Why Don't Old Firms Do New Things?[*]

Nicolas Crouzet[†] Zhiguo He[‡] Victor Lyonnet[§] Yueran Ma[¶]

June 2026

Abstract

Since at least Schumpeter (1942), new firms have been widely viewed as the driving force for implementing new technologies. Yet limited economics research explains when and why new technologies require new firms. We examine the view that incumbents struggle when new technologies lead to organizational conflicts, by requiring a shift towards occupations with different approaches for solving problems compared to the existing workforce. We measure these approaches using workstyles of each occupation from O*NET. In the data, young firms grow significantly faster than old firms when new technologies in the industry induce larger shifts. Technologies that generate larger shifts are also more likely to be first implemented by young firms. By contrast, the sheer volume of new technologies (e.g., the number of all or important patents) is not significantly related to the out-performance of young firms. These results indicate that organizational frictions shape firms' abilities to absorb and implement new technologies, and provide new perspectives for the Coase (1937) boundary of the firm question.

[*]We are grateful to Geneve Bullo, Jingyi Jia, Howard Guankun Li, Haomin Qin, Raymon Yue, and Mengdi Zhang for extraordinary research assistance. We thank Ashwini Agrawal, Tania Babina, Onur Bayar, Nick Bloom, Ed Glaeser, Oliver Hart, Greg Kaplan, Steve Kaplan, Josh Lerner, Gustavo Manso, Raffaella Sadun, Bryan Seegmiller, Andrei Shleifer, Chad Syverson, Rob Vishny, Ben Zhang, and seminar participants at Boston University, e61, Harvard, Stanford, Tsinghua, UCLA, and University of Kentucky Finance Conference for insightful comments. First draft: November 2025.

[†]Northwestern University Kellogg School of Management (n-crouzet@kellogg.northwestern.edu).

[‡]Stanford University Graduate School of Business and NBER (hezhg@stanford.edu).

[§]University of Michigan Ross School of Business (vlyonnet@umich.edu).

[¶]University of Chicago Booth School of Business and NBER (yueran.ma@chicagobooth.edu).

I Introduction

Nearly one century ago, Schumpeter (1942) highlighted the importance of new technologies for economic progress. In his view, the implementation of new technologies relies crucially on entrepreneurship, and the process of creative destruction involves *new firms* overtaking existing ones. This view has been widely held among academics, policymakers, and the general public ever since (Decker et al., 2014; Draghi, 2024), and entrepreneurship is a core topic of economics research today. However, some studies have found that existing firms contribute substantially to technological advancement (Lerner, 2006; Garcia-Macia, Hsieh, and Klenow, 2019; Braguinsky et al., 2024), and we can name plenty of new technologies that are successfully implemented by incumbents, such as polymers, smartphones, and cloud computing (which we discuss in more detail in Section II). The necessity of new firms for new technologies is not entirely self-evident.

When—and why—do new technologies require new firms to implement? Although this question is fundamental for understanding the essence of creative destruction, economics research has offered limited theoretical guidance and systematic analyses. Influential models of creative destruction such as Aghion and Howitt (1992) do not address whether an entrant using a new technology is a new firm or an existing firm. Meanwhile, research motivated by the classic boundary of the firm question following Coase (1937) has focused on vertical integration (Williamson, 1971; Grossman and Hart, 1986; Hart and Moore, 1990), or "which transactions can be implemented within a firm?" Yet another aspect of the Coasian problem has received much less attention: "which ideas can be implemented within a firm?"

Case studies in management and reflections by managers often come to the view that old firms can implement new technologies if they are compatible with existing organizational processes and priorities, but struggle if they require new ones (Christensen, 1997; Gerstner, 2002; O'Reilly and Tushman, 2021). As Microsoft CEO Satya Nadella remarked at Chicago Booth's 125th anniversary celebration: "this is one of the foundational challenges—when the new thing comes not only as a technology challenge but also a business model challenge, most companies can't make it." This view also echoes observations about the limits of organization in Arrow (1964), who postulates that organizations form "codes... in accordance with the best expectations of the firm's creation," which will be "modified only slowly over time," in which case new technologies that require a different set of codes can be difficult for incumbents compared to new firms that start from scratch. The incompatibility between new technologies and existing organizational codes is then central to the need for new organizations.

In this paper, we formalize this hypothesis theoretically, and conduct extensive empirical analyses to document its relevance for the performance of young vs old firms in the face of new technologies. Our

primary challenge—conceptually and empirically—is finding a way to capture the extent to which new technologies require new organizational codes. Our entry point is to study how different technologies require different occupations to implement, which are associated with different approaches to solving problems; we later measure these traits using O*NET data on occupations' workstyles, and use this term in the following. For example, in the car industry, technological advancements in software have proven especially challenging for old incumbents—they require software engineers who emphasize flexibility and creativity, which differ from the traditional core values of hardware engineers (whereas advancements in hardware were easy to integrate into their existing systems). Numerous news reports attest to the difficulty for legacy car makers to adapt to the different workstyles associated with software (George, 2022; Davis, 2023). Accordingly, we present a model that lays out how new technologies associated with larger changes in workstyles can lead to more severe organizational frictions, especially in old firms. We then develop a new and theoretically-grounded measure of workstyle changes induced by new technologies, which allows us to perform large-scale empirical tests about the performance of young vs old firms.

In our model, firms initially operate an existing business model A, which has a continuum of business units, each having a manager who is in charge of hiring its workers. The occupation composition among workers in A is pinned down by the productivity of the occupations given the technology associated with that business model. Later, a new business model B emerges, which has its own occupation-specific productivity that determines the desired occupation composition for B. Each occupation has a set of workstyles (i.e., characteristics it emphasizes), so the overall (employment-weighted average) workstyle under the new business model B can differ from that under the existing business model A.

When differences in workstyle are larger, there is a higher probability that workers in business model B need to solve a problem differently from what business model A is used to. This leads to a "conflict," in which case managers in business model B need to justify their approach, in light of existing rules laid down by business model A that codify how to solve problems according to business model A's past experiences. Doing so takes time and reduces the productivity of business model B, which shrinks its optimal size. Older firms have more existing rules, so the time and productivity losses are greater. Therefore, they end up with a smaller operation of business model B. Accordingly, they grow more slowly when new technologies require larger changes in workstyle.

To validate that older firms indeed have more rules and more meetings that waste time, we process employee reviews from Revelio to measure the extent to which employees mention the presence of rules and the excess of meetings. Across different companies at a given point in time, we find that the fraction of reviews mentioning the presence of rules and the fraction of reviews mentioning the excess of meetings are positively correlated with firm age.

We build on the model to develop a new measure of workstyle changes in an industry induced by new technologies, which can be applied systematically across industries and over time. We proceed in two steps. First, we predict future occupation composition in light of new technologies in an industry following Kogan et al. (2024). The key idea is to capture the exposure of the occupations' routine and nonroutine tasks to new technologies measured through patents, based on the textual similarity between occupation task descriptions and patent text. As Kogan et al. (2024) show, when new technologies in an industry are more similar to an occupation's routine (nonroutine) tasks, substitutability (complementarity) is stronger, and the employment of the occupation falls (rises) going forward (using data on employment by industry and occupation from the Bureau of Labor Statistics). Correspondingly, the similarity between patents in an industry-year and an occupation's tasks allows us to predict the occupation's future employment in the industry. Second, once we have future industry-level occupation composition predicted by new technologies, we can obtain the associated (employment-weighted average) overall future workstyle in the industry, relative to the current overall workstyle (using data on the workstyle associated with each occupation from O*NET). The predicted industry-level workstyle change (over the next five years) induced by the new technologies is the key independent variable in our empirical analyses. The reliance of the measurement on BLS and O*NET data restricts our sample period to 2003 onward.

We use three sets of data to test the core hypothesis that young firms grow more than old firms when new technologies are associated with greater changes in workstyle. First, we investigate venture capital (VC) investment volume in an industry-year, which captures the forward-looking valuation of startups that can reflect their growth potential. We find that a one standard deviation increase in the predicted industry-level workstyle change is associated with around 0.4 log points higher VC investment in the industry, controlling for the log of total market capitalization of Compustat firms in the industry to capture other factors that can affect the prospects of firms in the industry. Importantly, the volume of new technologies per se—measured as the total number of patents, breakthrough patents (Kelly et al., 2021), creative patents (Kalyani, 2025), or rapidly evolving patents (Bowen, Frésard, and Hoberg, 2023) in the industry—does not have a significant relationship with VC investment, or affect the coefficient on workstyle change—our key variable of interest.

Second, we investigate variation among Compustat firms, including their valuation, as well as realized sales growth and employment growth over the next five years. Specifically, we regress firm-level Tobin's Q or subsequent growth on firm age interacted with the predicted industry-level workstyle change (based on the new technologies in the industry as before). We find that young firms have significantly higher valuation and realized future growth relative to old firms when the technology-induced workstyle change is larger. Meanwhile, the number of total patents, breakthrough patents, creative patents, or rapidly evolving

patents in the industry is not associated with significantly different outcomes for young vs old firms. In other words, old firms are not invariably incapable when the amount of new technologies is greater, or when the new technologies are more important, influential, or creative. We also show that the effects of firm age are not due to age being correlated with size, or due to older firms being profitable so they suffer more existing businesses being replaced and cannibalized (Arrow, 1962).

Third, we investigate the population of firms in the Census Business Dynamics Statistics (BDS) dataset, which has broad coverage (although many young firms in this case may not be aspirational entrepreneurship). BDS provides total employment of firms by age group: 0, 1-5, 6-10, 11-15, 16-20, 21-25, and the remaining age groups cannot be consistently defined over our sample period because firms' precise age is unknown if they are born before 1976. Therefore we restrict to firms with age groups between 1-5 and 16-20 (which become age groups 6-10 to 21-25 five years later). We find that the young age groups grow significantly faster than the older ones when the technology-induced workstyle change is larger. Again, the number of patents is not associated with significantly different outcomes for young vs old firms.

Furthermore, we show that the main results hold if we measure technology-induced workstyle change using Wikipedia articles instead of patents. Specifically, we capture new technologies using Wikipedia articles on technological inventions, and predict future occupation employment in an industry based on the exposure of the routine and nonroutine tasks, measured using the similarity between occupation task descriptions and Wikipedia articles' text (rather than patent text). All of our main results hold using predicted workstyle change based on Wikipedia articles instead. Additionally, we can estimate the predicted workstyle change associated with each Wikipedia article, and independently query LLM about whether the technological invention in a given Wikipedia article is most successfully implemented by an old or young (less than ten years old) firm.[1] We find that technological inventions associated with larger workstyle change (due to their predicted effects on occupation composition) are significantly more likely to be implemented by young firms according to the independent LLM query.

Taken together, these results support the importance of organizational frictions in shaping companies' adaptability. They show that new technologies per se do not necessarily challenge old firms. Incumbents are not ubiquitously incompetent in light of new technologies due to universal lack of learning or fear of cannibalization. However, when new technologies require changes in organizational workstyles, old firms struggle and young firms rise, in which case venture investment and government policies that facilitate entrepreneurship can be especially useful. Both our model and our empirical analyses on the growth of young vs old firms focus on the extent to which companies can *implement* new technologies, and we take

[1]This exercise is much harder for patents which are more micro and technical, so we cannot easily ask LLM to determine whether the technologies associated with an individual patent are implemented by an old or young firm.

the emergence of new technologies as given (e.g., artificial intelligence emerges from decades of research spanning academic and commercial domains, and information technology has roots in both publicly and privately funded research). Once the new technology emerges, in principle any company can work to use it. The implementation of new technologies is foundational for the process of creative destruction and for business dynamism.

Literature Review Our work relates to three sets of literature. First, we contribute to research on creative destruction and the implementation of new technologies. We perform systematic analyses of the performance of young vs old firms in the face of new technologies, and highlight the role of organizational challenges. Prior work often sidesteps whether an entrant using a new technology is a new firm or an incumbent firm (Aghion and Howitt, 1992). Some analyses emphasize that young firms are important for technological advancement (Decker et al., 2014; Loderer, Stulz, and Waelchli, 2017; Acemoglu et al., 2018; Draghi, 2024; Ewens and Marx, 2024), while others show that incumbents contribute meaningfully (Garcia-Macia, Hsieh, and Klenow, 2019; Cohen, Gurun, and Nguyen, 2022; Braguinsky et al., 2024). A recent study by Bowen, Frésard, and Hoberg (2023) shows that a greater quantity of "rapidly evolving" patents is associated with startups exiting through initial public offering—thus remaining independent—instead of selling out. Caskurlu, Hoberg, and Phillips (2024) focus on firm size and show that a rise of new technologies highly correlated across multiple industries is associated with faster growth by small firms. Our focus is on the importance of organizational frictions for understanding the implications of technological change for young vs old firms.

Second, we contribute to research on the nature of the firm. Many studies since Coase (1937) have focused on vertical integration (Williamson, 1971; Grossman and Hart, 1986; Hart and Moore, 1990; Aghion and Tirole, 1994; Anton and Yao, 1995), using insights of incomplete contracts and property rights,[2] but other considerations can also shape the activities that firms engage in (Atalay, Hortaçsu, and Syverson, 2014). Our work concerns another dimension of the boundary of the firm: which ideas can be implemented within a firm? We highlight challenges incumbents face that are not directly related to vertical integration, and point to a different set of mechanisms that restrict what a firm can do. In this regard, the most closely related work is Hart and Holmstrom (2010) on firm scope, which builds on the framework of Hart and Moore (2008) where authority—rather than bargaining a la Hart and Moore (1990)—plays a central role in resolving ex post conflicts. The key concept in Hart and Holmstrom (2010) is shading, an action motivated by one party's grievance that imposes negative externalities on others,

[2]Strictly speaking, the core takeaway for Grossman and Hart (1986), which highlights that asset ownership can be viewed as residual control rights, can apply to lateral integration too. Rajan and Zingales (2001) is an example which studies firm's organization structure based on the framework of property rights, showing that flat hierarchies feature more distinctive technologies (than steep ones).

which becomes more severe with the integration of different business lines. This mechanism resembles the frictions between existing and new business models in our framework. However, instead of shading, we focus on conflicts among workers and managers when the existing business model insists on legacy rules that emerge endogenously from the organization's dynamic structure. In the appendix, we show that our key prediction continues to hold under a more traditional setting in which conflicts arise between different business divisions under the same headquarters. This formulation draws on the insight of Hart and Moore (2008), where ex post deviations of realized payoffs from the preferred outcome specified in an ex ante contract lead to shading by division managers.[3]

Third, we contribute to research on organizational economics. The extent to which incumbent firms can implement new technologies is a classic question in management, which has produced a large volume of qualitative assessments and cases (Penrose, 1959; Nelson and Winter, 1985; Tushman and Anderson, 1986; Henderson and Clark, 1990; March, 1991; Leonard-Barton, 1992; Teece, Pisano, and Shuen, 1997; Christensen, 1997; O'Reilly and Tushman, 2021). We draw on their insights to develop formal theoretical modeling and systematic empirical analyses. This research question is also related to Holmstrom (1989), who postulates that large firms may be disadvantaged in innovation if combining hard-to-measure activities (e.g., innovation) with easy-to-measure activities (e.g., routine tasks) is costly due to misallocation of attention and effort, using the complete contracting framework of Holmstrom and Milgrom (1991). Relatedly, Chassang (2010) and Garicano and Rayo (2016) show that routines can endogenously arise in environments governed by relational contracts, as proposals by an informed party to deviate may be indistinguishable from that party's rent-seeking. As a result, routines can lock organizations into suboptimal practices, much like the rules governing the incumbent business model A in our framework. Our work differs from these papers by emphasizing the role of workstyle differences and by providing accompanying empirical evidence. Another set of models study coordination and adaptation in organizations, emphasizing how firms use local information in different settings (Milgrom and Roberts, 1992);[4] our focus is organizational conflicts even without informational frictions. Finally, our modeling relates to work on knowledge hierarchy (Garicano, 2000; Garicano and Rossi-Hansberg, 2004). Our mechanism incorporates the importance of managerial time, but our focus is the time loss from the conflict between business models, rather than time saving from the knowledge hierarchy.

[3] Our paper is also related to the literature on internal capital markets following the incomplete contracting tradition (Shleifer and Vishny, 1989; Meyer, Milgrom, and Roberts, 1992; Stein, 1997; Scharfstein and Stein, 2000; Rajan, Servaes, and Zingales, 2000), which typically focuses on the conflict between division managers and firm headquarters. In comparison, the main focus of our model is conflicts arising from workers in different business units.

[4] A long strand of the organizational economics literature emphasizes the role of local information and corresponding agency frictions across layers of decision makers, including formal vs real authority (Aghion and Tirole, 1995, 1997), corporate culture (Gorton and Zentefis, 2024), the design of functional units (Qian, Roland, and Xu, 2006; Dessein, Garicano, and Gertner, 2010), and adaptation to changing environments (Dessein and Santos, 2006).

The rest of the paper proceeds as follows. Section II presents stylized facts about the extent to which new technologies are implemented by new vs existing firms. Section III provides a model to motivate and organize our main empirical analyses. Section IV explains the data and measurement. Sections V and VI show the main empirical results. Section VII concludes.

II New Technologies and New Firms: Some Stylized Facts

In this section, we present stylized facts to show that significant technological inventions are not invariably implemented by young firms. Old firms can implement new technologies, but not always. In other words, some new technologies may favor young firms, but others can be done by old firms. The fact that the implementation of new technologies is split between young and old firms motivates our subsequent analyses that investigate the circumstances where new technologies favor new firms.

II.A Technological Inventions Implemented by Young vs Old Firms

To obtain an intuitive sense of the extent to which new technologies are implemented by new firms, we proceed as follows. First, we collect a list of technological inventions over the 20th century by filtering Wikipedia articles, following Asirvatham (2024) and Kalyani et al. (2025) who showed their usefulness for building catalogs of technologies.[5] Second, we use a Large Language Model (LLM) to summarize basic information about these inventions, including the time, location, and type of inventor (i.e., private company, individual inventor, government, or university/non-profit), with the prompt shown in Appendix Section IA2.2. We also ask the LLM to report whether the company that was most successful in its initial implementation and commercialization was a young firm (less than ten years at the time) or an old incumbent, with the prompt shown in Appendix Section IA2.3. As explained in the introduction, our work focuses on implementation, not simply the invention. Indeed, several well-known examples show that even when incumbents are able to invent new technologies, organizational rigidity can prevent them from effective implementation (e.g., Bell Labs invented transistors but did not commercialize them, Xerox developed personal computing technologies but did not commercialize them). Ultimately, new technologies need to be implemented to have an impact.

Figure 1 takes the technological inventions in the 20th century from the Wikipedia-based dataset, and plots the fraction that is most successfully implemented by young vs old firms. The first two bars

[5] We download Wiki articles from Wikimedia English Wikipedia Dumps and use the steps in Appendix Section IA2.1 to filter for technological inventions.

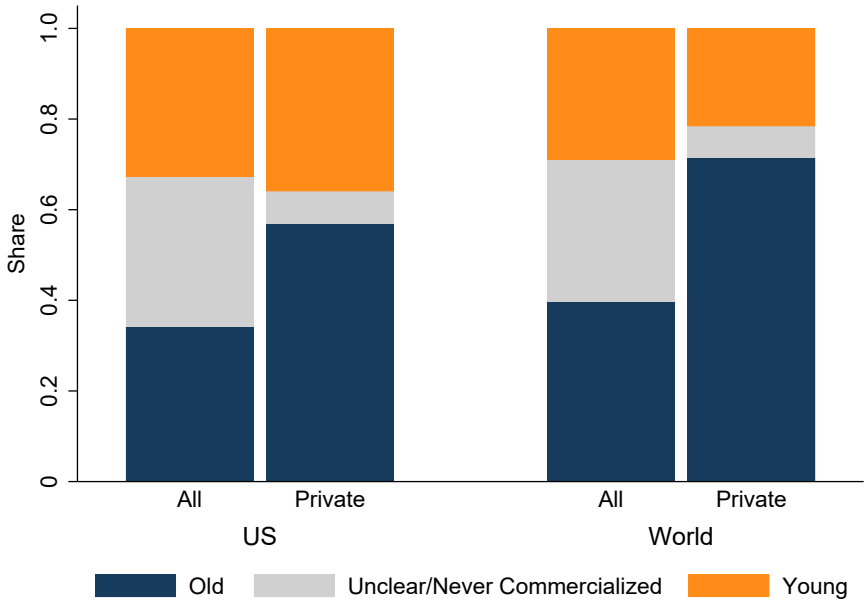

Figure 1. New Technologies Implemented by Young vs Old Firms (1900–2000)

Notes: This figure shows the fraction of technological inventions between 1900 and 2000 that are most successfully implemented by young firms (less than ten years) vs old firms (otherwise). The first two bars restrict to technological inventions in the U.S., where the left bar includes all inventions and the right bar includes only inventions by private companies. The second two bars show all technological inventions in the world, where the left bar includes all inventions and the right bar includes only inventions by private companies. The shares are normalized to one within each group. Colors correspond to firm age categories.

restrict to technological inventions in the U.S., where the left bar ("All") includes all inventions and the right bar ("Private") includes only inventions by private companies. The LLM cannot clearly determine the answer or identify a commercializing firm for a larger share of inventions by individuals, universities, and governments (included in the "All" bar but not in the "Private" bar) compared to inventions by private companies, which is understandable. The second two bars show the corresponding results for all technological inventions in the world.

Overall, we see a mix. Both young firms and old firms play a meaningful role in the implementation of new technologies. In the U.S., young firms have implemented close to 40% of new technologies among inventions by private companies, which is higher than their employment share according to Census Business Dynamics Statistics (BDS) dataset. This feature is consistent with the observation that young firms are more innovative than old firms (Kortum and Lerner, 2000; Ewens and Marx, 2024). Nonetheless, old firms are not entirely incapable of implementing new technologies (e.g., due to fear of cannibalization).

II.B Some Classic Examples

Business school classes have analyzed a number of cases featuring new technologies that incumbent firms succeeded or failed to implement. For example, Nylon (and polymers more generally) was invented and most successfully commercialized by DuPont, when the company was already more than 100 years old (Ndiaye, 2007). Antibiotics were most successfully commercialized by established companies such as Pfizer and Merck, which had existed for decades by that time. Television sets and jet engines were also produced most extensively by companies such as General Electric that had been founded long before their invention. More recently, Apple released the iPhone in 2007 when the company was over 30 years old.

Conversely, cars, aircraft, and semiconductors were most successfully produced by new companies at the time of their invention. Software was also mainly developed and commercialized by new companies, despite existing companies such as General Electric repeatedly trying to enter this domain (Cohan, 2022). Discount retail and discount airlines—which represent new business models though not necessarily technological inventions—are also associated with new firms (O'Reilly and Tushman, 2021).

A common view in business school classes is that the former group largely represents new technologies that were compatible with existing organizational processes. Nylon had similarities with the production of rayon for DuPont; antibiotics benefited from large-scale fermentation that Pfizer and Merck already used; TVs and jet engines built on GE's existing capabilities in manufacturing generators, turbines, and appliances; and iPhones follow from the production of other Apple devices. The latter group, however, led to organizational processes that differed from those in existing companies. The production of cars, aircraft, and semiconductors required a workflow that deviated from the procedures of their predecessors; software emphasized flexibility and maximizing upside, whereas hardware focused on attention to detail and minimizing downside; discount retailers and discount airlines relied on volume, whereas traditional retailers and airlines relied on margin.

How to capture the extent to which new technologies entail new organizational processes or priorities? These features are rich and difficult to measure in a uniform way. Our entry point is to extract information from the types of occupations required to implement new technologies. For example, as technological advancement leads to an increasing reliance on software in car manufacturing, the corresponding organizational processes will have to change given the different attributes of software vs hardware engineering that we later formalize as workstyle, which can be reflected by the shift in the occupation composition from hardware engineers to software engineers. Changes in the composition of occupations with different workstyles thus provide a window that can reveal corresponding changes in organizational processes. This angle has some limitations: subtle process changes may not be reflected in occupational composition, and

the reliance on such data confines our main empirical analysis to the past two decades. However, this approach allows us to develop measurement methods that apply broadly across industries, with which we can perform systematic empirical tests.

III Model

We set out our central hypothesis in a model with four ingredients. First, there is a continuum of firms in a single industry that produce under an existing business model, A. Second, within each firm, an owner hires managers, who in turn hire workers; workers are differentiated by occupation. Third, as a firm operates under business model A, it accumulates a stock of rules that codify how to handle the operational problems it has previously encountered. Fourth, a new business model, B, arrives at a single, industry-wide moment in time, and its different occupation composition exposes workers to a different mix of problems. As a result, business model B operates under a different set of organizational conventions, while the existing stock of A-rules can either substitute for, or interfere with, the construction of B-rules. The model formalizes the hypothesis introduced in the introduction, builds on the stylized facts of Section II, and motivates the empirical specifications taken to data in Section V.

III.A Setup

Industry, firms, business models Consider a single industry with a unit mass of firms, indexed by i. Each firm has an exogenous spanning capacity $\xi_i > 0$, summarizing its ability to operate many units in parallel, with industry total $\Xi \equiv \int_i \xi_i \, di$. The distribution of ξ_i is independent of other firm-level objects, in particular firm age.

There are two business models, indexed by $X \in \{A, B\}$. Business model A is the existing model while Business model B is new—it becomes available at a single, industry-wide Poisson time at which every firm has the option to expand into B. In the analysis below we focus on instantaneous responses to the arrival of B, taking the cross-section of firms at that moment as given.

Occupations and workstyles There are J occupations, indexed by $j \in \{1, \ldots, J\}$. Each occupation j is characterized by a productivity $\psi_{X,j}$ in business model X (so productivity in occupation j may differ between A and B). The productivity vectors $\{\psi_{X,j}\}_{j=1}^{J}$ are also exogenous, capturing the technology of each business model $X \in \{A, B\}$. To each occupation is also associated a workstyle vector $ws_j \in \mathbb{R}^D$ (which is independent of the business model). The ws_j vectors are exogenous, time-invariant primitives; in the data, $D = 16$ corresponds to the O*NET workstyle dimensions.

Operating units, managers, workers Each firm i in business model X operates $L_{X,i} \geq 0$ identical operating units, indexed by k. Each unit is run by one manager, who hires a bundle of workers from the J occupations to staff the unit. The owner of the firm chooses how many units of each business model to operate. The manager of each unit chooses the within-unit occupation composition. The friction in our microfoundation of Section III.C originates at the worker level, and is rooted in conflicts between the rules that govern how different business units respond to operational problems. But it also has a managerial dimension, since resolving these conflicts requires escalation to and arbitration by managers. The manager tier therefore plays two roles: it decentralizes occupation choice, keeping the workstyle distance Δ an industry-level object rather than a firm-specific one, and it is the level at which rule conflicts between business models are ultimately adjudicated.

Wages and profits Workers in any occupation are paid the exogenous wage w, which we normalize to $w = 1$ (numeraire). The total number of workers that can be employed in any operating unit is $\bar{l} = 1$, also a normalization. Managers in business model X are paid the wage $w_{X,t}$, which is endogenously determined by the clearing of the industry-level X-manager market. Per-unit baseline profit in business model X is

$$z_{X,t} \equiv Q_X - 1 - w_{X,t}, \tag{1}$$

where Q_X is the unit-level output of business model X (which will be time-invariant, as we explain shortly), and the wages of X-managers $w_{X,t}$ are derived in Sections III.B and III.C and may vary over time.

III.B Business model A

We first describe the firm's optimization under business model A, before B is available.

Manager's problem Within each A-unit, the manager allocates a unit measure of workers across the J occupations to maximize the unit's CES output,

$$Q_{A,i,k} = \left(\sum_{j=1}^{J} \psi_{A,j}^{\frac{1}{\sigma}} \left(l_{A,i,k,j} \right)^{\frac{\sigma-1}{\sigma}} \right)^{\frac{\sigma}{\sigma-1}}, \qquad \sum_{j=1}^{J} l_{A,i,k,j} = 1, \tag{2}$$

where σ is the elasticity of substitution across occupations. The solution is the standard CES one, with occupation shares and unit output common to all A-units in the industry:

$$l^*_{A,i,k,j} = \frac{\psi_{A,j}}{\sum_{j'=1}^{J} \psi_{A,j'}}, \qquad Q_{A,i,k} = Q_A \equiv \left(\sum_{j=1}^{J} \psi_{A,j} \right)^{\frac{1}{\sigma-1}}. \tag{3}$$

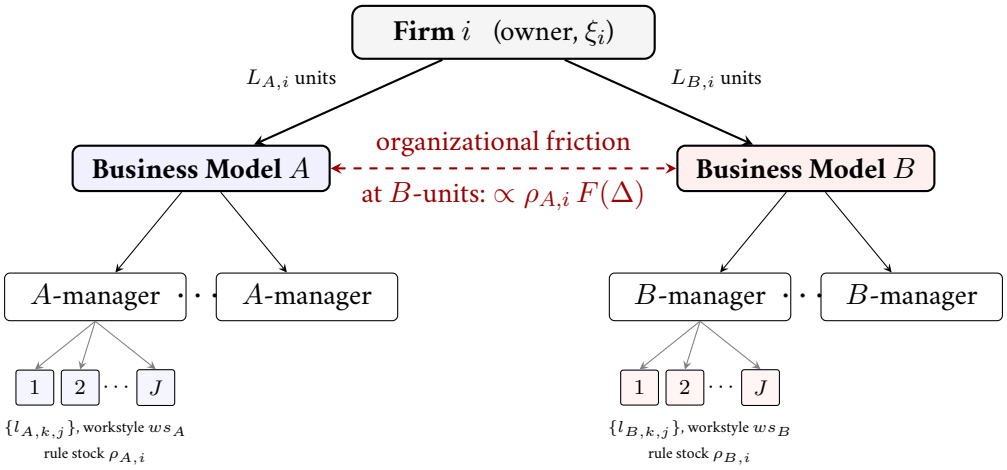

Figure 2. Model Structure

Notes: Firm i has spanning capacity ξ_i and operates under two business models: a pre-existing model A and a new model B. The owner chooses the number of operating units of each business model, $L_{A,i}$ and $L_{B,i}$. Each unit is led by a manager (an A- or B-manager) who hires workers across J occupations; $\{l_{A,k,j}\}$ and $\{l_{B,k,j}\}$ denote the employment of occupation j in unit k under each business model. The employment-weighted average workstyles are ws_A and ws_B, with workstyle distance $\Delta = \|ws_B - ws_A\|$. Each business model accumulates a rule stock, $\rho_{A,i}$ and $\rho_{B,i}$, that codifies the firm's procedures for resolving operational problems. When a problem hits a B-unit and no B-rule covers it, B-workers may consult the A-rulebook: with probability $1 - F(\Delta)$ the A-rule works, and with probability $F(\Delta)$ it conflicts. The conflict probability $F(\Delta)$ is increasing in workstyle distance.

