.461
69	Mining (except Oil and Gas)	212	39.406
70	Nursing and Residential Care Facilities	623	39.395
71	Rail Transportation	482	39.301
72	Hospitals	622	39.134
73	Water Transportation	483	38.970
74	Petroleum and Coal Products Manufacturing	324	38.784
75	Pipeline Transportation	486	38.354
76	Heavy and Civil Engineering Construction	237	38.143
77	Fishing, Hunting and Trapping	114	37.928
78	Forestry and Logging	113	37.907
79	Oil and Gas Extraction	211	37.815
80	Crop Production	111	37.642
81	Primary Metal Manufacturing	331	37.417

Notes: This table lists the 20 industries with the highest and the 20 industries the lowest AI Assistance Scores among 81 three-digit NAICS industries. Tasks are extracted from a corpus of entrepreneurship-focused texts and, after extraction and deduplication, a large language model (LLM) assesses each task's relevance to each industry and tailors the description accordingly. For each industry, the AI Assistance Score is the Importance Score-weighted average of task-level AI Assistance scores; higher values indicate greater potential for AI to assist industry tasks. A detailed explanation of the workflow to generate the variables is presented in the Subsection D.2 of the Online Appendix.

Table 2. Top and Bottom 20 Industries by Experimentation Intensity

Spot	Industry Name	NAICS Code	Value
1	Motion Picture and Sound Recording Industries	512	66.626
2	Broadcasting and Content Providers	516	66.194
3	Computing Infrastructure Providers, Data Processing, Web Hosting, and Related Services	518	66.165
4	Food Services and Drinking Places	722	65.983
5	Web Search Portals, Libraries, Archives, and Other Information Services	519	65.930
6	Clothing, Clothing Accessories, Shoe, and Jewelry Retailers	458	65.870
7	Sporting Goods, Hobby, Musical Instrument, Book, and Miscellaneous Retailers	459	65.820
8	Food and Beverage Retailers	445	65.700
9	Computer and Electronic Product Manufacturing	334	65.603
10	Professional, Scientific, and Technical Services	541	65.583
11	Educational Services	611	65.572
12	Accommodation	721	65.546
13	General Merchandise Retailers	455	65.544
14	Furniture, Home Furnishings, Electronics, and Appliance Retailers	449	65.498
15	Miscellaneous Manufacturing	339	65.497
16	Telecommunications	517	65.476
17	Publishing Industries	513	65.404
18	Apparel Manufacturing	315	65.263
19	Health and Personal Care Retailers	456	65.245
20	Rental and Leasing Services	532	64.932
⋮	⋮	⋮	⋮
62	Motor Vehicle and Parts Dealers	441	62.860
63	Lessors of Nonfinancial Intangible Assets (except Copyrighted Works)	533	62.839
64	Warehousing and Storage	493	62.804
65	Mining (except Oil and Gas)	212	62.791
66	Water Transportation	483	62.655
67	Rail Transportation	482	62.344
68	Construction of Buildings	236	62.241
69	Hospitals	622	61.873
70	Wholesale Trade Agents and Brokers	425	61.686
71	Chemical Manufacturing	325	61.605
72	Heavy and Civil Engineering Construction	237	60.863
73	Utilities	221	60.768
74	Paper Manufacturing	322	60.708
75	Primary Metal Manufacturing	331	60.696
76	Forestry and Logging	113	60.630
77	Nursing and Residential Care Facilities	623	60.216
78	Gasoline Stations and Fuel Dealers	457	59.166
79	Petroleum and Coal Products Manufacturing	324	59.143
80	Oil and Gas Extraction	211	58.639
81	Pipeline Transportation	486	57.516

Notes: This table lists the 20 industries with the highest and the 20 industries with the lowest Experimentation Intensity Scores among 81 three-digit NAICS industries. Tasks are extracted from a corpus of entrepreneurship-focused texts and, after extraction and deduplication, a large language model (LLM) assesses each task's relevance to each industry and tailors the description accordingly. For each industry, the Experimentation Intensity Score is the Importance Score-weighted average of task-level Experimentation Intensity Scores; higher values indicate greater industry reliance on trial-and-error, feedback-driven refinement and iterative learning. A detailed explanation of the workflow to generate the variables is presented in the Subsection D.2 of the Online Appendix.

Table 3. Summary Statistics

	(1) N. Obs	(2) Min	(3) p50	(4) Mean	(5) Max	(6) SD
Panel (a): Industry × Quarter						
N. New Firms	1,134	26.000	5,401.000	14,064.000	322,847.000	32,689.507
Investments ($ 000)	1,134	0.000	97.000	1,730.000	94,564.000	7,090.847
ROA	1,134	-5.844	0.005	-0.003	0.316	0.177
R&D Intensity	1,134	0.000	0.018	0.077	11.235	0.397
Capex Intensity	1,134	-36.034	0.013	0.003	17.816	1.346
Liquidity Ratio	1,134	0.004	0.096	0.127	0.726	0.103
Market Leverage Ratio	1,134	0.000	0.281	0.408	6.545	0.465
Panel (b): Industry						
AI Assistance Score	81	37.417	41.365	41.587	47.226	2.172
Experimentation Intensity	81	57.516	63.810	63.652	66.626	1.859

Notes: This table reports summary statistics for the variables used in the analysis over 2020Q4–2024Q1. Columns display the following: (1) number of observations, (2) minimum, (3) median, (4) mean, (5) maximum, and (6) standard deviation. Panel (a) presents summary statistics for variables varying at the industry-by-quarter level, while Panel (b) presents summary statistics for variables varying at the industry level. N. New Firms is the count of new firms created. Investments is the total amount (in thousands of dollars) of VCPE investments from Preqin. The following Compustat variables are industry medians. Return on Assets (ROA) is defined as income before extraordinary items divided by total assets. R&D Intensity is defined as R&D expenses over revenue. CapEx Intensity is defined as capital expenditures over revenue. Liquidity Ratio is cash holdings over total assets. Market Leverage Ratio is calculated as the sum of long-term and current debt divided by the market value of equity. AI Assistance Score is the industry-level average proxying for how much AI can help with task completion. Experimentation Intensity is the industry-level average proxying for how prominent experimentation is.

Table 4. Effect of AI Assistance on Firm Creation After the Release of ChatGPT

VARIABLES	(1) IHS N. New Firms	(2) IHS N. New Firms
AI Assistance × Post	0.105***	0.100***
	(0.025)	(0.025)
Observations	1,134	1,134
Industry 3D FE	Yes	Yes
Quarter FE	Yes	Yes
Industry 2D × Quarter FE	No	No
Controls	No	Yes
Adj. R. sq.	0.982	0.982

Notes: This table presents results testing the effect of AI's arrival on firm creation. The design is a difference-in-differences model at the industry-by-quarter level with continuous treatment. The sample includes 81 three-digit NAICS industries over 2020Q4–2024Q1. The empirical specification is

$$\text{IHS(N. New Firms)}_{it} = \alpha_0 + \alpha_1 \text{AI Assistance}_i \times \text{Post}_t + \text{Controls}_{it} + \mu_i + \mu_t + \epsilon_{it}$$

The dependent variable is the inverse hyperbolic sine (IHS) of the number of new firms created. Treatment intensity is proxied by AI Assistance in its standardized form. Treatment timing is proxied by the binary indicator Post, which equals 1 for all quarters after 2022Q3 and 0 otherwise. All specifications include industry and quarter fixed effects. Column (2) additionally includes the following controls: lag of VCPE Investments (IHS), median Return on Assets, median R&D Intensity, median CapEx Intensity, median Liquidity Ratio, and median Market Leverage. Standard errors are clustered at the three-digit NAICS industry level and reported in parentheses. Coefficients are statistically significant at the 10% (*), 5% (**), and 1% (***) levels.

Table 5. Horse Race Analysis of Competing AI Metrics

VARIABLES	(1) IHS N. New Firms	(2) IHS N. New Firms	(3) IHS N. New Firms
AI Assistance × Post	0.105***	0.087***	0.110***
	(0.025)	(0.029)	(0.027)
AI Task Exposure × Post		0.036	
		(0.033)	
AI Demand × Post			-0.014
			(0.019)
Observations	1,134	1,134	1,134
Industry 3D FE	Yes	Yes	Yes
Quarter FE	Yes	Yes	Yes
Industry 2D × Quarter FE	No	No	No
Controls	No	No	No
Adj. R. sq.	0.982	0.982	0.982

Notes: This table presents results probing the interpretation of the AI Assistance measure. The design is a horse race difference-in-differences model at the industry-by-quarter. The sample includes 81 three-digit NAICS industries over 2020Q4–2024Q1. Each horse race regression uses the empirical specification

$$\text{IHS(N. New Firms)}_{it} = \alpha_0' + \alpha_1' \text{AI Assistance}_i \times \text{Post}_t + \alpha_2' \text{Alternative Measure}_i \times \text{Post}_t + \mu_i + \mu_t + \epsilon_{it}$$

The dependent variable is the inverse hyperbolic sine (IHS) of the number of new firms created. In Column (1) we present the baseline results from Equation 5. In Column (2) AI Assistance variable is juxtaposed to the AI Task Exposure measure, both in standardized form. In Column (3) the same process is repeated, using the AI Demand measure instead. All specifications include industry and quarter fixed effects. Standard errors are clustered at the three-digit NAICS industry level and reported in parentheses. Coefficients are statistically significant at the 10% (*), 5% (**), and 1% (***) levels.

Table 6. Effects of the Cost of Experimentation on Firm Creation

VARIABLES	(1) IHS N. New Firms	(2) IHS N. New Firms
AI Assistance × Experimentation Intense × Post	0.261**	0.264**
	(0.111)	(0.112)
AI Assistance × Post	0.037	0.032
	(0.054)	(0.054)
Experimentation Intense × Post	-0.190	-0.193
	(0.119)	(0.118)
Observations	1,078	1,078
Industry 3D FE	Yes	Yes
Quarter FE	No	No
Industry 2D × Quarter FE	Yes	Yes
Controls	No	Yes
Adj. R sq.	0.981	0.981

Notes: This table presents results testing whether the reduced cost-of-experimentation mechanism operates. The design is a triple differences model at the industry-by-quarter level. The sample includes 81 three-digit NAICS industries over 2020Q4–2024Q1. The empirical specification is

$$\begin{aligned}\text{IHS(N. New Firms)}_{it} = \gamma_0 &+ \gamma_1 \text{AI Assistance}_i \times \text{Experimentation Intense}_i \times \text{Post}_t \\ &+ \gamma_2 \text{AI Assistance}_i \times \text{Post}_t + \gamma_3 \text{Experimentation Intense}_i \times \text{Post}_t \\ &+ \text{Controls}_{it} + \mu_i + \mu_{It} + \varepsilon_{it}\end{aligned}$$

The dependent variable is the inverse hyperbolic sine (IHS) of the number of new firms created. Treatment intensity is proxied by AI Assistance in its standardized form. Treatment timing is proxied by Post, which equals 1 for all quarters after 2022Q3 and 0 otherwise. Treatment heterogeneity comes from the indicator Experimentation Intense, which equals 1 if the industry's Experimentation Intensity Score is in the top quartile and 0 otherwise. The specifications include industry and sector-by-quarter fixed effects. Column (2) additionally includes the following controls: lagged VCPE Investments (IHS), median Return on Assets, median R&D Intensity, median CapEx Intensity, median Liquidity Ratio, and median Market Leverage. Standard errors are clustered at the three-digit NAICS industry level and reported in parentheses. Coefficients are statistically significant at the 10% (*), 5% (**), and 1% (***) levels.

Table 7. Mediation Analysis: Capital Allocation, AI Assistance and Firm Creation

VARIABLES	(1) IHS N. New Firms	(2) IHS Investments	(3) IHS N. New Firms
AI Assistance × Post	0.102**	-0.586**	0.102**
	(0.046)	(0.239)	(0.046)
IHS Investments			0.000
			(0.004)
Observations	1,078	1,078	1,078
Industry 3D FE	Yes	Yes	Yes
Quarter FE	No	No	No
Industry 2D × Quarter FE	Yes	Yes	Yes
Controls	No	No	No
Adj. R sq.	0.980	0.687	0.980

Notes: This table reports a mediation-style decomposition to study capital reallocation as a potential mechanism linking AI's arrival and firm creation. The analysis is at the industry-by-quarter level for 81 three-digit NAICS industries over 2020Q4–2024Q1. The mediator is total capital allocated, measured as the inverse hyperbolic sine (IHS) of total VCPE investments from Preqin (in $000).

Column (1) reports the results for the estimation of the total impact that the treatment has on firm formation. The specification is

$$\text{IHS(N. New Firms)}_{it} = \zeta_0 + \zeta_1 \text{AI Assistance}_i \times \text{Post}_t + \mu_i + \mu_{It} + \epsilon_{it}$$

Column (2) assesses the linkage between the treatment and the total capital allocation, according to the specification

$$\text{IHS(Investments)}_{it} = \eta_0 + \eta_1 \text{AI Assistance}_i \times \text{Post}_t + \mu_i + \mu_{It} + \epsilon_{it}$$

Column (3) reports the results for the estimation of the mediated effect of the treatment once we control for total allocated capital. The specification is

$$\text{IHS(N. New Firms)}_{it} = \theta_0 + \theta_1 \text{AI Assistance}_i \times \text{Post}_t + \theta_2 \text{IHS(Investments)}_{it} + \mu_i + \mu_{It} + \epsilon_{it}$$

All specifications include both industry and sector-by-quarter fixed effects. Treatment intensity is proxied by AI Assistance in its standardized form. Treatment timing is proxied by the binary indicator Post, which is equal to 1 for all quarters after 2022Q3, and 0 otherwise. Standard errors are clustered at the three-digit NAICS industry level and reported in parentheses. Coefficients are statistically significant at the 10% (*), 5% (**), and 1% (***) levels.

Table 8. AI Availability, Missing Skills and Entrepreneurial Entry - All Individuals

VARIABLES	(1) P(Entrepreneur)	(2) P(Entrepreneur)
Skill Gap × Post		0.715
		(0.907)
Skill Gap	-2.265***	-2.427***
	(1.042)	(1.054)
Observations	173,968,981	173,968,981
Sample	All Individuals	All Individuals
User × Quarter FE	Yes	Yes
User × Industry 3D FE	Yes	Yes
Quarter × Industry 3D FE	Yes	Yes
Adj. R sq.	0.119	0.119

Notes: The table reports the results for the test assessing whether AI promotes self-employment by bridging human capital gaps for potential entrepreneurs. The analysis is at the individual-by-industry-by-quarter level. The sample is composed on 200,000 individuals randomly sampled from the profiles on Revelio Labs. Industries are reported at the three-digit NAICS level. The dataset includes 81 industries over the period 2020Q4–2024Q1. The model used is

$$\text{P(Entrepreneur)}_{jit} = \iota_0 + \iota_1 \text{Skill Gap}_{jit} \times \text{Post}_t + \iota_2 \text{Skill Gap}_{jit} + \mu_{ji} + \mu_{jt} + \mu_{it} + \epsilon_{jit}$$

The dependent variable $\text{P(Entrepreneur)}_{jit}$ is an indicator reporting 10000 if the individual becomes an entrepreneur in a specific industry and quarter and 0 otherwise. Skill Gap is the measure of human capital gap calculated for each individual, in a given quarter, with respect a particular industry. We standardize Skill Gap using the pre-shock mean and standard deviation. The binary indicator Post, which is equal to 1 for all quarters after 2022Q3, and 0 otherwise, proxies for the availability of ChatGPT. The specifications include individual-by-industry, individual-by-quarter and industry-by-quarter fixed effects. Column (1) reports results assessing the predictive power for the measure of skill mismatch, while Column (2) present the results for the mechanism test. Standard errors are clustered at the individual level and reported in parentheses. Coefficients are statistically significant at the 10% (*), 5% (**), and 1% (***) levels.

Table 9. AI Availability, Missing Skills and Entrepreneurial Entry - Entrepreneurs Only

VARIABLES	(1) P(Entrepreneur)	(2) P(Entrepreneur)
Skill Gap × Post		1.939
		(3.081)
Skill Gap	-24.051***	-24.784***
	(6.214)	(6.288)
Observations	67,289,988	67,289,988
Sample	All Entrepreneurs	All Entrepreneurs
User × Quarter FE	Yes	Yes
User × Industry 3D FE	Yes	Yes
Quarter × Industry 3D FE	Yes	Yes
Adj. R sq.	0.096	0.096

Notes: The table reports the results for the test assessing whether AI determines the industry choice of entrepreneurs by bridging human capital gaps. The sample includes all invividuals who eventually became entrepreneurs between 2020Q4 and 2024Q1. The analysis is at the individual-by-industry-by-quarter level. Industries are reported at the three-digit NAICS level. The dataset includes 81 industries over the period 2020Q4–2024Q1. The model used is

$$\text{P(Entrepreneur)}_{jit} = \iota_0^1 + \iota_1^1 \text{Skill Gap}_{jit} \times \text{Post}_t + \iota_2^1 \text{Skill Gap}_{jit} + \mu_{ji} + \mu_{jt} + \mu_{it} + \epsilon_{jit}$$

The dependent variable $\text{P(Entrepreneur)}_{jit}$ is an indicator reporting 10000 if the individual becomes an entrepreneur in a specific industry and quarter and 0 otherwise. Skill Gap is the measure of human capital gap calculated for each individual, in a given quarter, with respect a particular industry. We standardize Skill Gap using the pre-shock mean and standard deviation. The binary indicator Post, which is equal to 1 for all quarters after 2022Q3, and 0 otherwise, proxies for the availability of ChatGPT. The specifications include individual-by-industry, individual-by-quarter and industry-by-quarter fixed effects. Column (1) reports results assessing the predictive power for the measure of skill mismatch, while Column (2) present the results for the mechanism test. Standard errors are clustered at the individual level and reported in parentheses. Coefficients are statistically significant at the 10% (*), 5% (**), and 1% (***) levels.

Table 10. Effect of AI Compatibility on New Firm Survival After ChatGPT

VARIABLES	(1) $\text{P(Survived)}^{(4)}$	(2) $\text{P(Survived)}^{(6)}$	(3) $\text{P(Survived)}^{(8)}$
AI Assistance × Post	0.003***	0.004***	0.003***
	(0.000)	(0.000)	(0.001)
Observations	1,021,651	945,087	826,018
Industry 3D	Yes	Yes	Yes
Registration Quarter FE	Yes	Yes	Yes
State FE	Yes	Yes	Yes
Controls	No	No	No
Adj. R sq.	0.014	0.018	0.018

Notes: The table reports the results for the test assessing whether firms established post the arrival of ChatGPT have better survival odds at different horizons (4, 6 and 8 quarters after registration). The analysis is at the firm level and we include all new firms created between 2020Q4 and 2024Q1 in any of the 81 three-digit NAICS industries of interest. The cross-sectional model specification is

$$\text{P(Survived)}_f^{(\tau)} = \kappa_0 + \kappa_1 \text{AI Assistance}_{i(f)} \times \text{Post}_{t(f)} + \mu_{i(f)} + \mu_{t(f)} + \mu_{s(f)} + \varepsilon_f$$

The dependent variable $\text{P(Survived)}_f^{(\tau)}$ is an indicator equal 1 if the firm is active τ quarters after establishment. The outcome variable is measure 4, 6 and 8 quarters from establishment. AI Assistance is measured at the three-digit NAICS industry level. The binary indicator Post, is equal to 1 for all firms created after 2022Q3, and 0 otherwise. The specification includes industry, quarter of registration and state fixed effects. Robust standard error are reported in parentheses. Coefficients are statistically significant at the 10% (*), 5% (**), and 1% (***) levels.

Table 11. Effect of AI Compatibility on New Firms' Employment Growth After ChatGPT

	(1)	(2)	(3)
VARIABLES	Emp. Growth$^{(4)}$	Emp. Growth$^{(6)}$	Emp. Growth$^{(8)}$
AI Assistance × Post	0.007***	0.009***	0.006**
	(0.001)	(0.002)	(0.003)
Observations	1,021,651	945,087	826,018
Industry 3D	Yes	Yes	Yes
Registration Quarter FE	Yes	Yes	Yes
State FE	Yes	Yes	Yes
Controls	No	No	No
Adj. R sq.	0.013	0.014	0.012

Notes: The table reports the results for the test assessing whether firms established post the arrival of ChatGPT grow faster at different horizons (4, 6 and 8 quarters after registration). The analysis is at the firm level and we include all new firms created between 2020Q4 and 2024Q1 in any of the 81 three-digit NAICS industries of interest. The cross-sectional model specification is

$$\text{Emp. Growth}_f^{(\tau)} = \kappa_0 + \kappa_1 \text{AI Assistance}_{i(f)} \times \text{Post}_{t(f)} + \mu_{i(f)} + \mu_{t(f)} + \mu_{s(f)} + \varepsilon_f$$

The dependent variable Emp. Growth$_f^{(\tau)}$ is the level change in employment between the quarter of establishment and τ quarters after establishment over the employment level in the quarter of establishment. The outcome variable is measure 4, 6 and 8 quarters from establishment. Employment growth for firms not surviving at a given horizon is coded as -1 (100%). AI Assistance is measured at the three-digit NAICS industry level. The binary indicator Post, is equal to 1 for all firms created after 2022Q3, and 0 otherwise. The specification includes industry, quarter of registration and state fixed effects. Robust standard error are reported in parentheses. Coefficients are statistically significant at the 10% (*), 5% (**), and 1% (***) levels.

Table 12. Effect of AI Compatibility on New Firms' Average Education Level After ChatGPT

VARIABLES	(1) Avg. Education$^{(4)}$	(2) Avg. Education$^{(6)}$	(3) Avg. Education$^{(8)}$
AI Assistance × Post	0.796***	0.812***	0.719**
	(0.205)	(0.263)	(0.219)
Observations	556,128	488,318	413,562
Industry 3D	Yes	Yes	Yes
Registration Quarter FE	Yes	Yes	Yes
State FE	Yes	Yes	Yes
Controls	No	No	No
Adj. R sq.	0.067	0.069	0.072

Notes: The table reports the results for the test assessing whether firms established post the arrival of ChatGPT are able to attract more qualified individuals at different horizons (4, 6 and 8 quarters after registration). The analysis is at the firm level and we include all new firms created between 2020Q4 and 2024Q1 in any of the 81 three-digit NAICS industries of interest. The cross-sectional model specification is

$$\text{Avg. Education}_f^{(\tau)} = \kappa_0 + \kappa_1 \text{AI Assistance}_{i(f)} \times \text{Post}_{t(f)} + \mu_{i(f)} + \mu_{t(f)} + \mu_{s(f)} + \varepsilon_f$$

The dependent variable Avg. Education$_f^{(\tau)}$ is the average number of years in education attained by workers active at the firm τ quarters after establishment. The outcome variable is measure 4, 6 and 8 quarters from establishment. AI Assistance is measured at the three-digit NAICS industry level. The binary indicator Post, is equal to 1 for all firms created after 2022Q3, and 0 otherwise. The specification includes industry, quarter of registration and state fixed effects. Robust standard error are reported in parentheses. Coefficients are statistically significant at the 10% (*), 5% (**), and 1% (***) levels.

Online Appendix for:

When AI Meets Entrepreneurship: Evidence from the Commercialization of ChatGPT

Table of Contents

Appendix A. Additional Results: Summary Statistics and Analysis

A.1. **The Effects of AI on Firm Creation.** The credibility of our main findings hinges on the validity of the identifying assumptions and on the robustness of our results to alternative specifications and interpretations. We therefore conduct a series of complementary analyses.

First, we verify that the AI Assistance measure is not simply capturing pre-existing industry characteristics or secular growth patterns by examining its relationship with a wide range of pre-period fundamentals.

Next, we assess the robustness of our estimates to alternative fixed-effects structures, treatment definitions, and timing falsification exercises, and we explore whether the effects display any meaningful nonlinearities or heterogeneity across industries.

We then show that our results are not driven by industries that experienced abnormal entry surges during the COVID-19 period or secular long-term growth in value added, which could otherwise confound the interpretation of post-ChatGPT dynamics.

Finally, we demonstrate that the AI Assistance measure captures an entrepreneur-specific dimension of AI exposure, distinct from the types of tasks prevalent among workers in incumbent firms.

A.1.1. *AI Assistance and Industry Characteristics.* A potential concern is that our measure of AI Assistance may simply capture pre-existing industry fundamentals—such as profitability, R&D intensity, or investment activity—rather than a distinct dimension of entrepreneurial relevance. If that were the case, the observed increase in firm creation in high–AI Assistance industries could reflect differences in baseline industry strength rather than the extent to which AI tools facilitate entrepreneurial tasks.

To address this concern, we examine how AI Assistance correlates with standard industry characteristics in the pre-ChatGPT period. Table OA1 reports results using three alternative definitions of the pre-shock window: Column (1) uses only 2022Q3 values, Column (2) averages over the last four quarters (2021Q4–2022Q3), and Column (3) averages over the last eight quarters (2020Q4–2022Q3). Across all definitions, the correlations remain modest, ranging from about -0.18 to 0.35, and are highly stable across panels. The strongest positive associations are with the number of investments and total investment, as well as with liquidity ratios. In contrast, AI Assistance is essentially uncorrelated with profitability, R&D intensity, and capex intensity, and is negatively but weakly related to market leverage.

Table OA1. Correlation of AI Assistance Score with Industry Characteristics

	(1) AI Assistance 2022Q3	(2) AI Assistance 2022Q4–2023Q3 Average	(3) AI Assistance 2021Q4–2022Q3 Average
Number of Investments	0.310	0.312	0.313
Total Investments ($000)	0.334	0.319	0.339
Average Investment ($000)	-0.025	0.021	0.128
ROA	0.028	0.025	0.098
R&D Intensity	0.024	-0.098	-0.093
Capex Intensity	0.073	0.092	0.075
Liquidity Ratio	0.260	0.317	0.357
Market Leverage Ratio	-0.138	-0.143	-0.178

Notes: This table presents the correlation between the AI Assistance measure and various industry characteristics, calculated on different time windows. The industry characteristics are the number of early stage equity investments, the total invested amount (in thousands of dollars), the average invested amount (in thousands of dollars), the Return on Assets (ROA, income before extraordinary items divided by total assets), R&D Intensity (R&D expenses over revenue), CapEx Intensity (capital expenditures over revenue), Liquidity Ratio (cash holdings over total assets), and Market Leverage Ratio (the sum of long-term and current debt divided by the market value of equity). Column (1) establishes industry characteristics using 2022Q3 values. Column (2) established the characteristics using the average industry value over the period 2022Q4 to 2022Q3. Column (3) established the characteristics using the average industry value over the period 2021Q4 to 2022Q3.

As an additional placebo-outcome check, we test whether the increase in firm creation in high-AI-Assistance industries coincides with differential changes in other industry fundamentals after ChatGPT's release. To do so, we regress pre-ChatGPT industry characteristics—ROA, R&D intensity, capex intensity, liquidity ratios, leverage ratios, and investment variables—on the interaction between AI Assistance and Post, controlling for industry and time fixed effects. This specification examines whether high-AI industries experienced disproportionate shifts in these outcomes in the post-ChatGPT period.

As seen in Table OA2, the estimated coefficients are small and statistically insignificant across all outcomes except for the liquidity ratio, where we find a negative but economically modest association (-0.012). This coefficient is an order of magnitude smaller than the effect of AI Assistance$\times$Post on new firm creation (0.105 in Table 4), implying that any relationship between AI Assistance and liquidity is negligible in economic terms. Moreover, the negative sign suggests that industries more compatible with AI exhibited slightly lower liquidity ratios, contrary to what one would expect if AI Assistance merely captured financially stronger sectors.

Table OA2. Effect of AI Assistance on Other Outcomes After the Release of ChatGPT

VARIABLES	(1) ROA	(2) R&D Intensity	(3) Capex Intensity	(4) Liquidity Ratio	(5) Market Leverage Ratio
AI Assistance × Post	-0.011	0.037	-0.015	-0.012***	0.003
	(0.008)	(0.032)	(0.022)	(0.003)	(0.019)
Observations	1,134	1,134	1,134	1,134	1,134
Industry 3D FE	Yes	Yes	Yes	Yes	Yes
Quarter FE	Yes	Yes	Yes	Yes	Yes
Industry 2D × Quarter FE	No	No	No	No	No
Controls	No	No	No	No	No
Adj. R. sq.	0.003	0.253	-0.077	0.919	0.800

Notes: This table presents the results for the test assessing whether the treatment used in the main analysis affects other industry level outcomes that are not entrepreneurship specific. The design is a difference-in-differences model at the industry-by-quarter level with continuous treatment. The sample includes 81 three-digit NAICS industries over 2020Q4–2024Q1. The empirical specification is

$$\mathrm{Y}_{it} = a_0 + a_1 \text{AI Assistance}_i \times \text{Post}_t + \text{Controls}_{it} + \mu_i + \mu_t + \epsilon_{it}$$

The dependent variable is either the Return on Assets (ROA, income before extraordinary items divided by total assets) in Column (1), R&D Intensity (R&D expenses over revenue) in Column (2), CapEx Intensity (capital expenditures over revenue) in Column (3), Liquidity Ratio (cash holdings over total assets) in Column (4), and Market Leverage Ratio (the sum of long-term and current debt divided by the market value of equity) in Column (5). Treatment intensity is proxied by AI Assistance in its standardized form. Treatment timing is proxied by the binary indicator Post, which equals 1 for all quarters after 2022Q3 and 0 otherwise. All specifications include industry and quarter fixed effects. Standard errors are clustered at the three-digit NAICS industry level and reported in parentheses. Coefficients are statistically significant at the 10% (*), 5% (**), and 1% (***) levels.

Overall, these results indicate that AI Assistance is not simply a proxy for standard industry fundamentals such as size, profitability, or investment intensity, but rather captures a distinct dimension of the entrepreneurial process—namely, the extent to which AI can provide support for early-stage entrepreneurial tasks.

A.1.2. *Richer Fixed Effects Structure.* A primary identification concern in our analysis is the presence of unobserved shocks at the industry-by-quarter level, which could simultaneously influence firm creation and correlate with our AI Assistance Score—omitted variable bias. Ideally, we would include Industry-by-Quarter fixed effects to completely absorb any shocks specific to individual industries each quarter. However, this approach would exhaust degrees of freedom and prevent identification of the treatment effect. Therefore, we adopt sector-by-quarter fixed effects as a robustness test, where Sectors are defined as industry groups at the 2-digits NAICS level. While not fully achieving the ideal scenario, these fixed effects significantly mitigate concerns by removing broader sector-level shocks each quarter, ensuring that the variation we exploit for identification arises predominantly from differences across finer-grained (3-digit) industries within the same sector and quarter. We present the results in Table OA3 and Figure OA1.

The robustness of our estimated coefficient, remaining positive and significant, enhances confidence that the observed effects are driven by industry-specific treatment intensity rather than broader sectoral shocks (Table OA3).

Table OA3. Effect of AI Assistance on Firm Creation After the Release of ChatGPT Using Richer Fixed Effects Structure

VARIABLES	(1) IHS N. New Firms	(2) IHS N. New Firms
AI Assistance × Post	0.102**	0.096**
	(0.046)	(0.046)
Observations	1,078	1,078
Industry 3D FE	Yes	Yes
Quarter FE	No	No
Industry 2D × Quarter FE	Yes	Yes
Controls	No	Yes
Adj. R. sq.	0.980	0.980

Notes: This table presents results testing the effect of AI's arrival on firm creation using a richer fixed effects structure. The design is a difference-in-differences model at the industry-by-quarter level with continuous treatment. The sample includes 81 three-digit NAICS industries over 2020Q4–2024Q1. The empirical specification is

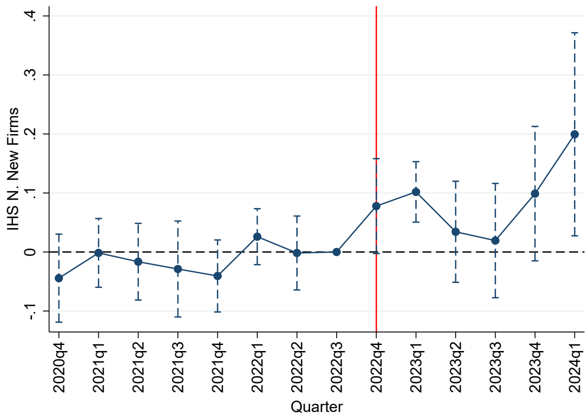
$$\text{IHS(N. New Firms)}_{it} = a_0 + a_1 \text{AI Assistance}_i \times \text{Post}_t + \text{Controls}_{it} + \mu_i + \mu_{It} + \epsilon_{it}$$

The dependent variable is the inverse hyperbolic sine (IHS) of the number of new firms created. Treatment intensity is proxied by AI Assistance in its standardized form. Treatment timing is proxied by the binary indicator Post, which equals 1 for all quarters after 2022Q3 and 0 otherwise. All specifications include industry and sector-by-quarter fixed effects. Column (2) additionally includes the following controls: lag of VCPE Investments (IHS), median Return on Assets, median R&D Intensity, median CapEx Intensity, median Liquidity Ratio, and median Market Leverage. Standard errors are clustered at the three-digit NAICS industry level and reported in parentheses. Coefficients are statistically significant at the 10% (*), 5% (**), and 1% (***) levels.

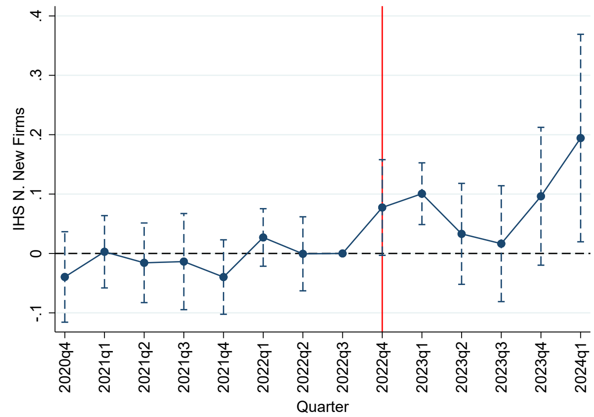
The event study analysis with the richer fixed effects structure (sector-by-quarter fixed effects) provides important insights (Figure OA1). Specifically, the coefficients on the lead terms—those representing periods prior to ChatGPT's release—are statistically insignificant, consistent with the validity of the parallel trends assumption. This indicates that industries with higher AI compatibility were not following differential pre-existing trends before the actual treatment event, strengthening the causal interpretation of our estimates.

However, coefficients on the lag terms—representing post-treatment periods—are generally positive but statistically insignificant. It is important to note that this statistical insignificance does not necessarily imply the absence of a meaningful economic effect. Instead, it arises primarily from the demanding structure imposed by the highly granular fixed effects. By including sector-by-quarter fixed effects, we remove substantial variation from the data at the broader sectoral level. Consequently, identification of treatment effects relies solely on variation across fine-grained (3-digit NAICS) industries within each 2-digit sector and each specific quarter. Given that some 2-digit sectors contain very few 3-digit industries—occasionally only two or three—the remaining variation available for estimating coefficients is severely limited. This sparsity of observations within each cell inflates standard errors and results in less stable, statistically noisier estimates.

Figure OA1. Event Study of AI Assistance on Firm Creation with Richer Fixed Effects Structure



(a) No Controls



(b) Yes Controls

Notes: The figure presents the event study for the sample of 81 three-digit NAICS industries in our analysis, over the time period 2020Q4–2024Q1. We plot the estimates from the dynamic difference-in-differences model for the interaction of the quarter-year dummies with the AI Assistance$_i$ from the specification

$$\text{IHS(N. New Firms)}_{it} = b_0 + \sum_{t \neq 2022q3} b_t \text{AI Assistance}_i \times 1\{D = t\}_t + \text{Controls}_{it} + \mu_i + \mu_{It} + \epsilon_{it}$$

The dependent variable is the inverse hyperbolic sine (IHS) of the number of new firms created. Treatment intensity is proxied by AI Assistance in its standardized form. The specifications include both industry and sector-by-quarter fixed effects. Standard errors are clustered at the three-digit NAICS industry level. The specification in Panel (a) does not include controls. The specification in Panel (b) includes the following controls: lag of VCPE Investments (IHS), median Return on Assets, median R&D Intensity, median CapEx Intensity, median Liquidity Ratio, and median Market Leverage.

Despite the statistical challenges entailed using a very rich fixed effects structure, the positive and significant coefficients in Table OA3, paired with the lack of pre-trends displayes in the plots of Figure OA1 provide reassuring qualitative support which align with our primary findings.

A.1.3. *Alternative Treatment Definitions.* The validity of our causal estimates relies critically on accurately measuring and defining treatment intensity, specifically industries' compatibility with generative AI technologies. Our main analysis in Section 4 employs a continuous measure (the AI Assistance Score), but this approach inherently risks being affected by measurement errors, sensitivity to extreme observations, and potential biases due to subtle inaccuracies or misclassification.

To address these potential concerns and to validate the robustness of our primary findings, we additionally test binary versions of our treatment definition—classifying industries simply as high or low compatibility based on whether their AI Assistance Scores exceed the mean or median. By using these binary definitions, we effectively reduce the sensitivity of our estimates to minor measurement errors, mitigate the influence of extreme values that might disproportionately affect our continuous measure, and clearly distinguish treated from control groups.

This simplified binary approach also helps alleviate attenuation bias that can arise from measurement noise in continuous variables, while simultaneously serving as an important check against assumptions of strict linearity or monotonicity in the treatment effect.

We re-run the specification in Equation 5, now using the two binary version of treatment. Table OA4 and Table OA5 presents the results. The positive and significant coefficients confirm that our findings are not dependent on a continuous scoring system, indicating robustness of our qualitative conclusions.

The larger estimates from binary definitions potentially suggest the presence of nonlinearities. Specifically, the effect of AI compatibility may be more pronounced once a certain threshold of compatibility is reached. In other words, industries that surpass a certain critical level of AI compatibility might experience disproportionately stronger entrepreneurial responses after ChatGPT's release compared to incremental increases at lower levels of compatibility.

Table OA4. Effect of Above-Mean AI Assistance on Firm Creation After the Release of ChatGPT

VARIABLES	(1) IHS N. New Firms	(2) IHS N. New Firms
AI Assistance Above Mean × Post	0.128**	0.112**
	(0.055)	(0.053)
Observations	1,134	1,134
Industry 3D FE	Yes	Yes
Quarter FE	Yes	Yes
Industry 2D × Quarter FE	No	No
Controls	No	Yes
Adj. R. sq.	0.981	0.981

Notes: This table presents results testing the effect of AI's arrival on firm creation using an alternative definition of treatment status. The design is a difference-in-differences model at the industry-by-quarter level with binary treatment. The sample includes 81 three-digit NAICS industries over 2020Q4–2024Q1. The empirical specification is

$$\text{IHS(N. New Firms)}_{it} = c_0 + c_1 \text{AI Assistance Above Mean}_i \times \text{Post}_t + \text{Controls}_{it} + \mu_i + \mu_t + \epsilon_{it}$$

The dependent variable is the inverse hyperbolic sine (IHS) of the number of new firms created. Treatment status is proxied by AI Assistance Above Mean, an indicator reporting 1 if the AI Assistance Score of the industry is above the sample mean, and 0 otherwise. Treatment timing is proxied by the binary indicator Post, which equals 1 for all quarters after 2022Q3 and 0 otherwise. All specifications include industry and quarter fixed effects. Column (2) additionally includes the following controls: lag of VCPE Investments (IHS), median Return on Assets, median R&D Intensity, median CapEx Intensity, median Liquidity Ratio, and median Market Leverage. Standard errors are clustered at the three-digit NAICS industry level and reported in parentheses. Coefficients are statistically significant at the 10% (*), 5% (**), and 1% (***) levels.

Table OA5. Effect of Above-Median AI Assistance on Firm Creation After the Release of ChatGPT

VARIABLES	(1) IHS N. New Firms	(2) IHS N. New Firms
AI Assistance Above Median × Post	0.143***	0.131**
	(0.054)	(0.052)
Observations	1,134	1,134
Industry 3D FE	Yes	Yes
Quarter FE	Yes	Yes
Industry 2D × Quarter FE	No	No
Controls	No	Yes
Adj. R. sq.	0.981	0.981

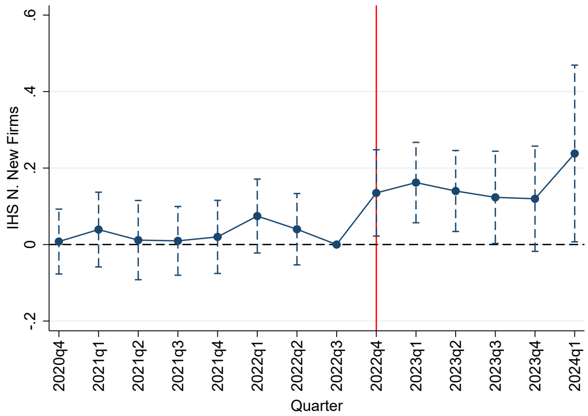
Notes: This table presents results testing the effect of AI's arrival on firm creation using an alternative definition of treatment status. The design is a difference-in-differences model at the industry-by-quarter level with binary treatment. The sample includes 81 three-digit NAICS industries over 2020Q4–2024Q1. The empirical specification is

$$\text{IHS(N. New Firms)}_{it} = d_0 + d_1 \text{AI Assistance Above Median}_i \times \text{Post}_t + \text{Controls}_{it} + \mu_i + \mu_t + \epsilon_{it}$$

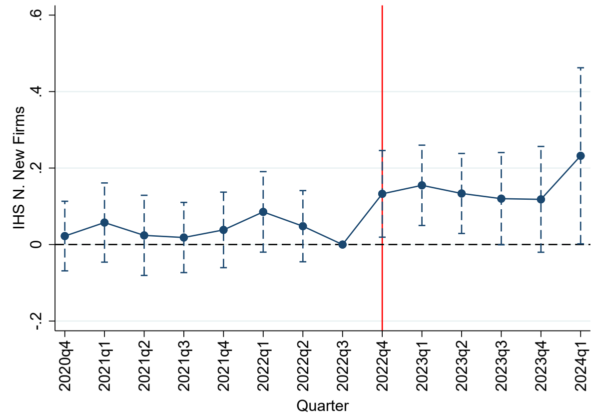
The dependent variable is the inverse hyperbolic sine (IHS) of the number of new firms created. Treatment status is proxied by AI Assistance Above Median, an indicator reporting 1 if the AI Assistance Score of an industry is above the sample median, and 0 otherwise. Treatment timing is proxied by the binary indicator Post, which equals 1 for all quarters after 2022Q3 and 0 otherwise. All specifications include industry and quarter fixed effects. Column (2) additionally includes the following controls: lag of VCPE Investments (IHS), median Return on Assets, median R&D Intensity, median CapEx Intensity, median Liquidity Ratio, and median Market Leverage. Standard errors are clustered at the three-digit NAICS industry level and reported in parentheses. Coefficients are statistically significant at the 10% (*), 5% (**), and 1% (***) levels.

In Figures OA2 and OA3 we present the event plot study. The results are largely unchanged from what we find in the main analysis.

Figure OA2. Event Study of AI Assistance Above Mean Group on Firm Creation



(a) No Controls



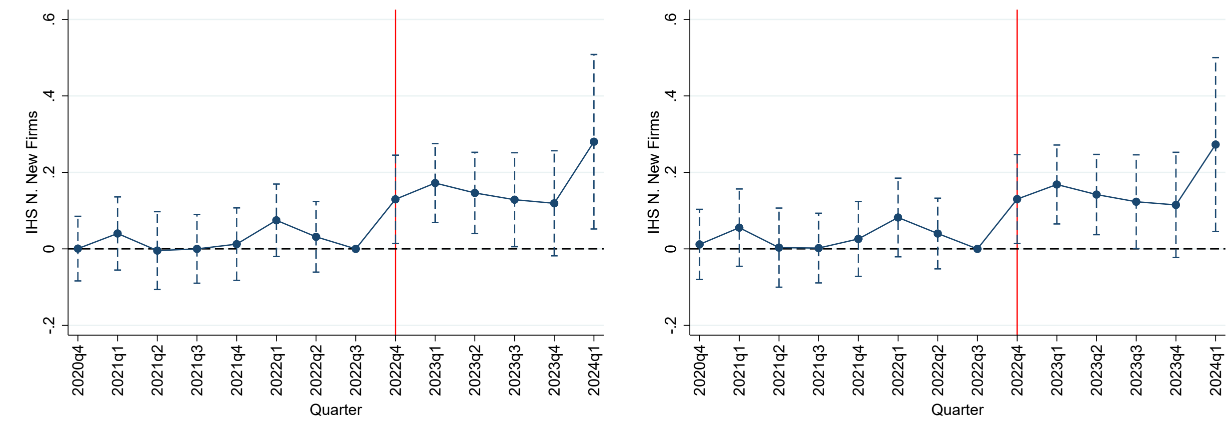
(b) Yes Controls

Notes: The figure presents the event study for the sample of 81 three-digit NAICS industries in our analysis, over the time period 2020Q4 to 2024Q1. We plot the estimates from the dynamic difference-in-differences model for the interaction of the quarter-year dummies with the AI Assistance Above Mean$_i$ from the specification

$$\text{IHS(N. New Firms)}_{it} = e_0 + \sum_{t \neq 2022q3} e_t \text{AI Assistance Above Mean}_i \times 1\{D = t\}_t + \text{Controls}_{it} + \mu_i + \mu_{It} + \epsilon_{it}$$

The dependent variable is the inverse hyperbolic sine (IHS) of the number of new firms created. Treatment intensity is proxied by AI Assistance in its binarized form (above versus below the sample mean). The specifications include both industry and quarter fixed effects. Standard errors are clustered at the three-digit NAICS industry level. The specification in Panel (a) does not include controls. The specification in Panel (b) includes the following controls: lag of VCPE Investments (IHS), median Return on Assets, median R&D Intensity, median CapEx Intensity, median Liquidity Ratio, and median Market Leverage.

Figure OA3. Event Study of AI Assistance Above Median Group on Firm Creation



(a) No Controls **(b)** Yes Controls

Notes: The figure presents the event study for the sample of 81 three-digit NAICS industries in our analysis, over the time period 2020Q4 to 2024Q1. We plot the estimates from the dynamic difference-in-differences model for the interaction of the quarter-year dummies with the AI Assistance Above Median$_i$ from the specification

$$\text{IHS(N. New Firms)}_{it} = f_0 + \sum_{t \neq 2022q3} f_t \text{AI Assistance Above Median}_i \times 1\{D = t\}_t + \text{Controls}_{it} + \mu_i + \mu_{It} + \epsilon_{it}$$

The dependent variable is the inverse hyperbolic sine (IHS) of the number of new firms created. Treatment intensity is proxied by AI Assistance in its binarized form (above versus below the sample median). The specifications include both industry and quarter fixed effects. Standard errors are clustered at the three-digit NAICS industry level. The specification in Panel (a) does not include controls. The specification in Panel (b) includes the following controls: lag of VCPE Investments (IHS), median Return on Assets, median R&D Intensity, median CapEx Intensity, median Liquidity Ratio, and median Market Leverage.

The consistent and significant positive results obtained from these binary treatment robustness checks reinforce the credibility and stability of our main results, by showing that they are not an artifact stemming from the choice of using a continuous variable as treatment definition.

A.1.4. *Falsification Test for Treatment Timing.* An essential condition for the validity of our difference-in-differences identification strategy is the parallel trends assumption. Specifically, we must assume that, in the absence of treatment (the release of ChatGPT), industries with different levels of AI compatibility would have followed parallel trajectories in firm creation over time. A critical concern is that industries classified as highly compatible with AI may have already been experiencing differential growth patterns unrelated to ChatGPT's release, thus violating the parallel trends assumption and potentially biasing our estimated effects.

To address this concern explicitly, we conduct a falsification (placebo) test. We artificially select a "placebo" event date—specifically, we choose an earlier date prior to the actual release of ChatGPT—and examine whether the differences between industries classified as high and low AI compatibility emerged at that placebo date. For this falsification test, we set the placebo event date to be 2020Q4, significantly before the actual treatment date (2022Q4). The sample time window is also rolled back accordingly to have the same number of pre and post periods in this exercises, as we have in the main analysis. Coincidentally, since we roll the treatment date back by 8 quarters—with the time window now being 2018Q4—2022Q1, this completely excludes the presence of post-treatment periods from the robustness check, allowing to better assess the potential presence of a reaction from some industries, which would be unrelated to the arrival of ChatGPT.

If our parallel trends assumption holds, we would expect no statistically significant divergence in firm creation between high and low AI-compatible industries at this earlier, placebo event date. Conversely, observing a significant divergence at this date would indicate pre-existing differential trends, weakening our causal interpretation.

The results from this falsification exercise, presented in Table OA6 and Figure OA4, indicate no significant differential effects at the placebo event date. Specifically, the coefficients from the compact DID in Table OA6 are small and statistically insignificant, while the event studies in Figure OA4 show no signs of diverging pre-trends or dynamic effects following the fictitious treatment.