Workstyle of business model A Define the workstyle of an A-unit as the share-weighted average of occupation workstyles:

$$ws_{A,i,k} \equiv \sum_{j=1}^{J} \frac{l^*_{A,i,k,j}}{\sum_{j'=1}^{J} l^*_{A,i,k,j'}}\, ws_j = \sum_{j=1}^{J} \frac{\psi_{A,j}}{\sum_{j'=1}^{J} \psi_{A,j'}}\, ws_j \equiv ws_A. \tag{4}$$

By (3), $ws_{A,i,k} = ws_A$ is identical across all units of all firms in the industry. The workstyle of business model A is therefore an industry-level object; it will play a central role when business model B arrives and the workstyle distance between A and B governs organizational frictions.

Rule accumulation under A The firm faces a stream of operational problems of distinct types (e.g., scheduling clashes, resource allocation conflicts, machine breakdowns, customer escalations, and so on). There is a continuum of distinct problem types, indexed by $\tau \in [0, R]$, with $R \in \mathbb{R}^+$. The firm encounters problems as a Poisson process at rate $\mu^p > 0$, with the type τ drawn uniformly from $[0, R]$ each time.

A rule for problem type τ is a codified procedure for resolving incidents of that type. The first time the firm encounters an uncovered type, its workers improvise a resolution at a one-shot downtime cost $\delta_p > 0$, and the procedure they devise is codified into a firm-level rule, available to all A-units thereafter at no further cost. Subsequent encounters of the same type are resolved at no cost using the existing rule.

Let $\rho_{A,i,t} \in [0,1]$ denote the fraction of problem types covered by firm i's A-rulebook at time t. Defining $\kappa \equiv \mu^p / R > 0$ as the rate at which the rule-coverage fraction grows when no rules yet exist, the law of motion for the rule stock is:

$$d\rho_{A,i,t} \;=\; \kappa\,(1-\rho_{A,i,t})\,dt, \qquad \Rightarrow \qquad \rho_{A,i,t} \;=\; 1 \;-\; e^{-\kappa\, a_{i,t}}, \tag{5}$$

where $a_{i,t}$ is firm i's age at time t. Rule coverage rises monotonically with age, with the rate of accumulation slowing as the firm has already learned more.[6] Section III.D returns to this prediction empirically.

Per-A-unit profit The expected per-unit downtime cost in business model A is the product of the problem arrival rate, the probability the problem is uncovered, and the per-incident cost: $\mu^p\,\delta_p\,(1-\rho_{A,i,t})$. Writing $\phi \equiv \mu^p\,\delta_p > 0$ for the maximum downtime cost per unit (when no rules exist), the expected per-A-unit flow profit is

$$\pi_{A,t}(\rho_{A,i,t}) \;=\; z_{A,t} \;-\; \phi\,(1-\rho_{A,i,t}), \tag{6}$$

where the time subscript records that this profit varies over time through the equilibrium A-manager wage $w_{A,t}$ embedded in $z_{A,t}$. We assume $\pi_{A,t} > 0$ for all $\rho_{A,i,t} \in [0,1]$ and at every date.[7] At birth, $\pi_{A,t}(0) = z_{A,t} - \phi$; at full coverage, $\pi_{A,t}(1) = z_{A,t}$. Older firms are more efficient at running A-units because they have learned to handle a larger fraction of routine problems without improvisation.

Owner's problem under A The owner of firm i chooses the number of A-units $L_{A,i,t}$ to maximize total flow profit, net of a quadratic spanning cost:

$$\max_{L_{A,i,t}\,\geq\,0} \quad \pi_{A,t}(\rho_{A,i,t})\,L_{A,i,t} \;-\; \frac{L_{A,i,t}^2}{2\,\xi_i}. \tag{7}$$

Firms with higher spanning capacity ξ_i pay a smaller marginal cost of an additional unit, so the first-order condition sets the number of units to the per-unit profit scaled by spanning capacity,

$$L^*_{A,i,t} \;=\; \pi_{A,t}(\rho_{A,i,t})\,\xi_i. \tag{8}$$

[6] The arrival rate μ^p is firm-level rather than per-unit, so rule accumulation depends only on firm age and not on firm size. Moreover, implicit in the per-unit form $\phi(1-\rho_A)$ is the assumption that each uncovered problem disrupts all A-units simultaneously, so that the firm's total downtime cost per uncovered problem is $L_{A,i,t}\,\delta_p$. An alternative microfoundation in which each operating unit independently encounters problems and pays its own downtime cost yields qualitatively similar cross-sectional predictions but loses the closed-form structure of the industry equilibrium.

[7] In Appendix IA6.1 we show that the single parameter restriction $L/\Xi > \max\{1,\chi\}\,\phi$ is sufficient to maintain $\pi_{A,t} > 0$ for all $\rho_A \in [0,1]$ at every date and, after B becomes available, $\pi_{B,t} > 0$ for all $(\rho_A,\rho_B) \in [0,1]^2$ at every date as well. When escalation costs are no larger than improvisation costs ($\chi \leq 1$), the condition reduces to $L/\Xi > \phi$.

Firm size in business model A rises in spanning capacity ξ_i and in rule coverage $\rho_{A,i,t}$, the latter of which increases with age.

Equilibrium wage of A-managers We close the model by clearing the industry-level A-manager market: $\int_i L^*_{A,i,t}\, di = L$.[8] Aggregating (8) across firms and using the fact that the spanning cost ξ_i and firm age are independent, the equilibrium per-A-unit profit is

$$z_{A,t} \;=\; \frac{L}{\Xi} \;+\; \phi\,(1-\bar{\rho}_{A,t})\,, \qquad \text{where} \quad \bar{\rho}_{A,t} \;\equiv\; \int_i \rho_{A,i,t}\, di. \tag{9}$$

The equilibrium A-manager wage is then $w_{A,t} = Q_A - 1 - z_{A,t}$. The intuition for why $z_{A,t}$ rises with the friction $\phi\ (1-\bar{\rho}_{A,t})$ is as follows. For each firm, the share of uncovered problems, $(1-\rho_{A,i,t})$, leads to operating costs that are borne by each A-unit. This depresses its willingness to pay for A-managers. In equilibrium, this pushes the managers' wage $w_{A,t}$ down. Because $z_{A,t} = Q_A - 1 - w_{A,t}$, this lower wage translates into a higher per-unit baseline profit $z_{A,t}$. Thus the effects of low rule coverage are borne not by firms but by A-managers in the form of lower wages.[9]

Substituting (9) back into (8) expresses firm size in business model A as

$$L^*_{A,i,t} \;=\; \left[\frac{L}{\Xi} \;+\; \phi\left(\rho_{A,i,t} - \bar{\rho}_{A,t}\right)\right]\xi_i. \tag{10}$$

Firms with bigger spanning capacity (ξ_i), and firms with above-average rule coverage $(\rho_{A,i,t} > \bar{\rho}_{A,t})$, have higher per-unit profitability, and therefore more units.

III.C Business model B

A new business model B becomes available at a single, industry-wide Poisson time. From this moment on, each firm may add B-units alongside its existing A-units. We focus in what follows on firms' taking the time $t = \tau$ of arrival as given, which we normalize to $\tau = 0$ for clarity.

Manager's problem Each B-unit's manager solves the same within-unit problem as under A, now with the B-productivities $\{\psi_{B,j}\}$ in place of the A-productivities. The solution is again common to all units

[8] The exogenous supplies of managers in business models A and B are normalized to $L_A = L_B \equiv L$. This normalization is purely a convenience and simplifies several of the cross-sectional expressions below.

[9] The average problem coverage $\bar{\rho}_{A,t}$ can be explicitly solved for under entry and exit dynamics for firms within an industry. We do not pursue this extension here as it does not affect our subsequent mapping from the model's predictions to the data.

and firms: unit output is $Q_B \equiv (\sum_j \psi_{B,j})^{1/(\sigma-1)}$, and the associated unit workstyle is

$$ws_B \equiv \sum_{j=1}^{J} \frac{\psi_{B,j}}{\sum_{j'=1}^{J} \psi_{B,j'}} ws_j, \tag{11}$$

the same at every B-unit in the industry.

Workstyle distance between A and B Define the workstyle distance between models A and B as:

$$\Delta \equiv \left\| ws_B - ws_A \right\| \geq 0, \tag{12}$$

where $\|\cdot\|$ denotes the Euclidean norm in $\mathbb{R}^D$. By (4) and (11), Δ captures both within-firm and industry-level distance in workstyle, since every unit in each firm operates using the same labor bundle in each of the two business models. The workstyle distance Δ is determined entirely by the technology vectors $\{\psi_{A,j}\}_{j=1}^{J}$ and $\{\psi_{B,j}\}_{j=1}^{J}$ and the workstyle matrix $\{ws_j\}_{j=1}^{J}$, both primitives of the model.

Conflict friction: a microfoundation Rules accumulated under business model A may not be usable by workers in units of business model B, in proportion to the workstyle distance Δ, and this leads to productivity losses at type-B units. The microfoundation is as follows. After B becomes available, B-units encounter operational problems at the same Poisson rate μ^p as A-units. Problems are drawn from the same set $[0, R]$ as A, so that there is potential overlap between the problems that a firm has encountered under the old business model and those it encounters in the new business model.

As before the firm builds a rulebook for its type-B units, and let $\rho_{B,i,t} \in [0, 1]$ denote the fraction of problem types in $[0, R]$ covered by firm i's B-rulebook at time t. When a problem of a given type arrives at the firm's B-side, three cases arise:

(i) With probability $\rho_{B,i,t}$, the firm's B-rulebook already covers the problem type. In this case, the rule is applied at no cost.

(ii) With probability $(1 - \rho_{B,i,t})\rho_{A,i,t}$, the B-rulebook does not cover the problem but the A-rulebook does. In this case, we assume that B-workers consult the A-rulebook. Two sub-cases arise:

(ii.a) The A rule does not conflict with the workstyle of business units of type B. We assume this happens with exogenous probability $1 - F(\Delta)$. (See the economic discussion of $F(\Delta)$ shortly.) In this case, the A rule is applied at no cost by B workers, and is codified into the B-rulebook.[10]

(ii.b) The A rule conflicts with the workstyle of business units of type B, and cannot be applied by B

[10] We will take this to mean: when B-workers consult the existing A-rule for the problem at hand, the proposed solution is consistent with the B-style way of doing things in a fraction $1 - F(\Delta)$ of cases.

workers. This occurs with exogenous probability $F(\Delta)$. In this case, B-workers must do two things before they can operate. First, the entrenched A-rule must be set aside: B-managers escalate the conflict and spend time persuading A-managers that the existing rule should not apply, at an *escalation cost* $\delta_c \equiv \chi\,\delta_p$, where $\chi > 0$ scales the escalation cost relative to the cost of improvising a fresh solution. Second, having cleared the conflict, B-workers improvise a B-style solution at downtime cost δ_p, and the new rule is added to the B-rulebook. The total downtime cost is therefore $\delta_p + \delta_c = (1+\chi)\,\delta_p$.

(iii) With probability $(1-\rho_{B,i,t})(1-\rho_{A,i,t})$, neither rulebook covers the problem. B-workers improvise from scratch at downtime cost δ_p, and the new rule is added to the B-rulebook. There is no A-rule to set aside, so no escalation cost is incurred.

The function $F(\Delta) \in [0,1]$ is the conflict probability: the probability that an existing A-rule, when applied to a B-style problem, is incompatible with the B-workers' way of operating. We assume $F(\Delta)$ is increasing and differentiable in Δ with $F(0) = 0$.

The escalation cost χ in case (ii.b) creates departures from a model in which A rules always have net positive spillovers for B units. An existing A-rule that conflicts is not merely useless, but, when $\chi > 0$, it is potentially *costly*, because overriding it consumes managerial time. Conflicts are resolved at every B-unit, so the per-B-unit downtime is δ_p in case (iii) and $(1+\chi)\,\delta_p$ in case (ii.b), consistent with the all-units-disruption assumption maintained in Section III.B.

The expected per-B-unit downtime is $\mu^p \delta_p = \phi$ in case (iii) and $\mu^p(1+\chi)\delta_p = (1+\chi)\phi$ in case (ii.b). Multiplying each case by its probability of arising, the expected per-B-unit flow profit is:

$$\pi_{B,t}(\rho_A,\,\rho_B;\,\Delta) \;=\; z_{B,t} \;-\; \underbrace{\phi\,(1-\rho_B)(1-\rho_A)}_{\substack{\text{problems never encountered}\\ \text{in } A \text{ or in } B}} - \underbrace{\phi\,(1-\rho_B)\rho_A\,(1+\chi)F(\Delta)}_{\substack{\text{problems encountered}\\ \text{in } A \text{ but not in } B}}. \tag{13}$$

As with $\pi_{A,t}$, the time subscript records the time-varying profit via the equilibrium B-manager wage $w_{B,t}$ embedded in $z_{B,t}$. We assume $\pi_{B,t} > 0$ for all $(\rho_A, \rho_B) \in [0,1]^2$ for all $t \geq 0$.

Three properties of (13) are worth highlighting. First, when workstyles coincide ($\Delta = 0$), Equation (13) reduces to $z_{B,t} - \phi(1-\rho_B)(1-\rho_A)$. Existing A-rules substitute fully for B-rules: an old firm with $\rho_A = 1$ effectively starts B with a complete rulebook. Rules can be translated from the A to B without friction, and the productivity gap between B units of high-ρ_A and low-ρ_A firms is the largest.

Second, when workstyles differ ($\Delta > 0$), it is convenient to define the threshold

$$F^* \;\equiv\; \frac{1}{1+\chi} \in (0,1), \quad \text{for any } \chi > 0. \tag{14}$$

In terms of F^*, the marginal value of an additional unit of ρ_A for B-profit is

$$\frac{\partial \pi_{B,t}}{\partial \rho_A} \;=\; \phi\,(1-\rho_B)\left(1-\frac{F(\Delta)}{F^*}\right). \tag{15}$$

When $F(\Delta) < F^*$, an additional A-rule raises B-profit: existing rules are a *net asset*, because they more often transfer successfully (saving an improvisation) than conflict (triggering an escalation). When $F(\Delta) > F^*$, the marginal value turns negative: existing rules become a *net liability*, because a conflicting A-rule costs the firm an escalation δ_c *on top of* the improvisation it would have paid anyway had no rule existed. As $\chi \to 0$ the escalation channel vanishes, $F^* \to 1$, and A-rules are always a net asset: they help for any $F(\Delta) < 1$.

Third, the cross-derivative of profits with respect to existing rules and workstyle distance is negative:

$$\frac{\partial^2 \pi_{B,t}}{\partial \rho_A\,\partial \Delta} \;=\; -\,\frac{\phi}{F^*}\,(1-\rho_B)\,F'(\Delta) \;\le\; 0. \tag{16}$$

The marginal benefit of additional A-rules at B-units declines with the workstyle distance. Equivalently, the productivity gap between high-ρ_A and low-ρ_A firms is smaller in industries undergoing larger workstyle change. This is the structural force behind the central empirical prediction in Section III.E, and holds regardless of whether $\chi = 0$ (no escalation costs) or $\chi > 0$ (escalation cost).

Owner's problem under both A and B Once B is available, the owner of firm i chooses both $L_{A,i,t}$ and $L_{B,i,t}$ to maximize total flow profit, net of separable quadratic spanning costs in each business model:

$$\max_{\{L_{A,i,t},L_{B,i,t}\}\ge 0} \quad \pi_{A,t}(\rho_{A,i,t})\,L_{A,i,t} \;+\; \pi_{B,t}(\rho_{A,i,t},\,\rho_{B,i,t};\Delta)\,L_{B,i,t} \;-\; \frac{L_{A,i,t}^2 + L_{B,i,t}^2}{2\,\xi_i}. \tag{17}$$

For each firm, the B-rule stock evolves following the microfoundation above. Cases (ii.a), (ii.b), and (iii) each add a new B-rule, so $d\rho_{B,i,t}/dt = \kappa(1-\rho_{B,i,t})$. All firms start from no rules under B and the arrival of B is a firm-wide event, yielding the closed form $\rho_{B,i,t} = 1 - e^{-\kappa t}$.

Equilibrium wage of B-managers The industry-level B-manager market clears: $\int_i L^*_{B,i,t}\,di \;=\; L$. Using the firm-level first-order condition for $L^*_{B,i,t}$ and aggregating across firms, this yields the following equilibrium per-B-unit productivity:

$$z_{B,t} \;=\; \frac{L}{\Xi} \;+\; \phi\,(1-\bar\rho_{B,t})\left[1 \;-\; \left(1-\frac{F(\Delta)}{F^*}\right)\bar\rho_{A,t}\right], \tag{18}$$

where $\bar{\rho}_{X,t} \equiv \int_i \rho_{X,i,t}\, di$ for $X \in \{A, B\}$. The rest of the equilibrium follows from the substituting this expression into the first-order condition for $L_{B,i,t}$.

III.D Discussion

Before deriving the empirical predictions of the model, we discuss two questions: whether older firms do accumulate more rules in the data, and whether a firm would want to insulate B from A.

Do older firms have more rules? The model associates older firms with more rules (Equation 5). To support this observation, we collect information through employee reviews from Revelio (Revelio Labs, 2025). First, we use LLM to screen for reviews that discuss rules at the employer company. Second, we use LLM to label the relevant reviews as those that mention the presence of rules vs those that mention the lack of rules; see Appendix Section IA3.1 for our prompt. For each firm (represented by GVKEY) in Compustat, we estimate its age following the procedure in Appendix IA5 (which is also used in empirical analyses in Section V and Section VI), and we calculate the intensity of rules in a GVKEY-year as:[11]

$$\text{Rule Index} = \frac{(\#\text{ of reviews mentioning the presence of rules - }\#\text{ of reviews mentioning the lack of rules})}{\#\text{ of total reviews}}. \tag{19}$$

Similarly, we use employee reviews to check whether old firms have more meetings, committees, and layers of approval that take time. We detail our prompt in Appendix Section IA3.2. First, we use LLM to screen for reviews that discuss meetings, committees, and layers of approval at the employer company. Second, we use LLM to label the relevant reviews as those complaining the company has too many such activities that waste time vs praising the company for being efficient and not having too many such activities. We calculate the intensity of meetings that waste time in a GVKEY-year as:

$$\text{Meeting Index} = \frac{(\#\text{ of reviews complaining too many - }\#\text{ of reviews praising not too many})}{\#\text{ of total reviews}}. \tag{20}$$

Figure 3 shows binscatter plots of the relationship between the Rule Index and the Meeting Index with respect to age. We see that both are strongly increasing with firm age.

Should B be insulated from A? Once the workstyle distance is large enough that a firm's existing rules become a net liability for B (Section III.C), it is natural to ask whether an incumbent can escape the conflict by insulating business model B from business model A. This is a question about the boundary of the firm

[11] We map Revelio company ID RCID to GVKEY in a given year using parent-subsidiary bridge provided by Revelio to us directly.

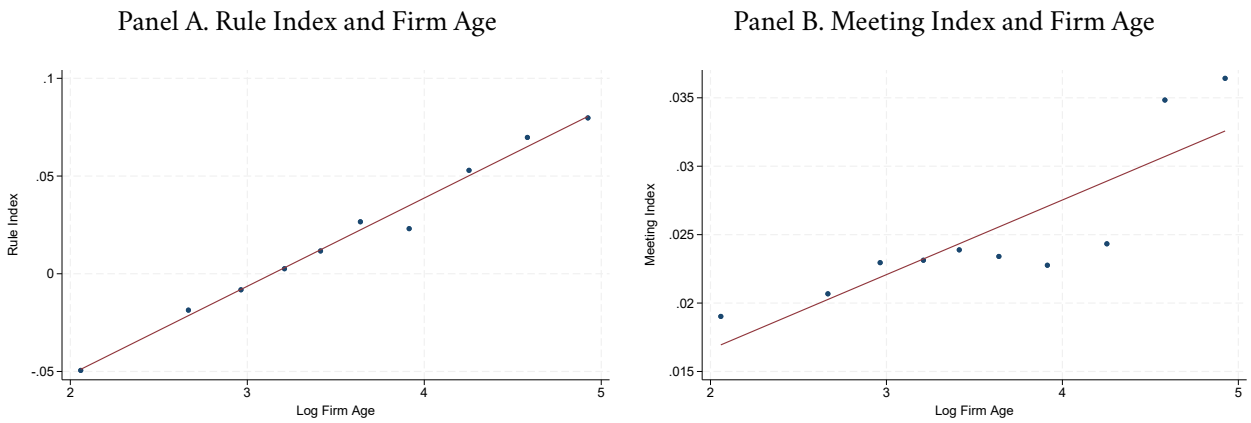

Figure 3. Firm Age, Rules, and Meetings that Waste Time

Notes: This figure shows binscatter plots of the Rule Index in (19) in Panel A and the Meeting Index in (20) in Panel B in 10 equal-sized bins based on log firm age for Compustat firms. The Rule Index and Meeting Index are constructed from employee reviews collected by Revelio. We absorb year fixed effects.

(Coase, 1937; Williamson, 1971; Grossman and Hart, 1986; Hart and Moore, 1990), and the answer hinges on what insulation can, and cannot, sever.

The option available to the incumbent *as a single firm* is to run B as a stand-alone division on its own rulebook, so that B-workers no longer consult the A-rulebook. This severs the worker-level conflict of Section III.C, but it does not make B independent. So long as the two divisions answer to a common headquarters that controls non-contractible shared resources, such as discretionary budgets, infrastructure, and shared services, the entrenched A-division manager retains a channel through which to impede B-division.

The division-level conflict, often modeled as a game between headquarters and managers, lies at the heart of the incomplete contracts literature (Hart and Moore, 2008; Scharfstein and Stein, 2000; Rajan, Servaes, and Zingales, 2000; Stein, 1997), and is one we abstract from in the baseline analyses earlier. Nevertheless, we argue that the key model implication in (16) carries through via this more canonical channel. Following Hart and Moore (2008), we model this friction as *shading*: a larger workstyle distance makes B the more valuable claimant on shared resources, so headquarters tilts its allocation toward B. The A-division, aggrieved relative to the allocation it regards as its due, retaliates by withholding cooperation and lobbying headquarters against B. We assume that it does so more effectively the longer and more codified its own rulebook, that is, the higher the ρ_A. Appendix IA6.8 provides the details. The per-B-unit profit under internal separation is the rule-free profit, net of this shading cost:

$$\pi_B^{\text{sep}} \;=\; \pi_B(0,\rho_B) \;-\; \theta\,\phi\,(1-\rho_B)\,\rho_A\,F(\Delta), \qquad \theta \geq 0, \tag{21}$$

with θ indexing the effectiveness of shading, equivalently the weakness of headquarters governance.

An incumbent that wants to implement B as its own compares internal separation with integration, and adopts the former only at a still-larger workstyle distance: a threshold $F^{**} \geq F^{*}$ derived in Appendix IA6.8. More importantly, whichever of the two it chooses, the division-level friction inherits the comparative statics of the baseline. It worsens with the workstyle distance and, because shading is more effective the larger the A-division's rule stock, bears down disproportionately on older firms:

$$\frac{\partial^2 \pi_B^{\text{sep}}}{\partial \rho_A \, \partial \Delta} \;=\; -\theta \, \phi \, (1 - \rho_B) \, F'(\Delta) \;<\; 0, \tag{22}$$

the same sign as the integrated cross-derivative of Section III.C. Whichever organizational form the incumbent adopts, integration or internal separation, the cross-derivative that drives our predictions stays negative, so the empirical results we take to the data in Section III.E are qualitatively unchanged.

Full insulation requires more than a division: it requires spinning B off as a legally separate firm, which removes the shading entirely and restores the rule-free profit $\pi_B(0, \rho_B)$. Read positively, this sheds light on observed organizational design. It helps explain why financial investors such as private-equity funds can hold portfolio companies that operate in markedly different styles, whereas a single operating company finds the same diversity hard to sustain internally (O'Reilly and Tushman, 2021): operationally and financially separate businesses avoid both the worker-level conflict we model above and the headquarters shading of Appendix IA6.8. There are, by the same logic, good reasons that ventures like X, Tesla, and SpaceX are organized as separate entities rather than as divisions of one operating company. The boundary of the firm here is drawn around shared real operations and control rights, not financial ownership: two businesses that merely share investors are not one firm, which is exactly what lets separate entities escape a friction that an integrated incumbent cannot. The flip side is that a spun-off B, with its own headquarters and residual control rights (Grossman and Hart, 1986; Hart and Moore, 1990), no longer contributes to firm i's employment, sales, or value, and so does not count as firm i implementing B. Whether the incumbent integrates, internally separates, or lets the business migrate to a separate firm, the prediction will be similar: older firms fall behind in industries undergoing large workstyle change.

III.E Empirical implications

We now derive the testable implications of the model for the growth of young and old firms when business model B arrives. As in the discussion above, we consider firms in multiple industries indexed by n, which differ in their workstyle distances Δ_n, per-unit outputs $Q_{A,n}, Q_{B,n}$, and equilibrium objects, and we continue to index firms by i.

Firm growth and change in workstyle We focus throughout on the firm's growth response from just before B arrives at the industry ($t = 0^-$) to just after ($t = 0^+$). Define the net growth rate of firm i in industry n as the proportional change in firm size over this window:

$$g_{i,n} \equiv \frac{L^*_{A,i,0^+} + L^*_{B,i,0^+}}{L^*_{A,i,0^-}} - 1 = \frac{L^*_{B,i,0^+}}{L^*_{A,i,0^-}}, \tag{23}$$

where the second equality uses the fact the firm's A-side is unaffected by the arrival of B.

Evaluating equation (18) at $t = 0^-$, and substituting into the firm's first order condition for the number of B units at time $t = 0^+$, yields:

$$L^*_{B,i,0^+} = \left[\frac{L_n}{\Xi_n} + \phi \left(1 - \frac{F(\Delta_n)}{F^*}\right) \left(\rho_{A,i,0^-} - \bar{\rho}_{A,n,0^-}\right) \right] \xi_i. \tag{24}$$

The effect of age on the absolute number of B-units opened at impulse depends on the regime identified by the threshold F^* of (14). When $F(\Delta) < F^*$ (existing rules are a net asset), the coefficient on $(\rho_{A,i,0^-} - \bar{\rho}_{A,0^-})$ is positive. Conditional on the spanning capacity ξ_i, older firms open *more* B-units than younger firms in absolute terms: their accumulated A-knowledge raises B-unit productivity. When $F(\Delta) > F^*$ (existing rules are a net liability), the coefficient turns negative, and older firms open *fewer* B-units, as the escalation cost swamp the productivity benefits of accumulated A-knowledge.

However, older firms always *grow* more slowly in proportional terms once B requires a different workstyle. Young firms are small in A and small in B (they have no rules to carry over). Without the friction, old firms are big in A and big in B (their rules carry over perfectly). In this case, proportional growth is the same for everyone. The friction shrinks old firms' B-side without touching its A-side: part of its knowledge conflicts with the B-workstyle and is lost, creating an asymmetry between small B units and large A ones at firms with an existing stock of rules. This does not occur for young firms with little existing rules. The resulting differential is therefore an effect of the friction, scales with the workstyle distance, and vanishes when workstyles coincide. The following result makes this precise.

Result 1. *Suppose that the workstyle friction takes the linear form:*

$$F(\Delta) = f\,\Delta, \tag{25}$$

where $f > 0$ is a primitive parameter capturing the strength of the workstyle–friction relationship.[12] *Then*

[12] Empirically, Δ has bounded support; we assume f is small enough that $F(\Delta) \in [0, 1]$ for all Δ realized in the data, equivalently $f \leq 1/\Delta_{\max}$ where $\Delta_{\max}$ is the largest workstyle distance observed across the industries we study. A non-linear specification of F does not affect any of the qualitative results below.

equilibrium impulse growth rate of firm i in industry n is

$$g_{i,n} \;=\; 1 \;-\; \frac{\phi\, f}{F^*}\,\frac{\left(\rho_{A,i,0^-} - \bar{\rho}_{A,n,0^-}\right)}{\left(L_n/\Xi_n\right) \;+\; \phi\left(\rho_{A,i,0^-} - \bar{\rho}_{A,n,0^-}\right)}\,\Delta_n, \tag{26}$$

where $\bar{\rho}_{A,n,0^-} \equiv \int_{i\in n} \rho_{A,i,0^-}\, di$ is the cross-firm sum of pre-arrival rule stocks in industry n (equivalently, the cross-firm average, since each industry has unit mass of firms). The change in observed industry-level employment-weighted average workstyle is proportional to Δ_n:

$$|d\overline{ws}|_n \equiv \left\| \sum_j \left(\frac{\bar{L}_{n,j,0^+}}{\bar{L}_{n,0^+}} - \frac{\bar{L}_{n,j,0^-}}{\bar{L}_{n,0^-}} \right) ws_j \right\| \;=\; \frac{1}{2}\,\Delta_n, \tag{27}$$

where $\bar{L}_{n,j,0^+}$ is total employment in occupation j, industry n immediately after B arrives, $\bar{L}_{n,0^+} = \sum_j \bar{L}_{n,j,0^+}$, $\bar{L}_{n,j,0^-}$ and $\bar{L}_{n,0^-}$ are defined similarly right before B arrives.[13]

Proof. See Appendix IA6. □

Three features of equations (26)–(27) matter for the empirical analysis. First, the spanning capacity ξ_i does not appear in $g_{i,n}$: in this baseline, the impulse growth rate is independent of firm size. Second, the proportionality (27) justifies using the observable industry-level workstyle change $|d\overline{ws}|_n$ as a proxy for the unobserved structural distance Δ_n. Third, the cross-derivative of $g_{i,n}$ with respect to the pre-arrival rule stock and the workstyle distance is unambiguously negative:

$$\frac{\partial^2 g_{i,n}}{\partial \rho_{A,i,0^-}\,\partial \Delta_n} \;<\; 0, \tag{28}$$

which means that the proportional growth gap between high-rule and low-rule firms (equivalently, between old and young firms) widens with the workstyle distance. The marginal benefit of additional A-rules at B-units declines with Δ_n: the conflict between A-rules and the workstyle of B workers erodes the substitutability that otherwise would benefit older firms.[14] Finally, the result is independent of whether existing rules are a net asset ($F(\Delta) < F^*$) or a net liability ($F(\Delta) > F^*$) for the new business model.