Table OA6. Effects of AI Assistance on Firm Creation using Falsified Treatment Date

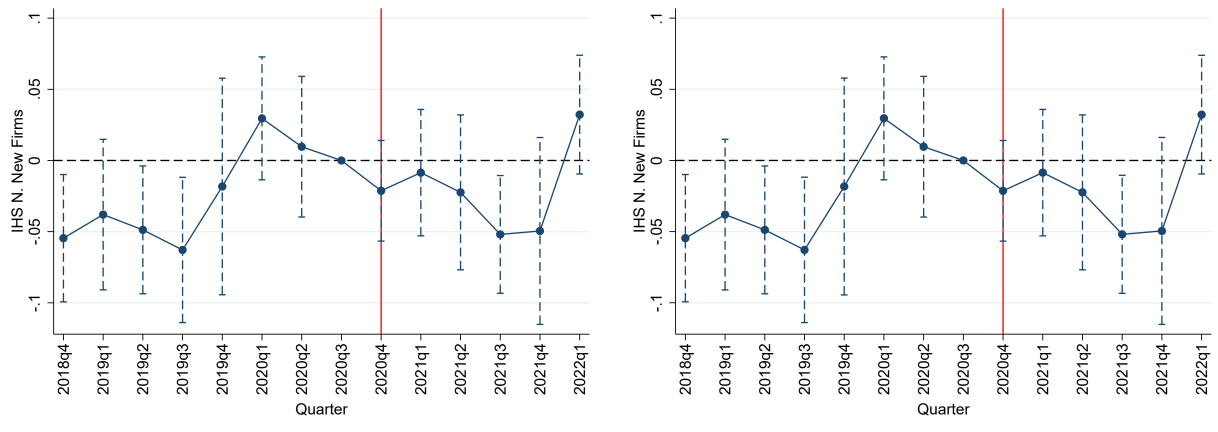
VARIABLES	(1) IHS N. New Firms	(2) IHS N. New Firms
AI Assistance × Placebo Event	0.014	0.011
	(0.013)	(0.013)
Observations	1,134	1,134
Industry 3D FE	Yes	Yes
Quarter FE	Yes	Yes
Industry 2D × Quarter FE	No	No
Controls	No	Yes
Adj. R. sq.	0.992	0.992

Notes: The table reports the results for the falsification exercise where the treatment date is moved to 2020Q4. The design is a difference-in-differences model at the industry-by-quarter level with continuous treatment. The sample includes 81 three-digit NAICS industries over the period 2018Q4–2022Q1. The empirical specification is

$$\text{IHS(N. New Firms)}_{it} = h_0 + h_1 \text{AI Assistance}_i \times \text{Placebo Event}_t + \text{Controls}_{it} + \mu_i + \mu_t + \epsilon_{it}$$

The dependent variable is the inverse hyperbolic sine (IHS) of the number of new firms created. Treatment intensity is proxied by AI Assistance in its standardized form. Treatment timing is proxied by the binary indicator Placebo Event, which equals 1 for all quarters after 2020Q3 and 0 otherwise. All specifications include industry and quarter fixed effects. Column (2) additionally includes the following controls: lag of VCPE Investments (IHS), median Return on Assets, median R&D Intensity, median CapEx Intensity, median Liquidity Ratio, and median Market Leverage. Standard errors are clustered at the three-digit NAICS industry level and reported in parentheses. Coefficients are statistically significant at the 10% (*), 5% (**), and 1% (***) levels.

Figure OA4. Falsified Event Study of AI Assistance on Firm Creation



(a) No Controls

(b) Yes Controls

Notes: This figure reports the event study plot for the sample of 81 three-digit NAICS industries in our analysis, over the time period 2018Q4 to 2022Q1. We falsify the event date by choosing as treatment quarter 2020Q3. We plot the estimates from the dynamic difference-in-differences model for the interaction of the quarter dummies with AI Assistance$_i$ from the falsified robustness specification

$$\text{IHS(N. New Firms)}_{it} = g_0 + \sum_{t \neq 2020q3} g_t \text{AI Assistance}_i \times 1\{\mathcal{D} = t\}_t + \text{Controls}_{it} + \mu_i + \mu_t + \epsilon_{it}$$

The dependent variable is the inverse hyperbolic sine (IHS) of the number of new firms created. Treatment intensity is proxied by AI Assistance in its standardized form. The specifications include both industry and quarter fixed effects. Standard errors are clustered at the three-digit NAICS industry level. The specification in Panel (a) does not include controls. The specification in Panel (b) includes the following controls: lag of VCPE Investments (IHS), median Return on Assets, median R&D Intensity, median CapEx Intensity, median Liquidity Ratio, and median Market Leverage.

The null findings strongly support the parallel trends assumption by demonstrating that the divergence in firm creation patterns observed post-ChatGPT does not reflect a continuation of pre-existing differential trajectories. Therefore, this falsification test significantly strengthens the credibility of our causal claims regarding the effects of generative AI compatibility on entrepreneurial activity.

A.1.5. *Heterogeneity and Non-linearity Analysis.* To further validate our findings and to explore possible nonlinearities, we perform a heterogeneity analysis. The motivation behind this analysis is twofold. First, our main analysis, employing a continuous measure of AI compatibility (AI Assistance Score), implicitly assumes a smooth, linear relationship between treatment intensity and firm creation. However, economically, the benefits of generative AI tools might differ significantly at higher levels of compatibility. Second, by explicitly examining subsets of industries—such as those in the top quartile or decile groups—we directly address concerns related to potential measurement error or misclassification bias in our continuous measure. Clearly defined groups based on explicit thresholds reduce sensitivity to subtle measurement inaccuracies, thus helping to clarify the robustness of our estimated effects.

In the first set of heterogeneity tests we redefine treatment by grouping industries based on whether their AI Assistance Score belongs to the top quartile of the distribution. Table OA7 strongly supports the existence of meaningful heterogeneity and nonlinearities in the treatment effect. Specifically, we find that industries in the top quartile of AI compatibility exhibit significantly larger effects compared to the overall continuous measure in Table 4. This result indicates that the entrepreneurial response to generative AI is notably stronger once a critical threshold of compatibility is surpassed.

Table OA7. Effect of Top Quartile AI Assistance on Firm Creation After the Release of ChatGPT

VARIABLES	(1) IHS N. New Firms	(2) IHS N. New Firms
AI Assistance Top Quartile × Post	0.240***	0.228***
	(0.057)	(0.058)
Observations	1,134	1,134
Industry 3D FE	Yes	Yes
Quarter FE	Yes	Yes
Industry 2D × Quarter FE	No	No
Controls	No	Yes
Adj. R. sq.	0.982	0.982

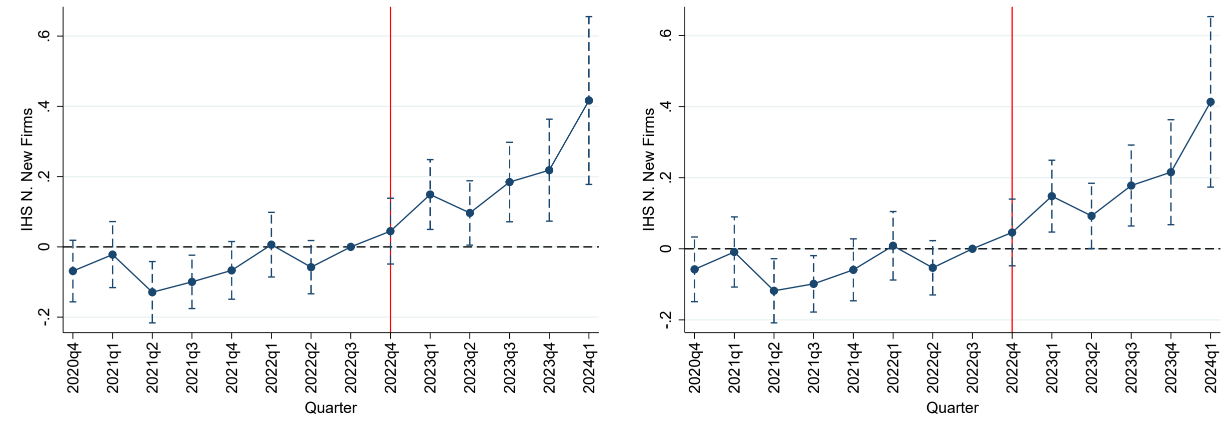
Notes: The table presents results testing the functional shape of the effect of AI's arrival on firm creation. The design is a difference-in-differences model at the industry-quarter level with a binary treatment. The sample includes 81 three-digit NAICS industries over the period 2020Q4–2024Q1. The empirical specification is

$$\text{IHS(N. New Firms)}_{it} = h_0 + h_1 \text{AI Assistance Top Quartile}_i \times \text{Post}_t + \text{Controls}_{it} + \mu_i + \mu_t + \epsilon_{it}$$

The dependent variable is the inverse hyperbolic sine (IHS) of the number of new firms created. Treatment status is proxied by AI Assistance Top Quartile, an indicator reporting 1 if the AI Assistance Score of an industry belongs to the top quartile of the sample distribution, and 0 otherwise. Treatment timing is proxied by the binary indicator Post, which equals 1 for all quarters after 2022Q3 and 0 otherwise. All specifications include industry and quarter fixed effects. Column (2) additionally includes the following controls: lag of VCPE Investments (IHS), median Return on Assets, median R&D Intensity, median CapEx Intensity, median Liquidity Ratio, and median Market Leverage. Standard errors are clustered at the three-digit NAICS industry level and reported in parentheses. Coefficients are statistically significant at the 10% (*), 5% (**), and 1% (***) levels.

In Figure OA5 we plot the event study for the specification where we use as definition of treatment belonging to the top quartile of the AI Assistance Score distribution and not only we do not observe diverging pre-trends, much like in the main analysis, but the significance of the lag coefficients is strong and the effects are incremental over time, consistent with idea of progressive take up.

Figure OA5. Event Study of Top Quartile AI Assistance Score on Firm Creation



(a) No Controls **(b)** Yes Controls

Notes: The figure presents the event study for the sample of 81 three-digit NAICS industries in our analysis, over the time period 2020Q4 to 2024Q1. We plot the estimates from the dynamic difference-in-differences model for the interaction of the quarter-year dummies with the AI Assistance Top Quartile$_i$ from the specification

$$\text{IHS(N. New Firms)}_{it} = j_0 + \sum_{t \neq 2022q3} j_t \text{AI Assistance Top Quartile}_i \times 1\{D = t\}_t + \text{Controls}_{it} + \mu_i + \mu_t + \epsilon_{it}$$

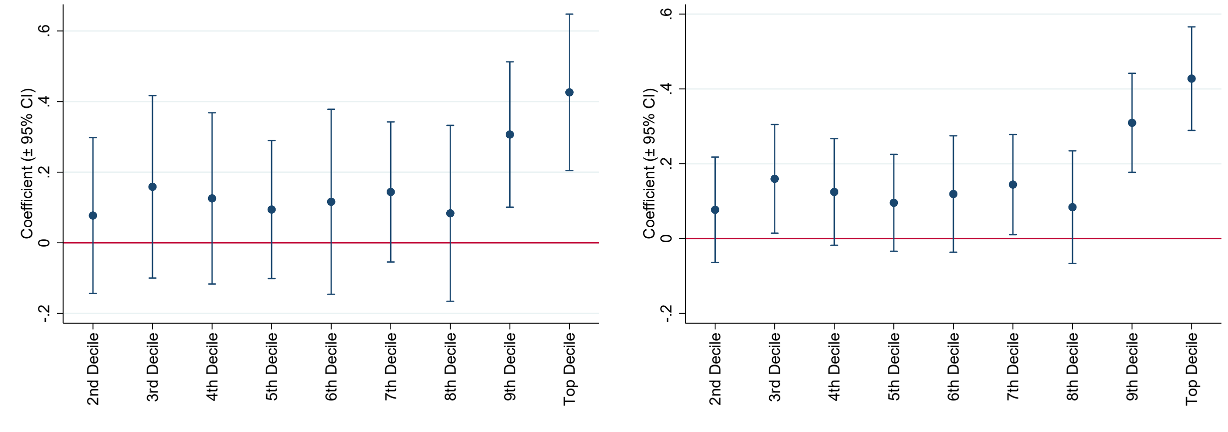
The dependent variable is the inverse hyperbolic sine (IHS) of the number of new firms created. Treatment intensity is proxied by AI Assistance in its standardized form. The specifications include both industry and quarter fixed effects. Standard errors are clustered at the three-digit NAICS industry level. The specification in Panel (a) does not include controls. The specification in Panel (b) includes the following controls: lag of VCPE Investments (IHS), median Return on Assets, median R&DIntensity, median CapEx Intensity, median Liquidity Ratio, and median Market Leverage.

Additionally, we plot the estimated treatment effects by decile to investigate how the impact of AI compatibility on entrepreneurial activity varies across the entire distribution of AI Assistance Scores. By dividing industries into ten equally-sized groups based on their AI compatibility, we can carefully assess whether the entrepreneurial responses increase smoothly and monotonically with higher compatibility levels or if distinct thresholds exist beyond which these effects intensify significantly. This analysis allows us to identify potential nonlinearities and provides valuable insights into whether the benefits of generative AI accrue gradually or primarily above specific compatibility thresholds.

Econometrically speaking, we run our difference-in-differences regression again, but this time we replace our continuous AI Assistance Score with a set of dummy variables for each decile. For instance, the 1st decile serves as the baseline reference group, and we estimate separate coefficients for the 2nd decile through the 10th decile, each interacted with the post-treatment indicator. Thus, each estimated coefficient captures how industries in a particular decile of AI compatibility differ in their entrepreneurial responses after ChatGPT's release relative to the lowest decile.

The decile analysis in Figure OA6 reveals a clear and systematic relationship between AI compatibility and entrepreneurial activity. Specifically, we observe that industries situated in lower deciles, representing those with comparatively low AI compatibility, exhibit modest or negligible increases in firm creation following ChatGPT's release. In contrast, as we move progressively towards higher deciles, the estimated effects become increasingly pronounced, demonstrating a clear and consistent monotonic relationship. Industries in the highest deciles, characterized by substantially greater AI compatibility, show considerably larger and statistically significant impacts on new firm creation. These findings underscore the presence of meaningful nonlinearities, indicating that entrepreneurial benefits associated with generative AI intensify substantially beyond certain critical compatibility thresholds. Thus, this decile analysis not only corroborates our main results but also enriches our understanding by highlighting nuanced dynamics that inform strategic decision-making and targeted policy interventions.

Figure OA6. Effect of the Release of ChatGPT by AI Assistance Deciles



(a) No Controls

(b) Yes Controls

Notes: The figure presents the heterogeneity study for the sample of 81 three-digit NAICS industries in our analysis, over the time period 2020Q4–2024Q1. We plot the estimates from a compact difference-in-differences for the vector of decile based categorical groups of AI Assistance

$$\text{IHS(N. New Firms)}_{it} = k_0 + \sum_{d=2}^{10} k_d \text{AI Assistance Decile}_i^d \times \text{Post}_t + \text{Controls}_{it} + \mu_i + \mu_t + \epsilon_{it}$$

The dependent variable is the inverse hyperbolic sine (IHS) of the number of new firms created. Treatment intensity groups are created by binning industries into ten equally-sized deciles according to their AI Assistance Score, with the first decile (lowest AI compatibility) serving as the baseline category. The specification includes both industry and quarter fixed effects. Standard errors are clustered at the three-digit NAICS industry level. The specification in Panel (a) does not include controls. The specification in Panel (b) includes the following controls: lag of VCPE Investments (IHS), median Return on Assets, median R&D Intensity, median Capex Intensity, median Liquidity Ratio and median Market Leverage.

Taken together, these heterogeneity analyses significantly enhance the interpretability and economic relevance of our primary findings by explicitly revealing that the entrepreneurial impact of generative AI compatibility does not merely increase uniformly across all industries. Instead, the results clearly indicate the presence of threshold-like or nonlinear effects, where the benefits of generative AI tools become especially pronounced beyond certain critical levels of compatibility. This finding has crucial implications: it suggests that policy interventions, resource allocations, or strategic industry investments aimed at fostering entrepreneurial activity through generative AI may be particularly effective when targeted at industries that surpass specific thresholds of AI compatibility, rather than applied broadly across all sectors. Furthermore, the monotonic increase in estimated effects across deciles provides a transparent and intuitive demonstration of a dose-response relationship, meaning industries progressively better suited to leveraging generative AI experience correspondingly larger entrepreneurial gains. These insights are economically meaningful as they clearly delineate where the potential for generative AI-driven entrepreneurial growth is highest, thereby offering practical guidance for policymakers and industry stakeholders seeking to maximize the economic benefits of AI technology adoption.

It is important to emphasize that the presence of nonlinearities revealed by our heterogeneity analyses does not contradict or invalidate our primary empirical strategy, which employs a continuous exposure variable (the AI Assistance Score). Rather, the identification of nonlinear effects enriches our understanding by explicitly highlighting thresholds and nuanced dynamics that are inherent to many economic relationships. The continuous measure remains a valid and informative way to capture the general and overall relationship between generative AI compatibility and entrepreneurial outcomes. The robustness and heterogeneity analyses complement this approach by demonstrating that entrepreneurial responses to AI compatibility intensify disproportionately at higher compatibility levels. Far from invalidating our continuous specification, these findings clarify and deepen its economic interpretation, offering valuable insights for policy targeting and strategic investment decisions.

A.1.6. *AI Assistance's Overlap with Other Industry Groups.* A natural concern is that the post-ChatGPT increase in firm creation we document could be driven by lingering dynamics from the COVID-19 pandemic. Haltiwanger (2022) shows that the surge in firm formation during COVID was highly concentrated in a small set of three-digit NAICS industries, in particular 41, 444, 445, 449, 455, 456, 457, 458, 459, 484, 541, 561, 721, and 812. We refer to these industries as "COVID-boom industries." If these sectors both expanded rapidly during the pandemic and happened to score highly on our AI Assistance measure, then our baseline results could be conflating pandemic-era entry dynamics with the effects of AI.

To address this concern, we extend our difference-in-differences specification in two ways. First, we augment the regression with an interaction between the COVID-boom industry indicator and the post-ChatGPT period. The results are reported in Table OA8. The coefficient on this interaction is statistically insignificant, while the coefficient on *AI Assistance* $\times$ *Post-ChatGPT* remains statistically significant and its magnitude is virtually unchanged from the baseline.

Table OA8. Disentangling Effect of Covid-Induced Firm Formation from AI Arrival

VARIABLES	(1) IHS N. New Firms	(2) IHS N. New Firms
AI Assistance × Post	0.105***	0.109***
	(0.025)	(0.025)
Covid Surging × Post		-0.040
		(0.072)
Observations	1,134	1,134
Industry 3D FE	Yes	Yes
Quarter FE	Yes	Yes
Industry 2D × Quarter FE	No	No
Controls	No	No
Adj. R. sq.	0.982	0.982

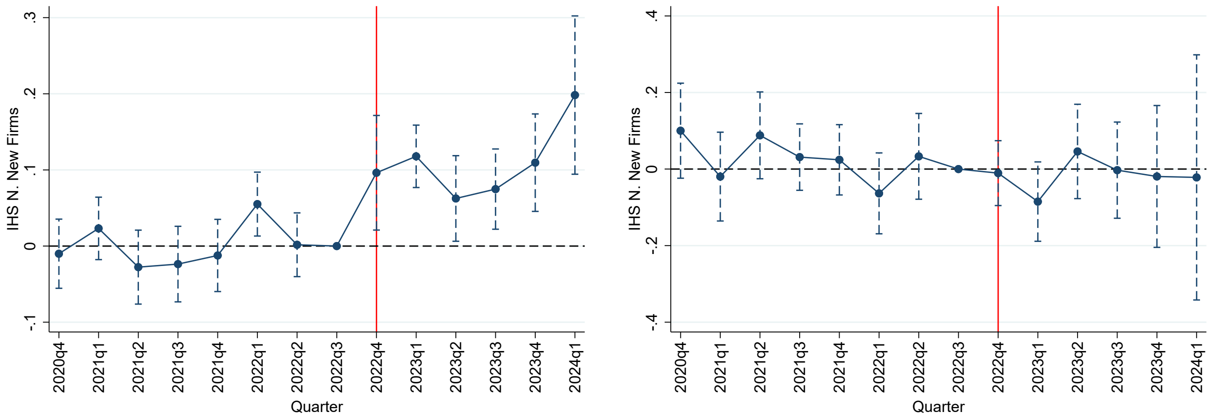
Notes: This table presents results isolating the effect of AI's arrival from COVID-induced firm creation on the number of new businesses. The design is a difference-in-differences model at the industry-by-quarter level with continuous treatment, where we control for the fact that some industries surged during COVID. The sample includes 81 three-digit NAICS industries over 2020Q4–2024Q1. In the first column we report the results from the main analysis where we use just AI Assistance as treatment level indicator. In the second column we test the effect of being a Covid Surging industry (441, 444, 445, 449, 455, 456, 457, 458, 459, 484, 541, 561, 721, 812) using empirical specification is

$$\text{IHS(N. New Firms)}_{it} = l_0 + l_1 \text{AI Assistance}_i \times \text{Post}_t + l_2 \text{Covid Surging}_i \times \text{Post}_t + \mu_i + \mu_t + \epsilon_{it}$$

The dependent variable is the inverse hyperbolic sine (IHS) of the number of new firms created. Treatment intensity is proxied by AI Assistance in its standardized form. Treatment timing is proxied by the binary indicator Post, which equals 1 for all quarters after 2022Q3 and 0 otherwise. Column (1) presents the results from the baseline model used in the main analysis, while Column (2) presents the results for the model accounting for Covid-related effects. All specifications include industry and quarter fixed effects. Standard errors are clustered at the three-digit NAICS industry level and reported in parentheses. Coefficients are statistically significant at the 10% (*), 5% (**), and 1% (***) levels.

Second, we estimate an event-study specification in which COVID-boom industries are allowed to have their own dynamic coefficients relative to the pre-ChatGPT period, and we also interact the AI Assistance variable with the dynamic coefficients. The results are presented in Figure OA7. While the coefficients for the AI Assistance variable do not deviate from the main results (Figure OA7, Panel (a)), the estimated coefficients for COVID-boom industries are statistically insignificant, and there is no evidence of systematic differences in their dynamics in the post-ChatGPT period (Figure OA7, Panel (b)).

Figure OA7. Event Study of Covid Surging Industries on Firm Creation Around the Release of ChatGPT



(a) AI Assistance **(b)** Covid Surging

Notes: In this panel of figures we report the coefficients from the event study including dynamic interactions for both covid-booming industries and industries by AI Assistance Score. The sample of 81 three-digit NAICS industries in our analysis, over the time period 2020Q4–2024Q1. We coefficients come from the following specification

$$\text{IHS(N. New Firms)}_{it} = m_0 + \sum_{t \neq 2022q3} m_t^1 \text{AI Assistance}_i \times 1\{D = t\}_t + \sum_{t \neq 2022q3} m_t^2 \text{Covid Surging}_i \times 1\{D = t\}_t + \mu_i + \mu_t + \epsilon_{it}$$

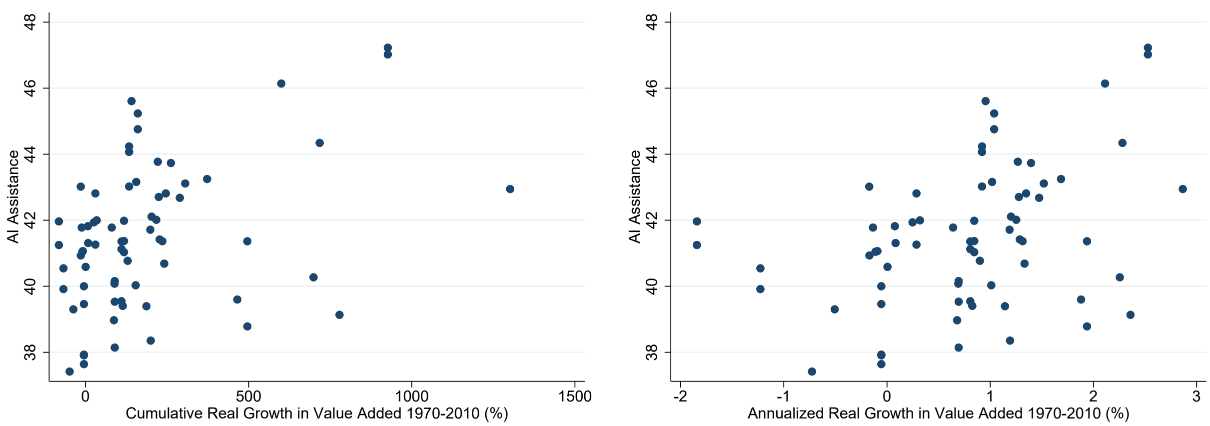
The dependent variable is the inverse hyperbolic sine (IHS) of the number of new firms created. We identify as Covid Surging the following industries: 441, 444, 445, 449, 455, 456, 457, 458, 459, 484, 541, 561, 721, 812. The specification also includes the interaction of AI Assistance (standardized) with the quarter indicators. The specifications include both industry and quarter fixed effects. Standard errors are clustered at the three-digit NAICS industry level.

All in all, industries which experienced exceptional growth during the Covid period, did not mantain the same growth trajectory over our sample period of study, confirming that our results are not driven by unaccounted factors.

A second concern is that the set of high AI Assistance industries could coincide with industries that, independently of AI, experienced strong secular growth in value added between 1970 and 2010. In other words, heightened firm creation in these industries might mechanically reflect long-run industry trajectories rather than a distinct dimension of entrepreneurial relevance. The worry is that industries with persistent increases in value added over several decades are precisely those that we rank highly on AI Assistance, so that our results could be interpreted as a continuation of long-run trends rather than a causal effect of the commercialization of AI.

To address this concern, we construct a measure of long-run industry growth based on the Bureau of Economic Analysis (BEA) data on Gross Value Added by industry . Specifically, for each 3-digit NAICS industry we compute the average annual log growth rate in real value added and the cumulative growth between the 1970s and the 2010s. We then correlate this measure of long-run growth with our AI Assistance ranking. The correlation for both is modest, at 0.39 (0.387 and 0.392), implying that only about 15% of the variation in AI Assistance across industries can be linearly explained by differences in long-run value added growth. This finding suggests that while industries that have grown more over the past decades tend to be somewhat more AI-assistable, the majority of the cross-industry variation in AI Assistance reflects other dimensions that are orthogonal to long-run secular growth trends. Additionally we scatter both measure of growth in value added against the AI Assistance variable in Figure OA8. to show that no striking patter emerges.

Figure OA8. Correlation between AI Assistance and Long-Term Growth in Value Added



(a) Cumulative Growth **(b)** Annualized Growth

Notes: This set of plots displays the correlation between the AI Assistance measure and estimates of long-term growth in value added. Plot (a) scatters the AI Assistance measure against the cumulative growth in value added over the period 1970-2010, while Plot (b) scatters it against the annualized log growth over the same period.

Taken together, this evidence alleviates the concern that our AI Assistance measure is simply a relabeling of historically "winning" industries. Moreover, our empirical design—with industry fixed effects and pre-trend tests—ensures that we identify the discrete shift in entrepreneurial activity that follows the release of ChatGPT, rather than a continuation of long-run trajectories.

A.2. **Alternative Specifications for Cost of Experimentation Mechanism.** A natural concern with our baseline specification is that the results could be sensitive to the particular functional form used to classify experimentation-intensive industries. To address the robustness of our cost of experimentation mechanism results presented in Subsection 5.1, we redefined the indicator for experimentation-intensive industries. In the baseline specification, this indicator equals one if an industry's experimentation intensity lies in the top quartile of the distribution.

In the first robustness test, we reclassify industries as experimentation-intensive when their Experimentation Intensity Score is above the sample median. Table OA9 displays the results for the specification using the new binary indicator. The triple interaction coefficient, used to test the mechanism still displays a positive sign, in line with out prediction. However, the results are attenuated both in terms of magnitude and significance.

Table OA9. Effects of the Cost of Experimentation Above Median on Firm Creation

VARIABLES	(1) IHS N. New Firms	(2) IHS N. New Firms
AI Assistance × Experimentation Above Median × Post	0.103	0.111
	(0.099)	(0.095)
AI Assistance × Post	0.070	0.062
	(0.066)	(0.063)
Experimentation Above Median × Post	-0.053	-0.057
	(0.074)	(0.073)
Observations	1,078	1,078
Industry 3D FE	Yes	Yes
Quarter FE	No	No
Industry 2D × Quarter FE	Yes	Yes
Controls	No	Yes
Adj. R sq.	0.980	0.980

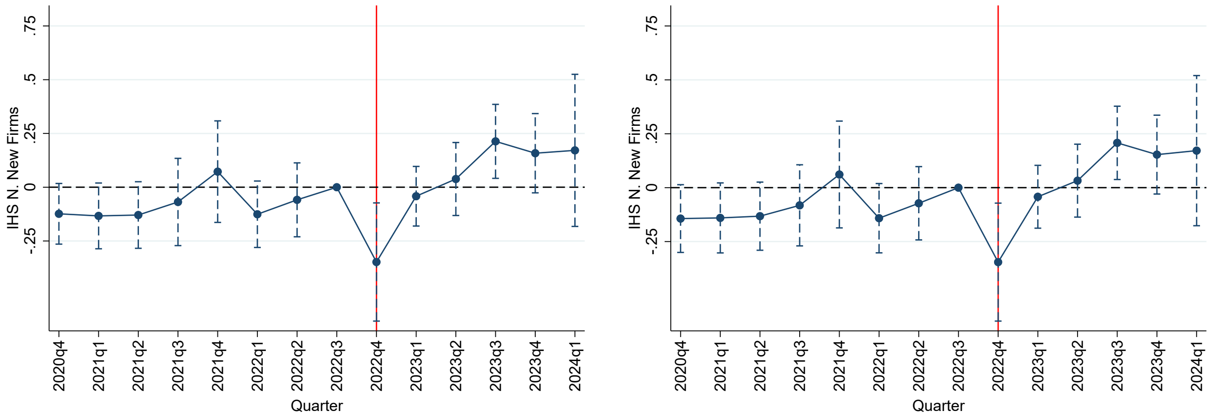
Notes: This table presents results testing whether the reduced cost-of-experimentation mechanism operates using an alternative definition of experimentation intense industries. The design is a triple differences model at the industry-by-quarter level with continuous treatment and a heterogeneity term. The sample includes 81 three-digit NAICS industries over 2020Q4–2024Q1. The empirical specification is

$$\begin{aligned}\text{IHS(N. New Firms)}_{it} = \gamma_0^1 &+ \gamma_1^1 \text{AI Assistance}_i \times \text{Experimentation Above Median}_i \times \text{Post}_t \\ &+ \gamma_2^1 \text{AI Assistance}_i \times \text{Post}_t + \gamma_3 \text{Experimentation Above Median}_i \times \text{Post}_t \\ &+ \text{Controls}_{it} + \mu_i + \mu_{It} + \varepsilon_{it}\end{aligned}$$

The dependent variable is the inverse hyperbolic sine (IHS) of the number of new firms created. Treatment intensity is proxied by AI Assistance in its standardized form. Treatment timing is proxied by Post, which equals 1 for all quarters after 2022Q3 and 0 otherwise. Treatment heterogeneity comes from the indicator Experimentation Above Median, which equals 1 if the industry's Experimentation Intensity Score is above the sample median and 0 otherwise. The specifications include industry and sector-by-quarter fixed effects. Column (2) additionally includes the following controls: lagged VCPE Investments (IHS), median Return on Assets, median R&D Intensity, median CapEx Intensity, median Liquidity Ratio, and median Market Leverage. Standard errors are clustered at the three-digit NAICS industry level and reported in parentheses. Coefficients are statistically significant at the 10% (*), 5% (**), and 1% (***) levels.

We also re-run the event study specification and plot the coefficients (Figure OA9), while using the alternative definition for experimentation-intense industries. The event study displays no pre-trends, in line with previous results and the post treatment coefficients are still positive, albeit not being significant.

Figure OA9. Event Study of Cost of Experimentation Above Median on Firm Creation Around the Release of ChatGPT



(a) No Controls **(b)** Yes Controls

Notes: The figure presents the event study for the sample of 81 three-digit NAICS industries in our analysis, over the time period 2020Q4–2024Q1. We plot the estimates from the dynamic triple differences model for the interaction of the quarter-year dummies with the AI Assistance$_i$ $\times$ Experimentation Above Median$_i$ from the specification

$$\begin{aligned}\text{IHS(N. New Firms)}_{it} = \delta_0^{a0} &+ \sum_{t\neq 2022q3} \delta_t^{a1}\text{AI Assistance}_i \times \text{Experimentation Above Median}_i \times 1\{D=t\}_t \\ &+ \sum_{t\neq 2022q3} \delta_t^{a2}\text{AI Assistance}_i \times 1\{D=t\}_t \\ &+ \sum_{t\neq 2022q3} \delta_t^{a3}\text{Experimentation Above Median}_i \times 1\{D=t\}_t + \text{Controls}_{it} + \mu_i + \mu_{It} + \varepsilon_{it}\end{aligned}$$

The dependent variable is the Inverse Hyperbolic Sine of the number of new firms created. Treatment intensity is proxied by AI Assistance in its standardized form. Treatment heterogeneity comes from the indicator variable Experimentation Above Median, which equal 1 if the industry reports a value of Experimentation Intensity above the sample median and 0 otherwise. The specifications include both industry and sector-by-quarter fixed effects. Standard errors are clustered at the three-digit NAICS industry level. The specification in Panel (a) does not include controls. The specification in Panel (b) includes the following controls: lag of VCPE Investments (IHS), median Return on Assets, median R&D Intensity, median CapEx Intensity, median Liquidity Ratio, and median Market Leverage.

In the second robustness test, we reclassify industries as experimentation-intensive when their Experimentation Intensity Score is above the sample mean. Table OA10 displays the results for the specification using the new binary indicator. Similar to the previous test, the triple interaction coefficient displays a positive sign, albeit being smaller in size and not significant when compared to the baseline result.

Table OA10. Effects of the Cost of Experimentation Above Mean on Firm Creation

VARIABLES	(1) IHS N. New Firms	(2) IHS N. New Firms
AI Assistance × Experimentation Above Mean × Post	0.085	0.095
	(0.091)	(0.087)
AI Assistance × Post	0.089	0.077
	(0.070)	(0.068)
Experimentation Above Mean × Post	-0.076	-0.075
	(0.072)	(0.069)
Observations	1,078	1,078
Industry 3D FE	Yes	Yes
Quarter FE	No	No
Industry 2D × Quarter FE	Yes	Yes
Controls	No	Yes
Adj. R sq.	0.980	0.980

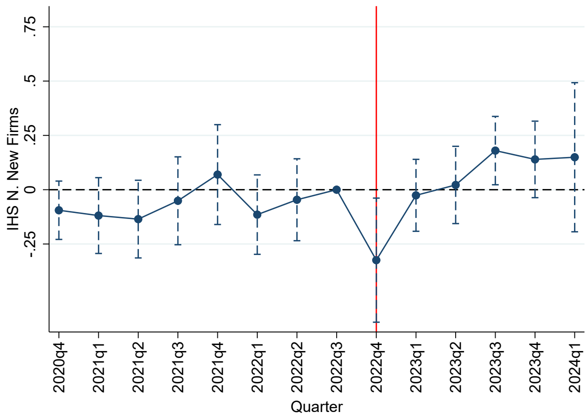
Notes: This table presents results testing whether the reduced cost-of-experimentation mechanism operates using an alternative definition of experimentation intense industries. The design is a difference-in-differences model at the industry-by-quarter level with continuous treatment and a heterogeneity term. The sample includes 81 three-digit NAICS industries over 2020Q4–2024Q1. The empirical specification is

$$\begin{aligned}\text{IHS(N. New Firms)}_{it} = \gamma_0^2 &+ \gamma_1^2 \text{AI Assistance}_i \times \text{Experimentation Above Mean}_i \times \text{Post}_t \\ &+ \gamma_2^2 \text{AI Assistance}_i \times \text{Post}_t + \gamma_3^2 \text{Experimentation Above Mean}_i \times \text{Post}_t \\ &+ \text{Controls}_{it} + \mu_i + \mu_{It} + \varepsilon_{it}\end{aligned}$$

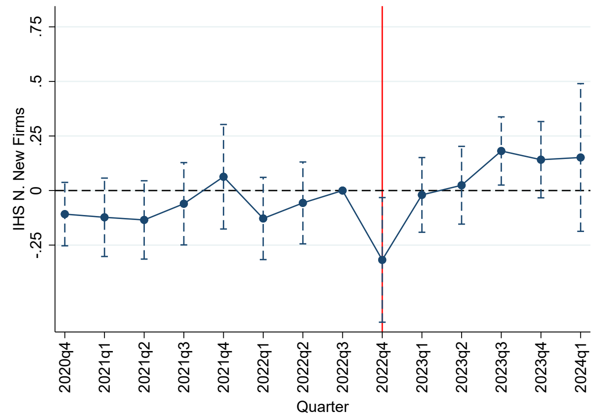
The dependent variable is the inverse hyperbolic sine (IHS) of the number of new firms created. Treatment intensity is proxied by AI Assistance in its standardized form. Treatment timing is proxied by Post, which equals 1 for all quarters after 2022Q3 and 0 otherwise. Treatment heterogeneity comes from the indicator Experimentation Intense, which equals 1 if the industry's Experimentation Intensity Score is above the sample mean and 0 otherwise. The specifications include industry and sector-by-quarter fixed effects. Column (2) additionally includes the following controls: lagged VCPE Investments (IHS), median Return on Assets, median R&D Intensity, median CapEx Intensity, median Liquidity Ratio, and median Market Leverage. Standard errors are clustered at the three-digit NAICS industry level and reported in parentheses. Coefficients are statistically significant at the 10% (*), 5% (**), and 1% (***) levels.

We also re-run the event study specification and plot the coefficients (Figure OA10), while using the alternative definition for experimentation-intense industries. The event study displays no pre-trends and positive (non-significant) post treatment coefficients, in line with previous test and the baseline results.

Figure OA10. Event Study of Cost of Experimentation Above Mean on Firm Creation Around the Release of ChatGPT



(a) No Controls



(b) Yes Controls

Notes: The figure presents the event study for the sample of 81 three-digit NAICS industries in our analysis, over the time period 2020Q4–2024Q1. We plot the estimates from the dynamic triple differences model for the interaction of the quarter-year dummies with the AI Assistance$_i$ $\times$ Experimentation Above Mean$_i$ from the specification

$$\begin{aligned}\text{IHS(N. New Firms)}_{it} = \delta_0^{b0} &+ \sum_{t \neq 2022q3} \delta_t^{b1} \text{AI Assistance}_i \times \text{Experimentation Above Mean}_i \times 1\{D = t\}_t \\ &+ \sum_{t \neq 2022q3} \delta_t^{b2} \text{AI Assistance}_i \times 1\{D = t\}_t \\ &+ \sum_{t \neq 2022q3} \delta_t^{b3} \text{Experimentation Above Mean}_i \times 1\{D = t\}_t + \text{Controls}_{it} + \mu_i + \mu_{It} + \varepsilon_{it}\end{aligned}$$

The dependent variable is the Inverse Hyperbolic Sine of the number of new firms created. Treatment intensity is proxied by AI Assistance in its standardized form. Treatment heterogeneity comes from the indicator variable Experimentation Above Mean, which equal 1 if the industry reports a value of Experimentation Intensity above the sample mean and 0 otherwise. The specifications include both industry and sector-by-quarter fixed effects. Standard errors are clustered at the three-digit NAICS industry level. The specification in Panel (a) does not include controls. The specification in Panel (b) includes the following controls: lag of VCPE Investments (IHS), median Return on Assets, median R&D Intensity, median CapEx Intensity, median Liquidity Ratio, and median Market Leverage.

Overall, these robustness tests qualitatively match our main findings. The fact that the coefficients are somewhat attenuated in size and significance level should not be interpreted as evidence that the mechanism is sensitive to functional form choice. Instead, it highlights that experimentation costs act as a barrier to entry primarily when experimentation is central to the operations of an industry.

A.3. **Alternative Interpretation of the Capital Allocation Mechanism.** An alternative explanation for the observed increase in firm creation is that the *capital response channel* may operate in the opposite direction of the standard capital reallocation story. Rather than attracting more financing, AI-compatible industries may require less capital per firm once AI raises productivity. In this case, average deal sizes would fall, and the rise in firm creation could be mediated by reduced capital needs rather than by increased capital inflows. This "efficiency channel" would imply that lower capital requirements per startup are what enable more entrepreneurs to enter.

To test this mechanism, we rerun the mediation analysis presented in Table 7 of the main text using the inverse hyperbolic sine (IHS) of the average investment size as the mediator, rather than total capital. The results are presented in Table OA11. For the sake of exposition, we present again th effect of the treatment on firm creation in Column (1). In column (2), we regress the IHS of average investment size on the treatment variable. Here the coefficient is negative, about –0.36, and statistically significant at the 10 percent level. This provides some, albeit weak, evidence that deals are smaller in AI-compatible industries after ChatGPT. Finally, we include the IHS of average investment size directly in the firm creation regression as a potential mediator. In this specification, the coefficient on average investment size is statistically insignificant, and the coefficient on AI Assistance $\times$ Post remains unchanged at approximately 0.10.

These results suggest that although there is a weak indication that average deal sizes decline in AI-compatible industries post-ChatGPT, this decline does not mediate the observed increase in firm creation. In other words, the treatment effect on entry is unaffected by controlling for changes in average investment size.

Table OA11. Mediation Analysis: Average Investment Size, AI Assistance and Firm Creation

VARIABLES	(1) IHS N. New Firms	(2) IHS Avg. Investment	(3) IHS N. New Firms
AI Assistance × Post	0.102**	-0.360*	0.102**
	(0.046)	(0.203)	(0.046)
IHS Avg. Investment			-0.000
			(0.004)
Observations	1,078	1,078	1,078
Industry 3D FE	Yes	Yes	Yes
Quarter FE	No	No	No
Industry 2D × Quarter FE	Yes	Yes	Yes
Controls	No	No	No
Adj. R sq.	0.980	0.451	0.980

Notes: This table reports a mediation-style decomposition to study an alternative version of the capital reallocation as a potential mechanism linking AI's arrival and firm creation. The analysis is at the industry-by-quarter level for 81 three-digit NAICS industries over 2020Q4–2024Q1. The mediator is average allocated capital, measured as the inverse hyperbolic sine (IHS) of the average VCPE investment size from Preqin (in $000).
Column (1) reports the results for the estimation of the total impact that the treatment has on firm formation. The specification is

$$\text{IHS(N. New Firms)}_{it} = \bar{\zeta}_0 + \bar{\zeta}_1 \text{AI Assistance}_i \times \text{Post}_t + \mu_i + \mu_{It} + \epsilon_{it}$$

Column (2) assesses the linkage between the treatment and the average size of VCPE investments, according to the specification

$$\text{IHS(Avg. Investment)}_{it} = \bar{\eta}_0 + \bar{\eta}_1 \text{AI Assistance}_i \times \text{Post}_t + \mu_i + \mu_{It} + \epsilon_{it}$$

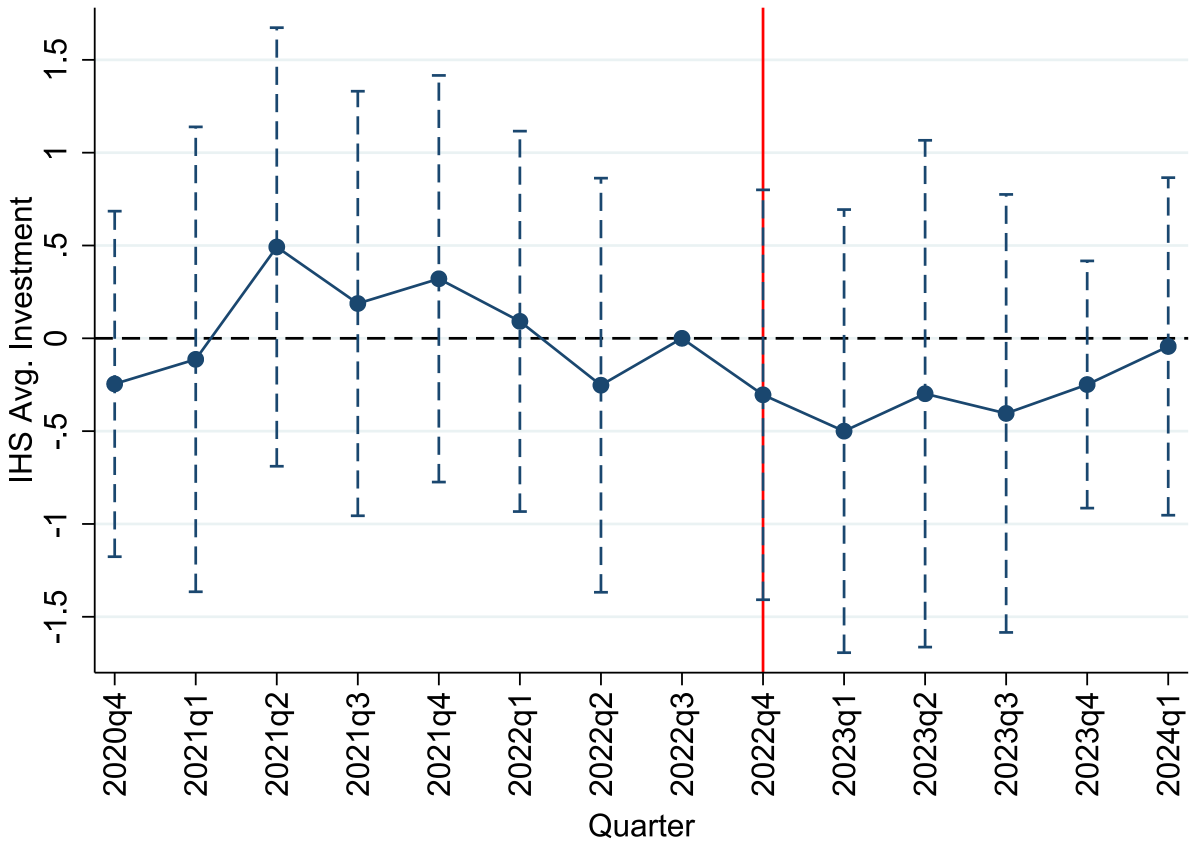
Column (3) reports the results for the estimation of the mediated effect of the treatment once we control for the average size of VCPE investments. The specification is

$$\text{IHS(N. New Firms)}_{it} = \bar{\theta}_0 + \bar{\theta}_1 \text{AI Assistance}_i \times \text{Post}_t + \bar{\theta}_2 \text{IHS(Avg. Investment)}_{it} + \mu_i + \mu_{It} + \epsilon_{it}$$

All specifications include both industry and sector-by-quarter fixed effects. Treatment intensity is proxied by AI Assistance in its standardized form. Treatment timing is proxied by the binary indicator Post, which is equal to 1 for all quarters after 2022Q3, and 0 otherwise. Standard errors are clustered at the three-digit NAICS industry level and reported in parentheses. Coefficients are statistically significant at the 10% (*), 5% (**), and 1% (***) levels.

To ensure comparability across differently treated industries, we run the event study version of the specification in Table OA11, Column (2). Table OA11 shows that neither anticipatory effects nor differences in pre-trends are much of a concern for average deal size, reinforcing the findings in Table OA11. Moreover, the post treatment effects appear to be weakly negative.

Figure OA11. Event Study of AI Assistance on Average Allocated Capital Around the Release of ChatGPT



Notes: The figure presents the event study for the sample of 81 three-digit NAICS industries in our analysis, over the time period 2020Q4–2024Q1. We plot the estimates from the dynamic difference-in-differences model for the interaction of the quarter-year dummies with the AI Assistance$_i$ from the specification

$$\text{IHS(Avg. Investment)}_{it} = \bar{\psi}_0 + \sum_{t \neq 2022q3} \bar{\psi}_t \text{AI Assistance}_i \times 1\{D = t\}_t + \text{Controls}_{it} + \mu_i + \mu_{It} + \epsilon_{it}$$

The dependent variable is the Inverse Hyperbolic Sine of the average size of VCPE investments, in thousands of dollars. Treatment intensity is proxied by AI Assistance in its standardized form. The specification includes both industry and sector-by-quarter fixed effects. Standard errors are clustered at the three-digit NAICS industry level.

Overall, the finding indicates that the efficiency channel—in the sense of reduced capital needs per firm leading to more entry—does not have much traction in the aggregate data. It remains possible that the channel operates for a subset of ventures, particularly those whose founders can directly exploit AI to substitute for costly inputs. However, when examining firm creation at the industry level, which aggregates across heterogeneous entrants, the effect does not appear clearly.

Appendix B. Robustness: Chen and Roth (2024)'s Critique

A potential concern with our baseline specification is the use of the inverse hyperbolic sine (IHS) transformation of the dependent variable, the number of newly created firms. Chen and Roth (2024) show that estimators based on log-like transformations such as IHS or $\log(1+y)$ can suffer from interpretational and robustness issues, particularly when the outcome variable can take the value zero. Although our outcome is strictly positive in the estimation sample—making the natural logarithm well defined—it is important to demonstrate that our findings do not hinge on the choice of functional form.

Since the dependent variables used in the main analysis (Table 4 and Figure 4) and in the cost of experimentation mechanism (Table 6 and Figure 5) are strictly positive, we can test whether our results are solely an artifact of the functional specification (IHS) chosen by replicating the results transforming the dependent variable using the natural logarithm.

In the capital reallocation test (Table 7 and Figure 6), we need to employ more nuanced techniques to test whether the results are driven the the IHS. We therefore conduct a series of robustness tests designed to directly address these concerns. Each test isolates a specific aspect of the Chen and Roth (2024) critique and shows that our substantive conclusions remain unchanged.

B.1. **Replicating Effects of AI on Firm Creation Results.** Table OA12 presents our baseline difference-in-differences estimates of the effect of AI Assistance on firm creation, replacing the inverse hyperbolic sine transformation with the natural logarithm of the number of new firms. We do this for two reasons. Since the number of newly created firms is strictly positive in our estimation sample, the logarithm is well defined and provides a clean percentage interpretation of the estimated coefficients. Therefore, the new specification directly address the critique raised by Chen and Roth (2024), who show that log-like transformations such as IHS or $\log(1+y)$ can yield estimators with problematic interpretation when zeros are present. By relying on Log(New Firms) rather than IHS, we demonstrate that our results are not driven by functional-form artifacts.

Across both specifications, with and without controls, the coefficient on *AI Assistance* × *Post* is stable at approximately 0.10 and statistically significant at the 1% level. The results are identical to the baseline estimates in Table 4.