[13] The factor $1/2$ is the employment share of B-units at $t = 0^+$, which equals $1/2$ under the normalization that the supply of managers of types A and B is the same in each industry, $L_{A,n} = L_{B,n} \equiv L_n$.

[14] In the baseline, owners delegate occupation choice to unit managers, who maximize per-unit output and do not internalize the cross-unit friction. Appendix IA6.7 shows that letting owners choose the labor composition directly induces a "composition wedge" that compresses the realized workstyle distance from Δ to some $\Delta^* < \Delta$, attenuating the magnitude of this cross-derivative but preserving its sign. The two edge cases bracket the effect: at $\sigma = 0$ (Leontief production, occupations perfect complements) the wedge collapses, the firm bears the friction in full, and the cross-derivative holds in its strongest form; at $\sigma \to \infty$ (occupations perfect substitutes) the firm rebalances occupations across A- and B-units to align workstyles exactly, $\Delta^* \to 0$, and the friction is neutralized. The prediction therefore survives whenever occupations are not perfect substitutes in production.

Empirical regression We now express this observation as a reduced-form empirical regression. Let $w_n \equiv |d\overline{ws}|_n$ denote the industry-level workstyle change and $a_i \equiv \text{age}_i$ denote firm i's age. We analyze the regression

$$g_{i,n} \;=\; \nu \;+\; \kappa_w\, w_n \;+\; \rho_a\, a_i \;+\; \zeta\,(a_i \times w_n) \;+\; \varepsilon_{i,n}, \tag{29}$$

where $g_{i,n}$ is the firm's impulse growth rate from (23).[15] The main coefficient of interest is ζ, which captures the differential growth response of old vs. young firms to industry-level workstyle change. To interpret ζ, we make two independence assumptions, both restated formally in Appendix IA6.

Result 2. *Suppose that:*

(i) the joint distribution of (a_i, ξ_i) is identical across industries n and is independent of all industry-level variables;

(ii) the ratio Ξ_n / L_n is independent of the workstyle distance Δ_n.

Then, to first order in firm-level rule deviations $\rho_{A,i,0^-} - \bar{\rho}_{A,n,0^-}$ around $f\Delta_n = 0$, the population OLS coefficient on $a_i \times w_n$ in regression (29) is

$$\hat{\zeta} \;=\; -\,\Phi \;\times\; \frac{\text{cov}(\rho_{A,i,0^-},\, a_i)}{\text{var}(a_i)}, \qquad \textit{where} \quad \Phi \;\equiv\; \frac{2\,\phi\, f}{F^*}\,\mathbb{E}\!\left[\frac{\Xi_n}{L_n}\right] \;>\; 0. \tag{30}$$

Proof. See Appendix IA6. □

The economic content of Result 2 is as follows. The coefficient $\hat{\zeta}$ is the product of two terms: the industry-level scalar Φ, which captures the strength of the rule–workstyle friction, and the firm-level covariance $\text{cov}(\rho_{A,i,0^-}, a_i)$ between rule coverage and age. Both factors are positive. The covariance is positive because the size of the rulebook, $\rho_{A,i,0}$ is monotonically increasing in the age of the firm at the time B arrives, so older firms have systematically more A-rules. The scalar Φ is positive by construction: $\mathbb{E}[\Xi_n/L_n]$, ϕ, and f are all strictly positive. Therefore $\hat{\zeta} < 0$: older firms grow less in industries undergoing larger workstyle change. This regression provides a direct test of our core hypothesis that legacy rules become a friction at B-units in proportion to the workstyle distance, with the friction borne disproportionately by firms with longer rule histories.[16]

In Appendix IA6, we provide a generalization of the regression in (29) to a multivariate specification that includes additional industry-level and firm-level controls. This generalization does not change the economic interpretation of the interaction coefficient described in Result 2.

[15] We focus on a single cross-sectional regression with growth rates computed from before to after the arrival of business model B. In the empirical analyses in Section V, we use firm growth and industry workstyle change over five years.

[16] Appendix IA6 also derives the more general expression for Φ in which Assumption (ii) is dropped. We show that it remains positive whenever Ξ_n/L_n does not covary too negatively with Δ_n.

Who implements B? We characterize firm i's share of total B-output in industry n at the time of the arrival of B. Since per-B-unit output is firm-invariant, the output share is equal to the managerial employment share. Appendix IA6.6 then shows that this share is given by:

$$\sigma_{B,i,0^+} = \frac{\xi_i}{\Xi}\omega_{B,n}\Big(\rho_{A,i,0^-};\Delta\Big), \tag{31}$$

where $\omega_B(\rho_A;\Delta)$ is a function whose analytical expression is reported in Appendix IA6.6, and whose monotonicity in ρ_A depends on the regime identified by the threshold F^* of (14): it is increasing in ρ_A when $F(\Delta) < F^*$ and decreasing in ρ_A when $F(\Delta) > F^*$. This result has two implications. Conditional on age $(\rho_{A,i,0^-})$, larger firms (those with higher ξ_i) implement more B in absolute terms; this spanning-capacity channel operates regardless of regime. Conditional on ξ_i, the effect of age is regime-dependent: when $F(\Delta) < F^*$, older firms (those with more existing rules $\rho_{A,i,0^-}$) implement more B in absolute terms, because their rule stock partially transfers to B-units; when $F(\Delta) > F^*$, older firms implement *less* B in absolute terms, because their conflicting rules raise escalation costs.

The pre-arrival A-share admits a similar solution, reported in Appendix IA6.6. The ratio of firm i's B-share to its pre-arrival A-share is independent of ξ_i, and equal to:

$$\frac{\sigma_{B,i,0^+}}{\sigma_{A,i,0^-}} = \frac{\omega_B\Big(\rho_{A,i,0^-};\Delta\Big)}{\omega_A\Big(\rho_{A,i,0^-}\Big)}. \tag{32}$$

This ratio is *strictly decreasing* in $\rho_{A,i,0^-}$ whenever $F(\Delta_n) > 0$, in both regimes of (14).[17] This implies that young firms capture a disproportionately large share of B-output relative to their pre-existing A-share, even if they are small in absolute terms. The size of this effect is governed by the workstyle distance: in industries with larger workstyle changes between A and B, young firms are more over-represented in their share of B output. When the workstyle change is large enough that $F(\Delta_n) > F^*$, the over-representation is starker still, since older firms then implement less B even in absolute terms. This prediction mirrors the motivating fact in Section II.A: young firms have implemented close to 40% of new technologies among 20th-century inventions by private companies, higher than their employment share in the Census Business Dynamics Statistics.

[17] The proof is reported in Appendix IA6.6.

IV Data and Measurement

Our empirical work builds a counterpart to the workstyle change $|d\overline{ws}|_n$ in the model at the industry-year level, which aims to capture the extent to which new technologies shift the workstyle that an industry's production requires. To do so, we combine two ingredients. First, we measure workstyle at the occupation level using the O*NET "Work Styles" dataset, which records the importance of 16 characteristics for each occupation (e.g., attention to detail, self control, stress tolerance, adaptability/flexibility, innovation). We define the industry-level workstyle as the predicted weighted average workstyle across occupations, based on the workstyle of each occupation and the occupation's employment share. Correspondingly, we define industry-level predicted workstyle change as the Euclidean difference between the current industry-level workstyle in year t and the industry-level workstyle based on industry occupation composition in year $t+5$ predicted by new technologies in the industry in year t. Second, to obtain the industry occupation composition in light of new technologies, we follow the method of Kogan et al. (2024): employment of an occupation in an industry declines (rises) when the occupation's routine (nonroutine) tasks are similar to new technologies in the industry, based on the text of breakthrough patents (Kelly et al., 2021). Taken together, new technologies in year t predict the future occupation mix in the industry, which in turn implies a predicted change in the industry's average workstyle.[18] Our approach builds on prior work linking technology to occupation composition (Bresnahan, Brynjolfsson, and Hitt, 2002; Autor, Levy, and Murnane, 2003; Acemoglu and Autor, 2011), and on recent research showing that occupation similarity across industries shapes firms' horizontal expansions and mergers (Lee, Mauer, and Xu, 2018; Beaumont, Hebert, and Lyonnet, 2025).

IV.A Measuring Workstyle

We first describe the measurement of workstyle at the occupation level using O*NET data by the U.S. Department of Labor (U.S. Department of Labor, Employment and Training Administration, 2025). The O*NET dataset describes the key characteristics of occupations. The information comes from surveys of workers and occupational experts. The occupations in O*NET use a slight variant of the Bureau of Labor Statistics' Standard Occupational Classification (SOC) codes.

[18] We focus on the predicted change based on new technologies rather than the realized change in workstyle for two reasons. First, it ties workstyle change to new technologies, which aligns with the force highlighted by the model. Second, the realized change may be affected by ex post shocks that could confound the analyses.

Workstyle for occupation j For each occupation, the O*NET Work Styles module provides the importance score of 16 characteristics, on a scale of 1 to 5.[19] The characteristics are achievement/effort, persistence, initiative, leadership, cooperation, concern for others, social orientation, self-control, stress tolerance, adaptability/flexibility, dependability, attention to detail, integrity, independence, innovation, and analytical thinking. We use $\mathbf{ws}_j$ to denote this 16×1 vector at the occupation level.

Workstyle in industry n and year t In the theoretical framework developed in Section III.E, one ingredient in (29) is the change of industry-level workstyle $|d\overline{ws}|_n$. To construct its empirical counterpart, we first define the industry-level workstyle as the employment-weighted average of occupation workstyles:

$$\mathbf{WS}_{n,t} = \sum_{j=1}^{J} \frac{l_{j,n,t}}{l_{n,t}} \mathbf{ws}_j, \tag{33}$$

where $l_{n,t} = \sum_j l_{j,n,t}$ is the industry's total employment. This definition takes the occupation composition as an input. We compute the industry-level workstyle using two sets of occupation compositions. First, we measure the current industry-level workstyle using the actual occupation composition based on the Occupational Employment and Wage Statistics (OEWS) dataset by the BLS (U.S. Bureau of Labor Statistics, 2025) at the three-digit NAICS level, which is consistently available since 2003. Second, we also measure the predicted industry-level workstyle in light of new technologies, using the occupation composition predicted by new technologies explained in Section IV.B. The Euclidean difference between the two would be the predicted workstyle change, which would be our key focus in the empirical analyses.

IV.B Predicting Industry Occupation Composition from New Technologies

We now describe how we obtain the year-$t+5$ industry occupation composition predicted by new technologies in year t, which follows the method of Kogan et al. (2024). The core idea is to predict future employment by occupation in an industry using the similarity between occupation tasks (from task description text) and new technologies (from patent text); Appendix Section IA4 provides further details on the similarity computation. We use 183,213 U.S. breakthrough patents granted between 2003 and 2016 (the end of breakthrough patent data).

Calculate patents' similarity with occupation j's routine & nonroutine tasks We obtain the tasks of each occupation according to the O*NET-SOC 2000 taxonomy (U.S. Department of Labor, Employment and Training Administration, 2005), so that the task descriptions predate the sample period of our analyses.

[19] For each occupation and characteristic, we take the average value across unique survey waves if there are multiple. If we only use the first value, then some occupations will have missing values in the early years.

We then query LLM to label them as routine tasks vs non-routine tasks.[20] For each occupation, we combine the descriptions of routine tasks and nonroutine tasks, and will use each set to calculate the similarity with patent text later. We also compute the share of tasks that are routine or non-routine for each occupation. The median occupation has 16 tasks and 33.3% are routine.[21]

As the first step, we follow Kogan et al. (2024) to represent each document—including each patent and each occupation's combined routine (nonroutine) task descriptions—as a weighted average of its word embeddings. Specifically, the embeddings of document i, which is denoted by $\mathcal{X}_i$, is calculated as:

$$\mathcal{X}_i = \sum_k q_{i,k} x_k, \tag{34}$$

where $q_{i,k}$ is the Term Frequency-Inverse Document Frequency (TF-IDF) weight,[22] and word embeddings x_k are obtained using the *GloVe* model.[23] Similar to Kogan et al. (2024), we pre-clean the patent and occupation text to keep only verbs and nouns and remove stop words before calculating the embeddings, as explained in detail in Appendix Section IA4.1.

We then calculate the cosine similarity $s^{\tau}_{p,j}$ between a patent p and the routine or nonroutine component of occupation j:

$$s^{\tau}_{p,j} = \frac{\mathcal{X}_p}{\|\mathcal{X}_p\|} \cdot \frac{\mathcal{X}^{\tau}_j}{\|\mathcal{X}^{\tau}_j\|}, \quad \tau \in \{R, NR\}. \tag{35}$$

We follow Kogan et al. (2024) to remove year fixed effects (for routine and nonroutine metrics and for each year separately), and impose sparsity on $s^{\tau}_{p,j}$ by setting values below the 80th percentile to zero (then rescaling the values to [0, 1]) to obtain an adjusted similarity measure $\tilde{s}^{\tau}_{p,j}$; for details, see Appendix Section IA4.1.

[20] We follow Kogan et al. (2024) in querying LLM (GPT-4o) for this classification. The prompt is: "A routine task can be defined as follows: A routine task involves carrying out a limited and well-defined set of work activities, those that can be accomplished by following explicit rules. These tasks require methodical repetition of an unwavering procedure, and they can be exhaustively specified with programmed instructions and performed by machines. Tell me whether the following task is primarily routine or primarily non-routine; and explain your reasoning in one sentence.

Task: *task statement text from O*NET*

Output your answer in JSON like the following format: {"answer": "primarily routine/primarily non-routine", "reasoning": "your reasoning"}"

[21] The 2000 O*NET classification contains 15,643 different tasks, 98.3% of which are specific to a particular occupation in the O*NET-SOC classification. Examples of tasks include: "Manage and treat common health problems, such as infections, influenza and pneumonia, as well as serious, chronic, and complex illnesses, in adolescents, adults, and the elderly"; "Measures and marks location of studs, leaders, and receptacle openings, using tape measure, template, and marker"; "Receive mortgage, loan, or public utility bill payments, verifying payment dates and amounts due". By contrast, as explained above, the workstyle data which we use in our measures are lower-dimensional and score all occupations on the same set of characteristics.

[22] The weight $q_{i,k}$ reflects the frequency of a term within a document relative to its frequency across all documents, thereby emphasizing words that are distinctive to that document. We compute TF-IDF separately for the patent and occupation text corpora.

[23] We use the `glove-wiki-gigaword-300` model, trained on the Wikipedia and Gigaword corpora, which provides 300-dimensional word vectors that preserve semantic relationships among words.

Sum over patents in industry n and year t to get occupation j's routine and nonroutine exposures. To obtain the occupation j's total exposures (for both routine and nonroutine) in industry n at year t, we sum over similarity between occupation j and all granted patents to U.S. assignees in the industry-year according to the mapping between Cooperative Patent Classification (CPC) and three-digit NAICS industries provided by Goldschlag, Lybbert, and Zolas (2016). We refer to the total similarity between occupation j's routine (nonroutine) tasks and patents in industry n year t as $\xi^{R}_{j,n,t}$ ($\xi^{NR}_{j,n,t}$). Specifically, the exposure of occupation j in industry n to technology at year t is calculated as:

$$\xi^{\tau}_{j,n,t} = \theta^{\tau}_{j} \log\left(1 + \sum_{p \in \mathcal{P}_t} \pi_{p,n}\, \tilde{s}^{\tau}_{p,j}\right), \quad \tau \in \{R, NR\}, \tag{36}$$

where θ^{τ}_{j} is the routine/nonroutine share of tasks for occupation j, $\mathcal{P}_t$ is the set of breakthrough patents granted in year t, and $\pi_{p,n}$ is the probability weight linking patent p to industry n from Goldschlag, Lybbert, and Zolas (2016).[24]

Predict future employment of occupation j in industry n We then follow Kogan et al. (2024) and use occupation exposure to technologies $\xi^{R}_{j,n,t}$ and $\xi^{NR}_{j,n,t}$ to predict future employment. Specifically, we run a regression of log employment of occupation j in industry n and year $t+5$ ($\log l_{j,n,t+5}$) on the occupation's exposure to year t technologies through routine tasks ($\xi^{R}_{j,n,t}$), the occupation's exposure to year t technologies through nonroutine tasks ($\xi^{NR}_{j,n,t}$), and current employment of occupation j in industry n ($\log l_{j,n,t}$). We obtain predicted log employment of occupation j in industry n and year $t+5$ ($\widehat{\log l_{j,n,t+5}}$):

$$\widehat{\log l_{j,n,t+5}} = \underset{(0.060)}{0.308^{***}} - \underset{(0.009)}{0.052^{***}} \cdot \xi^{R}_{j,n,t} + \underset{(0.004)}{0.031^{***}} \cdot \xi^{NR}_{j,n,t} + \underset{(0.006)}{0.953^{***}} \cdot \log l_{j,n,t}. \tag{37}$$

This regression shows that the employment of an occupation declines by more when new technologies are more similar to its routine tasks, indicating substitution and replacement, while the employment rises by more when new technologies are more similar to its nonroutine tasks, indicating complementarity and enhancement. These effects, in line with Kogan et al. (2024), capture substitution and replacement on routine tasks and complementarity and enhancement on nonroutine tasks.

IV.C Predicted Workstyle Change

Finally, we combine the ingredients in Section IV.A and IV.B to construct the predicted workstyle change in industry n from year t to $t+5$. We define $\widehat{l_{n,t+5}} \equiv \sum_j \widehat{l_{j,n,t+5}}$ and obtain the predicted workstyle

[24] The Goldschlag, Lybbert, and Zolas (2016) bridge maps each Cooperative Patent Classification (CPC) code to an industry with a probability weight. We follow Kelly et al. (2021) to make the probability weights associated with a patent sum to one.

in year $t+5$ using $\widehat{\log l_{j,n,t+5}}$:

$$\mathbf{WS}^{e}_{n,t+5} = \sum_{j=1}^{J} \frac{\widehat{l_{j,n,t+5}}}{\widehat{l_{n,t+5}}} \mathbf{ws}_j. \tag{38}$$

Here, the superscript e indicates "expected" or "predicted." Accordingly, the predicted workstyle change in industry n from year t to $t+5$ is calculated as:

$$|d\overline{ws}|^{e}_{n,t,5} \equiv \|\mathbf{WS}^{e}_{n,t+5} - \mathbf{WS}_{n,t}\|. \tag{39}$$

This predicted workstyle change based on technologies at the industry-year level, $|d\overline{ws}|^{e}_{n,t,5}$, is the empirical counterpart to $|d\overline{ws}|_n$ in the model and the key independent variable in our analysis in Section V. We standardize it to unit variance and zero median to facilitate the assessment of economic magnitude in the regressions.

V Empirical Results

In this section, we analyze how workstyle changes in an industry help us understand the growth of young vs old firms. The key prediction is that such changes are more difficult for old firms to accommodate, and they grow more slowly than young firms when new technologies induce greater shifts in workstyle. We test this hypothesis in several datasets. In Section V.A, we measure the strength of young firms using venture capital investment. In Section V.B, we measure the performance of young firms vs old firms using Compustat data, which allow for a variety of metrics including valuation, sales growth, and employment growth. In Section V.C, we measure the growth of young firms vs old firms using employment growth in Census Business Dynamics Statistics (BDS) data. The BDS dataset offers broad coverage of the economy, although many young firms in that case may be subsistence entrepreneurship rather than innovative entrepreneurship. In addition, firms' exact age is unknown for those formed before the start of the BDS dataset in 1976. All of our tests use three-digit nonfarm nonfinancial NAICS industries (i.e., excluding NAICS codes starting with 11, 52, 53, and 55).[25]

Here we focus on workstyle change in the industry, and examine how young vs old firms in the industry adapt to this change. Conceptually, this is consistent with the model setup in Section III where young and old firms in an industry all face the arrival of business model B. Empirically, measuring predicted workstyle change at the industry level is convenient because it is easiest to construct shifts in employment

[25] Since the Goldschlag, Lybbert, and Zolas (2016) crosswalk of patent technology classes to NAICS codes does not include 483 (water transportation) and 813 (religious, grantmaking, civic, professional, and similar organizations), our analyses also exclude these industries.

by occupation due to new technologies at the industry level as explained in Section IV.B. Since we rely on BLS employment composition data at the three-digit NAICS code level and O*NET workstyle data to construct the predicted workstyle change, our sample for the empirical analyses starts in 2003 when these data became consistently available; our sample ends in 2016 because several of the patent controls are available through 2016.

V.A Venture Capital Investment

We start with measuring the strength of young firms through venture capital (VC) investment, which captures forward-looking valuation of startups that can reflect their growth potential. First, we obtain VC investment from Refinitiv and calculate total VC investment in year t and industry n. Specifically, we retrieve a list of deals from Refinitiv's Private Equity database (Refinitiv, 2025*a*), and select Fund Investors Type to be Venture Capital. We then use a fuzzy matching procedure to standardize industry names in Refinitiv data (which are 2007 NAICS codes in text) and map them to numerical 2007 NAICS codes. We sum over VC investment value by three-digit NAICS and year to obtain the industry-year level VC investment. Second, we compare the VC investment with the market capitalization of public firms in the industry, which reflects the strengths of young firms relative to incumbents. This test assumes that investors understand conditions that favor the growth of young firms (e.g., through examining their competitive advantages).

In Table 1, we regress VC investment relative to Compustat market capitalization in industry n following year t ($\log VC/MCAP_{n,t+1}$) on $|d\overline{ws}|^e_{n,t,5}$, the predicted workstyle change over the next five years due to year t technologies; the dependent variable is in year $t+1$ so that the year t predicted workstyle change is within the information set of investors:

$$\log VC/MCAP_{n,t+1} = \alpha_t + \beta\ |d\overline{ws}|^e_{n,t,5} + \gamma\, z_{n,t} + \epsilon_{n,t}. \tag{40}$$

This regression is a variation of our key empirical specification in (29) and we include year fixed effects to remove fluctuations in VC investment due to general economic conditions. Since $|d\overline{ws}|^e_{n,t,5}$ is a generated regressor based on (37) and (39), we bootstrap standard errors in all of our regression analyses.[26] In each

[26]The bootstrap addresses the inference problem that the generated regressor creates, and Internet Appendix IA6.4 addresses the complementary question of bias. We show that for the first-stage construction error to do more than attenuate the interaction coefficient toward zero and spuriously generate its sign, our proxy would have to be negatively correlated with the true expected workstyle change. Because the proxy is strongly positively correlated with realized workstyle changes, this would require the true expected change to covary negatively with realized changes. This is implausible in any model in which technology shocks are an important driver of workstyle change, and at odds with the evidence in Kogan et al. (2024) that technology shocks predict realized worker-level outcomes.

bootstrap draw, we resample industries with replacement, and use the corresponding bootstrap weights to reestimate the technology-implied employment change by industry-occupation-year in (37) and recalculate the predicted workstyle change in Equation (39), which is then used to rerun the final regression of interest in the bootstrap sample.

In Table 1 column (1), we see that VC investment value is significantly higher relative to public firm valuations when an industry is hit by new technologies that are predicted to change its workstyle. In columns (2) to (5), we control for the volume of new technologies in the industry using several patent-based measures from prior work, with limited impact on the regression coefficients on $|\overline{dws}|^e_{n,t,5}$. In column (2), we control for the log number of total patents in industry n and year t, demonstrating that the sheer quantity of new technologies does not display a significant relationship with the strength of young firms. Column (3) controls for the log number of breakthrough patents in industry n and year t using data from Kelly et al. (2021); the breakthrough patents aim to capture transformative new technologies, which are distinct from patents that came before them and followed by many similar patents afterwards. This measure is available for patents granted through 2016; we use the measure based on backward and forward significance over five years (instead of the version based on backward and forward significance over ten years, which is only available for patents granted through 2011), and a patent counts as a breakthrough one if its significance score (based on its textual similarity to subsequent patents relative to its textual similarity to prior patents) is in the top 10%. Column (4) controls for the log number of creative patents using data from Kalyani (2025) available for patents filed through 2016. The creativity scores for patents are based on the share of their technical bigrams that did not appear in any patents filed over the past five years, and creative patents are those with the top 10% creativity scores. Column (5) controls for the log number of total rapidly evolving technology (RETech) patents in industry n and year t, using data from Bowen, Frésard, and Hoberg (2023). Patents using words that are contemporaneously surging across the patent corpus receive a higher score, and we count RETech patents as those with scores in the top 10%. Overall, the volume of new technologies itself does not seem to have a significant relationship with the strength of startups relative to incumbents.

V.B Growth of Young vs Old Firms in Compustat Data

We then examine the relative strength of young vs old firms among public companies using Compustat data (S&P Global Market Intelligence, 2025). We combine information from several sources to measure the age of Compustat firms, which is not directly available in Compustat. First, we obtain firms' IPO date from Compustat and incorporation date from Refinitiv (Refinitiv, 2025*b*). The limitation of these data is that firms do not always become public soon after their founding, and they may reincorporate later;

Table 1 – Venture Capital Investment

	Forward 1-Year Log(VC Investment/Compustat Market Cap)				
	(1)	(2)	(3)	(4)	(5)
Predicted Workstyle Change in Industry	0.438***	0.446***	0.387***	0.414***	0.403***
	(0.142)	(0.149)	(0.142)	(0.146)	(0.146)
Log Patents in Industry		-0.048	0.072	0.027	0.059
		(0.080)	(0.094)	(0.106)	(0.109)
Patent Type	-	All	Breakthrough	Creative	RETech
Year FE	X	X	X	X	X
Observations	754	754	754	754	754
Within R^2	0.048	0.049	0.055	0.049	0.052

Notes: This table presents regressions at the industry-year level in Equation (40). The outcome variable is log VC investment relative to Compustat market capitalization in industry n following year t ($\log VC/MCAP_{n,t+1}$). The independent variables include the predicted workstyle change in industry n over the next 5 years due to year t technologies ($|d\overline{ws}|^e_{n,t,5}$), constructed in Equation (39). We standardize $|d\overline{ws}|^e_{n,t,5}$ to unit variance and zero median. Column (2) controls for the log number of total patents in industry n and year t. Column (3) controls for the log number of breakthrough patents in industry n and year t using data from Kelly et al. (2021). Column (4) controls for the log number of creative patents in industry n and year t using data from Kalyani (2025). Column (5) controls for the log number of rapidly evolving patents (RETech) in industry n and year t using data from Bowen, Frésard, and Hoberg (2023). We include year fixed effects, and bootstrap standard errors. In each bootstrap draw, we resample industries with replacement, and use the corresponding bootstrap weights to reestimate the technology-implied employment change by industry-occupation-year in Equation (37) and recalculate the predicted workstyle change in Equation (39), which is then used to rerun the final regression of interest in the bootstrap sample. The coefficient is the bootstrap median. Asterisks denote significance levels (***=1%, **=5%, *=10%). Number of observations and within R^2 (without fixed effects) are based on the OLS regression. Sample years are 2003 to 2016.

these data are also available only for a subset of Compustat firms (Compustat only began to systematically record IPO dates for IPOs after 2002). Second, we use LLM to query the founding year of each Compustat name (based on their real-time historical name), as explained in more detail in Appendix Section IA5. This information is reliable for well known companies, but can be less accurate for companies that are not well recognized or those that no longer exist. Taken together, we calculate firm age as the number of years since the earliest of these dates.[27]

In Table 2, we perform regressions using Tobin's Q as the outcome variable as follows:

$$Q_{i,t+1} = \alpha_t + \zeta \left(|d\overline{ws}|^e_{n,t,5} \times \log \text{age}_{i,t} \right) + \eta_w \ |d\overline{ws}|^e_{n,t,5} + \eta_a \ \log \text{age}_{i,t} + \gamma \, z_{n,t} + \epsilon_{i,t}. \tag{41}$$

Tobin's Q is calculated as the market value of equity plus book value of debt, divided by book assets; it captures forward-looking valuation of young vs old firms regarding their growth potential, analogous to the VC investment analyses in Section V.A but now within the set of public companies (rather than startups relative to public companies). This test also assumes that investors understand conditions under which

[27] We set age to missing if the earliest of these dates is after the earliest year of the company's record in Compustat, which indicates that our founding year information may not be accurate.

young firms are especially powerful (e.g., through examining their competitive advantages). Specifically, the outcome variable is Tobin's Q of firm i in industry n and year $t+1$ (to make sure year t technologies are within information sets similar to the VC regressions in Equation (40)). The independent variables include the predicted workstyle change in industry n over the next five years due to year-t technologies ($|d\overline{ws}|^e_{n,t,5}$), and its interaction with log firm age. The controls include the patent variables in columns (2) to (5) of Table 1, as well as their interactions with log firm age. To differentiate the effects of age from firm size, we also control for firm size and its interactions with the predicted workstyle change (as well as its interactions with the patent controls).[28] We include year fixed effects, and bootstrap standard errors as explained earlier.

In Tables 3 and 4, we perform similar regressions using realized sales growth and employment growth in the five years after t as the outcome variables:

$$\Delta Y_{i,t\to t+5} = \alpha_t + \zeta \left(|d\overline{ws}|^e_{n,t,5} \times \log \text{age}_{i,t} \right) + \eta_w \ |d\overline{ws}|^e_{n,t,5} + \eta_a \ \log \text{age}_{i,t} + \gamma \, z_{n,t} + \epsilon_{i,t}. \tag{42}$$

The growth rates $\Delta Y_{i,t\to t+5}$ are calculated following Davis, Haltiwanger, and Schuh (1992), to be consistent with the specification we need to use in Census Business Dynamics Statistics dataset later in Section V.C. It is defined as $\Delta Y_{i,t\to t+5} = (Y_{i,t+5} - Y_{i,t})/(0.5 \times Y_{i,t+5} + 0.5 \times Y_{i,t})$, which avoids missing values if the starting level is zero or extreme values when the starting level is small. The independent variables include the predicted workstyle change in industry n over the next five years due to year-t technologies ($|d\overline{ws}|^e_{n,t,5}$), and its interaction with firm age, patent controls interacted with firm age, as well as firm size and the same set of interactions. We include year fixed effects and bootstrap standard errors.