Table OA12. Logarithmic Functional Form: Effect of AI Assistance on Firm Creation After the Release of ChatGPT

VARIABLES	(1) Log N. New Firms	(2) Log N. New Firms
AI Assistance × Post	0.105***	0.100***
	(0.025)	(0.025)
Observations	1,134	1,134
Industry 3D FE	Yes	Yes
Quarter FE	Yes	Yes
Industry 2D × Quarter FE	No	No
Controls	No	Yes
Adj. R. sq.	0.982	0.982

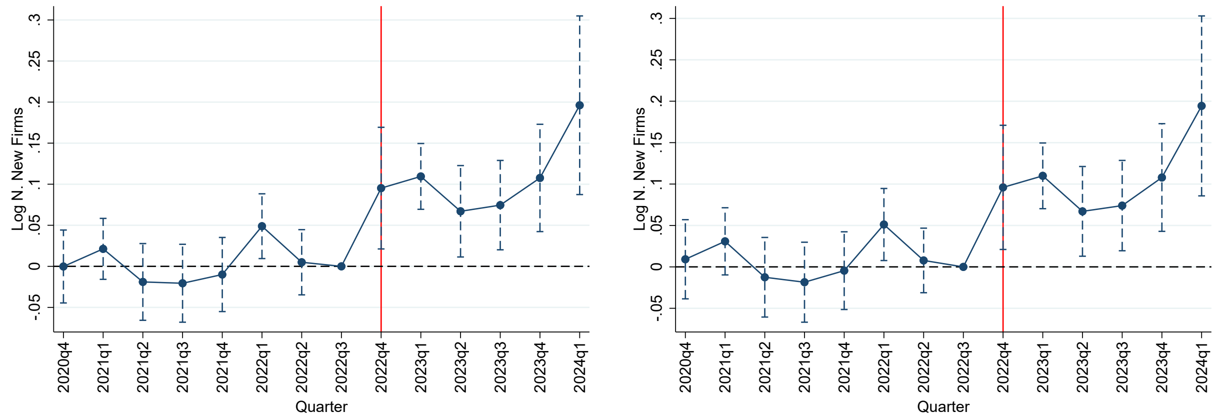
Notes: This table addresses the concerned raised in Chen and Roth (2024) regarding the use of the IHS transformation on the outcome variable. We re-run the estimation results in Table 4, now expressing the dependent variable in its logarithmic form. The design is a difference-in-differences model at the industry-by-quarter level with continuous treatment. The sample includes 81 three-digit NAICS industries over 2020Q4–2024Q1. The empirical specification is

$$\text{Log(N. New Firms)}_{it} = \alpha_0^* + \alpha_1^* \text{AI Assistance}_i \times \text{Post}_t + \text{Controls}_{it} + \mu_i + \mu_t + \epsilon_{it}$$

The dependent variable is the natural logarithm of the number of new firms created. Treatment intensity is proxied by AI Assistance in its standardized form. Treatment timing is proxied by the binary indicator Post, which equals 1 for all quarters after 2022Q3 and 0 otherwise. All specifications include industry and quarter fixed effects. Column (2) additionally includes the following controls: lag of VCPE Investments (IHS), median Return on Assets, median R&D Intensity, median CapEx Intensity, median Liquidity Ratio, and median Market Leverage. Standard errors are clustered at the three-digit NAICS industry level and reported in parentheses. Coefficients are statistically significant at the 10% (*), 5% (**), and 1% (***) levels.

Figure OA12 presents the dynamic difference-in-differences estimates of the effect of AI Assistance on firm creation, using Log(New Firms) as the dependent variable. The results are identical to the baseline event studies in Figure 4.

Figure OA12. Logarithmic Functional Form: Event Study of AI Assistance on Firm Creation around the Release of ChatGPT



(a) No Controls **(b)** Yes Controls

Notes: This panel of figures addresses the concerned raised in Chen and Roth (2024) regarding the use of the IHS transformation on the outcome variable. We re-run the estimation results in Figure 4, now expressing the dependent variable in its logarithmic form. The figures presents the event study for the sample of 81 three-digit NAICS industries in our analysis, over the time period 2020Q4–2024Q1. We plot the estimates from the dynamic difference-in-differences model for the interaction of the quarter-year dummies with the AI Assistance$_i$ from the specification

$$\text{Log(N. New Firms)}_{it} = \beta_0^* + \sum_{t \neq 2022q3} \beta_t^* \text{AI Assistance}_i \times 1\{D = t\}_t + \text{Controls}_{it} + \mu_i + \mu_t + \epsilon_{it}$$

The dependent variable is the natural logarithm new firms created. Treatment intensity is proxied by AI Assistance in its standardized form. The specifications include both industry and quarter fixed effects. Standard errors are clustered at the three-digit NAICS industry level. The specification in Panel (a) does not include controls. The specification in Panel (b) includes the following controls: lag of VCPE Investments (IHS), median Return on Assets, median R&D Intensity, median CapEx Intensity, median Liquidity Ratio, and median Market Leverage.

Overall, our findings support the idea that the main results of our analysis presented in Section 4 are not an artifact of the IHS specification.

B.2. **Replicating of Cost of Experimentation Mechanism Results.** The results on the reduction of the cost of experimentation mechanism (Subsection 5.1) are also subject to the critique presented in Chen and Roth (2024) regarding the use of the IHS transformation on the outcome variable. To address these concern we replicate the analysis using the natural logarithm of the dependent variable. We are able to do so because the dependent variable is strictly positive.

The results using the natural logarithm for the dependent variable are completely unvaried both in term of size and significance. The coefficient on the triple interaction variable for the specification not including control remain significant at the 5% level, with magnitude 0.26 (Table OA13), exactly as we find in Table 6, in the main text.

Table OA13. Logarithmic Functional Form: Effects of the Cost of Experimentation on Firm Creation

VARIABLES	(1) Log N. New Firms	(2) Log N. New Firms
AI Assistance × Experimentation Intense × Post	0.261**	0.264**
	(0.111)	(0.112)
AI Assistance × Post	0.037	0.032
	(0.054)	(0.054)
Experimentation Intense × Post	-0.190	-0.193
	(0.119)	(0.118)
Observations	1,078	1,078
Industry 3D FE	Yes	Yes
Quarter FE	No	No
Industry 2D × Quarter FE	Yes	Yes
Controls	No	Yes
Adj. R sq.	0.981	0.981

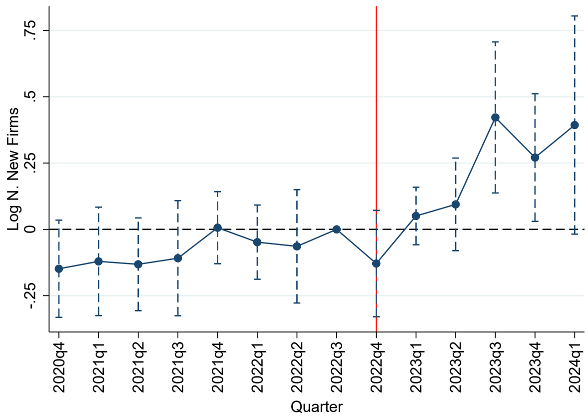
Notes: This table addresses the concerned raised in Chen and Roth (2024) regarding the use of the IHS transformation on the outcome variable. We re-run the estimation results in Table 6, now expressing the dependent variable in its logarithmic form. The design is a difference-in-differences model at the industry-by-quarter level with continuous treatment and a heterogeneity term. The sample includes 81 three-digit NAICS industries over 2020Q4–2024Q1. The empirical specification is

$$\begin{aligned}\text{Log(N. New Firms)}_{it} = \gamma_0^* &+ \gamma_1^* \text{AI Assistance}_i \times \text{Experimentation Intense}_i \times \text{Post}_t \\ &+ \gamma_2^* \text{AI Assistance}_i \times \text{Post}_t + \gamma_3 \text{Experimentation Intense}_i \times \text{Post}_t \\ &+ \text{Controls}_{it} + \mu_i + \mu_{It} + \varepsilon_{it}\end{aligned}$$

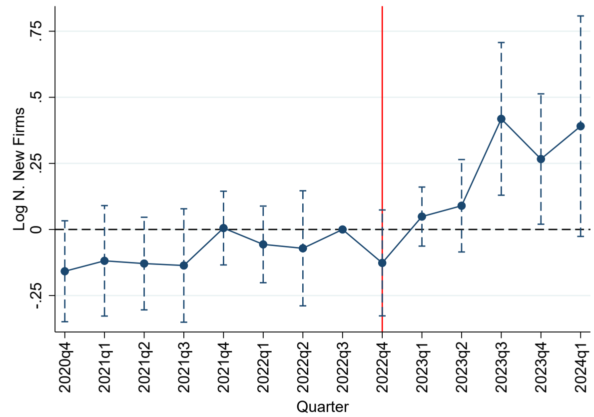
The dependent variable is the natural logarithm of the number of new firms created. Treatment intensity is proxied by AI Assistance in its standardized form. Treatment timing is proxied by Post, which equals 1 for all quarters after 2022Q3 and 0 otherwise. Treatment heterogeneity comes from the indicator Experimentation Intense, which equals 1 if the industry's Experimentation Intensity Score is in the top quartile and 0 otherwise. The specifications include industry and sector-by-quarter fixed effects. Column (2) additionally includes the following controls: lagged VCPE Investments (IHS), median Return on Assets, median R&D Intensity, median CapEx Intensity, median Liquidity Ratio, and median Market Leverage. Standard errors are clustered at the three-digit NAICS industry level and reported in parentheses. Coefficients are statistically significant at the 10% (*), 5% (**), and 1% (***) levels.

The event study results are also unchanged when we substitute the IHS form with the log one. Figure OA13 presents the results, showing that the coefficients in it display the same patterns observed in Figure 5.

Figure OA13. Logarithmic Functional Form: Event Study of Cost of Experimentation on Firm Creation around the Release of ChatGPT



(a) No Controls



(b) Yes Controls

Notes: This panel of figures addresses the concerned raised in Chen and Roth (2024) regarding the use of the IHS transformation on the outcome variable. We re-run the estimation results in Figure 5, now expressing the dependent variable in its logarithmic form. The figure presents the event study for the sample of 81 three-digit NAICS industries in our analysis, over the time period 2020Q4–2024Q1. We plot the estimates from the dynamic difference-in-differences model for the interaction of the quarter-year dummies with the AI Assistance$_i$ × Experimentation Intensity$_i$ from the specification

$$\begin{aligned}\text{Log(N. New Firms)}_{it} = \delta_0 + &\sum_{t\neq 2022q3} \delta_t^1 \text{AI Assistance}_i \times \text{Experimentation Intense}_i \times 1\{D=t\}_t \\ &+ \sum_{t\neq 2022q3} \delta_t^2 \text{AI Assistance}_i \times 1\{D=t\}_t + \sum_{t\neq 2022q3} \delta_t^3 \text{Experimentation Intense}_i \times 1\{D=t\}_t \\ &+ \text{Controls}_{it} + \mu_i + \mu_{It} + \varepsilon_{it}\end{aligned}$$

The dependent variable is the natural logarithm of the number of new firms created. Treatment intensity is proxied by AI Assistance in its standardized form. Treatment heterogeneity comes from the indicator variable Experimentation Intense, which equal 1 if the industry reports a value of Experimentation Intensity within the top quartile and 0 otherwise. The specifications include both industry and sector-by-quarter fixed effects. Standard errors are clustered at the three-digit NAICS industry level. The specification in Panel (a) does not include controls. The specification in Panel (b) includes the following controls: lag of VCPE Investments (IHS), median Return on Assets, median R&D Intensity, median CapEx Intensity, median Liquidity Ratio, and median Market Leverage.

Overall, our findings support the idea that the main results of our analysis presented in Subsection 5.1 are not an artifact of the IHS specification.

B.3. **Testing Capital Reallocation Mechanism Results.** A distinctive challenge in the capital reallocation tests is that the mediator, total VCPE investments, occasionally takes the value zero. This creates difficulties for conventional logarithmic specifications: log(0) is undefined, and ad hoc fixes such as $\log(1 + y)$ or IHS can blur the distinction between the extensive margin (whether there is any investment at all) and the intensive margin (the size of investment when positive). Chen and Roth (2024) show that such transformations can lead to estimators with problematic interpretations, particularly when zeros are common. For this reason, the tests of capital reallocation require more nuanced approaches than those for firm creation or experimentation, where the dependent variables are strictly positive.

We therefore conduct three complementary robustness exercises, each designed to address a specific aspect of the Chen and Roth (2024) critique. First, we check whether results hinge on using IHS versus the natural log, ensuring clean percentage interpretations when outcomes are strictly positive. Second, we employ a PPML estimation, which handles the presence of zeros in investment flows and avoids arbitrary log-like transformations by modeling the conditional mean directly. Lastly, we use a two-part (hurdle) model, which separates extensive (any investment) and intensive (log size when positive) margins, addressing the concern that log/IHS transforms conflate them.

B.3.1. *Log Specification Estimation.* First, we reproduce the mediation table replacing IHS transformations with natural logarithms (Table OA14). This specification is natural for strictly positive variables, such as the number of new firms in our sample, and allows a clean percentage interpretation of coefficients. If our findings were an artifact of using IHS, results would change materially under the log specification, due to dropping the zero-valued industry-quarter cells. Instead, we find that the coefficient on *AI Assistance* $\times$ *Post* remains around 0.10 and statistically significant, very similar to the IHS baseline (Table 7, Column (1)). In the investment regressions (Table OA14, Column (2)), the effect of AI compatibility remains negative and statistically significant, with size comparable to what we find in Table 7 in the main text. Moreover, when looking at Column (3) in Table OA14, not only we find that the coefficient on *AI Assistance* $\times$ *Post* remains around 0.10 and statistically significant, but the mediator continues to be uninformative. These results are almost identical to what we found in the main text (Table 7). The above confirms that our main results are robust to abandoning IHS and using the theoretically preferred logarithmic specification.

Table OA14. Logarithmic Functional Form: Capital Reallocation Mechanism

VARIABLES	(1) Log N. New Firms	(2) Log Total Investments	(3) Log N. New Firms
AI Assistance × Post	0.102**	-0.493	0.100*
	(0.046)	(0.350)	(0.053)
Log Total Investments			0.004
			(0.006)
Observations	1,078	811	811
Industry 3D FE	Yes	Yes	Yes
Quarter FE	No	No	No
Industry 2D × Quarter FE	Yes	Yes	Yes
Controls	No	No	No
Adj. R sq.	0.980	0.559	0.982

Notes: This table reports the robustness tests results for the mediation-style decomposition to study capital reallocation as a potential mechanism linking AI's arrival and firm creation. To address the critique presented Chen and Roth (2024), we replicated the analysis using the natural logarithm instead of the inverse hyperbolic sine transformation. The analysis is at the industry-by-quarter level for 81 three-digit NAICS industries over 2020Q4–2024Q1. The mediator is total capital allocated, measured as the natural logarithm of total VCPE investments from Preqin (in $000).

Column (1) reports the results for the estimation of the total impact that the treatment has on firm formation. The specification is

$$\text{Log(N. New Firms)}_{it} = \zeta_0^1 + \zeta_1^1 \text{AI Assistance}_i \times \text{Post}_t + \mu_i + \mu_{It} + \epsilon_{it}$$

Column (2) assesses the linkage between the treatment and the total capital allocation, according to the specification

$$\text{Log(Investments)}_{it} = \eta_0^1 + \eta_1^1 \text{AI Assistance}_i \times \text{Post}_t + \mu_i + \mu_{It} + \epsilon_{it}$$

Column (3) reports the results for the estimation of the mediated effect of the treatment once we control for total allocated capital. The specification is

$$\text{Log(N. New Firms)}_{it} = \theta_0^1 + \theta_1^1 \text{AI Assistance}_i \times \text{Post}_t + \theta_2^1 \text{Log(Investments)}_{it} + \mu_i + \mu_{It} + \epsilon_{it}$$

All specifications include both industry and sector-by-quarter fixed effects. Treatment intensity is proxied by AI Assistance in its standardized form. Treatment timing is proxied by the binary indicator Post, which is equal to 1 for all quarters after 2022Q3, and 0 otherwise. Standard errors are clustered at the three-digit NAICS industry level and reported in parentheses. Coefficients are statistically significant at the 10% (*), 5% (**), and 1% (***) levels.

B.3.2. *Estimating Investments with PPML.* To further address the short-comings of using dependent variables in their IHS form, we follow the recommendation in Chen and Roth (2024) to use the Poisson pseudo-maximum likelihood (PPML) estimator, which avoids outcome transformations altogether and is well defined in the presence of zeros. While our firm creation outcome never equals zero, investments occasionally do, making PPML a natural robustness check for the mediator equation. We therefore re-estimate the investment regression using PPML, while keeping Columns (1) and (3) in IHS form for comparability. The results are strikingly similar: the estimated coefficient on *AI Assistance* × *Post* on total investments remains negative (around -0.33) and statistically significant, in line with the results in Table 7. This exercise demonstrates that our conclusion about the irrelevance of capital reallocation does not hinge on log-like transformations of the investment variable.

Table OA15. Testing the Capital Reallocation Channel with PPML

VARIABLES	(1) IHS N. New Firms	(2) Total Investments	(3) IHS N. New Firms
AI Assistance × Post	0.102**	-0.329**	0.102**
	(0.046)	(0.165)	(0.046)
IHS Total Investments			0.000
			(0.004)
Observations	1,078	1,064	1,078
Industry 3D FE	Yes	Yes	Yes
Quarter FE	No	No	No
Industry 2D × Quarter FE	Yes	Yes	Yes
Controls	No	No	No
Adj. R sq.	0.980	0.774	0.980

Notes: This table reports the robustness tests results for the mediation-style decomposition to study capital reallocation as a potential mechanism linking AI's arrival and firm creation. To address the critique presented Chen and Roth (2024), we replicate the linkage between the treatment and total capital allocation using Poisson pseudo-maximum likelihood. The rest of the analysis is unchanged from the baseline estimation. The analysis is at the industry-by-quarter level for 81 three-digit NAICS industries over 2020Q4–2024Q1. The mediator is total capital allocated, measured as the natural logarithm of total VCPE investments from Preqin (in $000).
Column (1) reports the results for the estimation of the total impact that the treatment has on firm formation. The specification is

$$\text{Log(N. New Firms)}_{it} = \zeta_0^2 + \zeta_1^2 \text{AI Assistance}_i \times \text{Post}_t + \mu_i + \mu_{It} + \epsilon_{it}$$

Column (2) assesses the linkage between the treatment and total capital allocation. The dependent variable now is total investments, in levels, estimated using Poisson pseudo-maximum likelihood according to the specification

$$E[\text{Investments}_{it} \mid \text{AI Assistance}_i, \text{Post}_t, \mu_i, \mu_{It}] = \exp\Big(\nu_0^2 + \nu_1^2 \text{AI Assistance}_i \times \text{Post}_t + \mu_i + \mu_{It}\Big)$$

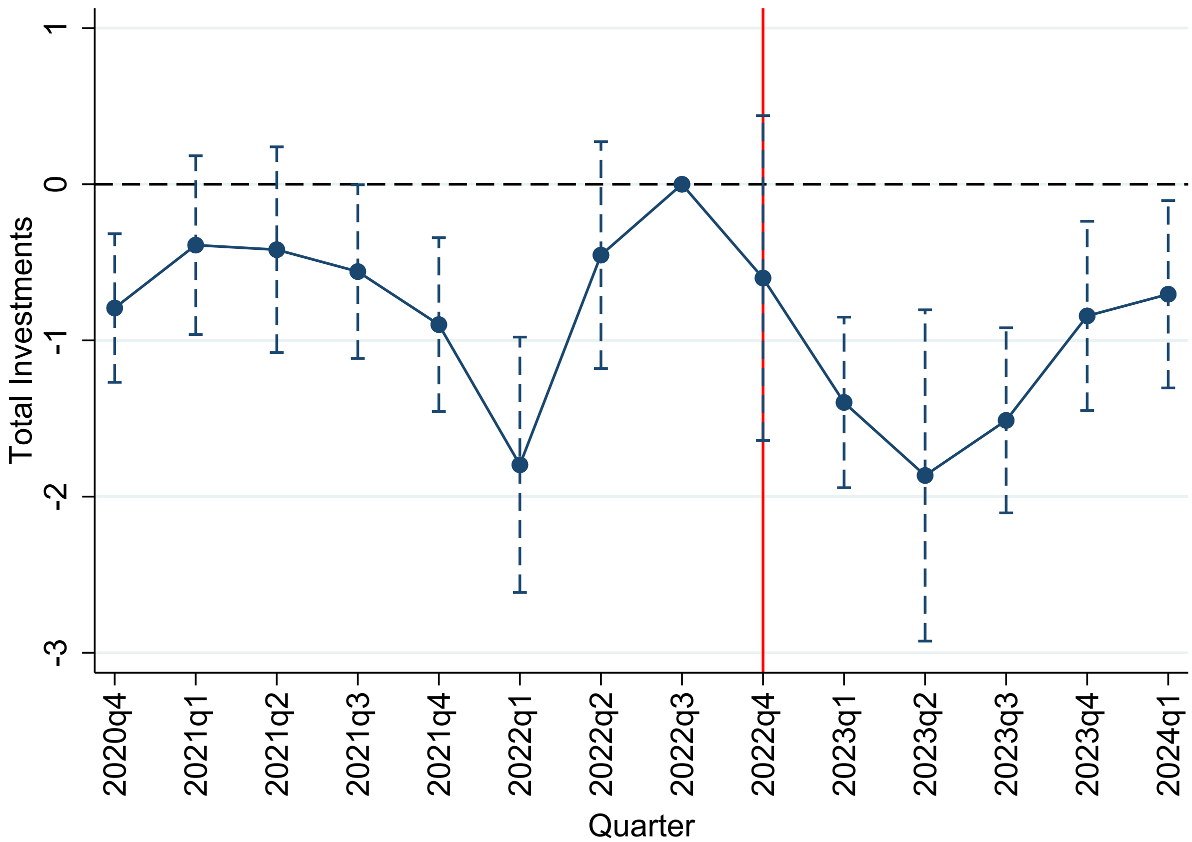
Column (3) reports the results for the estimation of the mediated effect of the treatment once we control for total allocated capital. The specification is

$$\text{IHS(N. New Firms)}_{it} = \theta_0^2 + \theta_1^2 \text{AI Assistance}_i \times \text{Post}_t + \theta_2^2 \text{IHS(Investments)}_{it} + \mu_i + \mu_{It} + \epsilon_{it}$$

All specifications include both industry and sector-by-quarter fixed effects. Treatment intensity is proxied by AI Assistance in its standardized form. Treatment timing is proxied by the binary indicator Post, which is equal to 1 for all quarters after 2022Q3, and 0 otherwise. Standard errors are clustered at the three-digit NAICS industry level and reported in parentheses. Coefficients are statistically significant at the 10% (*), 5% (**), and 1% (***) levels.

We also replicate the event study for total investments using PPML, to contrast it with the results in Figure 6. Figure OA14 presents the event study estimates for the model in Column (2) Table OA15. The results show no evidence of anticipatory investment: pre-treatment coefficients are generally statistically indistinguishable from zero. Following the release of ChatGPT, the coefficients become negative and significant, indicating that industries more more compatible with AI experienced a relative decline in VCPE investment flows. This dynamic pattern, not only corroborates the static PPML results in Table OA15, but is also largely in line with the results in the main text showing that capital does not flow towards more AI-compatible industries.

Figure OA14. Dynamic Effects of AI Assistance on Investments: PPML Event Study



Notes: The figure presents the event study for regression of total VCPE investments estimated using PPML. The sample is comprised of 81 three-digit NAICS industries, over the time period 2020Q4–2024Q1. We plot the estimates from the dynamic difference-in-differences model for the interaction of the quarter-year dummies with the AI Assistance$_i$ from the specification

$$E\left[\text{Investments}_{it}\,\middle|\,\text{AI Assistance}_i, \sum_{t\neq 2022q3} 1\{D=t\}_t, \mu_i, \mu_{It}\right] = \exp\Big(\psi_0^1 + \sum_{t\neq 2022q3} \psi_t^1 \text{AI Assistance}_i \times 1\{D=t\}_t + \mu_i + \mu_{It}\Big)$$

The dependent variable is the the dollar amount of VCPE investments, in thousands. Treatment intensity is proxied by AI Assistance in its standardized form. The specification includes both industry and sector-by-quarter fixed effects. Standard errors are clustered at the three-digit NAICS industry level.

B.3.3. *Two-Part Hurdle Specification: Extensive Versus Intensive Margin.* A key critique in Chen and Roth (2024) is that log or IHS transformations implicitly conflate the extensive margin (whether investment is zero or positive) with the intensive margin (the size of investment conditional on being positive). To address this issue. we explicitly decompose investments into extensive and intensive margins. To disentangle these effects, we estimate a two-part (hurdle) specification that includes (a) a dummy for any positive investment and (b) the log of investment conditional on being positive. This allows us to test whether either margin explains firm creation. The results in Table OA16 show that neither component is statistically significant: the extensive margin dummy is small and negative, and the intensive margin coefficient is close to zero. Importantly, the coefficient on *AI Assistance* × *Post* remains stable and significant at around 0.10, as in the baseline specification reported in the main text (Table 7, Column (3)). This confirms that our treatment effect is not mediated by either margin of investment activity.