Tables 2, 3, and 4 show that when $|d\overline{ws}|^e_{n,t,5}$ is high, older firms have significantly lower valuation, as well as lower sales and employment growth. Meanwhile, the number of patents $z_{n,t}$ (all, breakthrough, creative, or rapidly evolving; for their definitions, see Section V.B) in the industry does not have any significant interaction with firm age. The interaction coefficient on $|d\overline{ws}|^e_{n,t,5}$ and firm age is also not affected by the controls on the quantity of patents. In other words, the volume of new technologies itself does not seem to drive the strength of young firms relative to old firms, similar to the findings in Section V.A. Figures IA1 and IA2 visualize the implied coefficient on $|d\overline{ws}|^e_{n,t,5}$ for each level of firm age, using the specification in column (1) of Tables 2, 3, and 4. We transform the coefficient on log firm age to that on firm age for ease of illustration.

A natural question is whether the results on firm age reflect a correlation between age and size: perhaps

[28] Appendix Section IA6.5 shows that adding industry-level and firm-level controls does not change the economic interpretation of the interaction coefficient between predicted workstyle change and firm age described in Result 2. In the presence of additional controls, the estimated coefficient is the product of an industry-level constant, which becomes conditional on other industry-level controls, and the covariance of age and organizational frictions conditional on other firm-level observables.

Table 2 – Tobin's Q among Compustat Firms

	Tobin's Q				
	(1)	(2)	(3)	(4)	(5)
Predicted Workstyle Change in Industry × Log Age	-0.136**	-0.159***	-0.144***	-0.150***	-0.156***
	(0.049)	(0.040)	(0.034)	(0.037)	(0.040)
Predicted Workstyle Change in Industry × Log Size	0.039	0.048**	0.050**	0.049**	0.051**
	(0.021)	(0.018)	(0.020)	(0.018)	(0.020)
Predicted Workstyle Change in Industry	0.298	0.223	0.145	0.171	0.184
	(0.188)	(0.148)	(0.140)	(0.140)	(0.145)
Log Patents in Industry × Log Age		-0.002	-0.015	-0.016	-0.009
		(0.027)	(0.031)	(0.033)	(0.033)
Log Patents in Industry × Log Size		-0.006	-0.008	-0.009	-0.015
		(0.013)	(0.018)	(0.019)	(0.019)
Log Patents in Industry		0.081	0.164	0.188	0.219
		(0.146)	(0.146)	(0.166)	(0.156)
Log Age	-0.266***	-0.247	-0.199*	-0.197*	-0.210**
	(0.077)	(0.169)	(0.119)	(0.120)	(0.102)
Log Size	-0.183***	-0.150*	-0.152**	-0.148**	-0.130**
	(0.039)	(0.074)	(0.059)	(0.059)	(0.051)
Patent Type	-	All	Breakthrough	Creative	RETech
Year FE	X	X	X	X	X
Observations	27,978	27,978	27,978	27,978	27,978
Within R^2	0.079	0.082	0.082	0.083	0.085

Notes: This table presents regressions at the firm-year level in Equation (41). The independent variables include the predicted workstyle change in industry n over the next 5 years due to year t technologies ($|\overline{dws}|^e_{n,t,5}$), constructed in Equation (39), and its interaction with firm age and firm size (log book assets in year t). We standardize $|\overline{dws}|^e_{n,t,5}$ to unit variance and zero median. Column (2) controls for the log number of total patents in industry n and year t, as well as its interactions with firm age and firm size. Column (3) controls for the log number of breakthrough patents in industry n and year t using data from Kelly et al. (2021), as well as its interactions with firm age and firm size. Column (4) controls for the log number of creative patents in industry n and year t using data from Kalyani (2025), as well as its interactions with firm age and firm size. Column (5) controls for the log number of rapidly evolving patents (RETech) in industry n and year t using data from Bowen, Frésard, and Hoberg (2023), as well as its interactions with firm age and firm size. We include year fixed effects, and bootstrap standard errors. In each bootstrap draw, we resample industries with replacement, and use the corresponding bootstrap weights to reestimate the technology-implied employment change by industry-occupation-year in Equation (37) and recalculate the predicted workstyle change in Equation (39), which is then used to rerun the final regression of interest in the bootstrap sample. The coefficient is the bootstrap median. Asterisks denote significance levels (***=1%, **=5%, *=10%). Number of observations and within R^2 (without fixed effects) are based on the OLS regression. Sample years are 2003 to 2016.

young firms are more nimble in light of technology-induced style changes because they are small. In previous work, some find that small firms have innovative advantages (Akcigit and Kerr, 2018; Argente et al., 2025), while others find that large firms adopt technologies like IT and AI more effectively (Hsieh and Rossi-Hansberg, 2023; Lashkari, Bauer, and Boussard, 2024; Babina et al., 2024). In Tables 2, 3, and 4, we do not observe that size is negatively associated with how firm outcomes respond to workstyle change. The interaction between predicted workstyle change $|\overline{dws}|^e_{n,t,5}$ and firm size tends to be positive. It could be

Table 3 – Sales Growth among Compustat Firms

	Sales Growth (5-Year DHS Rate)				
	(1)	(2)	(3)	(4)	(5)
Predicted Workstyle Change in Industry × Log Age	-0.025**	-0.026**	-0.024***	-0.025***	-0.025**
	(0.009)	(0.010)	(0.009)	(0.009)	(0.010)
Predicted Workstyle Change in Industry × Log Size	0.011**	0.011**	0.009*	0.011*	0.010*
	(0.005)	(0.006)	(0.005)	(0.006)	(0.006)
Predicted Workstyle Change in Industry	0.026	0.015	0.020	0.015	0.013
	(0.031)	(0.037)	(0.035)	(0.037)	(0.036)
Log Patents in Industry × Log Age		0.004	0.001	0.002	0.001
		(0.007)	(0.009)	(0.010)	(0.009)
Log Patents in Industry × Log Size		-0.002	0.002	-0.001	-0.001
		(0.003)	(0.004)	(0.005)	(0.005)
Log Patents in Industry		-0.003	-0.011	-0.002	0.005
		(0.024)	(0.025)	(0.029)	(0.026)
Log Age	-0.096***	-0.120**	-0.101***	-0.107***	-0.101***
	(0.018)	(0.043)	(0.032)	(0.032)	(0.027)
Log Size	-0.025*	-0.014	-0.028**	-0.020*	-0.020**
	(0.013)	(0.014)	(0.013)	(0.012)	(0.010)
Patent Type	-	All	Breakthrough	Creative	RETech
Year FE	X	X	X	X	X
Observations	27,978	27,978	27,978	27,978	27,978
Within R^2	0.056	0.056	0.056	0.056	0.056

Notes: This table presents regressions at the firm-year level in Equation (42), where the outcome variable $\Delta Y_{i,t\to t+5}$ is sales growth between year t and $t+5$ calculated following Davis, Haltiwanger, and Schuh (1992) as $\Delta Y_{i,t\to t+5} = (Y_{i,t+5} - Y_{i,t})/(0.5 \times Y_{i,t+5} + 0.5 \times Y_{i,t})$. The independent variables include the predicted workstyle change in industry n over the next 5 years due to year t technologies ($|\overline{dws}|^e_{n,t,5}$), constructed in Equation (39), and its interaction with firm age and firm size (log sales in year t). We standardize $|\overline{dws}|^e_{n,t,5}$ to unit variance and zero median. Column (2) controls for the log number of total patents in industry n and year t, as well as its interactions with firm age and firm size. Column (3) controls for the log number of breakthrough patents in industry n and year t using data from Kelly et al. (2021), as well as its interactions with firm age and firm size. Column (4) controls for the log number of creative patents in industry n and year t using data from Kalyani (2025), as well as its interactions with firm age and firm size. Column (5) controls for the log number of rapidly evolving patents (RETech) in industry n and year t using data from Bowen, Frésard, and Hoberg (2023), as well as its interactions with firm age and firm size. We include year fixed effects, and bootstrap standard errors. In each bootstrap draw, we resample industries with replacement, and use the corresponding bootstrap weights to reestimate the technology-implied employment change by industry-occupation-year in Equation (37) and recalculate the predicted workstyle change in Equation (39), which is then used to rerun the final regression of interest in the bootstrap sample. The coefficient is the bootstrap median. Asterisks denote significance levels (***=1%, **=5%, *=10%). Number of observations and within R^2 (without fixed effects) are based on the OLS regression. Sample years are 2003 to 2016.

possible that, all else equal, larger firms can have more financial resources that help them with adjustment and adaptation.[29]

[29]The baseline model does not generate an interaction between size and workstyle change, since the growth rate in Eqution (26) is independent of spanning capacity. Two assumptions would deliver a positive interaction. The first is that costly conflict resolution occurs at the firm level, so that these costs act as operating leverage and do not scale with the number of units. The second is that downtime costs, which measure the difficulty of solving problems, decrease with spanning capacity ξ_i. Under

Table 4 – Employment Growth among Compustat Firms

	Employment Growth (5-Year DHS Rate)				
	(1)	(2)	(3)	(4)	(5)
Predicted Workstyle Change in Industry × Log Age	-0.020**	-0.020**	-0.019**	-0.019**	-0.018**
	(0.007)	(0.008)	(0.008)	(0.008)	(0.008)
Predicted Workstyle Change in Industry × Log Size	0.009*	0.008**	0.006**	0.007**	0.007**
	(0.004)	(0.003)	(0.003)	(0.003)	(0.003)
Predicted Workstyle Change in Industry	0.072**	0.072**	0.069**	0.068*	0.065*
	(0.025)	(0.030)	(0.030)	(0.031)	(0.031)
Log Patents in Industry × Log Age		-0.000	-0.001	-0.002	-0.004
		(0.006)	(0.007)	(0.007)	(0.006)
Log Patents in Industry × Log Size		0.001	0.002	0.002	0.002
		(0.002)	(0.003)	(0.003)	(0.003)
Log Patents in Industry		-0.001	0.005	0.007	0.014
		(0.021)	(0.024)	(0.027)	(0.024)
Log Age	-0.091***	-0.092**	-0.086***	-0.086***	-0.081***
	(0.014)	(0.034)	(0.025)	(0.026)	(0.021)
Log Size	-0.009	-0.013	-0.016*	-0.014	-0.013*
	(0.006)	(0.012)	(0.010)	(0.010)	(0.008)
Patent Type	-	All	Breakthrough	Creative	RETech
Year FE	X	X	X	X	X
Observations	27,978	27,978	27,978	27,978	27,978
Within R^2	0.041	0.042	0.042	0.042	0.042

Notes: This table presents regressions at the firm-year level in Equation (42), where the outcome variable $\Delta Y_{i,t\to t+5}$ is employment growth between year t and $t+5$ calculated following Davis, Haltiwanger, and Schuh (1992) as $\Delta Y_{i,t\to t+5} = (Y_{i,t+5} - Y_{i,t})/(0.5 \times Y_{i,t+5} + 0.5 \times Y_{i,t})$. The independent variables include the predicted workstyle change in industry n over the next 5 years due to year t technologies ($|\widehat{dws}|^e_{n,t,5}$), constructed in Equation (39), and its interaction with firm age and firm size (log employment in year t). We standardize $|\widehat{dws}|^e_{n,t,5}$ to unit variance and zero median. Column (2) controls for the log number of total patents in industry n and year t, as well as its interactions with firm age and firm size. Column (3) controls for the log number of breakthrough patents in industry n and year t using data from Kelly et al. (2021), as well as its interactions with firm age and firm size. Column (4) controls for the log number of creative patents in industry n and year t using data from Kalyani (2025), as well as its interactions with firm age and firm size. Column (5) controls for the log number of rapidly evolving patents (RETech) in industry n and year t using data from Bowen, Frésard, and Hoberg (2023), as well as its interactions with firm age and firm size. We include year fixed effects, and bootstrap standard errors. In each bootstrap draw, we resample industries with replacement, and use the corresponding bootstrap weights to reestimate the technology-implied employment change by industry-occupation-year in Equation (37) and recalculate the predicted workstyle change in Equation (39), which is then used to rerun the final regression of interest in the bootstrap sample. The coefficient is the bootstrap median. Asterisks denote significance levels (***=1%, **=5%, *=10%). Number of observations and within R^2 (without fixed effects) are based on the OLS regression. Sample years are 2003 to 2016.

Another common question is whether old firms can address the challenges they face by acquiring other firms that do well in implementing the new technologies. Conceptually, if old firms can solve the problem through acquisition, then ultimately they would not display slower growth that we see in Tables 2 and

either assumption, larger firms, namely those with higher spanning capacities, respond more positively to the shock, conditional on age.

3, where the outcome variable is the firm-wide growth that includes both organic growth and growth through acquisitions. In fact, the model in Section III allows for the growth of business model B internally or through acquisitions—in either case, the organizational frictions we analyze can emerge. In practice, it is also well documented in management cases that many acquisitions faced integration challenges and were not able to help incumbents succeed in new businesses (Christensen, 1997). Nevertheless, we perform the following test on this hypothesis. Table IA7 reports the regression results of (42) but now using the change in acquisition spending as the outcome variable $\Delta Y_{i,t\to t+5}$. In this case, a greater workstyle change is not associated with increases in acquisitions spending for old firms.

V.C Growth of Young vs Old Firms in BDS Data

Finally, we turn to Census Business Dynamics Statistics (BDS), which has broader coverage of firms. Here the primary measure of size is employment (United States Census Bureau, 2023). In particular, the publicly available BDS dataset provides total employment of firms by industry and age group: 0, 1-5, 6-10, 11-15, 16-20, 21-25, and the remaining age groups cannot be consistently defined over our sample period because firms' precise age is unknown if they are born before 1976. Therefore we restrict to firms with age groups between 1-5 and 16-20 (which become age groups 6-10 to 21-25 five years later). We perform regressions at the industry-year-age group level:

$$\Delta \text{emp}_{i,t\to t+5} = \alpha_t + \sum_{g\neq g_0} \zeta_g \left(|\overline{dws}|^e_{n,t,5} \times \mathbf{1}_{i=g} \right) + \eta_w \, |\overline{dws}|^e_{n,t,5} + \sum_{g\neq g_0} \eta_{a,g} \, \mathbf{1}_{i=g} + \gamma \, z_{n,t} + \epsilon_{i,t}. \quad (43)$$

The outcome variable is the growth rate of employment among firm age group i in industry n between year t and $t+5$ (e.g., age group 1-5 becomes 6-10 after five years), calculated following Davis, Haltiwanger, and Schuh (1992). The independent variables include the predicted workstyle change in industry n over the next five years due to year-t technologies ($|\overline{dws}|^e_{n,t,5}$), and its interaction with age group indicators $\mathbf{1}\{i=g\}$, where g indexes age groups (6-10, 11-15, 16-20) and g_0 (1-5) is the omitted reference group. The controls include the patent variables in columns (2) to (5) of Table 1, as well as their interactions with age group dummies. We include year fixed effects, and bootstrap standard errors. We see that when $|\overline{dws}|^e_{n,t,5}$ is high, employment growth is faster for the young firm groups, and less so for the older firm groups.

V.D Additional Tests

We perform several additional analyses in this section to further address alternative mechanisms.

Table 5 – Employment Growth among BDS Firms

	Employment Growth (5-Year DHS Rate)				
	(1)	(2)	(3)	(4)	(5)
Predicted Workstyle Change in Industry × Age 6–10	0.005	0.004	0.004	0.004	0.003
	(0.010)	(0.010)	(0.009)	(0.010)	(0.010)
Predicted Workstyle Change in Industry × Age 11–15	-0.007	-0.007	-0.009	-0.008	-0.007
	(0.018)	(0.018)	(0.018)	(0.018)	(0.018)
Predicted Workstyle Change in Industry × Age 16–20	-0.026*	-0.026*	-0.029**	-0.028*	-0.028**
	(0.014)	(0.014)	(0.014)	(0.014)	(0.014)
Predicted Workstyle Change in Industry	0.021	0.021	0.022	0.021	0.020
	(0.016)	(0.016)	(0.016)	(0.016)	(0.016)
Log Patents in Industry × Age 6–10		0.005	0.005	0.006	0.007
		(0.006)	(0.007)	(0.008)	(0.009)
Log Patents in Industry × Age 11–15		0.007	0.012	0.009	0.005
		(0.006)	(0.009)	(0.010)	(0.010)
Log Patents in Industry × Age 16–20		0.014**	0.018*	0.015	0.015
		(0.007)	(0.010)	(0.010)	(0.011)
Log Patents in Industry		-0.010	-0.011	-0.007	-0.002
		(0.009)	(0.012)	(0.013)	(0.012)
Age 6–10	0.016	-0.006	0.004	0.004	0.004
	(0.016)	(0.029)	(0.024)	(0.022)	(0.020)
Age 11–15	0.013	-0.017	-0.015	-0.004	0.007
	(0.021)	(0.035)	(0.032)	(0.029)	(0.026)
Age 16–20	0.034	-0.028	-0.009	0.002	0.009
	(0.022)	(0.038)	(0.032)	(0.031)	(0.028)
Patent Type	-	All	Breakthrough	Creative	RETech
Year FE	X	X	X	X	X
Observations	4,032	4,032	4,032	4,032	4,032
Within R^2	0.019	0.020	0.020	0.019	0.021

Notes: This table presents regressions at the industry-age group-year level in Equation (43), where the outcome variable $\Delta Y_{i,t\to t+5}$ is employment growth of age group i in industry n between year t and $t+5$, calculated following Davis, Haltiwanger, and Schuh (1992) as $\Delta Y_{i,t\to t+5} = (Y_{i,t+5} - Y_{i,t})/(0.5 \times Y_{i,t+5} + 0.5 \times Y_{i,t})$. The independent variables include the predicted workstyle change in industry n over the next 5 years due to year t technologies ($|d\overline{ws}|^e_{n,t,5}$), constructed in Equation (39), and its interaction with age group dummies. We standardize $|d\overline{ws}|^e_{n,t,5}$ to unit variance and zero median. Column (2) controls for the log number of total patents in industry n and year t, as well as its interaction with firm age groups. Column (3) controls for the log number of breakthrough patents in industry n and year t using data from Kelly et al. (2021), as well as its interaction with firm age groups. Column (4) controls for the log number of creative patents in industry n and year t using data from Kalyani (2025), as well as its interaction with firm age groups. Column (5) controls for the log number of rapidly evolving patents (RETech) in industry n and year t using data from Bowen, Frésard, and Hoberg (2023), as well as its interaction with firm age groups. We include year fixed effects, and bootstrap standard errors. In each bootstrap draw, we resample industries with replacement, and use the corresponding bootstrap weights to reestimate the technology-implied employment change by industry-occupation-year in Equation (37) and recalculate the predicted workstyle change in Equation (39), which is then used to rerun the final regression of interest in the bootstrap sample. The coefficient is the bootstrap median. Asterisks denote significance levels (***=1%, **=5%, *=10%). Number of observations and within R^2 (without fixed effects) are based on the OLS regression. Sample years are 2003 to 2016.

Uncertainty Recent work by Goldman, Slutzky, and Zeume (2025) using global data finds that uncertainty appears to hurt old firms more than young firms: older firms contract significantly more than younger firms during periods of high uncertainty. In Tables IA9, IA10, IA11, we check that our results on predicted workstyle change are not driven by this variable being correlated with uncertainty. We use uncertainty measured from World Uncertainty Index like Goldman, Slutzky, and Zeume (2025), as well as economic policy uncertainty (Baker, Bloom, and Davis, 2016) and firm-level stock return volatility. The uncertainty controls do not affect our main results.

Cannibalization A classic hypothesis is that incumbents may fail to implement new technologies because doing so will cannibalize their existing businesses (Arrow, 1962). When existing businesses are highly profitable, the net gain from replacing them with new technologies would be smaller, weakening incumbents' incentive to implement new technologies. Our core hypothesis of incumbents facing stronger organizational frictions in implementing new technologies that require organizational change may exist independently of Arrow's replacement effect, which is at work even without any agency frictions; there is no obvious reason to expect any interaction effect between profitability and workstyle change.

Nonetheless, it is useful to check empirically that the role of firm age in our main results is not just driven by older firms having more current profits that may be cannibalized. In Appendix Tables IA12, IA13, IA14, we perform additional analyses controlling for predicted workstyle change interacting with current profits, measured as net income over assets, net income over sales, and net income over stock market capitalization (effectively the inverse of P/E ratio). These empirical results confirm that our main findings are robust to these controls, and as expected we do not observe a clear interaction between profitability and the predicted workstyle change.

VI Analysis Based on Wikipedia Articles

In Section II, we used Wikipedia articles to present stylized facts about technologies implemented by young vs old firms. In this section, we return to Wikipedia articles as the measure of new technologies. In Section VI.A, we repeat the main empirical analyses in Section V using Wikipedia articles instead of patents. In Section VI.B, we further examine whether Wikipedia articles that imply greater workstyle change are more likely to be implemented by young firms, as measured in Section II. The advantage of Wikipedia articles is that they capture technologies that are well-defined, so we can trace whether the implementation and commercialization are achieved by young vs old firms as in Section II. This allows us to perform the additional test in Section VI.B, and verify that new technologies associated with greater workstyle changes are more compatible with young firms (and less compatible with old firms). In comparison, patents are

numerous and relatively micro, so it is rather difficult to trace out what types of firms implemented and commercialized the technologies in each patent.

VI.A Predicted Workstyle Change and Firm Outcomes using Wikipedia Articles

We first present the main analyses about the performance of young vs old firms using predicted workstyle change based on Wikipedia articles instead of patents. The data construction is similar to that in Section IV.B, which we explain in Section VI.A.1. We then present the results in Section VI.A.2.

VI.A.1 Data Construction: Predicted Workstyle Change using Wikipedia Articles

We compute predicted workstyle change in each industry-year using Wikipedia articles, analogous to the construction using patent data in Section IV.B. To align the sample coverage with the patent-based measures, we focus on the U.S. over the period 2003–2016, yielding a total of 1,255 Wikipedia articles.

Calculate Wikipedia articles' similarity with occupation j's routine and nonroutine tasks We start by extracting technical keywords from the Wikipedia article's summary using LLM (Gemini 2.5 Pro), which capture the core mechanisms and characteristics of the article, as explained in detail in Appendix Section IA4.2. We represent each Wikipedia article w as the mean of its keywords' word embeddings using the same *GloVe* model as in Section IV.B:

$$\mathcal{X}_w = \frac{1}{K_w} \sum_{k=1}^{K_w} x_k, \tag{44}$$

where K_w is the number of keyword words for Wikipedia article w and x_k are GloVe word vectors. We use keywords for Wikipedia articles (instead of TF-IDF as in the case with patents in Section IV.B) because Wikipedia articles contain more background information (e.g., time and place of invention) aside from descriptions about the functions of the technology. The occupation task representations $\mathcal{X}_j^\tau$ ($\tau \in \{R, NR\}$) are constructed the same way as in Section IV.B, using TF-IDF weighted GloVe embeddings of each occupation's routine and nonroutine task descriptions (keeping only verbs and nouns and removing stop words before calculating the embeddings). We calculate cosine similarity $s_{w,j}^\tau$ between Wikipedia article w and the routine or nonroutine component of occupation j, analogous to (35), and apply the same adjustments of removing year fixed effects and imposing sparsity (at the 80th percentile) to obtain $\tilde{s}_{w,j}^\tau$.

Sum over Wikipedia articles in industry n and year t to get occupation j's exposure $\xi^R_{j,n,t}$ and $\xi^{NR}_{j,n,t}$ We assign each Wikipedia article to industries using semantic similarity.[30] Specifically, we represent each three-digit NAICS industry description as a TF-IDF weighted GloVe embedding, and compute cosine similarity between each Wikipedia article's keyword embedding $\mathcal{X}_w$ and each NAICS industry embedding. We normalize the resulting similarities to probability weights $\pi_{w,n}$ that sum to one across industries for each technology (we impose sparsity by setting weights below the 80th percentile to zero before calculating the probability weights). The exposure of industry n and occupation j to technology at year t is then:

$$\xi^{\tau}_{j,n,t} = \theta^{\tau}_j \log\left(1 + \sum_{w \in \mathcal{W}_t} \pi_{w,n}\, \tilde{s}^{\tau}_{w,j}\right), \quad \tau \in \{R, NR\}, \tag{45}$$

where θ^{τ}_j is the routine/nonroutine share of tasks for occupation j, $\mathcal{W}_t$ is the set of Wikipedia articles with invention year t, and $\pi_{w,n}$ is the probability weight linking Wikipedia article w to industry n.

Predict future employment of occupation j in industry n We follow the same approach as in Section IV.B and use occupation exposure to technologies $\xi^R_{j,n,t}$ and $\xi^{NR}_{j,n,t}$ from (45) to predict future employment $\widehat{\log l_{j,n,t+5}}$, by running the same regression as Equation (37) of log employment on exposure through routine tasks, exposure through nonroutine tasks, and current employment.

$$\widehat{\log l_{j,n,t+5}} = \underset{(0.061)}{0.314^{***}} - \underset{(0.137)}{0.749^{***}} \cdot \xi^R_{j,n,t} + \underset{(0.064)}{0.194^{**}} \cdot \xi^{NR}_{j,n,t} + \underset{(0.006)}{0.954^{***}} \cdot \log l_{j,n,t}. \tag{46}$$

Consistent with the results in (37) and with intuition, we observe that the employment of an occupation declines by more when new technologies are more similar to its routine tasks, indicating substitution and replacement, while the employment rises by more when new technologies are more similar to its nonroutine tasks, indicating complementarity and enhancement.

Predicted workstyle change in industry n from year t to $t+5$ The predicted workstyle change $|d\overline{ws}|^e_{n,t,5}$ is computed from the predicted employment $\widehat{\log l_{j,n,t+5}}$ following the same procedure as in Equations (38) and (39). As before, we standardize $|d\overline{ws}|^e_{n,t,5}$ to unit variance and zero median.

[30] We obtain three-digit NAICS industry descriptions from the 2017 NAICS classification (U.S. Census Bureau, 2017) and clean them using LLM (Gemini 2.5 Pro) to remove clauses describing what the industry does *not* include, retaining only descriptions of what the industry covers. We preprocess the cleaned NAICS descriptions using the same text pipeline as for the occupation tasks (lowercase, tokenize, POS-tag, keep nouns and verbs, remove stopwords, and lemmatize). Each NAICS industry is then represented as a TF-IDF-weighted GloVe embedding of its processed description. We compute cosine similarity between each Wikipedia article's keyword embedding $\mathcal{X}_w$ and each NAICS industry embedding, impose sparsity by setting weights below the 80th percentile to zero, and normalize the remaining similarities to probability weights $\pi_{w,n}$ that sum to one across industries for each technology.

VI.A.2 Empirical Results

We perform the analyses in Section V now using the predicted workstyle change in (46). Table IA2 repeats the analysis in Table 1; Tables IA3, IA4, and IA5 repeat the analyses in Tables 2, 3, and 4; Table IA6 repeats the analysis in Table 5. The results are very similar.

VI.B Predicted Workstyle Change and Technology Implementation

The analyses in Sections V and VI.A examine how young and old firms perform differently in industries experiencing greater predicted workstyle change. Wikipedia articles allow a further test: for each individual technology in a Wikipedia article, we can ask whether technologies that imply greater workstyle disruption are more likely to be implemented by young firms, as identified in Section II.

We construct predicted workstyle change associated with each individual Wikipedia technology w in its primary industry (the industry with the highest probability weight $\pi_{w,n}$ from the mapping in Section VI.A.1). We follow the same procedure as in Section VI.A.1, but compute the workstyle change attributable to technology w alone, rather than summing over all technologies in the industry-year. Specifically, we calculate occupation j's exposure to technology w as $\theta_j^\tau \log(1+\tilde{s}_{w,j}^\tau)$ for $\tau \in \{R, NR\}$ (where θ_j^τ is the routine/nonroutine share of tasks τ for occupation j and $\tilde{s}_{w,j}^\tau$ is the similarity between Wikipedia article w and the task), similar to Equation (45) except for a single technology. We then use the estimated coefficients from Equation (46) to derive predicted employment in five years $\hat{l}_{j,n,t+5}^w$ for each occupation associated with technology w, and in turn the associated workstyle $\mathbf{WS}_{n,t+5}^{e,w}$. The technology-level predicted workstyle change is then defined as:

$$|d\overline{ws}|_{n,t,5}^{e,w} \equiv \|\mathbf{WS}_{n,t+5}^{e,w} - \mathbf{WS}_{n,t}\|, \tag{47}$$

with $\mathbf{WS}_{n,t}$ being the current workstyle in technology w's primary industry n. We standardize $|d\overline{ws}|_{n,t,5}^{e,w}$ to unit variance and zero median.

We run the following technology-level regression:

$$\mathbf{1}[\text{Technology } w \text{ Implemented by Young}] = \alpha_{t_w} + \beta\ |d\overline{ws}|_{n,t,5}^{e,w} + \epsilon_w, \tag{48}$$

where the independent variable is an indicator equal to one if the company most successful in the initial implementation and commercialization of the technology is a young firm (less than ten years old at the time), independently classified using LLM as described in Section II.A and Appendix Section IA2.3. The coefficient β captures whether technologies with greater predicted workstyle change are more likely to be implemented by young firms. We use U.S. technologies with invention years from 2003 to 2016 like in

Table 6 – Technology Implementation by Young vs Old Firms

	1[Implemented by Young]			
	All Technologies		Classified Technologies	
	(1)	(2)	(3)	(4)
Technology-level Predicted Workstyle Change	0.036***	0.036***	0.049***	0.047**
	(0.012)	(0.012)	(0.017)	(0.018)
Log Wikipedia Article Length		-0.011		-0.049
		(0.018)		(0.022)
Year FE	X	X	X	X
Observations	1,255	1,255	899	899
Within R^2	0.022	0.022	0.030	0.040

Notes: This table presents regressions at the technology level in Equation (48). The dependent variable is an indicator equal to one if the company most successful in the initial implementation and commercialization of the technology is a young firm (founded less than ten years before commercialization), classified using LLM as described in Section II.A and Appendix Section IA2.3. The key independent variable is the technology-level predicted workstyle change $|\overline{dws}|^{e,w}_{n,t,5}$ in Equation (47), standardized to unit variance and zero median. Columns (1) and (2) include all technologies; Columns (3) and (4) restrict to technologies with definitive firm type classifications, dropping those classified as ambiguous or never commercially implemented. All specifications include invention year fixed effects and use bootstrap standard errors. In each bootstrap draw, we resample industries with replacement, and use the corresponding bootstrap weights to reestimate the technology-level employment change (including reestimating coefficients from Equation (46)), and recalculate the predicted workstyle change in Equation (47), which is then used to rerun the final regression of interest in the bootstrap sample. The coefficient is the bootstrap median. Asterisks denote significance levels (***=1%, **=5%, *=10%). Number of observations and within R^2 (without fixed effects) are based on the OLS regression. Sample years are 2003 to 2016.

Section VI.A. We include year fixed effects based on invention year of technology w (in case Wikipedia coverage of technologies varies over time). We bootstrap standard errors as before. In each bootstrap draw, we resample industries with replacement, and use the corresponding bootstrap weights to reestimate the technology-level predicted employment change and recalculate the predicted workstyle change in Equation (47), which is then used to rerun the final regression of interest in the bootstrap sample.

Table 6 presents the results. We find that β is consistently positive and significant: technologies that predict larger workstyle changes are more likely to be implemented by young firms. Column (1) includes all technologies while Column (3) restricts to technologies with definitive classifications (young or old), excluding those classified as ambiguous or never commercially implemented. These technology-level findings provide direct evidence that new technologies that require greater change in workstyle are more likely to be implemented by young firms, consistent with our industry-level results.

Columns (2) and (4) in Table 6 further control for the length of the entire Wikipedia article (not just the summary), to account for the prominence or complexity of the technology. For example, one alternative hypothesis can be that more powerful technologies may lead to greater workstyle change, and they tend to be implemented by young firms. In the data, longer Wikipedia articles tend to be weakly associated with

implementation by old firms instead (not statistically significant), and this control does not affect our key coefficient of interest on the predicted workstyle change.

VII Conclusion

Every year, over 200 NBER working papers are written on entrepreneurship (National Bureau of Economic Research, 2025). Given the popularity of this subject, it is essential to address the fundamental question: why don't old firms do new things, and when do we need new firms to implement new technologies? Our work highlights the importance of organizational frictions for understanding organizational change and disruption. In particular, existing firms are not always incapable of implementing new technologies due to the slowness of learning or fear of cannibalization. However, they are especially vulnerable when new technologies require changes in organizational processes and priorities, which naturally favor new organizations that start fresh. With advances in data and measurement—thanks to more ways to extract information from texts—future research can make further progress in capturing organizational frictions and investigating their role in innovation, productivity, and growth.

References

Acemoglu, Daron, and David Autor. 2011. "Skills, tasks and technologies: Implications for employment and earnings." In *Handbook of Labor Economics*. Vol. 4, 1043–1171. Elsevier.

Acemoglu, Daron, Ufuk Akcigit, Harun Alp, Nicholas Bloom, and William Kerr. 2018. "Innovation, reallocation, and growth." *American Economic Review*, 108(11): 3450–3491.

Aghion, Philippe, and Jean Tirole. 1994. "The management of innovation." *Quarterly Journal of Economics*, 109(4): 1185–1209.

Aghion, Philippe, and Jean Tirole. 1995. "Some implications of growth for organizational form and ownership structure." *European Economic Review*, 39(3-4): 440–455.

Aghion, Philippe, and Jean Tirole. 1997. "Formal and real authority in organizations." *Journal of Political Economy*, 105(1): 1–29.

Aghion, Philippe, and Peter Howitt. 1992. "A model of growth through creative destruction." *Econometrica*, 323–351.

Akcigit, Ufuk, and William R Kerr. 2018. "Growth through heterogeneous innovations." *Journal of Political Economy*, 126(4): 1374–1443.

Anton, James J, and Dennis A Yao. 1995. "Start-ups, spin-offs, and internal projects." *Journal of Law, Economics, and Organization*, 11(2): 362–378.

Argente, David, Salomé Baslandze, Douglas Hanley, and Sara Moreira. 2025. "Patents to products: Product innovation and firm dynamics." Working Paper.

Arrow, Kenneth J. 1964. "Control in large organizations." *Management Science*, 10(3): 397–408.

Arrow, Kenneth Joseph. 1962. "Economic welfare and the allocation of resources for invention." In *Readings in Industrial Economics: Volume two: Private Enterprise and State Intervention*. 219–236. Springer.

Asirvatham, Hemanth. 2024. "The dynamo is not the computer." Working Paper.

Atalay, Enghin, Ali Hortaçsu, and Chad Syverson. 2014. "Vertical integration and input flows." *American Economic Review*, 104(4): 1120–1148.

Autor, David H, Frank Levy, and Richard J Murnane. 2003. "The skill content of recent technological change: An empirical exploration." *Quarterly Journal of Economics*, 118(4): 1279–1333.

Babina, Tania, Anastassia Fedyk, Alex He, and James Hodson. 2024. "Artificial intelligence, firm growth, and product innovation." *Journal of Financial Economics*, 151: 103745.

Baker, Scott R, Nicholas Bloom, and Steven J Davis. 2016. "Measuring economic policy uncertainty." *Quarterly Journal of Economics*, 131(4): 1593–1636.

Beaumont, Paul, Camille Hebert, and Victor Lyonnet. 2025. "Build or buy? Human capital and corporate diversification." *Review of Financial Studies*, 38(5): 1333–1367.

Bloom, Nicholas, and John Van Reenen. 2007. "Measuring and explaining management practices across firms and countries." *Quarterly Journal of Economics*, 122(4): 1351–1408.

Bowen, Donald E, Laurent Frésard, and Gerard Hoberg. 2023. "Rapidly evolving technologies and startup exits." *Management Science*, 69(2): 940–967.

Braguinsky, Serguey, Joonkyu Choi, Yuheng Ding, Karam Jo, and Seula Kim. 2024. "Megafirms and recent trends in the US innovation: Empirical evidence from the US patent data." Working Paper.

Bresnahan, Timothy F, Erik Brynjolfsson, and Lorin M Hitt. 2002. "Information technology, workplace organization, and the demand for skilled labor: Firm-level evidence." *Quarterly Journal of Economics*, 117(1): 339–376.

Caskurlu, Tolga, Gerard Hoberg, and Gordon M Phillips. 2024. "New technology sectoral disruptions." Working Paper.

Chassang, Sylvain. 2010. "Building routines: Learning, cooperation, and the dynamics of incomplete relational contracts." *American Economic Review*, 100(1): 448–465.

Christensen, Clayton M. 1997. *The Innovator's Dilemma: When New Technologies Cause Great Firms to Fail.* Harvard Business School Press.

Coase, Ronald Harry. 1937. "The nature of the firm." *Economica*, 4(16): 386–405.

Cohan, William D. 2022. *Power Failure: The Rise and Fall of an American Icon.* Penguin.

Cohen, Lauren, Umit G Gurun, and Quoc H Nguyen. 2022. "The ESG-innovation disconnect: Evidence from green patenting." Working Paper.

Davis, River. 2023. "The World's biggest carmaker made a huge bet on tech. Things went wrong fast." *The Wall Street Journal*. (Accessed November 9, 2025).

Davis, Steven J., John Haltiwanger, and Scott Schuh. 1992. "Gross job creation and destruction." *Journal of Labor Economics*, 10(3): 305–358.

Decker, Ryan, John Haltiwanger, Ron Jarmin, and Javier Miranda. 2014. "The role of entrepreneurship in US job creation and economic dynamism." *Journal of Economic Perspectives*, 28(3): 3–24.

Dessein, Wouter, and Tano Santos. 2006. "Adaptive organizations." *Journal of political Economy*, 114(5): 956–995.

Dessein, Wouter, Luis Garicano, and Robert Gertner. 2010. "Organizing for synergies." *American Economic Journal: Microeconomics*, 2(4): 77–114.

Draghi, Mario. 2024. "The future of European competitiveness - A competitiveness strategy for Europe." , https://commission.europa.eu/topics/eu-competitiveness/draghi-report_en.

Ewens, Michael, and Matt Marx. 2024. "Firm age and invention: An open-access dataset." Working Paper.

Garcia-Macia, Daniel, Chang-Tai Hsieh, and Peter J Klenow. 2019. "How destructive is innovation?" *Econometrica*, 87(5): 1507–1541.

Garicano, Luis. 2000. "Hierarchies and the organization of knowledge in production." *Journal of political economy*, 108(5): 874–904.

Garicano, Luis, and Esteban Rossi-Hansberg. 2004. "Inequality and the organization of knowledge." *American Economic Review*, 94(2): 197–202.

Garicano, Luis, and Luis Rayo. 2016. "Why organizations fail: Models and cases." *Journal of Economic Literature*, 54(1): 137–192.

George, Patrick. 2022. "Volkswagen has given a name to its pain, and it is 'software'." *The Verge*. (Accessed November 9, 2025).

Gerstner, Louis V. 2002. *Who Says Elephants Can't Dance? Inside IBM's Historic Turnaround.* HarperCollins Publishers.

Goldman, Jim, Pablo Slutzky, and Stefan Zeume. 2025. "A dark side to experience? Firm age and uncertainty." Working Paper.

Goldschlag, Nathan, Travis J Lybbert, and Nikolas Jason Zolas. 2016. "An 'algorithmic links with probabilities' crosswalk for uspc and cpc patent classifications with an application towards industrial technology composition." *US Census Bureau Center for Economic Studies paper no. CES-WP-16-15*.

Gorton, Gary B, and Alexander K Zentefis. 2024. "Corporate culture as a theory of the firm." *Economica*, 91(364): 1391–1423.

Grossman, Sanford J, and Oliver D Hart. 1986. "The costs and benefits of ownership: A theory of vertical and lateral integration." *Journal of Political Economy*, 94(4): 691–719.

Hart, Oliver, and Bengt Holmstrom. 2010. "A theory of firm scope." *Quarterly Journal of Economics*, 125(2): 483–513.

Hart, Oliver, and John Moore. 1990. "Property Rights and the Nature of the Firm." *Journal of political economy*, 98(6): 1119–1158.

Hart, Oliver, and John Moore. 2008. "Contracts as reference points." *Quarterly Journal of Economics*, 123(1): 1–48.

Henderson, Rebecca M, and Kim B Clark. 1990. "Architectural innovation: The reconfiguration of existing product technologies and the failure of established firms." *Administrative science quarterly*, 9–30.

Holmstrom, Bengt. 1989. "Agency costs and innovation." *Journal of Economic Behavior & Organization*, 12(3): 305–327.

Holmstrom, Bengt, and Paul Milgrom. 1991. "Multitask principal–agent analyses: Incentive contracts, asset ownership, and job design." *The Journal of Law, Economics, and Organization*, 7(special_issue): 24–52.

Hsieh, Chang-Tai, and Esteban Rossi-Hansberg. 2023. "The industrial revolution in services." *Journal of Political Economy: Macroeconomics*, 1(1): 3–42.

Kalyani, Aakash. 2025. "The creativity decline: Evidence from US patents." Working Paper.

Kalyani, Aakash, Nicholas Bloom, Marcela Carvalho, Tarek Hassan, Josh Lerner, and Ahmed Tahoun. 2025. "The diffusion of new technologies." *Quarterly Journal of Economics*, 140(2): 1299–1365.

Kelly, Bryan, Dimitris Papanikolaou, Amit Seru, and Matt Taddy. 2021. "Measuring technological innovation over the long run." *American Economic Review: Insights*, 3(3): 303–320.

Kogan, Leonid, Dimitris Papanikolaou, Lawrence DW Schmidt, and Bryan Seegmiller. 2024. "Technology and labor displacement: Evidence from linking patents with worker-level data." Working Paper.

Kortum, Samuel, and Josh Lerner. 2000. "Assessing the contribution of venture capital to innovation." *RAND journal of Economics*, 674–692.

Lashkari, Danial, Arthur Bauer, and Jocelyn Boussard. 2024. "Information technology and returns to scale." *American Economic Review*, 114(6): 1769–1815.

Lee, Kyeong Hun, David C Mauer, and Emma Qianying Xu. 2018. "Human capital relatedness and mergers and acquisitions." *Journal of Financial Economics*, 129(1): 111–135.

Leonard-Barton, Dorothy. 1992. "Core capabilities and core rigidities: A paradox in managing new product development." *Strategic Management Journal*, 13(S1): 111–125.

Lerner, Josh. 2006. "The new new financial thing: The origins of financial innovations." *Journal of Financial Economics*, 79(2): 223–255.

Levitt, Steven D, John A List, and Chad Syverson. 2013. "Toward an understanding of learning by doing: Evidence from an automobile assembly plant." *Journal of Political Economy*, 121(4): 643–681.

Loderer, Claudio, René Stulz, and Urs Waelchli. 2017. "Firm rigidities and the decline in growth opportunities." *Management Science*, 63(9): 3000–3020.

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

Meyer, Margaret, Paul Milgrom, and John Roberts. 1992. "Organizational prospects, influence costs, and ownership changes." *Journal of Economics & Management Strategy*, 1(1): 9–35.

Milgrom, Paul R, and John Roberts. 1992. *Economics, Organization and Management.* Vol. 7, Prentice-hall Englewood Cliffs, NJ.

Minni, Virginia M. L. 2026. "Making the invisible hand visible: Managers and the allocation of workers to jobs." *Quarterly Journal of Economics*. Forthcoming.

National Bureau of Economic Research. 2025. "NBER working papers and chapters metadata." https://www.nber.org/research/data/nber-working-papers-and-chapters-metadata Updated weekly; available in CSV, TSV, and Stata (DTA) formats.

Ndiaye, Pap. 2007. *Nylon and Bombs: DuPont and the March of Modern America.* JHU Press.

Nelson, Richard R, and Sidney G Winter. 1985. *An evolutionary theory of economic change.* Harvard University Press.

O'Reilly, Charles A, and Michael L Tushman. 2021. *Lead and Disrupt: How to Solve the Innovator's Dilemma.* Stanford University Press.

Penrose, Edith T. 1959. *The Theory of the Growth of the Firm.* Oxford:Basil Blackwell.

Qian, Yingyi, Gerard Roland, and Chenggang Xu. 2006. "Coordination and experimentation in M-form and U-form organizations." *Journal of Political Economy*, 114(2): 366–402.

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

Rajan, Raghuram G, and Luigi Zingales. 2001. "The firm as a dedicated hierarchy: A theory of the origins and growth of firms." *Quarterly Journal of Economics*, 116(3): 805–851.

Rajan, Raghuram, Henri Servaes, and Luigi Zingales. 2000. "The cost of diversity: The diversification discount and inefficient investment." *Journal of Finance*, 55(1): 35–80.

Refinitiv. 2025*a*. "Refinitiv / LSEG workspace private equity screener." https://workspace.refinitiv.com/web/Apps/ScreenerApp/?universe=PVTInvestment Accessed through Refinitiv LSEG (Refinitiv Workspace) platform.

Refinitiv. 2025*b*. "Refinitiv Eikon / Workspace Database." https://www.refinitiv.com/en/products/eikon-trading-software Accessed through Refinitiv Eikon (Refinitiv Workspace) platform.

Revelio Labs. 2025. "Individual reviews (sentiment) dataset." https://wrds-www.wharton.upenn.edu/pages/get-data/revelio-labs/revelio-sentiment/individual-reviews/ Accessed November 6, 2025.

Scharfstein, David S, and Jeremy C Stein. 2000. "The dark side of internal capital markets: Divisional rent-seeking and inefficient investment." *Journal of Finance*, 55(6): 2537–2564.

Schumpeter, Joseph Alois. 1942. *Capitalism, Socialism and Democracy.* Allen & Unwin London.

Shleifer, Andrei, and Robert W Vishny. 1989. "Management entrenchment: The case of manager-specific investments." *Journal of Financial Economics*, 25(1): 123–139.

S&P Global Market Intelligence. 2025. "Compustat North America Fundamentals Annual Database." https://wrds-www.wharton.upenn.edu/ Accessed via Wharton Research Data Services (WRDS).

Stein, Jeremy C. 1997. "Internal capital markets and the competition for corporate resources." *Journal of Finance*, 52(1): 111–133.

Teece, David J, Gary Pisano, and Amy Shuen. 1997. "Dynamic capabilities and strategic management." *Strategic Management Journal*, 18(7): 509–533.

Tushman, Michael L, and Philip Anderson. 1986. "Technological discontinuities and organizational environments." *Administrative Science Quarterly*, 31(3): 439–65.

United States Census Bureau. 2023. "Business Dynamics Statistics Datasets." https://www.census.gov/data/datasets/time-series/econ/bds/bds-datasets.html Accessed September 25, 2025".

U.S. Bureau of Labor Statistics. 2025. "Occupational Employment and Wage Statistics (OEWS) Tables." https://www.bls.gov/oes/tables.htm Accessed November 1, 2025.

U.S. Census Bureau. 2017. "2017 NAICS Descriptions." , https://www.census.gov/naics/2017NAICS/2017_NAICS_Descriptions.xlsx (Accessed February 2, 2026).

U.S. Department of Labor, Employment and Training Administration. 2005. "O*NET Database, Version 9.0 (December 2005)." https://www.onetcenter.org/dl_files/ncsc/onet90ac.zip Accessed November 8, 2025.

U.S. Department of Labor, Employment and Training Administration. 2025. "O*NET® Database Releases Archive." https://www.onetcenter.org/db_releases.html Accessed September 30, 2025.

Williamson, Oliver E. 1971. "The vertical integration of production: Market failure considerations." *American Economic Review*, 61(2): 112–123.

Internet Appendix

IA1 Additional Figures and Tables

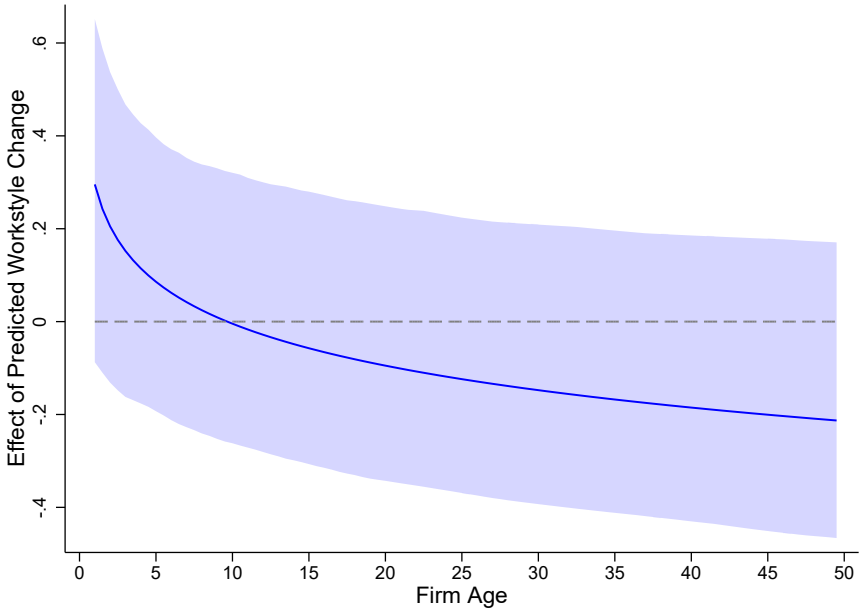

Figure IA1. Tobin's Q among Compustat Firms: Implied Coefficient on $|\overline{dws}|^{e}_{n,t,5}$

Notes: This figure shows the implied coefficient on $|\overline{dws}|^{e}_{n,t,5}$ for different levels of firm age in Table 2, column (1).

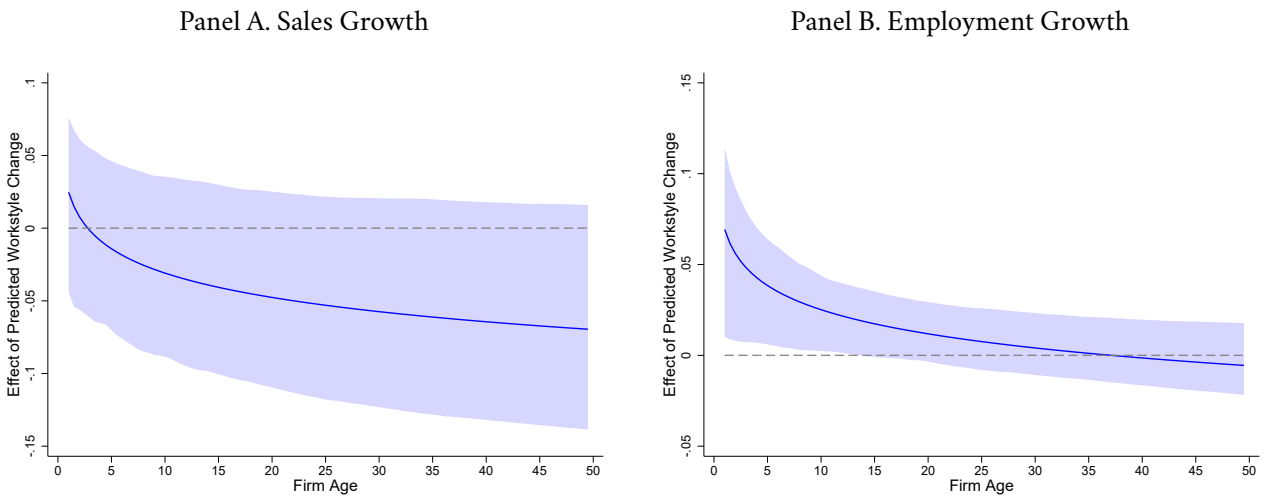

Figure IA2. Sales and Employment Growth among Compustat Firms: Implied Coefficient on $|d\overline{ws}|^e_{n,t,5}$

Notes: Panel A shows the implied coefficient on $|d\overline{ws}|^e_{n,t,5}$ for different levels of firm age in Table 3, column (1). Panel B shows the implied coefficient on $|d\overline{ws}|^e_{n,t,5}$ for different levels of firm age in Table 4, column (1).

Table IA1 – Summary Statistics

	P25	P50	P75	Mean	SD	N
Panel A: Venture Capital Sample						
Predicted Workstyle Change in Industry (Patent)	-0.15	-0.00	0.30	0.33	1.02	754
Predicted Workstyle Change in Industry (Wiki)	-0.13	-0.01	0.25	0.32	1.02	754
Log VC Investment in Industry	16.72	18.12	19.68	18.15	2.30	754
Log Compustat Market Cap in Industry	24.09	25.22	26.43	25.24	1.44	754
Log(VC Investment/Market Cap)	-8.27	-6.84	-5.64	-7.09	2.02	754
Log Total Patents in Industry	2.82	5.07	7.16	4.85	2.73	754
Log Breakthrough Patents in Industry	0.37	2.63	4.61	2.73	2.31	754
Log Creative Patents in Industry	0.54	2.35	4.02	2.44	2.00	754
Log RETech Patents in Industry	0.11	1.36	3.44	1.99	1.99	754
Panel B: Compustat Sample						
Predicted Workstyle Change in Industry (Patent)	-0.07	0.21	0.80	0.56	1.02	27,978
Predicted Workstyle Change in Industry (Wiki)	-0.11	0.08	0.59	0.44	0.97	27,978
Log Firm Age	2.89	3.47	4.19	3.51	0.87	27,978
Firm Tobin's Q	0.91	1.33	2.15	1.91	2.10	27,978
Sales Growth Next 5 Years	-0.03	0.25	0.58	0.27	0.56	27,978
Employment Growth Next 5 Years	-0.11	0.14	0.44	0.16	0.48	27,978
Log Firm Assets	4.52	6.13	7.65	6.08	2.26	27,978
Log Firm Sales	4.30	6.09	7.56	5.85	2.48	27,978
Log Firm Employment	-1.31	0.45	1.93	0.33	2.25	27,978
Log Total Patents in Industry	3.81	7.60	8.81	6.33	3.21	27,978
Log Breakthrough Patents in Industry	1.15	4.86	6.31	4.25	2.91	27,978
Log Creative Patents in Industry	1.22	4.63	5.76	3.84	2.54	27,978
Log RETech Patents in Industry	0.78	3.63	5.66	3.43	2.60	27,978
Panel C: BDS Sample						
Predicted Workstyle Change in Industry (Patent)	-0.14	-0.00	0.30	0.33	1.01	4,032
Predicted Workstyle Change in Industry (Wiki)	-0.13	-0.00	0.25	0.33	1.01	4,032
Employment Growth Next 5 Years	-0.24	-0.09	0.02	-0.10	0.26	4,032
Log Employment in Industry	12.79	13.43	14.21	13.46	1.26	4,032
Log Total Patents in Industry	1.99	4.24	6.58	4.40	2.68	4,032
Log Breakthrough Patents in Industry	0.31	2.01	4.31	2.44	2.23	4,032
Log Creative Patents in Industry	0.33	1.65	3.45	2.13	1.93	4,032
Log RETech Patents in Industry	0.07	1.00	3.11	1.71	1.91	4,032

Notes: This table reports summary statistics for the main variables used in the analysis. Panel A presents statistics for the venture capital sample at the industry-year level. Panel B presents statistics for the Compustat sample at the firm-year level. Panel C presents statistics for the BDS sample at the industry-age group-year level. Reported values include the 25th, 50th, and 75th percentiles (P25, P50, P75), mean, and standard deviation (SD).

Table IA2 – Venture Capital Investment: Wikipedia Article Results

	Forward 1-Year Log(VC Investment/Compustat Market Cap)				
	(1)	(2)	(3)	(4)	(5)
Predicted Workstyle Change in Industry	0.428***	0.419***	0.388***	0.405***	0.400***
	(0.154)	(0.157)	(0.151)	(0.154)	(0.155)
Log Patents in Industry		-0.035	0.088	0.047	0.074
		(0.080)	(0.093)	(0.104)	(0.107)
Patent Type	-	All	Breakthrough	Creative	RETech
Year FE	X	X	X	X	X
Observations	754	754	754	754	754
Within R^2	0.045	0.045	0.056	0.048	0.052

Notes: This table presents regressions at the industry-year level in Equation (40). The outcome variable is log VC investment relative to Compustat market capitalization in industry n following year t ($\log VC/MCAP_{n,t+1}$). The independent variables include the predicted workstyle change in industry n over the next 5 years due to year t Wikipedia technologies ($|\overline{dws}|^e_{n,t,5}$), constructed in Equation (39). We standardize $|\overline{dws}|^e_{n,t,5}$ to unit variance and zero median. Column (2) controls for the log number of total patents in industry n and year t. Column (3) controls for the log number of breakthrough patents in industry n and year t using data from Kelly et al. (2021). Column (4) controls for the log number of creative patents in industry n and year t using data from Kalyani (2025). Column (5) controls for the log number of rapidly evolving patents (RETech) in industry n and year t using data from Bowen, Frésard, and Hoberg (2023). We include year fixed effects, and bootstrap standard errors. In each bootstrap draw, we resample industries with replacement, and use the corresponding bootstrap weights to reestimate the technology-implied employment change by industry-occupation-year in Equation (46) and recalculate the predicted workstyle change in Equation (39), which is then used to rerun the final regression of interest in the bootstrap sample. The coefficient is the bootstrap median. Asterisks denote significance levels (***=1%, **=5%, *=10%). Number of observations and within R^2 (without fixed effects) are based on the OLS regression. Sample years are 2003 to 2016.

Table IA3 – Tobin's Q among Compustat Firms: Wikipedia Article Results

	Tobin's Q				
	(1)	(2)	(3)	(4)	(5)
Predicted Workstyle Change in Industry × Log Age	-0.122**	-0.146***	-0.131***	-0.137***	-0.143***
	(0.049)	(0.041)	(0.036)	(0.037)	(0.040)
Predicted Workstyle Change in Industry × Log Size	0.040*	0.045**	0.047**	0.046**	0.047**
	(0.019)	(0.018)	(0.018)	(0.018)	(0.020)
Predicted Workstyle Change in Industry	0.199	0.174	0.095	0.143	0.169
	(0.205)	(0.159)	(0.152)	(0.147)	(0.151)
Log Patents in Industry × Log Age		-0.009	-0.024	-0.026	-0.020
		(0.029)	(0.032)	(0.035)	(0.034)
Log Patents in Industry × Log Size		-0.003	-0.005	-0.006	-0.012
		(0.013)	(0.017)	(0.019)	(0.019)
Log Patents in Industry		0.091	0.175	0.200	0.233
		(0.148)	(0.146)	(0.166)	(0.158)
Log Age	-0.287***	-0.222	-0.183	-0.183	-0.198**
	(0.081)	(0.183)	(0.124)	(0.126)	(0.106)
Log Size	-0.181***	-0.158*	-0.156**	-0.152**	-0.133**
	(0.039)	(0.073)	(0.059)	(0.059)	(0.052)
Patent Type	-	All	Breakthrough	Creative	RETech
Year FE	X	X	X	X	X
Observations	27,978	27,978	27,978	27,978	27,978
Within R^2	0.079	0.081	0.082	0.083	0.084

Notes: This table presents regressions at the firm-year level in Equation (41). The independent variables include the predicted workstyle change in industry n over the next 5 years due to year t Wikipedia technologies ($|\overline{dws}|^e_{n,t,5}$), constructed in Equation (39), and its interaction with firm age and firm size (log book assets in year t). We standardize $|\overline{dws}|^e_{n,t,5}$ to unit variance and zero median. Column (2) controls for the log number of total patents in industry n and year t, as well as its interactions with firm age and firm size. Column (3) controls for the log number of breakthrough patents in industry n and year t using data from Kelly et al. (2021), as well as its interactions with firm age and firm size. Column (4) controls for the log number of creative patents in industry n and year t using data from Kalyani (2025), as well as its interactions with firm age and firm size. Column (5) controls for the log number of rapidly evolving patents (RETech) in industry n and year t using data from Bowen, Frésard, and Hoberg (2023), as well as its interactions with firm age and firm size. We include year fixed effects, and bootstrap standard errors. In each bootstrap draw, we resample industries with replacement, and use the corresponding bootstrap weights to reestimate the technology-implied employment change by industry-occupation-year in Equation (46) and recalculate the predicted workstyle change in Equation (39), which is then used to rerun the final regression of interest in the bootstrap sample. The coefficient is the bootstrap median. Asterisks denote significance levels (***=1%, **=5%, *=10%). Number of observations and within R^2 (without fixed effects) are based on the OLS regression. Sample years are 2003 to 2016.