Table OA16. Decomposing the Capital Reallocation Channel: Extensive and Intensive Margins of Investment

VARIABLES	(1) Log N. New Firms
AI Assistance × Post	0.106**
	(0.046)
Extensive Margin	-0.052
	(0.038)
Intensive Margin	0.006
	(0.005)
Observations	1,078
Industry 3D FE	Yes
Quarter FE	No
Industry 2D × Quarter FE	Yes
Controls	No
Adj. R. sq.	0.980

Notes: This table tests whether the effect of AI Assistance on new firm creation is mediated through the extensive or intensive margins of investment, following the indication of Chen and Roth (2024). The analysis is at the industry-by-quarter level for 81 three-digit NAICS industries over 2020Q4–2024Q1. The specification is

$$\text{Log(N. New Firms)}_{it} = \upsilon_0 + \upsilon_1 \text{AI Assistance}_i \times \text{Post}_t + \upsilon_2 \text{Extensive Margin}_{it} + \upsilon_3 \text{Intensive Margin}_{it} + \mu_i + \mu_{It} + \epsilon_{it}$$

The dependent variable is the natural logarithm of the number of new firms created. Treatment intensity is proxied by AI Assistance in standardized form, and treatment timing is proxied by Post, which equals 1 for all quarters after 2022Q3 and 0 otherwise. Investments are decomposed into two components: (i) an Extensive Margin indicator equal to 1 if total investments in the industry-quarter cell are positive and 0 otherwise, and (ii) an Intensive Margin measure equal to the log of total investment, conditional on being positive. The specification includes industry and sector-by-quarter fixed effects. Standard errors are clustered at the three-digit NAICS industry level and reported in parentheses. Coefficients are statistically significant at the 10% (*), 5% (**), and 1% (***) levels.

Appendix C. Additional Information on Revelio Labs

C.1. **Sourcing of Revelio Labs Data.** Revelio Labs is a workforce intelligence provider whose primary data source is public information drawn from LinkedIn. By systematically scraping and parsing professional profiles and job postings, the company builds a structured dataset that captures employment histories, job descriptions, and firm characteristics. Although LinkedIn is the main input, Revelio further integrates auxiliary data streams such as employee reviews and official layoff notices, but these are secondary relative to the core profile-based data.

A central feature of Revelio Labs' methodology is the standardization and homogenization of raw information. Professional profiles and postings often contain highly heterogeneous job titles, firm names, and skill descriptions. To address this, Revelio applies proprietary taxonomies and entity-resolution techniques that map companies to unique identifiers, collapse synonymous or overlapping job titles into coherent occupational categories, and cluster skills into standardized hierarchies. This harmonization makes the data usable for systematic analysis at scale.

The structured database that results can be analyzed at several levels of aggregation. At the company level, Revelio provides measures such as total headcount, inflows and outflows of employees, hiring rates, and job posting dynamics. At the position level, the data track individual job postings and role descriptions, including salary information and required skills, deduplicated across multiple sources. At the individual level, the dataset captures workers' career histories, including transitions across employers, tenure, and educational backgrounds.

Because the underlying source is LinkedIn, coverage is not uniform across occupations, industries, and geographies. White-collar and professional workers are more likely to maintain profiles than blue-collar workers, for example. To correct for this selection bias, Revelio Labs constructs and distributes sampling weights at the occupation level. These weights are designed to re-weight observed profiles so that the aggregate distributions more closely match official labor force statistics. For example, if a particular occupation is less likely to appear in online profiles, each observed instance may be weighted to represent multiple actual workers in the population. This sampling adjustment ensures the dataset reflects a more accurate workforce composition. This adjustment enhances the representativeness of the dataset and supports credible empirical analysis.

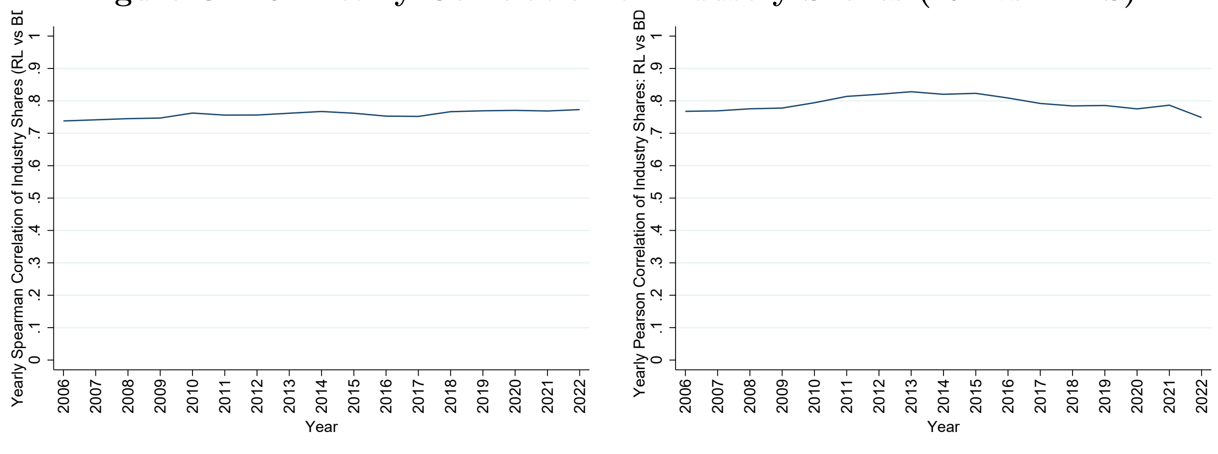
C.2. **Representativeness of Revelio Labs Data.** As discussed in the main paper (Section 2), the *Revelio Labs* (RL) dataset provides an alternative source of information on new firm creation, which we compare here to the official Business Dynamics Statistics (BDS) from the U.S. Census[16]. Due to the fact that BDS is available exclusively at the yearly level and it is reported with a lag, our validation exercise focuses on yearly firm formation data from 2006-2022[17]. Overall, the evidence in this appendix indicates that RL is broadly representative of the general economy in terms of new firm formation. In particular, RL data closely tracks the patterns and distribution of firm births observed in BDS. We emphasize that our representativeness claims pertain to firm creation statistics *specifically* — RL may not be as representative for other economic measures, but it proves to be an acceptably reliable proxy for studying new firm formation. Finally, we note that while the BDS dataset occasionally suppresses firm creation counts in industry-year cells where the true number is very small (to protect confidentiality), this occurs infrequently in our validation exercise, which uses a macro-level aggregation (3-digit NAICS by year) where such censoring is rare and inconsequential for our comparisons.

C.2.1. *Industry Distribution of New Firms in RL vs. BDS.* We first examine whether RL captures a similar *industry composition* of new firms as in the official data. We plot the Spearman and Pearson coefficients over time in Figure OA15. The plots display the correlation between the industry share of new firms in RL and the corresponding industry share in BDS. Figure OA15, Panel (a) uses Spearman's rank correlation (measuring similarity in the rank-order of industry importance), while Figure OA15, Panel (b) uses Pearson's correlation (measuring linear alignment of the share percentages). In both cases, the yearly correlations are very high and have improved over time. In the mid-2000s (when digital coverage was more limited), the correlations were moderately high; by the 2010s and beyond, they approach or exceed 0.8. This indicates that in each year RL data produces a very similar distribution of new firms across industries as the economy-wide distribution captured by BDS. In other words, the types of businesses being started (by industry) in RL match those in the broader economy. This alignment gives confidence that RL is not missing whole categories of firms: new firms in manufacturing, services, tech, etc., all appear in roughly the correct proportions.

[16]BDS is the benchmark source for annual counts of new employer firms in the United States. RL data, derived from online workforce records, offers more granular and timely information; this appendix evaluates how well RL mirrors BDS in the context of firm creation.

[17]While most of the analysis in the main paper is done for the period 2020Q4-2024Q1, we are forced to change the time aggregation and window in order to draw any sort of meaningful comparison.

Figure OA15. Yearly Correlation of Industry Shares (RL vs. BDS)



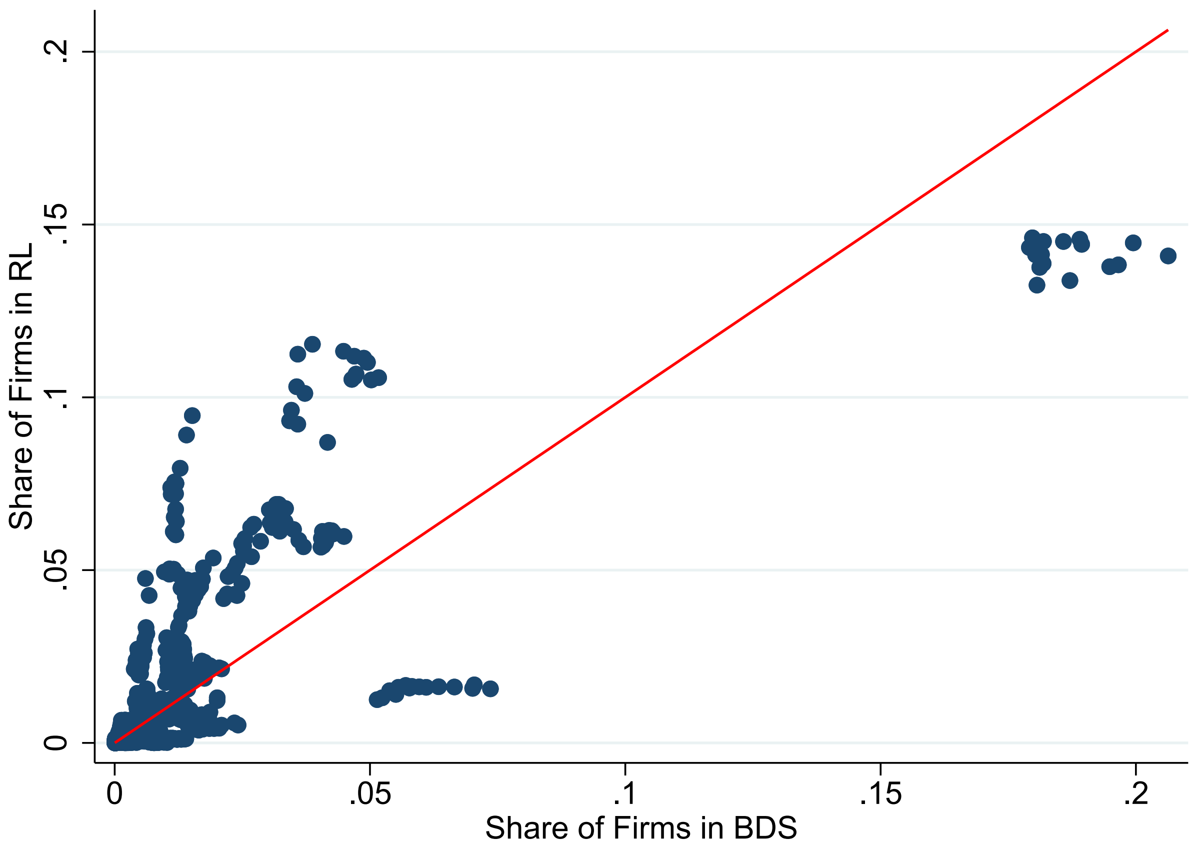
(a) Spearman **(b)** Pearson

Notes: This panel of plots displays the Spearman (Panel (a)) and Pearson (Panel (b)) correlation measures between the industry share vectors of new firms in RL and BDS for each year (2006–2022). Each point on the curve reflects how closely the fraction of new firms in each industry, as recorded by RL, aligns with the corresponding fractions from BDS in that year.

Another way to assess industry coverage is to compare the share of new firms by industry directly between datasets. Figure OA16 presents a scatter plot of industry shares in BDS versus RL. Industries are plotted for the period 2008 to 2022. Each point represents a single industry, plotting its percentage share of total new firms in BDS on the horizontal axis against the same share in RL on the vertical axis. The points cluster tightly around the 45-degree line, confirming that the two sources give very similar prominence to each industry. Industries that account for a large share of new firm births in the official data also tend to account for a large share in RL, and vice versa. For instance, if an industry constitutes about 10% of all new firms according to BDS, RL typically records a very similar share for that industry.

Most industries lie close to this line, indicating that their share of new firm activity is nearly the same in RL as in BDS. Industries above the line are represented slightly more in RL than in BDS, whereas industries below the line have a somewhat smaller share in RL. For example, RL shows a somewhat higher share of new firms in certain professional/technical sectors (above the line) and a slightly lower share in some traditional sectors like construction or food services (below the line). Overall, however, the deviations are minor and the high density of points near the diagonal demonstrates RL's strong fidelity to the true industry distribution of new firms.

Figure OA16. Industry Share of New Firms: RL vs. BDS



Notes: This scatter plot compares the distribution of new firms across industries in RL and BDS. Each dot is an industry-year observation (using a consistent industry classification in both datasets), with the horizontal coordinate being that industry's share of total new firms in BDS and the vertical coordinate its share in RL. The dashed 45-degree line represents equality.

While the overall industry distribution in RL is close to that of BDS, there are some intuitive differences in the relative shares of specific sectors. Table OA17 lists the top 10 industries (by share of total new firms) in the RL data and compares them with the top 10 in BDS. We see that many of the same industries dominate new firm formation in both datasets (indeed, six industries appear in the top ten of both RL and BDS). In particular, Professional, Scientific & Technical Services (NAICS 541) is the largest category of new firms in both datasets; however, RL attributes an even larger fraction of startups to this sector (about 24.8% of all new firms) compared to BDS (15.3%). This suggests that RL may over-sample new firms in highly knowledge-oriented or white-collar sectors, likely because those firms (and their employees) have a greater digital footprint. Conversely, some industries that rank highly in the official data have a somewhat smaller presence in RL's top ranks. For example, NAICS 722 (Food Services & Drinking Places, which includes restaurants and bars) accounts for around 11.3% of new firms in BDS (making it the second-largest category in BDS), but constitutes only about 5.8% of new firms in RL. Similarly, construction-related industries (e.g., NAICS 236 and 238, various construction contractors) and certain local service industries (e.g., Repair & Maintenance, NAICS 811) feature prominently in BDS's startup activity but are less represented in RL's top ten. These differences are understandable: new restaurants, construction firms, and repair shops are often smaller and less likely to have employees with online professional profiles, meaning RL is more prone to miss them. Nonetheless, RL still captures many new firms in these sectors (they do appear in RL, just at slightly lower ranks), and importantly, the broad composition is aligned. The rank-order correlation of industry importance between RL and BDS remains extremely high (as seen in Figure OA15), even if the exact shares differ a slightly for certain industries.

Table OA17. Top 10 Industries by New Firm Share: RL vs. BDS

Spot	Revelio Labs	Business Dynamics Statistics
1	541 (24.751%)	541 (15.283%)
2	611 (8.099%)	722 (11.250%)
3	722 (5.783%)	238 (7.878%)
4	561 (5.518%)	621 (6.962%)
5	621 (4.276%)	561 (6.459%)
6	531 (3.271%)	531 (5.696%)
7	522 (2.466%)	236 (4.912%)
8	512 (2.238%)	812 (4.454%)
9	423 (2.103%)	484 (3.111%)
10	812 (2.000%)	811 (3.072%)

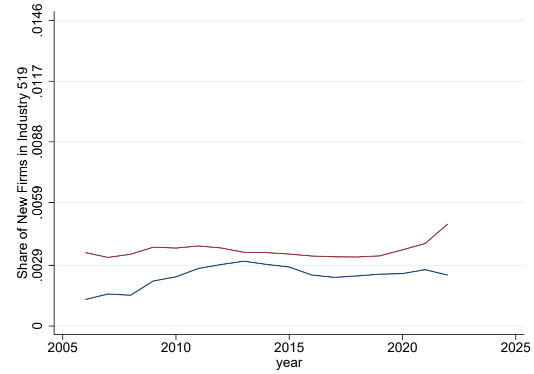
Notes: This table compares the leading industries for new firm formation in RL and BDS. For each rank (1 = highest share of new firms), we report the NAICS industry code and its percentage share of total new firms in that dataset.

To elaborate further, we also verify that RL adequately represents the industries that are particularly relevant to our study on AI and entrepreneurship. The main paper focuses on the impact of AI, which may vary by industry, so it is important that RL provides good coverage in both highly AI-exposed sectors and less AI-exposed ones. Figure OA17 and Figure OA18 address this by comparing RL and BDS in the subsets of industries with the highest and lowest *AI Assistance* scores, respectively (these scores, defined in the main text, measure the potential for AI to assist with tasks in each industry). Figure OA17 looks at the top industries (those most likely to benefit from AI). We focus on the top industries in this AI-relevance ranking to ensure that RL appropriately captures new firm activity in the sectors most pertinent to our analysis of AI spillovers. The comparison reveals that RL and BDS assign very similar shares to each high-AI industry. For example, if an AI-intensive industry constitutes $X\%$ of all new firms in BDS, the corresponding share in RL is also around $X\%$, with only minor deviations. This indicates that even in the technologically advanced sectors where one might particularly want to track entrepreneurship (due to AI's prominence), RL provides a representative sample of new firm creation.

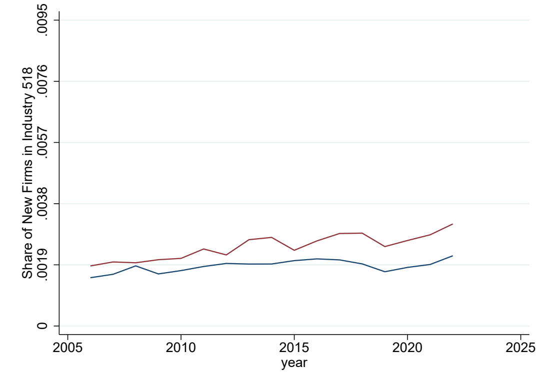
Figure OA18 covers the bottom industries (those least likely to be affected by AI in the near term). Generally, the shares are comparable, indicating that RL does capture new firm formation even in sectors that are not technology-intensive. There are some modest gaps for certain industries in this group: for instance, the Food Services industry (a low-AI sector) constitutes a smaller share of new firms in RL than in BDS (consistent with Table OA17, where NAICS 722 is underrepresented in RL relative to BDS). Likewise, some construction and manual-service industries show lower shares in RL. These differences reflect the limitations of RL in covering very traditional sectors, but importantly they are not large enough to change the overall picture. Even for industries where AI plays little role, RL data still registers significant numbers of new firms and tracks the general distribution of new firm activity seen in the official data. This suggests that our use of RL data is not omitting entire classes of low-tech entrepreneurship.

All in all, in both groups, RL's representation of new firms closely matches that of BDS.

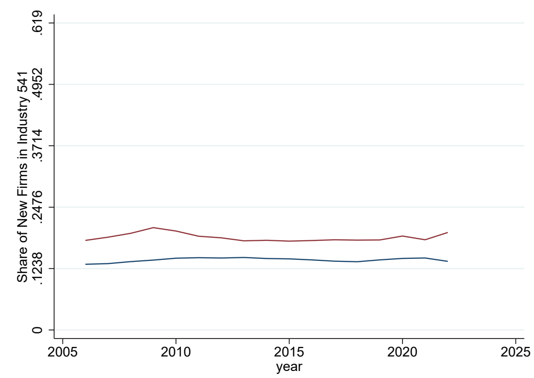
Figure OA17. New Firm Shares in Top AI Assistance Industries



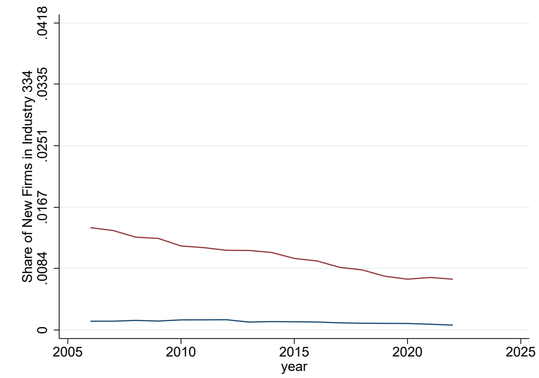
(a) Web Search Portals, Libraries, Archives, and Other Information Services



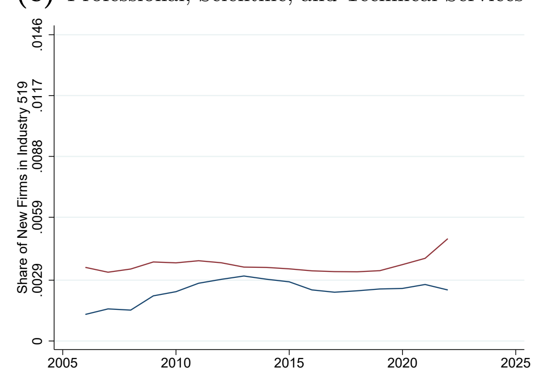
(b) Computing Infrastructure Providers, Data Processing, Web Hosting, and Related Services



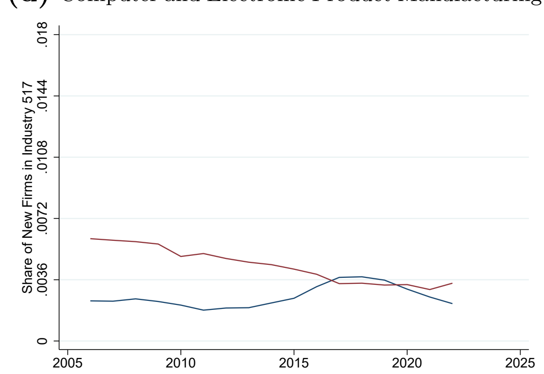
(c) Professional, Scientific, and Technical Services



(d) Computer and Electronic Product Manufacturing



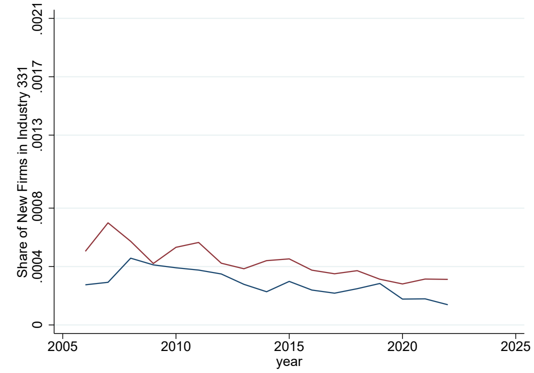
(e) Broadcasting and Content Providers



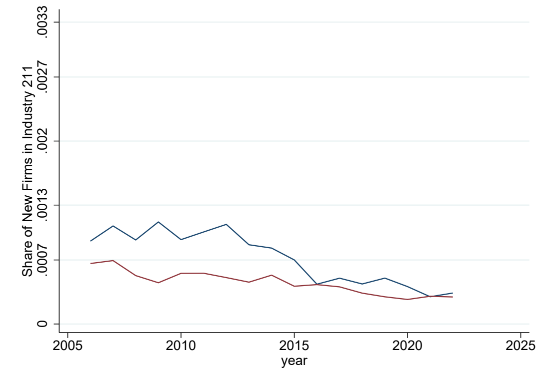
(f) Telecommunications

Notes: This figure considers the set of industries with the highest AI-assistance scores (i.e., industries where a large portion of occupational tasks can potentially be augmented by AI, as defined in the main paper). It shows the share of total new firms accounted for by each of these AI-intensive industries, as recorded in BDS (blue bars) and in RL (red bars).

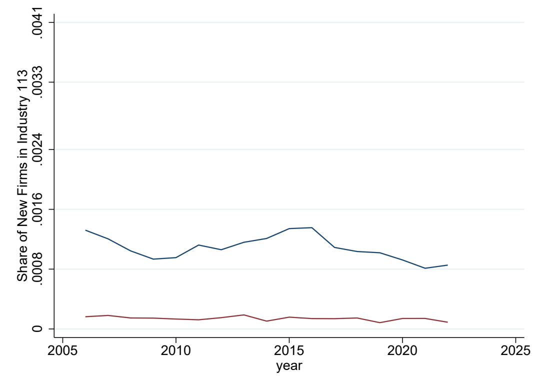
Figure OA18. New Firm Shares in Bottom AI Assistance Industries



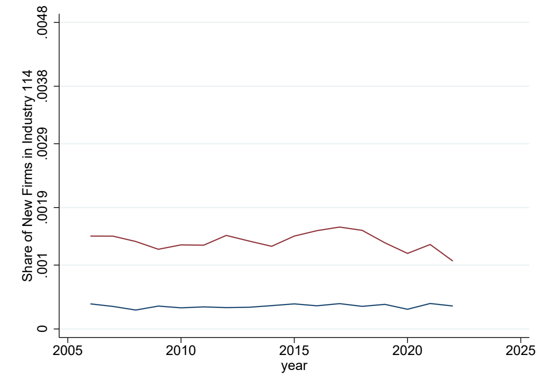
(a) Primary Metal Manufacturing



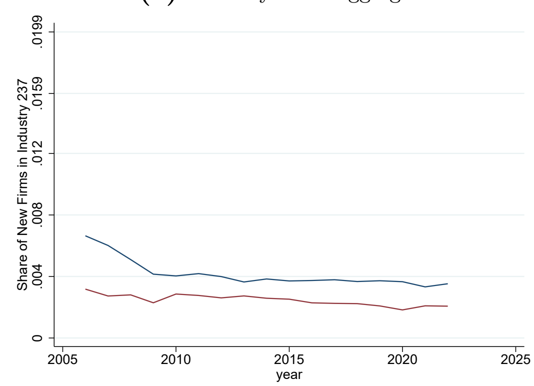
(b) Oil and Gas Extraction



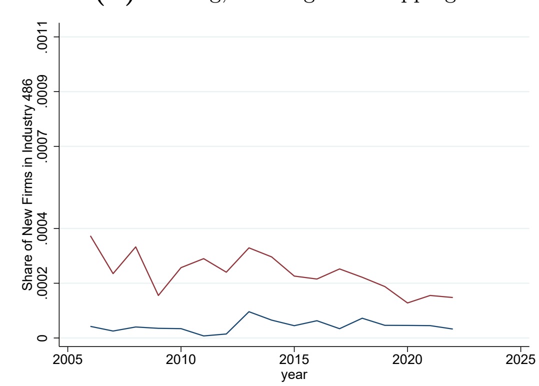
(c) Forestry and Logging



(d) Fishing, Hunting and Trapping



(e) Heavy and Civil Engineering Construction



(f) Pipeline Transportation

Notes: This figure examines industries with the lowest AI-assistance scores (industries where AI is expected to have minimal immediate impact on tasks). We compare the share of new firms in these low-AI industries between BDS (blue) and RL (red).