Table IA4 – Sales Growth among Compustat Firms: Wikipedia Article Results

	Sales Growth (5-Year DHS Rate)				
	(1)	(2)	(3)	(4)	(5)
Predicted Workstyle Change in Industry × Log Age	-0.025**	-0.024**	-0.023**	-0.024**	-0.024**
	(0.009)	(0.010)	(0.009)	(0.009)	(0.010)
Predicted Workstyle Change in Industry × Log Size	0.013**	0.012**	0.011*	0.012**	0.011*
	(0.006)	(0.006)	(0.006)	(0.006)	(0.006)
Predicted Workstyle Change in Industry	0.014	0.004	0.006	0.005	0.005
	(0.036)	(0.039)	(0.039)	(0.040)	(0.038)
Log Patents in Industry × Log Age		0.002	-0.001	0.001	-0.000
		(0.007)	(0.009)	(0.009)	(0.009)
Log Patents in Industry × Log Size		-0.001	0.002	-0.000	-0.000
		(0.003)	(0.003)	(0.004)	(0.004)
Log Patents in Industry		-0.003	-0.008	-0.000	0.006
		(0.023)	(0.025)	(0.028)	(0.025)
Log Age	-0.098***	-0.115**	-0.098***	-0.103***	-0.099***
	(0.018)	(0.044)	(0.032)	(0.032)	(0.027)
Log Size	-0.025*	-0.016	-0.028***	-0.021*	-0.021**
	(0.013)	(0.014)	(0.012)	(0.011)	(0.009)
Patent Type	-	All	Breakthrough	Creative	RETech
Year FE	X	X	X	X	X
Observations	27,978	27,978	27,978	27,978	27,978
Within R^2	0.056	0.056	0.057	0.056	0.056

Notes: This table presents regressions at the firm-year level in Equation (42), where the outcome variable $\Delta Y_{i,t\to t+5}$ is sales growth between year t and $t+5$ calculated following Davis, Haltiwanger, and Schuh (1992) as $\Delta Y_{i,t\to t+5} = (Y_{i,t+5} - Y_{i,t})/(0.5 \times Y_{i,t+5} + 0.5 \times Y_{i,t})$. The independent variables include the predicted workstyle change in industry n over the next 5 years due to year t Wikipedia technologies ($|\overline{dws}|^e_{n,t,5}$), constructed in Equation (39), and its interaction with firm age and firm size (log sales in year t). We standardize $|\overline{dws}|^e_{n,t,5}$ to unit variance and zero median. Column (2) controls for the log number of total patents in industry n and year t, as well as its interactions with firm age and firm size. Column (3) controls for the log number of breakthrough patents in industry n and year t using data from Kelly et al. (2021), as well as its interactions with firm age and firm size. Column (4) controls for the log number of creative patents in industry n and year t using data from Kalyani (2025), as well as its interactions with firm age and firm size. Column (5) controls for the log number of rapidly evolving patents (RETech) in industry n and year t using data from Bowen, Frésard, and Hoberg (2023), as well as its interactions with firm age and firm size. We include year fixed effects, and bootstrap standard errors. In each bootstrap draw, we resample industries with replacement, and use the corresponding bootstrap weights to reestimate the technology-implied employment change by industry-occupation-year in Equation (46) and recalculate the predicted workstyle change in Equation (39), which is then used to rerun the final regression of interest in the bootstrap sample. The coefficient is the bootstrap median. Asterisks denote significance levels (***=1%, **=5%, *=10%). Number of observations and within R^2 (without fixed effects) are based on the OLS regression. Sample years are 2003 to 2016.

Table IA5 – Employment Growth among Compustat Firms: Wikipedia Article Results

	Employment Growth (5-Year DHS Rate)				
	(1)	(2)	(3)	(4)	(5)
Predicted Workstyle Change in Industry × Log Age	-0.020**	-0.020**	-0.019**	-0.019**	-0.019**
	(0.008)	(0.008)	(0.008)	(0.008)	(0.008)
Predicted Workstyle Change in Industry × Log Size	0.009**	0.008**	0.007**	0.007**	0.007**
	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)
Predicted Workstyle Change in Industry	0.070**	0.070**	0.066*	0.067*	0.066**
	(0.026)	(0.029)	(0.030)	(0.030)	(0.030)
Log Patents in Industry × Log Age		-0.001	-0.002	-0.003	-0.005
		(0.006)	(0.007)	(0.007)	(0.006)
Log Patents in Industry × Log Size		0.001	0.003	0.002	0.002
		(0.002)	(0.003)	(0.003)	(0.003)
Log Patents in Industry		0.002	0.008	0.011	0.017
		(0.021)	(0.024)	(0.026)	(0.024)
Log Age	-0.093***	-0.088**	-0.085***	-0.084***	-0.080***
	(0.014)	(0.034)	(0.025)	(0.026)	(0.021)
Log Size	-0.008	-0.014	-0.017*	-0.015	-0.014*
	(0.006)	(0.012)	(0.010)	(0.010)	(0.008)
Patent Type	-	All	Breakthrough	Creative	RETech
Year FE	X	X	X	X	X
Observations	27,978	27,978	27,978	27,978	27,978
Within R^2	0.040	0.042	0.042	0.042	0.042

Notes: This table presents regressions at the firm-year level in Equation (42), where the outcome variable $\Delta Y_{i,t\to t+5}$ is employment growth between year t and $t+5$ calculated following Davis, Haltiwanger, and Schuh (1992) as $\Delta Y_{i,t\to t+5} = (Y_{i,t+5} - Y_{i,t})/(0.5 \times Y_{i,t+5} + 0.5 \times Y_{i,t})$. The independent variables include the predicted workstyle change in industry n over the next 5 years due to year t Wikipedia technologies ($|\overline{dws}|^e_{n,t,5}$), constructed in Equation (39), and its interaction with firm age and firm size (log employment in year t). We standardize $|\overline{dws}|^e_{n,t,5}$ to unit variance and zero median. Column (2) controls for the log number of total patents in industry n and year t, as well as its interactions with firm age and firm size. Column (3) controls for the log number of breakthrough patents in industry n and year t using data from Kelly et al. (2021), as well as its interactions with firm age and firm size. Column (4) controls for the log number of creative patents in industry n and year t using data from Kalyani (2025), as well as its interactions with firm age and firm size. Column (5) controls for the log number of rapidly evolving patents (RETech) in industry n and year t using data from Bowen, Frésard, and Hoberg (2023), as well as its interactions with firm age and firm size. We include year fixed effects, and bootstrap standard errors. In each bootstrap draw, we resample industries with replacement, and use the corresponding bootstrap weights to reestimate the technology-implied employment change by industry-occupation-year in Equation (46) and recalculate the predicted workstyle change in Equation (39), which is then used to rerun the final regression of interest in the bootstrap sample. The coefficient is the bootstrap median. Asterisks denote significance levels (***=1%, **=5%, *=10%). Number of observations and within R^2 (without fixed effects) are based on the OLS regression. Sample years are 2003 to 2016.

Table IA6 – Employment Growth among BDS Firms: Wikipedia Article Results

	Employment Growth (5-Year DHS Rate)				
	(1)	(2)	(3)	(4)	(5)
Predicted Workstyle Change in Industry × Age 6–10	0.005	0.005	0.005	0.005	0.005
	(0.009)	(0.009)	(0.009)	(0.009)	(0.009)
Predicted Workstyle Change in Industry × Age 11–15	-0.009	-0.009	-0.010	-0.009	-0.009
	(0.017)	(0.018)	(0.018)	(0.018)	(0.017)
Predicted Workstyle Change in Industry × Age 16–20	-0.030**	-0.028**	-0.030**	-0.029**	-0.030**
	(0.014)	(0.014)	(0.014)	(0.014)	(0.014)
Predicted Workstyle Change in Industry	0.023	0.021	0.022	0.022	0.022
	(0.016)	(0.016)	(0.016)	(0.016)	(0.016)
Log Patents in Industry × Age 6–10		0.005	0.005	0.006	0.008
		(0.006)	(0.007)	(0.008)	(0.009)
Log Patents in Industry × Age 11–15		0.007	0.012	0.009	0.004
		(0.006)	(0.009)	(0.010)	(0.010)
Log Patents in Industry × Age 16–20		0.013*	0.016*	0.014	0.014
		(0.007)	(0.010)	(0.010)	(0.011)
Log Patents in Industry		-0.009	-0.010	-0.006	-0.001
		(0.009)	(0.012)	(0.013)	(0.012)
Age 6–10	0.016	-0.007	0.003	0.004	0.003
	(0.016)	(0.029)	(0.024)	(0.022)	(0.020)
Age 11–15	0.014	-0.015	-0.014	-0.003	0.008
	(0.020)	(0.035)	(0.033)	(0.030)	(0.026)
Age 16–20	0.035	-0.024	-0.006	0.006	0.012
	(0.022)	(0.038)	(0.032)	(0.031)	(0.028)
Patent Type	-	All	Breakthrough	Creative	RETech
Year FE	X	X	X	X	X
Observations	4,032	4,032	4,032	4,032	4,032
Within R^2	0.019	0.020	0.020	0.020	0.022

Notes: This table presents regressions at the industry-age group-year level in Equation (43), where the outcome variable $\Delta Y_{i,t\to t+5}$ is employment growth of age group i in industry n between year t and $t+5$, calculated following Davis, Haltiwanger, and Schuh (1992) as $\Delta Y_{i,t\to t+5} = (Y_{i,t+5} - Y_{i,t})/(0.5 \times Y_{i,t+5} + 0.5 \times Y_{i,t})$. The independent variables include the predicted workstyle change in industry n over the next 5 years due to year t Wikipedia technologies ($|\overline{dws}|^{e}_{n,t,5}$), constructed in Equation (39), and its interaction with age group dummies. We standardize $|\overline{dws}|^{e}_{n,t,5}$ to unit variance and zero median. Column (2) controls for the log number of total patents in industry n and year t, as well as its interaction with firm age groups. Column (3) controls for the log number of breakthrough patents in industry n and year t using data from Kelly et al. (2021), as well as its interaction with firm age groups. Column (4) controls for the log number of creative patents in industry n and year t using data from Kalyani (2025), as well as its interaction with firm age groups. Column (5) controls for the log number of rapidly evolving patents (RETech) in industry n and year t using data from Bowen, Frésard, and Hoberg (2023), as well as its interaction with firm age groups. We include year fixed effects, and bootstrap standard errors. In each bootstrap draw, we resample industries with replacement, and use the corresponding bootstrap weights to reestimate the technology-implied employment change by industry-occupation-year in Equation (46) and recalculate the predicted workstyle change in Equation (39), which is then used to rerun the final regression of interest in the bootstrap sample. The coefficient is the bootstrap median. Asterisks denote significance levels (***=1%, **=5%, *=10%). Number of observations and within R^2 (without fixed effects) are based on the OLS regression. Sample years are 2003 to 2016.

Table IA7 – Acquisition Growth among Compustat Firms

	Acquisition Growth (5-Year DHS Rate)				
	(1)	(2)	(3)	(4)	(5)
Predicted Workstyle Change in Industry × Log Age	-0.016	-0.023	-0.019	-0.021	-0.022
	(0.024)	(0.026)	(0.027)	(0.026)	(0.026)
Predicted Workstyle Change in Industry × Log Size	-0.009	-0.010	-0.012	-0.011	-0.011
	(0.011)	(0.011)	(0.012)	(0.011)	(0.011)
Predicted Workstyle Change in Industry	0.148	0.173	0.168	0.169	0.172
	(0.109)	(0.117)	(0.116)	(0.114)	(0.116)
Log Patents in Industry × Log Age		0.015*	0.008	0.013	0.017
		(0.010)	(0.013)	(0.014)	(0.014)
Log Patents in Industry × Log Size		-0.001	0.002	0.001	0.000
		(0.004)	(0.004)	(0.004)	(0.004)
Log Patents in Industry		-0.048	-0.039	-0.050	-0.066*
		(0.042)	(0.050)	(0.054)	(0.054)
Log Age	0.057*	-0.039	0.024	0.006	0.002
	(0.031)	(0.062)	(0.054)	(0.054)	(0.045)
Log Size	-0.048***	-0.038	-0.052**	-0.046**	-0.046**
	(0.014)	(0.029)	(0.024)	(0.025)	(0.022)
Patent Type	-	All	Breakthrough	Creative	RETech
Year FE	X	X	X	X	X
Observations	14,766	14,766	14,766	14,766	14,766
Within R^2	0.005	0.006	0.006	0.006	0.006

Notes: This table presents regressions at the firm-year level in Equation (42), where the outcome variable $\Delta Y_{i,t\to t+5}$ is acquisition growth between year t and $t+5$ calculated following Davis, Haltiwanger, and Schuh (1992) as $\Delta Y_{i,t\to t+5} = (Y_{i,t+5} - Y_{i,t})/(0.5 \times Y_{i,t+5} + 0.5 \times Y_{i,t})$. The independent variables include the predicted workstyle change in industry n over the next 5 years due to year t technologies ($|\overline{dws}|^e_{n,t,5}$), constructed in Equation (39), and its interaction with firm age and firm size (log employment in year t). We standardize $|\overline{dws}|^e_{n,t,5}$ to unit variance and zero median. Column (2) controls for the log number of total patents in industry n and year t, as well as its interactions with firm age and firm size. Column (3) controls for the log number of breakthrough patents in industry n and year t using data from Kelly et al. (2021), as well as its interactions with firm age and firm size. Column (4) controls for the log number of creative patents in industry n and year t using data from Kalyani (2025), as well as its interactions with firm age and firm size. Column (5) controls for the log number of rapidly evolving patents (RETech) in industry n and year t using data from Bowen, Frésard, and Hoberg (2023), as well as its interactions with firm age and firm size. We include year fixed effects, and bootstrap standard errors. In each bootstrap draw, we resample industries with replacement, and use the corresponding bootstrap weights to reestimate the technology-implied employment change by industry-occupation-year in Equation (37) and recalculate the predicted workstyle change in Equation (39), which is then used to rerun the final regression of interest in the bootstrap sample. The coefficient is the bootstrap median. Asterisks denote significance levels (***=1%, **=5%, *=10%). Number of observations and within R^2 (without fixed effects) are based on the OLS regression. Sample years are 2003 to 2016.

Table IA8 – Acquisition Growth among Compustat Firms: Wikipedia Article Results

	Acquisition Growth (5-Year DHS Rate)				
	(1)	(2)	(3)	(4)	(5)
Predicted Workstyle Change in Industry × Log Age	-0.016	-0.019	-0.017	-0.018	-0.018
	(0.024)	(0.026)	(0.026)	(0.026)	(0.026)
Predicted Workstyle Change in Industry × Log Size	-0.009	-0.010	-0.011	-0.011	-0.010
	(0.010)	(0.010)	(0.011)	(0.011)	(0.011)
Predicted Workstyle Change in Industry	0.156	0.154	0.158	0.156	0.155
	(0.104)	(0.116)	(0.112)	(0.112)	(0.115)
Log Patents in Industry × Log Age		0.014	0.007	0.012	0.016
		(0.010)	(0.012)	(0.014)	(0.014)
Log Patents in Industry × Log Size		-0.002	0.001	0.000	-0.000
		(0.003)	(0.004)	(0.004)	(0.004)
Log Patents in Industry		-0.040	-0.030	-0.040	-0.055
		(0.043)	(0.050)	(0.054)	(0.055)
Log Age	0.056*	-0.036	0.026	0.009	0.003
	(0.030)	(0.063)	(0.054)	(0.055)	(0.046)
Log Size	-0.049***	-0.037	-0.051**	-0.046*	-0.045**
	(0.013)	(0.030)	(0.024)	(0.025)	(0.023)
Patent Type	-	All	Breakthrough	Creative	RETech
Year FE	X	X	X	X	X
Observations	14,766	14,766	14,766	14,766	14,766
Within R^2	0.005	0.006	0.006	0.006	0.006

Notes: This table presents regressions at the firm-year level in Equation (42), where the outcome variable $\Delta Y_{i,t\to t+5}$ is acquisition growth between year t and $t+5$ calculated following Davis, Haltiwanger, and Schuh (1992) as $\Delta Y_{i,t\to t+5} = (Y_{i,t+5} - Y_{i,t})/(0.5 \times Y_{i,t+5} + 0.5 \times Y_{i,t})$. The independent variables include the predicted workstyle change in industry n over the next 5 years due to year t Wikipedia technologies ($|\overline{dws}|^e_{n,t,5}$), constructed in Equation (39), and its interaction with firm age and firm size (log employment in year t). We standardize $|\overline{dws}|^e_{n,t,5}$ to unit variance and zero median. Column (2) controls for the log number of total patents in industry n and year t, as well as its interactions with firm age and firm size. Column (3) controls for the log number of breakthrough patents in industry n and year t using data from Kelly et al. (2021), as well as its interactions with firm age and firm size. Column (4) controls for the log number of creative patents in industry n and year t using data from Kalyani (2025), as well as its interactions with firm age and firm size. Column (5) controls for the log number of rapidly evolving patents (RETech) in industry n and year t using data from Bowen, Frésard, and Hoberg (2023), as well as its interactions with firm age and firm size. We include year fixed effects, and bootstrap standard errors. In each bootstrap draw, we resample industries with replacement, and use the corresponding bootstrap weights to reestimate the technology-implied employment change by industry-occupation-year in Equation (46) and recalculate the predicted workstyle change in Equation (39), which is then used to rerun the final regression of interest in the bootstrap sample. The coefficient is the bootstrap median. Asterisks denote significance levels (***=1%, **=5%, *=10%). Number of observations and within R^2 (without fixed effects) are based on the OLS regression. Sample years are 2003 to 2016.

Table IA9 – Tobin's Q among Compustat Firms, Controlling for Uncertainty

	Tobin's Q					
	Patent			Wiki		
	(1)	(2)	(3)	(4)	(5)	(6)
Predicted Workstyle Change in Industry × Log Age	-0.157***	-0.157***	-0.151**	-0.148***	-0.149***	-0.148**
	(0.045)	(0.045)	(0.062)	(0.046)	(0.046)	(0.060)
Predicted Workstyle Change in Industry × Log Size	0.046*	0.043*	0.027*	0.048**	0.046**	0.030*
	(0.020)	(0.019)	(0.016)	(0.018)	(0.018)	(0.016)
Predicted Workstyle Change in Industry	0.282*	0.293*	0.385*	0.190	0.205	0.317
	(0.174)	(0.175)	(0.229)	(0.192)	(0.194)	(0.275)
Uncertainty × Log Age	-0.050	0.002**	-0.154	-0.057	0.002**	-0.146
	(0.136)	(0.001)	(0.157)	(0.135)	(0.001)	(0.156)
Uncertainty × Log Size	-0.099	-0.001**	-0.141***	-0.100	-0.001**	-0.143***
	(0.073)	(0.001)	(0.047)	(0.073)	(0.001)	(0.047)
Uncertainty	1.014**	0.004	0.606	1.033**	0.004	0.579
	(0.566)	(0.005)	(0.593)	(0.560)	(0.005)	(0.593)
Log Age	-0.245***	-0.509***	-0.145	-0.267***	-0.555***	-0.171
	(0.066)	(0.131)	(0.110)	(0.071)	(0.140)	(0.108)
Log Size	-0.174***	-0.039	-0.109**	-0.169***	-0.027	-0.107**
	(0.034)	(0.064)	(0.043)	(0.035)	(0.063)	(0.043)
Uncertainty Measure	WUI	EPU	Return Vol	WUI	EPU	Return Vol
Year FE	X	X	X	X	X	X
Observations	27,978	27,978	16,811	27,978	27,978	16,811
Within R^2	0.092	0.092	0.117	0.092	0.092	0.115

Notes: This table presents regressions at the firm-year level in Equation (41). The independent variables include the predicted workstyle change in industry n over the next 5 years due to year t technologies ($|\overline{dws}|^e_{n,t,5}$), and its interaction with firm age and firm size (log book assets in year t). We standardize $|\overline{dws}|^e_{n,t,5}$ to unit variance and zero median. Columns (1) to (3) use the predicted workstyle change based on breakthrough patents as in Tables 2 to 4, and columns (4) to (6) use the predicted workstyle change based on Wikipedia articles as in Tables IA3 to IA5. All columns additionally control for an uncertainty measure and its interaction with firm age and firm size, where uncertainty is proxied by the World Uncertainty Index (WUI) in columns (1) and (4), economic policy uncertainty (EPU) from Baker, Bloom, and Davis (2016) in columns (2) and (5), and firm-level stock return volatility (Return Vol) in columns (3) and (6). We include year fixed effects, and bootstrap standard errors. In each bootstrap draw, we resample industries with replacement, and use the corresponding bootstrap weights to reestimate the technology-implied employment change by industry-occupation-year and recalculate the predicted workstyle change, which is then used to rerun the final regression of interest in the bootstrap sample. The coefficient is the bootstrap median. Asterisks denote significance levels (***=1%, **=5%, *=10%). Number of observations and within R^2 (without fixed effects) are based on the OLS regression. Sample years are 2003 to 2016.

Table IA10 – Sales Growth among Compustat Firms, Controlling for Uncertainty

	Sales Growth (5-Year DHS Rate)					
	Patent			Wiki		
	(1)	(2)	(3)	(4)	(5)	(6)
Predicted Workstyle Change in Industry × Log Age	-0.026**	-0.026**	-0.021**	-0.027**	-0.026**	-0.021**
	(0.009)	(0.009)	(0.008)	(0.009)	(0.009)	(0.008)
Predicted Workstyle Change in Industry × Log Size	0.011*	0.011*	0.009	0.013**	0.012**	0.011
	(0.005)	(0.005)	(0.005)	(0.006)	(0.005)	(0.006)
Predicted Workstyle Change in Industry	0.018	0.017	0.007	0.007	0.007	-0.012
	(0.032)	(0.031)	(0.035)	(0.037)	(0.035)	(0.042)
Uncertainty × Log Age	0.026	0.001***	0.071**	0.024	0.001***	0.071***
	(0.024)	(0.000)	(0.029)	(0.024)	(0.000)	(0.029)
Uncertainty × Log Size	0.007	0.000	-0.047**	0.007	0.000	-0.047**
	(0.011)	(0.000)	(0.016)	(0.011)	(0.000)	(0.016)
Uncertainty	-0.145	-0.003***	-0.256***	-0.139	-0.003***	-0.258***
	(0.092)	(0.001)	(0.107)	(0.092)	(0.001)	(0.106)
Log Age	-0.102***	-0.167***	-0.127***	-0.105***	-0.173***	-0.129***
	(0.018)	(0.032)	(0.022)	(0.018)	(0.033)	(0.022)
Log Size	-0.027**	-0.047**	-0.024*	-0.026**	-0.044**	-0.024*
	(0.012)	(0.017)	(0.013)	(0.012)	(0.017)	(0.012)
Uncertainty Measure	WUI	EPU	Return Vol	WUI	EPU	Return Vol
Year FE	X	X	X	X	X	X
Observations	27,978	27,978	16,811	27,978	27,978	16,811
Within R^2	0.077	0.077	0.095	0.077	0.077	0.095

Notes: This table presents regressions at the firm-year level in Equation (42), where the outcome variable $\Delta Y_{i,t\to t+5}$ is sales growth between year t and $t+5$ calculated following Davis, Haltiwanger, and Schuh (1992) as $\Delta Y_{i,t\to t+5} = (Y_{i,t+5} - Y_{i,t})/(0.5 \times Y_{i,t+5} + 0.5 \times Y_{i,t})$. The independent variables include the predicted workstyle change in industry n over the next 5 years due to year t technologies ($|\overline{dws}|^e_{n,t,5}$), and its interaction with firm age and firm size (log sales in year t). We standardize $|\overline{dws}|^e_{n,t,5}$ to unit variance and zero median. Columns (1) to (3) use the predicted workstyle change based on breakthrough patents as in Tables 2 to 4, and columns (4) to (6) use the predicted workstyle change based on Wikipedia articles as in Tables IA3 to IA5. All columns additionally control for an uncertainty measure and its interaction with firm age and firm size, where uncertainty is proxied by the World Uncertainty Index (WUI) in columns (1) and (4), economic policy uncertainty (EPU) from Baker, Bloom, and Davis (2016) in columns (2) and (5), and firm-level stock return volatility (Return Vol) in columns (3) and (6). We include year fixed effects, and bootstrap standard errors. In each bootstrap draw, we resample industries with replacement, and use the corresponding bootstrap weights to reestimate the technology-implied employment change by industry-occupation-year and recalculate the predicted workstyle change, which is then used to rerun the final regression of interest in the bootstrap sample. The coefficient is the bootstrap median. Asterisks denote significance levels (***=1%, **=5%, *=10%). Number of observations and within R^2 (without fixed effects) are based on the OLS regression. Sample years are 2003 to 2016.

Table IA11 – Employment Growth among Compustat Firms, Controlling for Uncertainty

	Employment Growth (5-Year DHS Rate)					
	Patent			Wiki		
	(1)	(2)	(3)	(4)	(5)	(6)
Predicted Workstyle Change in Industry × Log Age	-0.022**	-0.022**	-0.019**	-0.022**	-0.022**	-0.021**
	(0.007)	(0.007)	(0.008)	(0.008)	(0.007)	(0.009)
Predicted Workstyle Change in Industry × Log Size	0.009**	0.009**	0.007	0.009**	0.009**	0.007
	(0.004)	(0.004)	(0.004)	(0.003)	(0.003)	(0.004)
Predicted Workstyle Change in Industry	0.078**	0.078**	0.075***	0.076**	0.078**	0.080***
	(0.026)	(0.026)	(0.026)	(0.027)	(0.026)	(0.026)
Uncertainty × Log Age	0.011	0.000	0.053**	0.011	0.000	0.054**
	(0.025)	(0.000)	(0.030)	(0.024)	(0.000)	(0.030)
Uncertainty × Log Size	-0.015	0.000	-0.015	-0.015	0.000	-0.016
	(0.009)	(0.000)	(0.013)	(0.009)	(0.000)	(0.013)
Uncertainty	-0.050	-0.001	-0.369***	-0.048	-0.001	-0.371***
	(0.095)	(0.001)	(0.123)	(0.095)	(0.001)	(0.124)
Log Age	-0.093***	-0.141***	-0.110***	-0.095***	-0.147***	-0.111***
	(0.015)	(0.037)	(0.022)	(0.015)	(0.037)	(0.022)
Log Size	-0.006	-0.018	-0.015	-0.005	-0.015	-0.014
	(0.006)	(0.014)	(0.010)	(0.006)	(0.014)	(0.009)
Uncertainty Measure	WUI	EPU	Return Vol	WUI	EPU	Return Vol
Year FE	X	X	X	X	X	X
Observations	27,978	27,978	16,811	27,978	27,978	16,811
Within R^2	0.048	0.048	0.049	0.048	0.048	0.049

Notes: This table presents regressions at the firm-year level in Equation (42), where the outcome variable $\Delta Y_{i,t\to t+5}$ is employment growth between year t and $t+5$ calculated following Davis, Haltiwanger, and Schuh (1992) as $\Delta Y_{i,t\to t+5} = (Y_{i,t+5} - Y_{i,t})/(0.5 \times Y_{i,t+5} + 0.5 \times Y_{i,t})$. The independent variables include the predicted workstyle change in industry n over the next 5 years due to year t technologies ($|\overline{dws}|^e_{n,t,5}$), and its interaction with firm age and firm size (log employment in year t). We standardize $|\overline{dws}|^e_{n,t,5}$ to unit variance and zero median. Columns (1) to (3) use the predicted workstyle change based on breakthrough patents as in Tables 2 to 4, and columns (4) to (6) use the predicted workstyle change based on Wikipedia articles as in Tables IA3 to IA5. All columns additionally control for an uncertainty measure and its interaction with firm age and firm size, where uncertainty is proxied by the World Uncertainty Index (WUI) in columns (1) and (4), economic policy uncertainty (EPU) from Baker, Bloom, and Davis (2016) in columns (2) and (5), and firm-level stock return volatility (Return Vol) in columns (3) and (6). We include year fixed effects, and bootstrap standard errors. In each bootstrap draw, we resample industries with replacement, and use the corresponding bootstrap weights to reestimate the technology-implied employment change by industry-occupation-year and recalculate the predicted workstyle change, which is then used to rerun the final regression of interest in the bootstrap sample. The coefficient is the bootstrap median. Asterisks denote significance levels (***=1%, **=5%, *=10%). Number of observations and within R^2 (without fixed effects) are based on the OLS regression. Sample years are 2003 to 2016.

Table IA12 – Tobin's Q among Compustat Firms, Controlling for Profitability

	Tobin's Q					
	Patent			Wiki		
	(1)	(2)	(3)	(4)	(5)	(6)
Predicted Workstyle Change in Industry × Log Age	-0.122**	-0.134***	-0.160***	-0.115**	-0.127***	-0.148***
	(0.042)	(0.040)	(0.045)	(0.040)	(0.041)	(0.047)
Predicted Workstyle Change in Industry × Log Size	0.047**	0.047*	0.038	0.049**	0.048**	0.042*
	(0.017)	(0.020)	(0.020)	(0.017)	(0.019)	(0.018)
Predicted Workstyle Change in Industry × Profitability	-0.192	-0.033	0.067	-0.268	-0.038	0.037
	(0.300)	(0.040)	(0.063)	(0.296)	(0.042)	(0.062)
Predicted Workstyle Change in Industry	0.165	0.196	0.342*	0.089	0.129	0.227
	(0.164)	(0.187)	(0.189)	(0.177)	(0.205)	(0.219)
Log Age	-0.190***	-0.234***	-0.275***	-0.208***	-0.253***	-0.301***
	(0.062)	(0.075)	(0.074)	(0.065)	(0.078)	(0.078)
Log Size	-0.114***	-0.165***	-0.196***	-0.108***	-0.160***	-0.192***
	(0.024)	(0.034)	(0.036)	(0.024)	(0.035)	(0.037)
Profitability	-1.034***	-0.048***	0.309***	-1.067***	-0.058***	0.330***
	(0.248)	(0.069)	(0.055)	(0.229)	(0.064)	(0.059)
Profitability Measure	NI/Assets	NI/Sales	E/P	NI/Assets	NI/Sales	E/P
Year FE	X	X	X	X	X	X
Observations	27,084	27,204	27,007	27,084	27,204	27,007
Within R^2	0.143	0.108	0.098	0.142	0.108	0.097

Notes: This table presents regressions at the firm-year level in Equation (41). The independent variables include the predicted workstyle change in industry n over the next 5 years due to year t technologies ($|\overline{dws}|^e_{n,t,5}$), and its interaction with firm age and firm size (log book assets in year t). We standardize $|\overline{dws}|^e_{n,t,5}$ to unit variance and zero median. Columns (1) to (3) use the predicted workstyle change based on breakthrough patents as in Tables 2 to 4, and columns (4) to (6) use the predicted workstyle change based on Wikipedia articles as in Tables IA3 to IA5. All columns additionally control for the interaction of $|\overline{dws}|^e_{n,t,5}$ and a measure of current profitability to capture concerns of cannibalization, which uses net income over assets (NI/Assets) in columns (1) and (4), net income over sales (NI/Sales) in columns (2) and (5), and the earnings-to-price ratio (E/P) in columns (3) and (6). We include year fixed effects, and bootstrap standard errors. In each bootstrap draw, we resample industries with replacement, and use the corresponding bootstrap weights to reestimate the technology-implied employment change by industry-occupation-year and recalculate the predicted workstyle change, which is then used to rerun the final regression of interest in the bootstrap sample. The coefficient is the bootstrap median. Asterisks denote significance levels (***=1%, **=5%, *=10%). Number of observations and within R^2 (without fixed effects) are based on the OLS regression. Sample years are 2003 to 2016.

Table IA13 – Sales Growth among Compustat Firms, Controlling for Profitability

	Sales Growth (5-Year DHS Rate)					
	Patent			Wiki		
	(1)	(2)	(3)	(4)	(5)	(6)
Predicted Workstyle Change in Industry × Log Age	-0.025**	-0.024**	-0.023*	-0.025**	-0.024**	-0.023*
	(0.008)	(0.008)	(0.009)	(0.009)	(0.008)	(0.009)
Predicted Workstyle Change in Industry × Log Size	0.010*	0.009*	0.010*	0.011*	0.008*	0.012*
	(0.005)	(0.004)	(0.005)	(0.005)	(0.004)	(0.006)
Predicted Workstyle Change in Industry × Profitability	0.019	0.003	0.001	0.023	0.003	-0.000
	(0.039)	(0.006)	(0.013)	(0.041)	(0.007)	(0.013)
Predicted Workstyle Change in Industry	0.015	0.030	0.011	0.010	0.030	-0.002
	(0.033)	(0.028)	(0.034)	(0.034)	(0.028)	(0.040)
Log Age	-0.092***	-0.099***	-0.099***	-0.094***	-0.102***	-0.101***
	(0.018)	(0.017)	(0.017)	(0.018)	(0.017)	(0.018)
Log Size	-0.019*	-0.006	-0.026**	-0.018	-0.005	-0.025**
	(0.010)	(0.007)	(0.013)	(0.010)	(0.007)	(0.013)
Profitability	-0.107*	-0.023**	0.094***	-0.101*	-0.022**	0.094***
	(0.038)	(0.007)	(0.023)	(0.034)	(0.006)	(0.023)
Profitability Measure	NI/Assets	NI/Sales	E/P	NI/Assets	NI/Sales	E/P
Year FE	X	X	X	X	X	X
Observations	27,084	27,204	27,007	27,084	27,204	27,007
Within R^2	0.078	0.075	0.076	0.078	0.075	0.076

Notes: This table presents regressions at the firm-year level in Equation (42), where the outcome variable $\Delta Y_{i,t\to t+5}$ is sales growth between year t and $t+5$ calculated following Davis, Haltiwanger, and Schuh (1992) as $\Delta Y_{i,t\to t+5} = (Y_{i,t+5} - Y_{i,t})/(0.5 \times Y_{i,t+5} + 0.5 \times Y_{i,t})$. The independent variables include the predicted workstyle change in industry n over the next 5 years due to year t technologies ($|\overline{dws}|^e_{n,t,5}$), and its interaction with firm age and firm size (log sales in year t). We standardize $|\overline{dws}|^e_{n,t,5}$ to unit variance and zero median. Columns (1) to (3) use the predicted workstyle change based on breakthrough patents as in Tables 2 to 4, and columns (4) to (6) use the predicted workstyle change based on Wikipedia articles as in Tables IA3 to IA5. All columns additionally control for the interaction of $|\overline{dws}|^e_{n,t,5}$ and a measure of current profitability to capture concerns of cannibalization, which uses net income over assets (NI/Assets) in columns (1) and (4), net income over sales (NI/Sales) in columns (2) and (5), and the earnings-to-price ratio (E/P) in columns (3) and (6). We include year fixed effects, and bootstrap standard errors. In each bootstrap draw, we resample industries with replacement, and use the corresponding bootstrap weights to reestimate the technology-implied employment change by industry-occupation-year and recalculate the predicted workstyle change, which is then used to rerun the final regression of interest in the bootstrap sample. The coefficient is the bootstrap median. Asterisks denote significance levels (***=1%, **=5%, *=10%). Number of observations and within R^2 (without fixed effects) are based on the OLS regression. Sample years are 2003 to 2016.

Table IA14 – Employment Growth among Compustat Firms, Controlling for Profitability

	Employment Growth (5-Year DHS Rate)					
	Patent			Wiki		
	(1)	(2)	(3)	(4)	(5)	(6)
Predicted Workstyle Change in Industry × Log Age	-0.024**	-0.023**	-0.019**	-0.025**	-0.023**	-0.020**
	(0.007)	(0.007)	(0.007)	(0.008)	(0.007)	(0.008)
Predicted Workstyle Change in Industry × Log Size	0.008*	0.009**	0.008*	0.007*	0.010**	0.008*
	(0.003)	(0.003)	(0.004)	(0.003)	(0.003)	(0.004)
Predicted Workstyle Change in Industry × Profitability	0.047	0.000	0.004	0.065*	0.001	-0.004
	(0.034)	(0.004)	(0.020)	(0.040)	(0.005)	(0.018)
Predicted Workstyle Change in Industry	0.085***	0.080**	0.070**	0.087***	0.079**	0.069**
	(0.025)	(0.025)	(0.025)	(0.025)	(0.025)	(0.027)
Log Age	-0.089***	-0.090***	-0.094***	-0.092***	-0.092***	-0.096***
	(0.014)	(0.014)	(0.014)	(0.014)	(0.014)	(0.014)
Log Size	-0.010*	-0.011*	-0.012*	-0.009	-0.009	-0.011
	(0.006)	(0.006)	(0.006)	(0.006)	(0.006)	(0.007)
Profitability	0.014	0.003*	0.154***	0.023	0.003**	0.159***
	(0.059)	(0.008)	(0.024)	(0.052)	(0.008)	(0.024)
Profitability Measure	NI/Assets	NI/Sales	E/P	NI/Assets	NI/Sales	E/P
Year FE	X	X	X	X	X	X
Observations	27,084	27,204	27,007	27,084	27,204	27,007
Within R^2	0.050	0.050	0.057	0.050	0.049	0.057

Notes: This table presents regressions at the firm-year level in Equation (42), where the outcome variable $\Delta Y_{i,t\to t+5}$ is employment growth between year t and $t+5$ calculated following Davis, Haltiwanger, and Schuh (1992) as $\Delta Y_{i,t\to t+5} = (Y_{i,t+5} - Y_{i,t})/(0.5 \times Y_{i,t+5} + 0.5 \times Y_{i,t})$. The independent variables include the predicted workstyle change in industry n over the next 5 years due to year t technologies ($|\overline{dws}|^e_{n,t,5}$), and its interaction with firm age and firm size (log employment in year t). We standardize $|\overline{dws}|^e_{n,t,5}$ to unit variance and zero median. Columns (1) to (3) use the predicted workstyle change based on breakthrough patents as in Tables 2 to 4, and columns (4) to (6) use the predicted workstyle change based on Wikipedia articles as in Tables IA3 to IA5. All columns additionally control for the interaction of $|\overline{dws}|^e_{n,t,5}$ and a measure of current profitability to capture concerns of cannibalization, which uses net income over assets (NI/Assets) in columns (1) and (4), net income over sales (NI/Sales) in columns (2) and (5), and the earnings-to-price ratio (E/P) in columns (3) and (6). We include year fixed effects, and bootstrap standard errors. In each bootstrap draw, we resample industries with replacement, and use the corresponding bootstrap weights to reestimate the technology-implied employment change by industry-occupation-year and recalculate the predicted workstyle change, which is then used to rerun the final regression of interest in the bootstrap sample. The coefficient is the bootstrap median. Asterisks denote significance levels (***=1%, **=5%, *=10%). Number of observations and within R^2 (without fixed effects) are based on the OLS regression. Sample years are 2003 to 2016.

IA2 Processing Wikipedia Articles

This appendix describes in detail how we obtain and process Wikipedia articles on technological inventions used in Section II.

IA2.1 Wikipedia Articles on Technological Inventions

We extract English Wikipedia articles from the Wikimedia English Wikipedia Dumps and filter them in two steps to identify articles related to technological inventions.

Step 1 We first do batch filtering to obtain a broad set of technology-related articles using GPT-4o-mini. The LLM is instructed to identify Wikipedia articles' titles corresponding to technological inventions. Wikipedia articles' titles are processed in batches of 50, and the LLM returns only those classified as technological inventions.

```
Instructions:
- You are tasked with filtering a list of Wikipedia article titles to include
only those of technological inventions. A technological invention is a novel
device, method, process, or system that applies scientific or engineering
principles to solve a problem or improve efficiency. These inventions
typically introduce new functionalities, enhance existing technologies, or
create entirely new categories of tools or systems.
- Focus ONLY on technological inventions, which are devices, systems, and
components with novel technological advancements, such as ''wheels'',
''artificial intelligence'', ''electric vehicle'', ''mobile phone'', ''Watt
steam engine'', etc.
- EXCLUDE names of people, movies, books, songs, food, animals, places,
companies, organizations, social movements, slogans, cultural phrases, and any
non-technology items.
- EXCLUDE specific models, product names, versions of weapons and vehicles,
unless they represent a significant technological advancement (e.g.,
''iPhone'' is a technological invention, but ''iPhone 13'' is not. Airplane is
a technological invention, but ''Boeing 747'' is not).
- Do NOT include scientific phenomena unless they describe a concrete
technological invention or device.
- If multiple titles are extremely similar (such as variations in
capitalization, punctuation, or wording), ONLY include ONE representative
title.
```

```
For each title, first assess if it represents a technological invention. Then,
return ONLY the valid titles (keeping the original format of input, including
the ''_'' character).

Respond in a JSON array of strings like this:
[''Title 1'', ''Title 2'', ''Title 3'']

Titles:
{titles_list}
```

Step 2 We validate each Wikipedia article's title from Step 1 using its Wikipedia summary to obtain the final set of technological inventions. For each Wikipedia articles' title, we use the Wikipedia API library to retrieve the corresponding English Wikipedia summary. We then supply both the title and the summary to Deepseek-Reasoner model, which is prompted to determine whether the article indeed describes a technological invention, following the same definition used in Step 1. The LLM's responses indicate whether the Wikipedia article should be retained.

```
Instructions:
- You are tasked with determining whether a given Wikipedia article is about a
technological invention based on the title and summary provided. A
technological invention is a novel device, method, process, or system that
applies scientific or engineering principles to solve a problem or improve
efficiency. These inventions typically introduce new functionalities, enhance
existing technologies, or create entirely new categories of tools or systems.
- EXCLUDE specific models, product names, versions of weapons and vehicles,
unless they represent a significant technological advancement (e.g.,
''iPhone'' is a technological invention, but ''iPhone 13'' is not. Airplane is
a technological invention, but ''Boeing 747'' is not). Example of titles that
should be EXCLUDED: ''AJS_Model_16'', ''AMD_2900'', ''A340_(aircraft)'',
''A.T_Mine_E.P._Mark_II''.

Wikipedia Article Title: {title}

Wikipedia Article Summary:
\''\''\''{summary}\''\''\''

Based on the above summary, determine if this article is about a technological
invention.
```

```
Respond ONLY with a JSON object in the following format:
{{''is_valid'': true}}
or
{{''is_valid'': false}}
```

IA2.2 Technology Information

We use the following prompt to summarize basic information about the technological inventions with Gemini-2.0-Flash, including the time and location of the invention, as well as the type of inventor (i.e., private company, individual inventor, government, or university/non-profit).

```
Instructions:
You are an expert in the history of technologies. Your task is to answer
questions about when, where, who invented the technology, and which category
of application it best fits into, based on the given information about the
technology and your domain knowledge.

Provide the following information along with a brief reason (less than 20
words) for each:

1) When the technology first became workable. If the exact year is unavailable
or uncertain, provide the closest approximation. Be as accurate as possible.
Include ''BCE'' in your response if the date is before 0 CE/very ancient.

2) Where it was first created (country where the first workable version was
created).

3) Who invented it (choose one of these categories by the type of inventor):
- ''Private Company'': For-profit company/companies or people working for
these companies developed the technology (e.g. Apple, Google);
- ''University/Non-Profit'': Non-profit institution or organization developed
it, such as universities, non-profit institutions, open-source projects (e.g.
UC Berkeley, SRI, Mozilla);
- ''Government Project'': Developed via government-funded or military project
(e.g. NASA, the Apollo program, the Manhattan Project);
- ''Individual Inventor'': Developed by an individual or a small team of
inventors unaffiliated with large institutions or companies (e.g. Nikola
Tesla, Philo Farnsworth);
```

Note: Some inventions involve collaboration. When multiple parties are involved, choose the category that best reflects the majority of developers or the most responsible entity. Aim for the most accurate classification based on available evidence. When inventors are not explicitly mentioned, use your best judgment to classify the technology based on the context and purpose of the invention.
For example:
- A collaboration between private companies is still classified as ''Private Company''.
- A small team of unaffiliated inventors is still ''Individual Inventor''.
- If an individual works under a government-funded initiative, classify it as ''Government Project''.
- If it was published by an industry association, it should be either ''Private Company'' or ''University/Non-Profit'' depending on the inventor and purpose of the technology. For example, SAE (Society of Automotive Engineers) is mostly run by engineers from private companies, so it should be classified as ''Private Company''. W3C (World Wide Web Consortium) is run by university researchers, so it should be classified as ''University/Non-Profit''.
- If it was designed for commercial use or production efficiency, it's likely ''Private Company''.
- If it served military, national defense, or public interest, it's likely ''Government Project''.
- If it was developed for academic research or knowledge advancement, it's likely ''University/Non-Profit''.
- If it was created by individuals not working for large institutions, it's likely ''Individual Inventor''.

4) The 'domain_of_application' (choose exactly one from the list below):
   [''Agriculture, Forestry, Fishing and Hunting'', ''Mining, Quarrying, and Oil and Gas Extraction'', ''Utilities'', ''Construction'', ''Manufacturing'', ''Wholesale Trade'', ''Retail Trade'', ''Transportation and Warehousing'', ''Information'', ''Finance and Insurance'', ''Real Estate and Rental and Leasing'', ''Professional, Scientific, and Technical Services'', ''Health Care and Social Assistance'', ''Military and Defense'', ''Basic Research'']

```
Importantly: Use only the listed categories above. Do not invent new
categories.

Technology Name: {title}

Wikipedia Page:
\''\''\''{wiki_text[:5000]}\''\''\''

Return ONLY a JSON object in the following format:
{{
  ''technology'': ''Title'',
  ''when'': year or ''Not Available'',
  ''when_reason'': ''Brief reason for chosen year'',
  ''where'': country name or ''Not Available'',
  ''where_reason'': ''Brief reason for chosen country'',
  ''who'': one of the 4 listed categories,
  ''who_reason'': ''Brief reason for chosen category'',
  ''domain_of_application'': one of the 15 listed categories,
  ''domain_reason'': ''Brief reason for chosen domain''
}}

Example Output #1:
{{
  ''technology'': ''iPhone'',
  ''when'': ''2007'',
  ''when_reason'': ''Apple released the first iPhone in 2007.'',
  ''where'': ''USA'',
  ''where_reason'': ''Apple developed the iPhone in the United States.'',
  ''who'': ''Private Company'',
  ''who_reason'': ''Developed by Apple, which is a for profit company.'',
  ''domain_of_application'': ''Information'',
  ''domain_reason'': ''Primarily used for communication and computing.''
}}

Example Output #2:
{{
  ''technology'': ''CRISPR'',
  ''when'': ''2012'',
```

```
  ''when_reason'': ''First practical demonstration of gene editing was in
  2012.'',
  ''where'': ''USA'',
  ''where_reason'': ''Initial research conducted at UC Berkeley.'',
  ''who'': ''University Lab'',
  ''who_reason'': ''Developed at UC Berkeley by academic researchers.'',
  ''domain_of_application'': ''Health Care and Social Assistance'',
  ''domain_reason'': ''Primarily applied to medicine and biotechnology.''
}}

Use this JSON schema:
{{
  ''technology'': ''str'',
  ''when'': ''str'',
  ''when_reason'': ''str'',
  ''where'': ''str'',
  ''where_reason'': ''str'',
  ''who'': ''str'',
  ''who_reason'': ''str'',
  ''domain_of_application'': ''str'',
  ''domain_reason'': ''str''
}}
```

IA2.3 Technology Implementation and Commercialization

We use the following prompt to ask web-enabled Gemini-2.0-Flash whether the company that was most successful in its initial implementation and commercialization was a young firm, or an old incumbent.

```
You are a historian of technology and innovation.
You must use web search to verify the answer.

Here is a technological invention.
Your task is to determine whether the given technology was ever
commercialized,
identify the company that was most successful in its initial implementation
and commercialization,
```

```
and classify whether that company was a young firm or an old firm at the time
of the commercialization.

------------------------------------------------
INSTRUCTIONS
------------------------------------------------

Step 1: Identify the implementing company if the technology was ever
commercialized
- Return the name of the company that was most successful in the initial
implementation and commercialization of the technology.
- Always use the name of the parent company and do not use the name of the
division or subsidiary.
- If multiple companies were involved, return the primary company most widely
credited with initial commercialization.
- If the technology has never been commercialized, return "Never
Commercialized".

Step 2: Classify firm type

Determine whether the implementing company was young or old at the time of the
commercialization, based on the age of the parent company:

- **1** → Young firm (founded less than 10 years before commercialization)
- **-1** → Old firm (existed long before the invention)
- **0** → Not sure or ambiguous
- **2** → Non-commercialized (never commercialized)

Step 3: Provide short reasoning and approximate period

------------------------------------------------
INPUT
------------------------------------------------

Technology: {title}

Wikipedia summary:
\"\"\"{summary_text}\"\"\"
```

```
------------------------------------------------
OUTPUT FORMAT (JSON ONLY)
------------------------------------------------

Return only valid JSON:

{{
  "technology": "{title}",
  "implementing_firm_name": "Name of company or 'Never Commercialized'",
  "firm_type": 1 or -1 or 0 or 2,
  "firm_type_reason": "Short explanation (<30 words)",
  "year_or_period": "Approximate year/decade of first commercialization if
  available, else 'Unknown'"
}}

Do not include any explanation outside the JSON.

------------------------------------------------
EXAMPLES
------------------------------------------------

Dropbox
{{
  "technology": "Dropbox",
  "implementing_firm_name": "Dropbox",
  "firm_type": 1,
  "firm_type_reason": "Founded and launched Dropbox in 2007.",
  "year_or_period": "2007"
}}

iPhone
{{
  "technology": "iPhone",
  "implementing_firm_name": "Apple",
  "firm_type": -1,
  "firm_type_reason": "Apple founded in 1976; launched iPhone in 2007.",
  "year_or_period": "2007"
```

```
}}

Ethernet
{{
  "technology": "Ethernet",
  "implementing_firm_name": "3Com",
  "firm_type": 1,
  "firm_type_reason": "3Com founded in 1979; commercialized Ethernet hardware
  in the early 1980s.",
  "year_or_period": "1980s"
}}

Asteroid Impact Avoidance
{{
  "technology": "Asteroid Impact Avoidance",
  "implementing_firm_name": "Never Commercialized",
  "firm_type": 2,
  "firm_type_reason": "Remained research; no market commercialization.",
  "year_or_period": "Unknown"
}}
```

IA3 Processing Employee Reviews

We collect information through employee reviews from Revelio (Revelio Labs, 2025). We process the reviews using the prompts below.

IA3.1 Prevalence of Rules

Step 1 We process all employee reviews with Gemini-2.0-Flash to screen for those that discuss rules.

```
 You are analyzing employee reviews of companies. Your task is to determine
 whether this review discusses rules, protocols, procedures, policies, or
 norms that the employee's company has. Rules include formal policies and
 processes as well as strong unwritten norms.

- Ignore comments about pay, culture, or leadership that do not explicitly
mention rules or norms.
- We are looking for reviews about rules in the reviewer's employer company,
not rules imposed by the regulators, suppliers, or clients of the reviewer's
employer company.

Here is a piece of employee review:
''{text}''

Does the review discuss rules, protocols, procedures, policies, or norms that
the employee's company has?

- Yes (return 1)
- No (return -1)
- Unclear (return 0)

Respond ONLY with a JSON object in the following format:
{{''flag'': 0, ''confidence'': 0.5}}
```

Step 2 We process reviews labeled as 1 or 0 from Step 1 with Gemini-2.0-Flash to determine whether they mention the presence or the lack of rules.

```
 Here is a piece of employee review for a company:
```

```
''{text}''

Please determine which of the following best describes this review:
- This review discusses the presence of rules and processes in the company
(return 1)
- This review discusses the lack of rules and processes in the company (return
-1)
- This review does not discuss rules and processes in the company (return 0)

Respond ONLY with a JSON object in the following format:
{{
    ''label'': 0,
    ''confidence'': 0.5,
    ''snippets'': (if the label is not 0, report the snippets from the review
    that was most important for your determination of the label)
}}
```

Rule index We use the results from Step 2 in the numerator of the rule index, and calculate the intensity of rules in a GVKEY-year as:

$$\text{Rule Index} = \frac{(\#\text{ of reviews mentioning the presence of rules - }\#\text{ of reviews mentioning the lack of rules})}{\#\text{ of total reviews}}.$$

We map the Revelio company ID RCID to GVKEY in a given year using the parent-subsidiary bridge provided by Revelio to us directly.

Examples of employee review discussions

- Presence of rules:
  - "It's a very structured place with a lot of rules, hierarchy, etc."
  - "Stick to the rules and not flexible."
  - "Lots of rules and guidelines."
  - "A stable and well structured corporation that follows the rules."
  - "Company follows strict rules in all type of scenario."
- Absence of rules:
  - "Unstructured work environment. Lack of leadership. Lack of accountability - Job Functions are not very well defined and will change constantly."

- "Lack of, or nonexistent enforcement of policy results in extremely inconsistent standards."
- "Make many changes at the last minute...schedules, resource guide, rules. etc."
- "Need to improve culture and standardize HR rules."
- "Rules of engagement and process were not always clear."

IA3.2 Too Many Meetings

Step 1 We process all employee reviews with Gemini-2.0-Flash to screen for those that discuss meetings, committees, and layers of approval.

```
You are analyzing employee reviews of companies. Your task is to determine
whether this review discusses meetings, committees, and layers of approval.
- We are looking for reviews about such activities in the reviewer's employer
company, not those imposed by the regulators, suppliers, or clients of the
reviewer's employer company or social meetings and parties.
- Do not include personal, non-work related activities, such as smoking,
drinking, eating, or playing games.
- Do not include general discussions about office politics, slow promotion, or
bad management.


Here is a piece of employee review:
''{text}''

Does the review discuss meetings, committees, and layers of approval at the
employee's company?

- Yes (return 1)
- No (return -1)
- Unclear (return 0)

Respond ONLY with a JSON object in the following format:
{{''flag'': 0, ''confidence'': 0.5}}
```

Step 2 We process reviews labeled as 1 or 0 from Step 1 with Gemini-2.0-Flash to determine whether they complain that the company has too many meetings, committees, and layers of approval or praise the company for not having too many.

```
Here is a piece of employee review for a company:

''{text}''

Please determine which of the following best describes this review:
- This review complains about too many meetings, committees, and layers of
approval that waste time. (return 1)
- This review praises the company for being efficient and not having too many
meetings, committees, and layers of approval that waste time. (return -1)
- It is difficult to tell (return 0)

Respond ONLY with a JSON object in the following format:
{{
    ''label'': 0,
    ''confidence'': 0.5,
    ''snippets'': (if the label is not 0, report the snippets from the review
    that was most important for your determination of the label)
}}
```

Meeting index We use the results from Step 2 in the numerator of the meeting index, and calculate the intensity of meetings that waste time in a GVKEY-year as:

$$\text{Meeting Index} = \frac{(\#\text{ of reviews complaining too many - }\#\text{ of reviews praising not too many})}{\#\text{ of total reviews}}.$$

We map Revelio company ID RCID to GVKEY in a given year using parent-subsidiary bridge provided by Revelio to us directly.

Examples of employee review discussions

- Too many meetings:
  - "Way too many meetings. Took forever to make changes and/or decisions."
  - "More meetings and games than work."
  - "Too many internal mandatory calls which takes away time best served somewhere else."
  - "A lot of meaningless meetings."
  - "Too much internal fighting, more competition with folks inside rather than outside."
- Not too many meetings:

- "Meeting-light: this was a HUGE unlock for me. My previous job had me back to back in meetings from 9a-5p so the only time I had to get work done was AFTER work."
- "Good work environment, not too many meetings during the day."
- "The weekly meetings were not excessive."
- "There are lots of resources and projects going on, and very few fiefdoms or silos to get in the way of you contributing."
- "Not being micro-managed, feeling challenged, being a part of something bigger than just a job, but a disruption of an industry, with camaraderie along the way are things I value."

IA4 Measuring Similarity between Technologies and Occupation Tasks

This appendix provides additional details on the computation of textual similarity between new technologies and occupations' routine and nonroutine task descriptions, as used in Sections IV.B and VI.A.1. The overall approach follows Kogan et al. (2024): we represent each technology document and each occupation's task descriptions as vectors in embedding space, and compute cosine similarity between them. Section IA4.1 describes the text preprocessing and embedding construction for the patent-based measures. Section IA4.2 describes the analogous procedure for Wikipedia articles, including the extraction of technical keywords used to construct Wikipedia article embeddings.

IA4.1 Similarity between Breakthrough Patents and Occupations' Task Descriptions

This subsection describes the text preprocessing steps applied to patent text and O*NET occupation task descriptions before computing the document embeddings $\mathcal{X}_p$ and $\mathcal{X}_j^\tau$ in Equation (34) and the cosine similarity $s_{p,j}^\tau$ in Equation (35) in Section IV.B.

Patent text We obtain patent text from PatentsView. For each utility patent, we combine the abstract, detailed description, and all claims into a single text document. We then apply the following preprocessing steps:

1. Remove non-alphabetic characters.
2. Convert to lowercase.
3. Tokenize using NLTK.
4. Part-of-speech (POS) tag using NLTK.
5. Keep only nouns and verbs.
6. Remove stopwords (using a combined stopword list from multiple sources, including NLTK, spaCy, Snowball, SMART, Lextek, ONIX, and patent-specific stopwords from Kogan et al. (2024) such as "abstract," "claim," "invention," "patent," "united," and "states").
7. Lemmatize using WordNet.

Occupation task text We apply the same preprocessing pipeline to O*NET occupation task descriptions. As described in Section IV.B, for each occupation, tasks are classified as routine or nonroutine using LLM (GPT-4o). We then concatenate all routine task descriptions into one document and all nonroutine task descriptions into another, producing two text documents per occupation.

Embeddings As described in Section IV.B, each document $\mathcal{X}_i$—whether a patent or an occupation's combined routine (nonroutine) task descriptions—is represented as:

$$\mathcal{X}_i = \sum_k q_{i,k}\, x_k,$$

where x_k is the GloVe word vector (`glove-wiki-gigaword-300`) for token k and $q_{i,k}$ is the TF-IDF weight. We compute TF-IDF weights separately for the patent text corpus (fitted separately for each grant year) and the occupation task text corpus. For each token in the document that appears in both the TF-IDF vocabulary and the GloVe model, we multiply the word vector by its TF-IDF weight, sum the weighted vectors, and divide by the number of valid weighted tokens.

Post-processing of similarity scores After computing the cosine similarity $s^{\tau}_{p,j}$ between patent p and occupation j's routine or nonroutine tasks as in Equation (35), we apply two adjustments following Kogan et al. (2024). First, we remove within-year means from the raw cosine similarity scores (separately for routine and nonroutine similarities and for each grant year). Second, we impose sparsity by setting values below the 80th percentile to zero, and rescale the remaining values to $[0, 1]$ by dividing by the within-year maximum above the 80th percentile, to obtain the adjusted similarity $\tilde{s}^{\tau}_{p,j}$ used in Equation (36).

IA4.2 Similarity between Wikipedia Articles and Occupations' Task Descriptions

This subsection describes how we construct embeddings for Wikipedia articles used to compute the cosine similarity $s^{\tau}_{w,j}$ between Wikipedia article w and occupation j's routine or non-routine tasks in Section VI.A.1. The key difference from the patent-based approach in Appendix Section IA4.1 is that we use LLM-extracted technical keywords to represent each Wikipedia article, rather than TF-IDF over the full text. This is because Wikipedia articles contain substantial background information (e.g., time and place of invention) beyond the technology's functional characteristics, and keywords allow us to focus the embedding on the core technical content.