C.2.2. *New Firm Counts and Trends: RL vs. BDS.* Beyond industry composition, we assess whether RL can replicate the overall volume and trends of new firm creation found in official data. In the main text, we highlighted that RL data can be used to analyze year-to-year changes in startup activity. Here we confirm that the ups and downs in new firm counts observed in BDS are reflected in RL as well. Figure OA19 presents the correlation between RL and BDS in terms of the *number of new firms by industry*, year by year. In these plots, we take each industry's count of new firms in a given year and measure how well RL's counts align with BDS's. The patterns are similar to the share-based correlations shown earlier. The results indicate that RL is capturing not just the proportions but the actual magnitudes of variation in new firm births across industries.

Figure OA19. Yearly Correlation of New Firm Counts by Industry (RL vs. BDS)

(a) Spearman **(b)** Pearson

Notes: This panel of plots displays the Spearman (Panel (a)) and Pearson (Panel (b)) correlation measures between the industry new firm count vectors of new firms in RL and BDS for each year (2006–2022). Each point on the curve reflects how closely the number of new firms in each industry, as recorded by RL, aligns with the corresponding count from BDS in that year.

To quantify the relationship between RL and BDS new firm counts more directly, we regress the RL counts on the BDS counts. Table OA18 reports the results of this exercise, using a panel of industry-by-year observations. Column(1) shows a simple regression of log(New Firms RL) on log(New Firms BDS) (with no fixed effects). The coefficient is about 0.57 and highly significant, implying that a 10% increase in the number of new firms in the official data is associated with roughly a 5.7% increase in the number of new firms captured by RL. This coefficient being below 1.0 is expected, since RL does not capture every new firm—for example, if BDS counts double (often driven by many very small firms), RL counts will increase but by a somewhat smaller proportion, reflecting its incomplete coverage of the smallest or least visible startups. In Column (2), we add industry and year fixed effects to account for any systematic level differences across industries and trends over time. With these controls, the coefficient falls to about 0.21, suggesting that within a given industry and year, deviations in RL from the average are still positively correlated with deviations in BDS, but much of the variance is absorbed by fixed effects. The very high R^2 (0.992) in Column (2) indicates that once we control for industry and year, RL's counts line up almost perfectly with BDS for explaining the remaining variation. In practical terms, this means that apart from a stable industry-specific under- or over-counting factor, RL captures fluctuations in new firm creation very reliably. This is reassuring information, as all specifications employed in the main text leverage both industry and time fixed effects, which will absorb these immutable discrepancies.

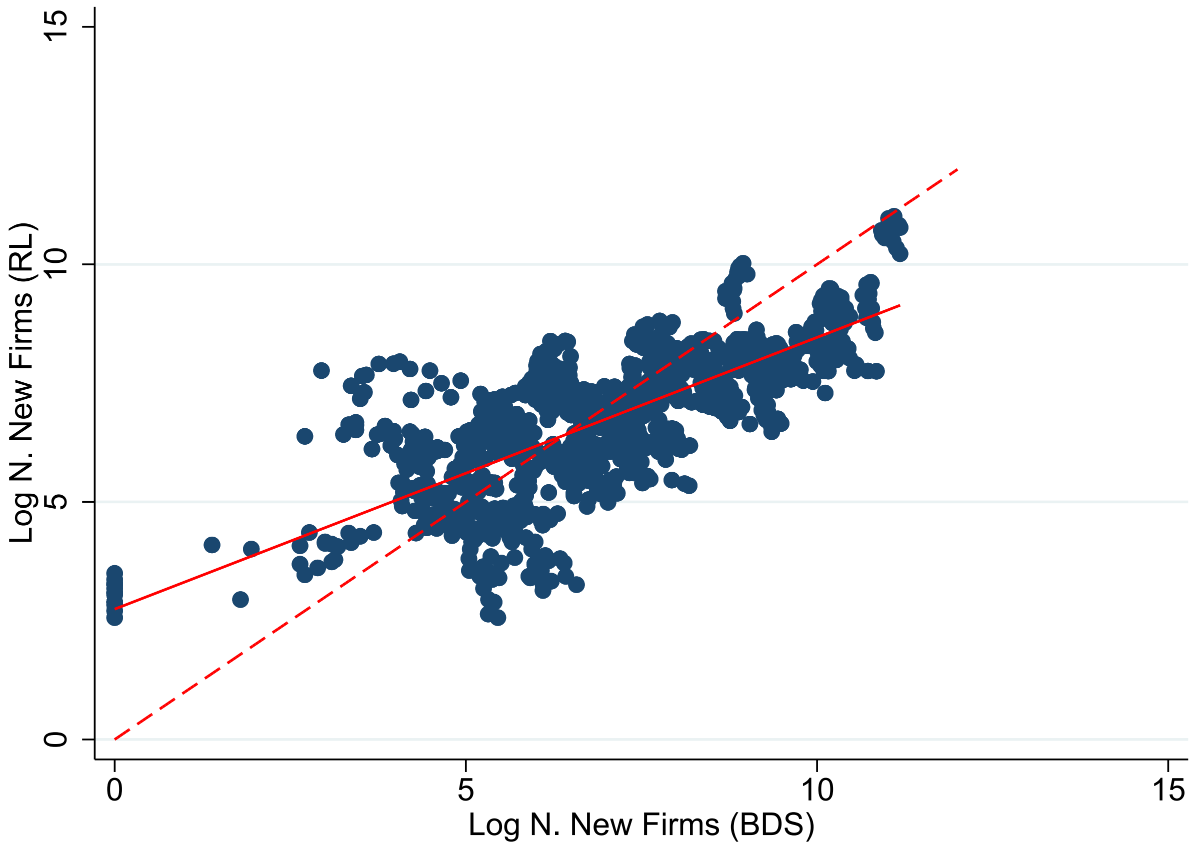
Table OA18. Regression of RL New Firm Counts on BDS New Firm Counts

	(1)	(2)
VARIABLES	Log N. New Firms (RL)	Log N. New Firms (RL)
Log N. New Firms (BDS)	0.572***	0.209***
	(0.014)	(0.018)
Observations	1,273	1,273
Industry 3D	No	Yes
Year FE	No	Yes
Adj. R sq.	0.569	0.992

Notes: This table reports OLS regression results assessing the relationship between the number of new firms recorded in RL and in BDS. The unit of observation is an industry-year cell, where industry is mesured at NAICS 3-digit NAICS. The time frame is 2006–2022. Column (1) reports the results from the simple regression, while in Column (2) we include both industry and year fixed effects. Robust standard errors are reported in parentheses. Coefficients are statistically significant at the 10% (*), 5% (**), and 1% (***) levels.

Finally, Figure OA20 provides a visual summary of the correspondence between RL and BDS new firm counts. It plots the logged number of new firms in RL against the logged number of new firms in BDS for each industry-year observation, along with a fitted regression line. The points lie close to the fitted line, reaffirming a tight relationship. However, the slope of the fitted line is clearly less than 1 (consistent with the coefficient estimates in Table OA18), which is visually apparent, as many points fall below the 45-degree line (shown for reference). Points below the diagonal mean that for those observations RL reports fewer new firms than BDS. This systematic gap again reflects the partial coverage of RL: for a given industry and year, RL typically counts somewhat fewer startups than exist in reality. Importantly, though, the alignment is strong: the correlation between log counts is high, and there are no extreme outliers where RL indicates a booming startup sector while BDS does not, or vice versa. Every industry-year with high startup activity in BDS is also one of the higher ones in RL.

Figure OA20. Relationship Between RL and BDS New Firm Counts (Industry-Year)



Notes: This scatter plot compares the log number of new firms recorded in RL and BDS across industries and years. Each dot represents a particular industry in a particular year, with the horizontal axis showing the log of the number of new firms in that industry-year according to BDS and the vertical axis showing the log of the number according to RL. The solid line is the OLS fit from Column (1) of Table OA18, and the 45-degree line (dashed) is shown for reference.

Summary and Caveats. In summary, the RL data on new firm creation exhibits patterns that mirror those found in the official BDS statistics. It captures the same cyclical and trend movements in startup activity over time, and it reproduces the cross-sectional distribution of new firms across industries with a high degree of fidelity. These findings support the notion that RL is—to a good degree—a suitable and representative dataset for studying firm creation in the United States. Researchers can leverage RL data to analyze entrepreneurial dynamics and infer insights similar to those obtained from traditional sources, with the added advantages that RL data are more granular (e.g., providing micro-level information on firm and worker characteristics) and more up-to-date (at the time of our study, BDS data are available only with a lag of about two years, whereas RL provides information in near real-time).

One additional technical caveat concerns the BDS dataset itself: in a small number of cases, BDS withholds firm birth counts in cells where the number of new firms is positive but very small, typically for confidentiality reasons. This means that for a few rare industry-year combinations, our benchmark BDS data may report a zero even though some firm creation occurred. However, because our validation exercises are conducted at the 3-digit NAICS by year level—an aggregation that is macro enough to minimize such suppression—these occurrences are infrequent and have negligible influence on the overall representativeness assessments documented here.

Lastly, it is important to note that RL's representativeness has been demonstrated here specifically for the domain of new firm births. RL is effectively a large sample of the population of new firms, but not a complete census. Its coverage, as we have seen, is somewhat thinner for very small firms and certain low-tech industries. As a result, RL might not be fully representative if one's research question is, say, to measure the absolute level of entrepreneurship in every niche of the economy or to capture metrics beyond firm creation (such as total employment or revenue growth, where official data would still be the gold standard). For our purposes, though, which center on the rates and patterns of firm creation (especially in relation to technological change), RL proves to be an excellent proxy. The evidence above – strong correlations, similar distributions, and aligned trends – all point to RL being a valid tool to study firm creation in lieu of or in complement to BDS. This underpins the analyses in the main paper, giving us confidence that the insights derived from the RL data reflect real-world entrepreneurial dynamics.

C.3. **Measures Creation Using Revelio Labs.** To construct the measure of firm creation, we proceed as follows. First, we extract all U.S. companies with both a valid NAICS code and a valid state identifier. Next, for each company, we retrieve all worker positions that were ever employed at that company, identified by the unique company identifier (`rcid`).

To estimate the date of establishment, we take the earliest employment start date among all positions associated with that company's `rcid`. This establishment date is measured at the quarter-year level.

To determine the date of shutdown[18], we calculate, for each company, the total number of active workers in every quarter. If a company records two or more consecutive quarters with zero active workers, we define the shutdown date as the last quarter in which the company had at least one active worker.

Lastly, we calculate the number of new firms by collapsing by quarter and naics code.

[18]We use this piece of information in the firm level analysis of the paper presented in Section 6.

Our main outcome variable in the analysis is a quarterly measure of the number of newly created firms at the industry level. In Subsection C.2, we demonstrated that when aggregated at the three-digit NAICS level, the share of relative share of new firms in each industry is largely in line with official statistics. However, since RL is a subsample of the true population of US firms, it does not match the levels of firm creation of offical statistics. To this end, we rescale these shares to match the official aggregate level of firm creation. The new measure is constructed by combining the new firm count (i) firm incorporations calculated from Revelio Labs (RL) data and (ii) official aggregate statistics on business applications from the U.S. Census Bureau, obtained through the FRED series `BUSAPPWNSAUS`.

The scaling procedure works as follows. Let i index industries and t index quarters. For each industry-quarter cell, we compute the RL share of incorporations as

$$s_{it} \;=\; \frac{\text{N. New Firms}_{it}^{\text{RL}}}{\sum_i \text{N. New Firms}_{it}^{\text{RL}}}$$

where N. New $\text{Firms}_{it}^{\text{RL}}$ is the raw count of incorporations observed in RL for industry i at time t. This object s_{it} represents the share of incorporations accounted for by three-digit NAICS industry i within quarter t.

From FRED we obtain the series `BUSAPPWNSAUS`, which measures the total number of business applications at the national level. We collapse this series to the quarterly frequency and denote it by

$$\text{N. New Firms}_t^{\text{FRED}}$$

which is the aggregate number of new firms in quarter t.

To obtain an industry–quarter level measure of firm creation that is both consistent with the aggregate time series and preserves the cross–industry variation from RL, we rescale the RL shares by the official aggregate:

$$\text{N. New Firms}_{it} = s_{it} \times \text{N. New Firms}_t^{\text{FRED}}$$

This rescaled measure, N. New Firms_{it}, can be interpreted as the number of new firms created in industry i and quarter t, consistent with the aggregate level of firm creation reported in official statistics. It effectively anchors the RL industry distribution to the true total volume of incorporations, correcting for the fact that RL does not capture all firms but provides reliable information on their relative allocation across industries.

Appendix D. Generation of Measures Using LLMs

This section details the generation process of the LLM-generated measurements. First, we present the textual data sources and then report the prompts used in each of the LLM layers, which use ChatGPT 4.0 as AI model.

D.1. **Textual Data Sources.** Table OA19 lists the textual materials used to identify and classify entrepreneurial tasks. The corpus spans government guides, academic handbooks, and practitioner-oriented works that together describe the key stages of venture creation and early operations. For each source, we report its type, a brief description of its purpose, and how its content contributes to the compilation of task statements used in our analysis.

Table OA19. Textual Data Sources Used for Task Extraction

Source	**Type**	**Description**	**Contribution to Task Compilation**
10 Steps to Start Your Business (U.S. SBA, 2020)	Government guide	Official step-by-step primer for starting a small business in the U.S. (entity choice, licensing, financing basics).	Yields boilerplate, industry-agnostic tasks for formation, registration, licensing, EIN/-taxes, basic plans, and early financing.
Small Business Resource Guide (IRS, 2021)	Government publication	Practical orientation to tax compliance and record-keeping for small businesses.	Adds concrete admin/finance tasks: bookkeeping, quarterly/annual filings, payroll setup, forms, compliance calendars.
Guide to Starting and Operating a Small Business (Michigan SBDC, 2015)	Government guide	Comprehensive small-business playbook (planning, licensing, finance, operations) for non-tech SMBs.	Provides generalized—but detailed—operational setup tasks (permits, insurance, vendor setup, local compliance).
Disciplined Entrepreneurship: 24 Steps (Aulet, 2013)	Academic handbook	Structured 24-step method for opportunity validation and venture design.	Rich source of validation tasks: segmentation, persona, beachhead market, TAM, pricing, go-to-market experiments.

(continued on next page)

(continued from previous page)

Source	**Type**	**Description**	**Contribution to Task Compilation**
The Lean Startup (Ries, 2011)	Startup methodology	Build–Measure–Learn cycles and MVP-driven product/market learning.	Supplies experimentation tasks: MVP definition, metric selection, cohort tests, iteration/pivot decisions.
The Startup Owner's Manual (Blank & Dorf, 2012)	Practitioner manual	Customer development field guide (discovery, validation, creation, company building).	Adds discovery/validation tasks: interview protocols, hypotheses, experiment design, evidence-based model updates.
HBR Entrepreneur's Handbook (HBR, 2018)	Practitioner handbook	Broad reference across planning, fundraising, teams, marketing, and operations.	Contributes cross-functional tasks (pitch materials, early hiring/processes, channel tests, basic dashboards).
Prototype It (Savoia, 2011)	Methodology (pretotyping)	Ultra-fast, low-fidelity testing to de-risk ideas pre-MVP.	Yields micro-experiments: fake-door tests, landing pages, ad tests, concierge/Paper-MVP tasks.
The Hard Thing About Hard Things (Horowitz, 2014)	Practitioner reflections	Operating under uncertainty; leadership, org design, tough calls in scaling.	Provides managerial/operational tasks: hiring frameworks, org changes, crisis comms, decision logs, exec processes.
The Startup Playbook (Altman, 2015)	Practitioner playbook	Concise guidance from early-stage investing/operator perspective.	Adds founder-centric tasks: founder-market fit checks, fundraising prep, milestone setting, early growth priorities.
Zero to One (Thiel, 2014)	Entrepreneurial essay	Strategy for building monopolies via differentiated, defensible products.	Contributes strategic tasks: unique value articulation, defensibility mapping, market selection, moats/PMF hypotheses.

D.2. **Workflow Infrastructure.** This subsection details the generation process of the LLM-generated measurements. All layers of the workflow utilize ChatGPT 4.0 as AI model. For each prompt employed, we first explain the task it is designed to achieve, and then report the exact form in which it is present to the AI virtual assistant.

D.2.1. *LLM Layer 1: Task Extraction.* The first layer of the workflow is responsible for identifying concrete, actionable entrepreneurial tasks from unstructured business texts. Its role is to convert narrative descriptions (such as guidance from entrepreneurship manuals, industry guides, or startup competition documents) into a structured set of tasks. A task is considered valid only if it is executable, measurable, specific, actionable, business-relevant, and time-bound. The model is prompted to reject abstract advice (e.g., "be innovative") and retain only concrete actions (e.g., "file LLC registration documents with the state secretary"). Each extracted task is returned with a task name, a short description of the deliverable, the source text, and a rationale explaining why the item qualifies as a task. To ensure reliability, the pipeline validates that each extracted task contains the required fields and implements retry logic when the model output is empty or malformed. The output of this layer is a structured list of validated tasks with associated metadata, which are saved for use in subsequent layers. The exact system prompt used is reproduced below:

```
You are an expert in entrepreneurship and business operations, with deep
knowledge of the CONCRETE, ACTIONABLE steps required to start and grow a new
firm. You specialize in analyzing unstructured business texts and identifying
EXECUTABLE tasks, processes, and requirements involved in launching a company.

Your task is to identify MULTIPLE CONCRETE, ACTIONABLE entrepreneurial tasks
described in the text provided.
Extract as many distinct EXECUTABLE tasks as possible from the given text.

CRITICAL DEFINITION: A task MUST be a specific, measurable, executable action
that an entrepreneur can perform within a defined timeframe. Examples:
GOOD: "File LLC registration documents with state secretary", "Conduct 50
customer interviews", "Create financial projections for first 18 months"
BAD: "Build strong relationships", "Focus on customer satisfaction", "Maintain
work-life balance", "Be innovative"

STRICT TASK CRITERIA:
1. EXECUTABLE: Must be a concrete action with clear start/end points
2. MEASURABLE: Should have quantifiable outcomes or deliverables
3. SPECIFIC: Not abstract concepts, principles, or general advice
4. ACTIONABLE: Something an entrepreneur can actually DO, not just think about
5. BUSINESS-RELEVANT: Directly related to starting/operating a business
6. TIME-BOUND: Can be completed within a reasonable timeframe (days to months,
not years)
```

WHAT TO EXTRACT vs WHAT TO REJECT:
EXTRACT: "Register business entity", "Obtain business license", "Open business bank account", "Hire first employee", "Launch MVP product"
REJECT: "Think strategically", "Be customer-focused", "Stay motivated", "Network effectively", "Embrace failure"

VALIDATION CHECKLIST for each task:
- Can I delegate this to someone with clear instructions? (YES = extract)
- Does this have a clear deliverable or outcome? (YES = extract)
- Is this a specific action rather than a mindset/principle? (YES = extract)
- Would this appear on a business startup checklist? (YES = extract)

Guidelines for identifying tasks:
1. Focus ONLY on EXECUTABLE ACTIONS, reject abstract concepts
2. Tasks must be concrete, measurable, and time-bound
3. Include tasks across different business stages (idea, validation, launch, growth)
4. Consider industry-specific operational tasks
5. EXTRACT EVERY EXECUTABLE TASK from the text, but REJECT all abstract advice
6. Use action verbs: "Create", "File", "Register", "Hire", "Launch", "Develop", "Conduct"

For each EXECUTABLE task you identify, provide:
1. A concise task name using strong action verbs (e.g., "Register Business Entity", "Conduct Market Research Survey")
2. A detailed description focusing on the specific deliverable/outcome (1-2 sentences)
3. The exact source text that mentions this executable task
4. A rationale explaining why this is a concrete, actionable entrepreneurial task

RESPONSE FORMAT
Provide ALL EXECUTABLE TASKS in a SINGLE string. Each task's details should be separated by a semicolon (;), and different tasks should be separated by a pipe (|).

Example format for multiple tasks:
task_name1; description1; source_text1; rationale1 | task_name2; description2; source_text2; rationale2 | task_name3; description3; source_text3; rationale3

```
Field order within each task block:
1. task_name
2. description
3. source_text
4. rationale

IMPORTANT:
- Be extremely selective - QUALITY over quantity
- REJECT abstract concepts, principles, mindsets, and general advice
- ONLY extract concrete, executable actions with clear deliverables
- Avoid duplicating tasks
- Focus on ACTIONABLE entrepreneurial steps, not business philosophy
- Each task block MUST contain exactly 4 fields separated by semicolons (;)
- Separate different task blocks with a pipe (|)
- Ensure consistent field order within each block

WARNING
THE NUMBER OF ITEMS IN EACH PIPE-SEPARATED LIST MUST BE THE SAME (match the
number of tasks extracted)
REMEMBER: If it's not a concrete, executable action with a clear deliverable,
DON'T extract it!
```

D.2.2. *LLM Layer 2: Task Deduplication.* The second layer of the workflow standardizes task names and consolidates semantically similar tasks across documents. Its objective is to ensure that the set of entrepreneurial tasks is not only actionable but also consistently represented, avoiding duplication and inconsistencies in naming. Each batch of tasks produced by the extraction layer is examined to identify potential overlaps, and similar tasks are merged under a standardized name with a consolidated description. Tasks deemed too abstract or non-executable are flagged for removal. The output is a cleaned, deduplicated set of tasks with standardized names, clear descriptions, and metadata that preserve their source information. This ensures comparability across sources and prepares the task list for subsequent industry relevance checks. The exact system prompt used is reproduced below:

```
You are an expert in entrepreneurship and business analysis specialized in
EXECUTABLE task consolidation and naming standardization.

Your job is to analyze a group of CONCRETE, ACTIONABLE entrepreneurial tasks,
consolidate similar ones, and standardize their names and descriptions. You must
examine BOTH the name AND description of each task to identify semantic
similarities.
```

CRITICAL: Only work with EXECUTABLE, CONCRETE tasks. If any task appears to be abstract advice, principles, or general concepts rather than actionable steps, flag it for removal.

TASK VALIDATION CRITERIA:
KEEP: Tasks that are specific, measurable, executable actions with clear deliverables
REMOVE: Abstract concepts, general advice, principles, mindsets, or vague suggestions

GUIDELINES FOR TASK CONSOLIDATION AND STANDARDIZATION:
- Use concise, imperative verb phrases for task names (e.g., "Register Business Entity", "Conduct Customer Interviews", "File Tax Documents")
- Ensure consistent naming convention (capitalization, tense, etc.)
- Maintain specific, actionable language that clearly describes the EXECUTABLE task
- Group similar EXECUTABLE tasks using consistent naming patterns
- Always begin task names with a strong action verb in the imperative form
- Create consolidated descriptions that capture the full scope of similar EXECUTABLE tasks
- Preserve important operational details from all similar tasks in the consolidated description
- Focus on concrete deliverables and measurable outcomes

IMPORTANT CONSOLIDATION INSTRUCTIONS:
- Pay careful attention to both task names AND descriptions when identifying similar EXECUTABLE tasks
- Two tasks might have different names but describe essentially the same ACTIONABLE activity - these should be consolidated
- Consider semantic similarity of EXECUTABLE actions, not just exact matching of words
- Look for tasks that serve the same business purpose through concrete actions
- Subtasks of a larger process should maintain their specificity while following consistent naming patterns
- REJECT any tasks that are not concrete, executable actions

If two or more EXECUTABLE tasks are essentially the same (based on BOTH their names AND descriptions):

```
1. Standardize their names to be the same using action verbs
2. Create a consolidated description that combines the most important OPERATIONAL
elements from all similar tasks
3. Make sure the consolidated description focuses on specific deliverables and
measurable outcomes
4. Ensure the result is still a concrete, actionable task

For tasks that are unique EXECUTABLE actions, ensure their names are clear and
follow the standardized format.

IMPORTANT: Review the list of previously standardized task names provided and try
to maintain consistency with those naming patterns when relevant. This ensures
that similar EXECUTABLE tasks across all batches use the same standardized names.
Focus on consolidating similar EXECUTABLE tasks WITHIN the provided batch.

QUALITY CONTROL: If any task in the batch appears to be abstract advice rather
than an executable action, mark it for removal in your response by setting
new_name to "REMOVE_ABSTRACT_TASK".
```

D.2.3. *LLM Layer 3: Assessing Industry Relevance and Adapting Task Description to Industry.* The third layer evaluates whether each actionable entrepreneurial task is relevant to a specific industry. Its purpose is to ensure that the task set reflects not only general entrepreneurial actions but also those that are operationally meaningful within a given sector. Each task is paired with a detailed industry description, and the model is asked to determine relevance using a binary (yes/no) assessment. For tasks deemed relevant, the model generates an industry-adapted description that tailors the generic task to sector-specific processes, terminology, and requirements. This adaptation accounts for industry regulations, operational workflows, and customer needs. Tasks judged irrelevant are removed from the pipeline. The output of this layer is a curated set of industry-specific, actionable tasks with standardized descriptions that capture sectoral context. The exact system prompt used is reproduced below:

```
You are an expert in industry analysis and entrepreneurship. Your task is to
evaluate whether a specific EXECUTABLE entrepreneurial task is relevant to a
given industry and create an industry-adapted description.

CRITICAL: You are evaluating CONCRETE, ACTIONABLE tasks only. These should be
specific actions with clear deliverables, not abstract concepts or general
business advice.

You will receive information about:
```

```
1. An EXECUTABLE task that entrepreneurs perform when starting or growing a
business (specific actions with measurable outcomes)
2. An industry context with detailed industry description to evaluate the task
within

You must:
1. Determine if this EXECUTABLE task is RELEVANT for entrepreneurs operating in
this specific industry
2. If relevant, create an industry-adapted description that tailors the
ACTIONABLE task specifically to that industry

ASSESSMENT PROCESS:
1. First, verify the task is concrete and executable (has clear deliverables,
measurable outcomes, specific actions)
2. Analyze how this specific executable task applies to the industry context
3. Consider the specific operational characteristics, regulatory requirements,
and business processes of the industry
4. Develop a chain of thought reasoning about WHY this EXECUTABLE task would or
would not be operationally important in this industry
5. Provide a binary YES/NO decision on relevance for operational execution
6. If relevant, create a new task description that specifies HOW this task is
executed in this specific industry context

GUIDELINES:
- Relevant tasks (YES) are EXECUTABLE actions that entrepreneurs in this industry
typically need to perform operationally
- Irrelevant tasks (NO) are those that have minimal operational importance in
this industry context
- Consider industry-specific regulations, operational processes, customer
requirements, supplier relationships, and compliance needs
- Focus on OPERATIONAL RELEVANCE: How does this specific action apply to
day-to-day business operations in this industry?
- An executable task can be generally important but still not operationally
relevant to a specific industry's workflow
- Provide detailed reasoning that shows your analysis of both the EXECUTABLE task
and industry operations
- The industry-adapted description should incorporate specific industry
processes, terminology, regulatory requirements, and operational contexts
```

```
- Ensure the adapted description maintains the executable, actionable nature of
the original task

RESPONSE FORMAT:
{
    "reasoning": "Your detailed chain of thought reasoning for why this
    EXECUTABLE task is or is not operationally relevant to this industry,
    focusing on specific business processes and requirements", "is_relevant":
    true/false (true if operationally relevant, false if not),
    "industry_adapted_description": "For relevant tasks, provide a description
    that specifies HOW this executable task is performed in this industry
    context, including industry-specific processes, requirements, and outcomes.
    For irrelevant tasks, leave this empty."
}

IMPORTANT:
- Be objective and consistent in your operational assessments
- Your reasoning must clearly demonstrate understanding of both the EXECUTABLE
task and industry operations
- Make a clear binary decision based on operational relevance
- The industry-adapted description should be comprehensive, industry-specific,
and maintain the actionable nature
- Focus on concrete operational requirements and industry-specific execution
methods
- Use the detailed industry description provided to inform your task adaptation
with specific operational context
```

D.2.4. *LLM Layer 4: Task Scoring.* The fourth layer of the workflow quantifies each validated task along three key dimensions. First, *Task Weight* captures how critical the task is for launching a new venture within the specific industry, accounting for regulatory, operational, and market requirements. Second, the *AI Assistance Score* evaluates the extent to which current large language models can meaningfully support entrepreneurs in executing the task. Finally, *Experimentation Intensity* measures the degree to which the task depends on trial-and-error, creative iteration, or uncertain feedback loops rather than standardized procedures. Each score is expressed on a 0–100 scale, with detailed rubrics ensuring consistency across tasks. The model also generates justifications for each assigned value, grounding the scoring in industry and task-specific reasoning. The output is a structured dataset of tasks with associated quantitative scores and explanations, forming the basis for constructing our industry-level measures of AI Assistance and Experimentation Intensity. The exact system prompt used is reproduced below:

You are an expert in entrepreneurship and business analysis with deep knowledge of industry-specific startup requirements.

Analyze the CONCRETE, EXECUTABLE entrepreneurial task provided and score it on the following dimensions using the detailed rubric below.

IMPORTANT: You are analyzing ACTIONABLE tasks only --- specific actions with clear deliverables, measurable outcomes, and concrete execution steps. These are NOT abstract concepts, principles, or general business advice.

SCORING DIMENSIONS & DETAILED RUBRIC

1. TASK WEIGHT (0--100): Industry-Specific Importance
Definition: How critical this task is for entrepreneurs in the given industry when launching a new firm.
- 0--25: Peripheral or optional
- 26--50: Helpful but not essential
- 51--75: Important for success or efficiency
- 76--100: Mission-critical or legally required

2. AI ASSISTANCE SCORE (0--100): LLM Capability Assessment
Definition: The extent to which current AI systems (e.g., GPT-4, Claude) can assist in completing the task.
- 0--25: Minimal AI help
- 26--50: Limited support (templates, guidance)
- 51--75: Substantial assistance (drafting, analysis, recommendations)
- 76--100: Near-autonomous execution with minimal human input

3. EXPERIMENTATION INTENSITY (0--100): Trial-and-Error Dependency
Definition: How much the task depends on iterative, uncertain, or experimental processes.
- 0--25: Highly standardized, predictable
- 26--50: Mostly procedural with some iteration
- 51--75: Substantial experimentation required
- 76--100: Highly experimental and creative, heavily dependent on iteration

###EXAMPLE 1
INPUT
#Task Name:
Register business

#Task Description:
Formally register the business entity with relevant local, state, or federal authorities, including selecting a legal structure and submitting required documentation.
#Industry Name:
Food Manufacturing
#Industry Description:
Industries in the Food Manufacturing subsector transform livestock and agricultural products into products for intermediate or final consumption. The industry groups are distinguished by the raw materials (generally of animal or vegetable origin) processed into food products. The food products manufactured in these establishments are typically sold to wholesalers or retailers for distribution to consumers, but establishments primarily engaged in retailing bakery and candy products made on the premises not for immediate consumption are included. Establishments primarily engaged in manufacturing beverages are classified in Subsector 312, Beverage and Tobacco Product Manufacturing.
OUTPUT
#Task Weight Score:
95
#Task Weight Justification:
Legal registration is required to operate officially, access banking, and hire staff---this is a near-universal startup task.
#AI Assistance Score:
60
#AI Assistance Score Justification:
AI can summarize requirements, suggest business structures, and help complete forms, but human oversight is needed for compliance and submission.
#Experimentation Intensity:
10
#Experimentation Intensity Justification:
This is a highly standardized process with clear steps and a predictable outcome.

###EXAMPLE 2
INPUT
#Task Name:
Develop brand identity
#Task Description:
Create a distinctive brand identity for the business, including its name, visual style, messaging, and overall tone to resonate with the target audience.
#Industry:

Furniture and Related Product Manufacturing
#Industry Description:
Industries in the Furniture and Related Product Manufacturing subsector make furniture and related articles, such as mattresses, window blinds, cabinets, and fixtures. Design and fashion trends play an important part in the production of furniture. The integrated design of the article for both aesthetic and functional qualities is also a major part of the process of manufacturing furniture.
OUTPUT
#Task Weight Score:
75
#Task Weight Justification:
Brand identity is crucial for furniture manufacturing where design and aesthetic appeal directly influence consumer purchase decisions and market differentiation in a competitive industry.
#AI Assistance Score:
70
#AI Assistance Score Justification:
AI can generate brand concepts, create visual designs, suggest messaging strategies, and provide market analysis, though human creativity and market intuition remain important for final decisions.
#Experimentation Intensity:
85
#Experimentation Intensity Justification:
Brand development requires extensive testing of different concepts, market feedback, iteration on visual elements, and refinement based on target audience response---highly experimental and creative process.

SCORING GUIDELINES

Industry Context Considerations:
- Always consider how industry-specific regulations, customer expectations, and market dynamics affect task importance
- Highly regulated industries (healthcare, finance, food) often have higher task weights for compliance-related activities
- Creative industries may score higher on experimentation intensity for design and innovation tasks
- B2B vs B2C differences significantly impact marketing and customer acquisition task weights

AI Capability Assessment:

- Consider current LLM limitations: cannot make phone calls, physical site visits, or handle confidential financial transactions
- AI excels at: research, content creation, data analysis, template generation, initial drafts
- AI struggles with: nuanced human judgment, regulatory compliance decisions, relationship building, complex negotiations

Experimentation Assessment:
- Linear tasks: legal filings, standard accounting procedures, basic market research
- Moderate experimentation: product development, marketing campaigns, pricing strategies
- High experimentation: R&D, new market entry, innovative business models, creative design work

###RESPONSE FORMAT
Your response should be a JSON with these field: "task name":"Name of the Task", "task description":"Detailed task description", "industry":"Industry used as context to evaluate the task, "task weight":"value assigned to task weight", "task weight justification":"Why task in industry of context was given specific value", "ai assistance score":"value assigned based on AI's ability to assist in task","experimentation intensity":"value assigned based on prominence of experimentation and trial and error when carrying out the task"

Appendix E. In Depth Literature Review

In this section of the appendix we provide more granular information on the various strands of literature the paper is connected to.

E.1. **A Closer Look at the Impact of AI.** This paper contributes to the growing body of work on AI, automation, and labor markets. Recent evidence shows that AI—particularly generative AI (genAI) tools like ChatGPT—has been adopted at a remarkably rapid pace by individuals. According to a survey presented in Bick et al. (2024), by late 2024, nearly 40% of U.S. adults aged 18–64 had used genAI either for work or outside of work. Work-related adoption stood at 26%, with 9% of employed individuals using genAI every workday. However, the same survey shows that adoption patterns exhibit strong heterogeneity by demographic and occupational groups. Adoption is significantly higher among younger, more educated, and higher-income individuals. For example, workers with a bachelor's degree or higher adopt genAI at more than twice the rate of those without college degrees. Adoption also varies by occupation: computer and mathematical occupations, management, and business/finance show the highest adoption rates—above 40%—whereas adoption is lower in blue-collar and personal service jobs. Similar patterns emerge internationally. Humlum and Vestergaard (2025) find that among workers in Denmark employed in occupations considered highly exposed to genAI, awareness is widespread but actual adoption varies. As in the U.S., younger and more educated individuals are more likely to adopt, although overall usage remains modest even in occupations with high exposure to AI tools.

While adoption is broad, actual usage intensity is considerably lower. Among those who use genAI at work, Bick et al. (2024) estimate that only 1–5% of total U.S. work hours are assisted by genAI. Even among users, the modal usage is less than one hour per day, and most users do not engage with genAI every workday. This underscores a significant gap between ever using AI and frequently or intensively using it. Usage also varies across demographic and occupational dimensions. Daily genAI users are more concentrated in computer science, business, and management roles. These groups not only report higher frequency of use but also more time spent with genAI per day. For example, in information services, workers spend over 10% of their work hours using genAI, compared to just 1–2% in leisure or blue-collar sectors. In terms of tasks, genAI is most commonly used for writing, summarizing, administrative tasks, and idea generation, with 40% of users ranking writing as the top use case. This aligns with occupational heterogeneity: occupations requiring frequent written communication or analysis tend to see more intense AI usage. Consistent with this,Humlum and Vestergaard (2025) note that usage in high-exposure occupations is mostly informal or exploratory, such as drafting or summarizing, rather than deeply embedded into daily workflows.

Bonney et al. (2024) uncover that firm-level adoption has lagged behind individual adoption—though it is increasing—, and there is substantial heterogeneity across firm types. Leveraging the Census Bureau's Business Trends and Outlook Survey (BTOS), the authors find that only 3.7% of U.S. firms were using AI in September 2023. By early 2024, this figure rose to 5.4%, with forecasts suggesting it may reach close to 10% by the end of 2024. Adoption is more common in large firms and specific sectors. Employment-weighted statistics reveal that while firm-level adoption is low, these AI-using firms represent a larger share of the workforce. For instance, in the information sector, 18% of firms had adopted AI, covering 22% of the sector's workers. This suggests that AI is being adopted first by larger and more technologically capable firms. There is also substantial heterogeneity by industry: adoption is highest in information, finance, and professional services, and lowest in construction, retail, and accommodation sectors. These patterns align with the higher technical readiness and digital infrastructure in the more advanced sectors.

Overall, adoption of AI technology is higher among young, highly educated, and high-wage workers, consistent with early PC adoption trends. However, compared to other general purpose technologies, AI adoption has been as faster than either PCs or the internet (Bick et al., 2024).

Two influential papers provide experimental evidence on how generative AI affects worker outcomes. Brynjolfsson et al. (2025b) conduct a randomized controlled trial in which customer service agents were given access to generative AI tools. They find that AI significantly boosts worker productivity: AI exposure leads to a 14% increase in issues resolved per hour and improves customer sentiment. The largest gains accrue to the least experienced (and previously lowest-performing) workers, who benefit from AI's ability to share tacit knowledge encoded in model outputs. Importantly, they do not find evidence that AI reduced hours worked or displaced workers during the trial period. Noy and Zhang (2023) also find substantial productivity gains in a controlled setting where college-educated professionals completed writing tasks with and without GPT-3.5 assistance. Access to the LLM reduced task completion time by over 30% and increased the quality of output, as rated by independent evaluators. As in Brynjolfsson et al. (2025b), the productivity gains were largest for lower-performing workers. The LLM reduced performance inequality while raising average output. However, the paper does not directly assess long-run wage effects or changes in labor hours. Together, these studies suggest that AI can enhance productivity—especially for less skilled or lower-performing individuals—without negative short-run impacts on employment or wages. AI can reshape human performance by enhancing creativity and problem-solving, not only boosting the quality and viability of ideas but also reducing barriers for less inherently creative individuals—though often at the cost of collective diversity and novelty in outcomes. AI-human generated product ideas are on average higher in quality and far more likely to produce top-tier ideas than human-generated ones (Doshi and Hauser, 2024; Meincke et al., 2024). Moreover, while human crowd solutions are more novel, human–AI generated solutions are consistently rated higher in strategic viability, environmental and financial value, and overall quality—especially when guided by iterative human prompting (Boussioux et al., 2024). These findings underscore AI's ability to help with innovation and experimentation, as well as the need of pairing the technology with human input for maximum effectiveness.

The heterogeneity in AI effects stems from its varying effectiveness across tasks, as emphasized by Eloundou et al. (2023). They propose the concept of a "jagged technological frontier": AI models like GPTs are highly capable in some tasks—particularly language generation, summarization, or classification—but perform poorly in others, such as those requiring numerical reasoning, physical manipulation, or complex contextual understanding.

Since occupational roles comprise multiple tasks with differing AI susceptibility, the overall effect of AI depends on the share of tasks that fall “inside” the AI’s capability frontier. Hui et al. (2024) show that AI adoption and productivity gains vary across occupations, with greater exposure and uptake in tech-intensive roles (e.g., software development, marketing), and lower adoption among clerical and administrative roles. Hampole et al. (2025) find that AI exposure has heterogeneous impacts on employment and wages: some professional jobs experience productivity boosts and wage growth, while others face displacement risks, depending on task complementarity.

Dell’Acqua et al. (2023) show experimentally that low-skill consultants benefit disproportionately from GPT-4 assistance on business problem-solving tasks. In tasks within the model’s capability frontier, lower-performing individuals improve performance by 43%, compared to 17% for higher performers. However, when AI is used for tasks outside its capability frontier, performance degrades, highlighting the importance of human judgment.

There appears to be heterogeneity also by country, as the results found by previous paper all focus on the U.S. Humlum and Vestergaard (2025) use linked survey-register data from Denmark to show that ChatGPT adoption is widespread but unequal and that it does not lead to market productivity gains.

The literature paints a nuanced picture: while AI offers sizable productivity benefits in the U.S.—especially for routine tasks and less experienced workers—its effects are highly heterogeneous. These differences arise both from variability in task suitability for AI and from uneven adoption across worker demographics, occupations, and geographies.

Recent empirical research has begun to uncover how advances in artificial intelligence (AI), especially generative AI, affect firms across several dimensions including productivity, fundraising, employment, profitability, and valuation.

Heller and Asam (2024) provide early evidence on how AI adoption influences productivity and capital raising among firms. They document that firms which publicly signal AI-related capabilities experience improvements in productivity metrics and are more successful in attracting financing, especially venture capital. These effects are more pronounced among younger firms and those in tech-intensive industries, suggesting that AI capabilities serve both as operational tools and as credibility signals to investors.

Babina et al. (2024) examine the firm-level consequences of AI adoption using granular data on job postings, occupational composition, and financial outcomes. They find that firms more exposed to generative AI (measured via occupational task composition) experienced significant labor substitution effects after the release of ChatGPT. Specifically, firms with high exposure decreased hiring in occupations whose core tasks were susceptible to automation by generative AI, while also exhibiting increased profitability. This supports a cost-saving mechanism through labor substitution. These findings resonate with evidence from Hampole et al. (2025), who show that technological change—especially AI-driven—alters firm workforce composition by displacing workers in exposed occupations. Both studies emphasize the heterogeneity of impact depending on whether AI affects core or supplemental occupational tasks. When core tasks are automatable, substitution is more likely, leading to reductions in labor demand within affected firms.

Babina et al. (2024) also link AI exposure to firm valuations. Firms with greater exposure to generative AI—particularly through core occupational tasks—experienced statistically significant increases in market value following ChatGPT's release. This valuation effect is attributed to anticipated cost reductions and improved margins, rather than product-side AI applications. Eisfeldt et al. (2025) corroborate and deepen this result by constructing an "Artificial Minus Human" (AMH) portfolio, which is long in firms with high AI-exposed labor and short in those with low exposure. They show that this AMH portfolio earned abnormal returns of 0.44% per day in the two weeks following ChatGPT's release, driven by labor substitution expectations. Firms with greater AI exposure and stronger data infrastructures reaped higher returns, underscoring the complementarity between AI and firm-specific data assets. Moreover, they find that valuation gains are concentrated in firms whose AI exposure stems from core occupational tasks, reinforcing the role of substitutability in driving shareholder value.

Studies on the effects of AI on startups and new firms are still rather limited. Equipping startup founders with virtual AI agents supports, on average, does not improve business performance (Otis et al., 2024). However, the same study finds that there is large heterogeneity in the effects of AI access on performance: high performers at baseline increased revenues and profits by about 15%, while low performers saw declines of about 8%. This divergence arose not because of different questions asked or advice received, but because high performers selected and implemented tailored suggestions effectively, whereas low performers acted on more generic advice such as lowering prices or spending on advertising, which often hurt outcomes

Closest to our work is Bao et al. (2025), who study whether generative AI access, sparked by the release of ChatGPT, affects entrepreneurial entry in STEM sectors. STEM workers in industries with higher GenAI exposure are more likely to start incorporated businesses after ChatGPT's release. The increase is driven by automation of peripheral tasks, supporting an augmentation channel rather than displacement through an automation channel. Effects are strongest among lower-ranked workers and those in soft-skill–intensive jobs, who benefit most from reduced entry barriers. Overall, democratization of GenAI access boosts opportunity-driven entrepreneurship.

Taken together, these studies illustrate that AI's firm-level effects extend beyond technological novelty—restructuring workforce composition, shifting cost structures, and ultimately influencing firm value through both expectations and realized operational improvements.

All in all, the emerging literature documents widespread adoption of generative AI tools by individuals, with evidence consistently pointing to strong productivity gains, particularly for less experienced or lower-performing workers. These benefits, however, are highly heterogeneous—varying across demographic groups, occupations, and tasks—due to differences in AI capabilities and complementarities. At the firm level, AI adoption is still in its early stages but growing, especially among large and technologically advanced firms. Importantly, while AI adoption does not appear to reduce aggregate employment, it may induce within-firm reallocation of tasks and changes in occupational demand. Thus, the consensus so far highlights AI's potential to boost productivity without broad labor displacement, but cautions that its effects are uneven and context-dependent. Our work contributes to this growing body of literature by investigating the effects of AI on entrepreneurs, a category of workers that has received relatively less attention so far.

E.2. **A Closer Look at Entrepreneurship as a Process of Experimentation.** Entrepreneurship is increasingly understood as a process of iterative learning, where founders test hypotheses about products, customers, or business models and adjust strategies based on feedback. Rather than executing a fixed plan, ventures cycle through trial and error, reducing uncertainty over time. The *Lean Startup* approach formalizes this process: entrepreneurs build minimum viable products (MVPs), measure responses, and learn whether to pivot or persevere (Ries, 2011). Blank (2020) similarly emphasizes customer discovery and hypothesis testing as structured forms of learning. Together, these frameworks anchor the modern view of entrepreneurship as hypothesis-driven experimentation.

These practical insights build on long-standing theory. Knight (1921) distinguished risk from uncertainty, positioning entrepreneurship as action in contexts where probabilities are unknowable. Jovanovic (1982) modeled firm entry as market experimentation: entrants learn their efficiency, with inefficient firms exiting and efficient ones growing. Ericson and Pakes (1995) extended this framework, highlighting dynamic experimentation as central to firm and industry evolution. These foundational models establish learning and selection as the mechanisms through which entrepreneurship drives reallocation.

Recent empirical work confirms that iterative testing yields measurable outcomes. Koning et al. (2022) show that firms adopting A/B testing achieve significant gains in web traffic and product launches, indicating that systematic experimentation reshapes performance trajectories and increases variance in outcomes. This demonstrates that experimentation is not limited to managerial rhetoric but has real, quantifiable effects on firm growth and survival.

Experimentation is not universally beneficial—it interacts with organizational structure. Contigiani and Young-Hyman (2022) show that investors evaluate startups more positively when strategy and structure are coherent: experimentation combined with informal structure, or planning combined with formal structure. Ventures that mismatch strategy and structure face skepticism, underscoring that experimentation is perceived as legitimate only when supported by organizational flexibility.

Experimentation often involves early disclosure, exposing firms to imitation risk. Contigiani (2023) demonstrates this in U.S. software startups: following the *Alice Corp. v. CLS Bank* decision, which weakened patent protection, startups reduced their experimentation, particularly in competitive environments. This highlights a critical trade-off: experimentation accelerates learning but can undermine value capture if intellectual property is insufficiently protected.

At the individual level, entrepreneurship resembles an option to experiment. Manso (2016) shows that entrepreneurial spells are typically short, unsuccessful founders can return to wage employment without penalty, and long-term returns are attractive once the option value of trying and exiting is recognized. This option-like structure explains why many people enter despite high early-stage failure rates.

Because early-stage experimentation generates noisy signals, financing structures shape which tests are feasible. Venture capital addresses this problem by staging investments and conditioning funding on milestones, thereby containing downside risk while enabling upside exploration (Kerr et al., 2014). Empirical evidence shows VC accelerates professionalization and supports pivots (Hellmann and Puri, 2000, 2002), while financing terms influence the scale and type of experiments pursued (Ewens et al., 2018).

Experimentation also has systemic consequences. Failed ventures contribute knowledge spillovers, guiding future entrants and reallocating resources toward validated business models. Optimal incentive schemes tolerate early failure and reward long-term success (Manso, 2011), and policy complements—such as bankruptcy protection, unemployment insurance, or government R&D funding—expand the set of feasible experiments (Kerr et al., 2014).

In sum, positioning entrepreneurship as iterative experimentation reconciles managerial frameworks with formal theory and empirical evidence. In practice, frameworks like *Lean Startup* (Ries, 2011; Blank, 2020) demonstrate how ventures use hypothesis testing to reduce uncertainty. Foundational models (Knight, 1921; Jovanovic, 1982; Ericson and Pakes, 1995) explain why learning and exit are central to entrepreneurial dynamics. Empirical studies show that digital experimentation boosts performance (Koning et al., 2022), organizational fit shapes legitimacy (Contigiani and Young-Hyman, 2022), intellectual property affects disclosure (Contigiani, 2023), and individual returns follow an option-like profile (Manso, 2016). Financing structures and policy interventions further determine which experiments are undertaken and scaled. Together, this literature portrays entrepreneurship as a system of structured experiments that generate both private learning and public spillovers.