Keyword extraction from Wikipedia articles For each Wikipedia articles' title and summary, we use Gemini 2.5 Pro to extract 12 to 25 technical keywords or phrases that capture the core mechanism or principle, key components or materials, main process verbs or operations, and relevant technical domain terms of the technology. The LLM is instructed to prefer domain-specific technical terms over generic words (such as "method," "system," or "process") and to exclude names of people, companies, places, and

dates. The prompt is as follows:

```
You are a technical analyst.

I will provide a Wikipedia title and its summary describing an
invention/technology.
Your task is to output a set of technical keywords/phrases that capture:
- core mechanism / principle
- key components / materials
- main process verbs / operations
- relevant technical domain terms

Rules:
- Output ONLY a JSON array of strings. No extra text.
- Include 12 to 25 items.
- Prefer domain-specific technical terms over generic words
  (avoid: "method", "system", "various", "process", "technology").
- Remove names of people, companies, places, and dates.
- Do NOT invent new information not implied by the input.
- Include both unigrams and short multi-word phrases when appropriate.

Input:
Title: {technology}

Summary:
{summary}

Output:
```

Wikipedia article embeddings As described in Section VI.A.1, each Wikipedia article w is represented as the simple mean of its keywords' GloVe word vectors:

$$\mathcal{X}_w = \frac{1}{K_w} \sum_{k=1}^{K_w} x_k,$$

where K_w is the number of keyword words for Wikipedia article w and x_k are GloVe word vectors (`glove-wiki-gigaword-300`). For multi-word keyword phrases, each constituent word is included separately in the average.

Occupation task embeddings The occupation task embeddings $\mathcal{X}_j^\tau$ ($\tau \in \{R, NR\}$) are constructed the same way as in the patent-based approach (Appendix Section IA4.1), using TF-IDF-weighted GloVe embeddings of each occupation's routine and nonroutine task descriptions (keeping only verbs and nouns and removing stopwords before calculating the embeddings).

Similarity computation and post-processing We compute cosine similarity $s_{w,j}^\tau$ between the unweighted Wikipedia article embedding $\mathcal{X}_w$ and the TF-IDF weighted occupation task embedding $\mathcal{X}_j^\tau$, analogous to Equation (35). We then apply the same post-processing as in the patent-based approach: remove within-year means (separately for routine and nonroutine similarities and for each year), impose sparsity at the 80th percentile, and rescale the remaining values to $[0, 1]$ to obtain the adjusted similarity $\tilde{s}_{w,j}^\tau$ used in Equation (45).

IA5 Measuring the Age of Compustat Firms

For the analyses in Sections III.D, V, and VI, we combine several sources to measure the age of Compustat firms, which we explain in detail below.

IPO date from Compustat We obtain the IPO date (ipodate variable) from Compustat. According to S&P Global Support, Compustat started collecting IPO date in 2002, after which this variable is systematically recorded for companies that have an Initial Public Offering on the NYSE, NYSE America, NYSE ARCA, or Nasdaq. For IPOs before 2002, the date may be missing, or it may be based on other sources or the date when daily pricing began.

Incorporation date from Refinitiv We collect incorporation date from Refinitiv. We merge this information into Compustat based on CUSIP.

LLM query of firm founding year For all unique historical real-time firm names in Compustat (conmh variable), we use Gemini 2.5 Pro to search for the company's founding year with the following query:

```
You are an expert in company history.
Return only the *founding year* of this company as a 4-digit number.
If unknown, return "NA".

Company: {company}

Output JSON ONLY:
{{
 "company": "{company}",
 "founding_year": "YYYY or NA"
}}
```

Estimated firm age We estimate firm age as the number of years since the earliest year among the available dates above. We set firm age to missing if the earliest of these dates is after the earliest year of the company's record in Compustat, which indicates that our founding year information may not be accurate.

IA6 Additional proofs

We work within a single industry and suppress the industry subscript n where no confusion arises.

IA6.1 Equilibrium positivity of per-unit profits

The model maintains $\pi_{A,t} > 0$ for all $\rho_A \in [0,1]$ and at every date t (Section III.B) and $\pi_{B,t} > 0$ for all $(\rho_A, \rho_B) \in [0,1]^2$ at every date $t \geq 0^+$ (Section III.C). We show that the single parameter restriction

$$\frac{L}{\Xi} \;>\; \max\{1,\chi\}\,\phi \tag{A}$$

is sufficient for both, at every date and for every firm. When escalation costs are no larger than improvisation costs ($\chi \leq 1$), the restriction reduces to the simpler $L/\Xi > \phi$.

Per-A-unit profit Substituting the equilibrium expression (9) into (6):

$$\pi_{A,t}(\rho_{A,i,t}) \;=\; \frac{L}{\Xi} \;+\; \phi\left(\rho_{A,i,t} - \bar{\rho}_{A,t}\right).$$

The right-hand side is increasing in $\rho_{A,i,t} \in [0,1]$ and decreasing in $\bar{\rho}_{A,t} \in [0,1]$, so the worst case is the newborn firm in an industry where $\bar{\rho}_{A,t} = 1$. Even at this worst case, $\pi_{A,t} \geq L/\Xi - \phi > 0$ under (A).

Per-B-unit profit Substituting the equilibrium expression (18) into (13), and using that the post-arrival B-rule stock is firm-invariant ($\rho_{B,i,t} = \bar{\rho}_{B,t} = 1 - e^{-\kappa t}$):

$$\pi_{B,t}(\rho_{A,i,t}, \rho_{B,t}; \Delta) \;=\; \frac{L}{\Xi} \;+\; \phi\,(1-\rho_{B,t})\left(1-(1+\chi)F(\Delta)\right)\left(\rho_{A,i,t} - \bar{\rho}_{A,t}\right).$$

The friction loading $(1-\rho_{B,t})\big(1-(1+\chi)F(\Delta)\big)$ now lies in $[-\chi, 1]$ rather than $[0,1]$, because the escalation cost can make existing rules a net liability when $F(\Delta) > F^* = 1/(1+\chi)$. Combined with $(\rho_{A,i,t} - \bar{\rho}_{A,t}) \in [-1,1]$, the product $(1-\rho_{B,t})(1-(1+\chi)F(\Delta))(\rho_{A,i,t} - \bar{\rho}_{A,t})$ attains its minimum of $-\max\{1,\chi\}$ at one of two corners: a newborn firm in a saturated industry facing no friction ($\rho_{A,i,t} = 0$, $\bar{\rho}_{A,t} = 1$, $F = 0$), or the oldest firm in a fresh industry facing maximal conflict ($\rho_{A,i,t} = 1$, $\bar{\rho}_{A,t} = 0$, $F = 1$). Even at this worst case, $\pi_{B,t} \geq L/\Xi - \max\{1,\chi\}\phi > 0$ under (A).

Economic interpretation Condition (A) requires that L/Ξ exceed the maximum per-unit downtime cost, $\max\{1,\chi\}\phi$. On the A-side this maximum is ϕ, incurred when none of the firm's problems are covered by rules; on the B-side it is $\max\{1,\chi\}\phi$, since a conflicting A-rule can cost as much as $(1+\chi)\phi$ in downtime. In equilibrium, L/Ξ is the marginal product of an additional unit absent any friction. The assumption therefore rules out the case in which a firm would prefer not to operate at all. When escalation costs are modest ($\chi \leq 1$), the binding constraint is the A-side one and the condition is simply $L/\Xi > \phi$:

in an industry where $L/\Xi \leq \phi$, the equilibrium wage w_A would have to be negative for the youngest firm to break even on its A-units.

IA6.2 Proof of Result 1

Closed-form impulse growth rate From Section III.B, the firm-level FOC for A-units is $L^*_{A,i,t} = \pi_{A,t}(\rho_{A,i,t})\xi_i$ with $\pi_{A,t}(\rho_A) = z_{A,t} - \phi(1-\rho_A)$. Substituting the equilibrium $z_{A,t} = L/\Xi + \phi(1-\bar{\rho}_{A,t})$ from (9), where $\bar{\rho}_{A,t} = \int_i \rho_{A,i,t}\, di$ is the cross-firm sum (equivalently, average) of rule stocks,

$$\pi_{A,t}(\rho_{A,i,t}) \;=\; \frac{L}{\Xi} \;+\; \phi\,(\rho_{A,i,t} - \bar{\rho}_{A,t}).$$

At the moment B arrives, this evaluates at $t = 0^-$, with $\bar{\rho}_{A,n,0^-} \equiv \int_{i\in n} \rho_{A,i,0^-}\, di$:

$$\pi_{A,0^-}(\rho_{A,i,0^-}) \;=\; \frac{L_n}{\Xi_n} \;+\; \phi\,(\rho_{A,i,0^-} - \bar{\rho}_{A,n,0^-}).$$

Similarly, from Section III.C, the B-side FOC is $L^*_{B,i,t} = \pi_{B,t}(\rho_{A,i,t}, \rho_{B,i,t}; \Delta)\xi_i$ with $\pi_{B,t}(\rho_A, \rho_B; \Delta) = z_{B,t} - \phi(1-\rho_B)[1 - \rho_A(1 - F(\Delta)/F^*)]$ from (13). At impulse, $\rho_{B,i,0^+} = 0$ and $\bar{\rho}_{B,0^+} = 0$, so the equilibrium $z_{B,t}$ from (18) reduces to $z_{B,0^+} = L/\Xi + \phi - \phi(1 - F(\Delta)/F^*)\bar{\rho}_{A,0^-}$, and $\pi_{B,0^+} = z_{B,0^+} - \phi + \phi(1 - F(\Delta)/F^*)\rho_{A,i,0^-}$. Substituting:

$$\pi_{B,0^+}\Big(\rho_{A,i,0^-}, 0; \Delta_n\Big) \;=\; \frac{L_n}{\Xi_n} \;+\; \phi\,\Big(1 - F(\Delta_n)/F^*\Big)\,\Big(\rho_{A,i,0^-} - \bar{\rho}_{A,n,0^-}\Big).$$

The impulse growth rate is the ratio:

$$g_{i,n} \;=\; \frac{L^*_{B,i,0^+}}{L^*_{A,i,0^-}} \;=\; \frac{\pi_{B,0^+}\Big(\rho_{A,i,0^-}, 0; \Delta_n\Big)\,\xi_i}{\pi_{A,0^-}\left(\rho_{A,i,0^-}\right)\xi_i} \;=\; \frac{\pi_{B,0^+}\Big(\rho_{A,i,0^-}, 0; \Delta_n\Big)}{\pi_{A,0^-}\left(\rho_{A,i,0^-}\right)},$$

in which ξ_i cancels. Substituting the linear friction $F(\Delta_n) = f\Delta_n$ delivers (26).

Industry-level workstyle change Pre-arrival, the industry-level employment-weighted average workstyle is $\overline{ws}_{n,0^-} = ws_{A,n}$, since every firm operates only A-units and $ws_{A,n}$ is firm-invariant. At impulse, the share-weighted average across A- and B-units is

$$\overline{ws}_{n,0^+} \;=\; \frac{\sum_i L^*_{A,i,0^+}\, ws_{A,n} + \sum_i L^*_{B,i,0^+}\, ws_{B,n}}{\sum_i\Big(L^*_{A,i,0^+} + L^*_{B,i,0^+}\Big)} \;=\; ws_{A,n} \;+\; \tfrac{1}{2}\,(ws_{B,n} - ws_{A,n}),$$

where the last equality uses industry-level market clearing $\sum_i L^*_{X,i,t} = L_n$ for $X \in \{A, B\}$ under the normalization $L_{A,n} = L_{B,n} \equiv L_n$. Hence $|d\overline{ws}|_n = \frac{1}{2}\,\Delta_n$. □

IA6.3 Proof of Result 2

To make notation lighter, write $w_n \equiv |d\overline{ws}|_n$, $a_i \equiv \text{age}_i$, and $\rho_i \equiv \rho_{A,i,0^-} = 1 - e^{-\kappa a_i}$. We start by linearizing the impulse growth rate (26) around the benchmark $f\Delta_n = 0$ and $\rho_i = \bar{\rho}_{A,n,0^-}$, where $\bar{\rho}_{A,n,0^-} \equiv \int_{i\in n} \rho_i \, di$ is the cross-firm sum (equivalently, average) of pre-arrival rule stocks in industry n. Under the normalization $L_{A,n} = L_{B,n} \equiv L_n$, the numerator and denominator of (26) both equal L_n/Ξ_n at the benchmark and have identical coefficients on $(\rho_i - \bar{\rho}_{A,n,0^-})$ when $f\Delta_n = 0$, so the standalone first partial derivatives in $(\rho_i - \bar{\rho}_{A,n,0^-})$ and in Δ_n both vanish at the benchmark and the leading non-vanishing term is the cross-derivative:

$$g_{i,n} \;\approx\; 1 \;-\; \Lambda_n \,(\rho_i - \bar{\rho}_{A,n,0^-}) \;+\; O(\cdot)^2, \qquad \text{where} \quad \Lambda_n \;\equiv\; \frac{\Xi_n \,\phi\, f}{L_n \,F^*}\,\Delta_n. \tag{IA1}$$

The remainder $O(\cdot)^2$ collects own quadratic terms in $(\rho_i - \bar{\rho}_{A,n,0^-})$ and Δ_n, both of which contribute zero to the FWL covariance below. The change in industry-level employment-weighted workstyle is $w_n = \frac{1}{2}\Delta_n$ from (27).

We restate Assumption 1 from the body of Result 2 for convenience.

Assumption IA1 (Assumption 1, restated)**.** *The joint distribution of (a_i, ξ_i) is identical across industries n and is independent of all industry-level variables (including $Q_{A,n}$, $Q_{B,n}$, Δ_n, w_n, L_n, and Ξ_n).*

Under Assumption IA1, the cross-firm average $\bar{\rho}_{A,n,0^-}$ is industry-invariant: $\bar{\rho}_{A,n,0^-} = \bar{\rho} \equiv \int_i \rho_i \, di$.

Result IA3 (Result 2, restated and generalized)**.** *Under Assumption IA1 and to first order in $f\Delta_n$ and in $(\rho_i - \bar{\rho})$, the population OLS coefficient on $w_n \times a_i$ in regression (29) is*

$$\hat{\zeta} \;=\; -\,\Phi \;\times\; \frac{\text{cov}(\rho_i, a_i)}{\text{var}(a_i)}, \tag{IA2}$$

where

$$\Phi \;\equiv\; \frac{2\,\phi\, f}{F^*} \;\times\; \frac{\text{cov}(\Xi_n \Delta_n / L_n,\; \Delta_n)}{\text{var}(\Delta_n)} \tag{IA3}$$

is an industry-level scalar. The body's $\Phi = 2(\Xi/L)\,\phi f/F^*$ is the special case in which the spanning-to-supply ratio Ξ_n/L_n is industry-invariant. Sufficient conditions for $\Phi > 0$ are given in Section IA6.3.

Proof. We apply the Frisch–Waugh–Lovell (FWL) theorem to the regression

$$g_{i,n} \;=\; \nu \;+\; \kappa_w\, w_n \;+\; \rho_a\, a_i \;+\; \zeta\,(w_n \times a_i) \;+\; \varepsilon_{i,n}. \tag{IA4}$$

The FWL theorem states that the coefficient ζ on $w_n \times a_i$ in (IA4) can be obtained equivalently by: (i) regressing $w_n \times a_i$ on the other regressors $(1, w_n, a_i)$ and collecting the residual; (ii) regressing $g_{i,n}$ on the same regressors $(1, w_n, a_i)$ and collecting the residual; and (iii) running a univariate regression of the residual from (ii) on the residual from (i). We carry out each of these explicitly.

Regression (i): projecting $w_n \times a_i$ onto $(1, w_n, a_i)$. Consider the population regression

$$w_n \times a_i \;=\; \alpha_1 \;+\; \beta_2\, w_n \;+\; \beta_1\, a_i \;+\; r^{(1)}_{i,n}, \tag{IA5}$$

where $r^{(1)}_{i,n}$ is the residual. By Assumption IA1, $\operatorname{cov}(a_i, w_n) = 0$, so the normal equations for (β_1, β_2) reduce to the diagonal system

$$\begin{pmatrix} \operatorname{var}(w_n) & 0 \\ 0 & \operatorname{var}(a_i) \end{pmatrix} \begin{pmatrix} \beta_2 \\ \beta_1 \end{pmatrix} \;=\; \begin{pmatrix} \operatorname{cov}(w_n \times a_i,\, w_n) \\ \operatorname{cov}(w_n \times a_i,\, a_i) \end{pmatrix}.$$

Using firm–industry independence to evaluate the right-hand side:

$$\operatorname{cov}(w_n \times a_i,\, w_n) \;=\; \mathbb{E}(a_i)\operatorname{var}(w_n), \qquad \operatorname{cov}(w_n \times a_i,\, a_i) \;=\; \mathbb{E}(w_n)\operatorname{var}(a_i),$$

so $\beta_2 = \mathbb{E}(a_i)$ and $\beta_1 = \mathbb{E}(w_n)$. The intercept is $\alpha_1 = \mathbb{E}(w_n \times a_i) - \beta_2 \mathbb{E}(w_n) - \beta_1 \mathbb{E}(a_i) = -\mathbb{E}(w_n)\mathbb{E}(a_i)$, where we used $\mathbb{E}(w_n \times a_i) = \mathbb{E}(w_n)\mathbb{E}(a_i)$ (by independence). Substituting into (IA5), the residual is

$$r^{(1)}_{i,n} \;=\; \big(w_n - \mathbb{E}(w_n)\big)\big(a_i - \mathbb{E}(a_i)\big) \;\equiv\; v_n\, \tilde{a}_i, \tag{IA6}$$

with $v_n \equiv w_n - \mathbb{E}(w_n)$ and $\tilde{a}_i \equiv a_i - \mathbb{E}(a_i)$.

Regression (ii): projecting $g_{i,n}$ onto $(1, w_n, a_i)$. Consider the population regression

$$g_{i,n} \;=\; \alpha_2 \;+\; \kappa_2\, w_n \;+\; \rho_2\, a_i \;+\; r^{(2)}_{i,n}, \tag{IA7}$$

where $r^{(2)}_{i,n}$ is the residual. We do not need to solve for $(\alpha_2, \kappa_2, \rho_2)$ explicitly. What matters is that, by definition of population OLS, $r^{(2)}_{i,n}$ is uncorrelated with each of the regressors $(1, w_n, a_i)$.

Regression (iii): univariate regression of $r^{(2)}_{i,n}$ on $r^{(1)}_{i,n}$. By the FWL theorem, the coefficient ζ in (IA4) equals the coefficient from the univariate regression of $r^{(2)}_{i,n}$ on $r^{(1)}_{i,n} = v_n \tilde{a}_i$:

$$\zeta \;=\; \frac{\operatorname{cov}(r^{(2)}_{i,n},\, v_n \tilde{a}_i)}{\operatorname{var}(v_n \tilde{a}_i)}. \tag{IA8}$$

We now show that $\operatorname{cov}(r^{(2)}_{i,n}, v_n \tilde{a}_i) = \operatorname{cov}(g_{i,n}, v_n \tilde{a}_i)$, so that $r^{(2)}_{i,n}$ can be replaced by $g_{i,n}$ in the numerator. From (IA7), $r^{(2)}_{i,n} = g_{i,n} - \alpha_2 - \rho_2 a_i - \kappa_2 w_n$, so

$$\operatorname{cov}(r^{(2)}_{i,n},\, v_n \tilde{a}_i) \;=\; \operatorname{cov}(g_{i,n},\, v_n \tilde{a}_i) \;-\; \rho_2 \operatorname{cov}(a_i,\, v_n \tilde{a}_i) \;-\; \kappa_2 \operatorname{cov}(w_n,\, v_n \tilde{a}_i).$$

Both $\operatorname{cov}(a_i, v_n \tilde{a}_i)$ and $\operatorname{cov}(w_n, v_n \tilde{a}_i)$ vanish, because $v_n \tilde{a}_i = r^{(1)}_{i,n}$ is the residual from regression (i) and is

therefore orthogonal to $(1, w_n, a_i)$. Hence

$$\zeta \;=\; \frac{\operatorname{cov}(g_{i,n},\, v_n \tilde{a}_i)}{\operatorname{var}(v_n \tilde{a}_i)}. \tag{IA9}$$

It remains to evaluate the numerator and denominator of (IA9) using (IA1) and Assumption IA1. For the denominator, since v_n depends only on industry n and $\tilde{a}_i$ only on firm i, firm–industry independence gives

$$\operatorname{var}(v_n\, \tilde{a}_i) \;=\; \operatorname{var}(w_n)\operatorname{var}(a_i).$$

For the numerator, substituting (IA1):

$$\operatorname{cov}(g_{i,n},\, v_n \tilde{a}_i) \;=\; \operatorname{cov}(1 - \Lambda_n(\rho_i - \bar{\rho}),\, v_n\, \tilde{a}_i)\,.$$

The constant 1 contributes zero. For the second term, Λ_n depends only on industry n, $(\rho_i - \bar{\rho})$ depends only on firm i with mean zero, and $v_n\, \tilde{a}_i$ has mean zero. Firm–industry independence gives

$$\begin{aligned}
\operatorname{cov}(\Lambda_n(\rho_i - \bar{\rho}),\, v_n \tilde{a}_i) \;&=\; \mathbb{E}[\Lambda_n(\rho_i - \bar{\rho})\, v_n \tilde{a}_i] - \mathbb{E}[\Lambda_n(\rho_i - \bar{\rho})]\; \mathbb{E}[v_n \tilde{a}_i] \\
&=\; \mathbb{E}[\Lambda_n\, v_n]\, \mathbb{E}[(\rho_i - \bar{\rho})\, \tilde{a}_i] \;-\; 0 \\
&=\; \operatorname{cov}(\Lambda_n,\, w_n)\operatorname{cov}(\rho_i, a_i),
\end{aligned}$$

where the last line uses $\mathbb{E}[v_n] = 0 \Rightarrow \mathbb{E}[\Lambda_n v_n] = \operatorname{cov}(\Lambda_n, w_n)$ and $\mathbb{E}[(\rho_i - \bar{\rho})\tilde{a}_i] = \operatorname{cov}(\rho_i, a_i)$. Substituting $\Lambda_n = (\Xi_n \phi f / L_n F^*)\Delta_n$ and $w_n = \frac{1}{2}\Delta_n$:

$$\operatorname{cov}(\Lambda_n,\, w_n) \;=\; \operatorname{cov}\Big(\tfrac{\Xi_n \phi f}{L_n F^*}\Delta_n,\, \tfrac{1}{2}\Delta_n\Big) \;=\; \frac{\phi\, f}{2F^*}\operatorname{cov}(\Xi_n \Delta_n / L_n,\, \Delta_n)\,,$$

so the numerator of (IA9) is

$$\operatorname{cov}(g_{i,n},\, v_n \tilde{a}_i) \;=\; -\,\frac{\phi\, f}{2F^*}\operatorname{cov}(\Xi_n \Delta_n / L_n,\, \Delta_n)\, \operatorname{cov}(\rho_i, a_i). \tag{IA10}$$

Combining (IA10) with $\operatorname{var}(w_n) = \frac{1}{4}\operatorname{var}(\Delta_n)$:

$$\zeta \;=\; \frac{-\frac{\phi f}{2F^*}\operatorname{cov}(\Xi_n \Delta_n / L_n,\, \Delta_n)\operatorname{cov}(\rho_i, a_i)}{\frac{1}{4}\operatorname{var}(\Delta_n)\operatorname{var}(a_i)} \;=\; -\,\frac{2\,\phi\, f}{F^*}\,\frac{\operatorname{cov}(\Xi_n \Delta_n / L_n,\, \Delta_n)}{\operatorname{var}(\Delta_n)}\,\frac{\operatorname{cov}(\rho_i, a_i)}{\operatorname{var}(a_i)},$$

which is (IA2) with Φ as defined in (IA3). □

Sufficient conditions for $\Phi > 0$ The covariance form (IA3) of Φ admits the following clean sufficient condition.

Assumption IA2. *The industry-level spanning-to-supply ratio Ξ_n / L_n is independent of the workstyle distance*

Δ_n.

Economically, Assumption IA2 says that industries in which the new business model induces a larger workstyle change are not systematically those in which firms are more (or less) constrained by their managerial supply. The assumption holds trivially when $\Xi_n/L_n = \Xi/L$ is industry-invariant, but is also consistent with substantial cross-industry heterogeneity in Ξ_n/L_n provided that heterogeneity is not specifically aligned with Δ_n.

Result IA4. *Under Assumption IA2,*

$$\Phi \;=\; \frac{2\,\phi\,f}{F^*}\,\mathbb{E}\!\left[\frac{\Xi_n}{L_n}\right] \;>\; 0. \tag{IA11}$$

In particular, when $\Xi_n/L_n = \Xi/L$ is industry-invariant, $\Phi = 2(\Xi/L)\,\phi f/F^*$.

Proof. Under Assumption IA2, Ξ_n/L_n and Δ_n are independent, so

$$\operatorname{cov}(\Xi_n\Delta_n/L_n,\;\Delta_n) \;=\; \mathbb{E}[\Xi_n/L_n]\;\operatorname{var}(\Delta_n).$$

Substituting into (IA3) gives $\Phi = 2\phi f\,\mathbb{E}[\Xi_n/L_n]/F^* > 0$, since $\phi, f > 0$, $F^* > 0$, and $\Xi_n/L_n > 0$ in every industry. □

Assumption IA2 is sufficient but not necessary. From (IA3), $\Phi > 0$ if and only if

$$\operatorname{cov}(\Xi_n\Delta_n/L_n,\;\Delta_n) \;>\; 0, \tag{IA12}$$

since $2\phi f/F^* > 0$. The left-hand side is the population covariance, across industries, between the firm-level rule-deviation loading $\Xi_n\Delta_n/L_n$ in (IA1) and the workstyle distance Δ_n itself. To see that (IA12) holds even without full independence, define $\widetilde{\Delta}_n \equiv \Delta_n - \mathbb{E}(\Delta_n)$ and decompose

$$\operatorname{cov}\!\left(\frac{\Xi_n}{L_n}\Delta_n,\;\Delta_n\right) \;=\; \underbrace{\mathbb{E}\!\left[\frac{\Xi_n}{L_n}\right]\operatorname{var}(\Delta_n)}_{>\,0} \;+\; \underbrace{\operatorname{cov}\!\left(\frac{\Xi_n}{L_n},\;\Delta_n\,\widetilde{\Delta}_n\right)}_{\gtrless\,0}. \tag{IA13}$$

The first term is strictly positive. The correction term vanishes under Assumption IA2 and is small when Ξ_n/L_n and Δ_n are weakly dependent. The condition $\Phi > 0$ therefore fails only if industries with the largest workstyle distances happen to have systematically smaller spanning-to-supply ratios, by an amount large enough to outweigh the positive baseline. This is the economically implausible case in which the new business models that require the most organizational adjustment also happen to arise in industries that are least able to absorb the friction at the manager-wage margin.

IA6.4 Generated-regressor bias in the single-interaction specification

In the data the model workstyle change $w_n = |d\overline{ws}|_n$ in regression (29) is replaced by the constructed proxy $|d\overline{ws}|^e_{n,t,5}$, estimated in a first stage (the occupation-employment prediction (37) and the workstyle aggregation (39)) and hence a *generated* regressor. We ask how this affects the OLS estimate of the interaction coefficient ζ. Throughout we use population (probability-limit) OLS coefficients, keep the intercept and both main effects, and derive rather than assume any centering.

Setup. Write the structural reduced form (29) as

$$g_{i,n} = \nu + \kappa_w w_n + \rho_a a_i + \zeta\,(a_i w_n) + \varepsilon_{i,n}, \qquad \mathbb{E}[\varepsilon_{i,n} \mid a_i, w_n] = 0, \tag{IA14}$$

with proxy

$$\widehat{w}_n = w_n + e_n \tag{IA15}$$

for first-stage construction error e_n. The feasible regression replaces w_n by $\widehat{w}_n$,

$$g_{i,n} = \nu + \kappa_w \widehat{w}_n + \rho_a a_i + \zeta\,(a_i \widehat{w}_n) + \eta_{i,n}, \tag{IA16}$$

where substituting (IA15) into (IA14) gives the error $\eta_{i,n} \equiv \varepsilon_{i,n} - (\kappa_w + \zeta a_i)\, e_n$,[31] and $\widehat{\zeta}$ denotes the probability limit of the OLS coefficient on $a_i \widehat{w}_n$.

Assumption IA3. *(i) Firm age is independent of the industry-level pair: $a_i \perp (w_n, e_n)$. (ii) The structural error in (IA14) is mean-independent of age and the proxy: $\mathbb{E}[\varepsilon_{i,n} \mid a_i, \widehat{w}_n] = 0$.*

Part (i) combines Assumption IA1 (a_i independent of all industry-level variables, so $a_i \perp w_n$) with the fact that the proxy is built only from industry×occupation×year objects, so its error e_n is an industry-level object that cannot track firm age ($a_i \perp e_n$). We impose nothing on the relationship between e_n and w_n; in particular we do not assume the error is classical. We also make no assumption on $\mathbb{E}[e_n]$: any nonzero mean of the proxy is absorbed by the intercept in (IA16) and, in the interaction, by the age main effect ρ_a (since $a_i \mathbb{E}[e_n]$ is collinear with a_i). Part (ii) is not implied by the conditional mean stated in (IA14); both follow from the primitive $\mathbb{E}[\varepsilon_{i,n} \mid a_i, w_n, e_n] = 0$.

Parts (i)–(ii) pin down the error $\eta_{i,n}$ in (IA16). Since $(e_n, \widehat{w}_n)$ is a function of (w_n, e_n), part (i) implies $a_i \perp (\widehat{w}_n, e_n)$ and hence $\mathbb{E}[e_n \mid a_i, \widehat{w}_n] = \mathbb{E}[e_n \mid \widehat{w}_n]$, so

$$\mathbb{E}[\eta_{i,n} \mid a_i, \widehat{w}_n] = -(\kappa_w + \zeta a_i)\, \mathbb{E}[e_n \mid \widehat{w}_n].$$

In words, any violation of mean-independence for the error term $\eta_{i,n}$ in the feasible regression (IA16) must run through the construction error e_n. Result IA5 formalizes its consequence for the OLS coefficient $\widehat{\zeta}$ in

[31] Not to be confused with the main-effect coefficients η_w and η_a in the paper's empirical specifications (41) and (42): $\eta_{i,n}$, indexed by firm and industry, is the composite error term of the feasible regression.

(IA16).

Result IA5. *Under Assumption IA3, the probability limit of the OLS interaction coefficient in* (IA16) *is*

$$\widehat{\zeta} = \zeta \times \frac{\operatorname{cov}(w_n, \widehat{w}_n)}{\operatorname{var}(\widehat{w}_n)}, \tag{IA17}$$

the population regression coefficient of w_n on the proxy $\widehat{w}_n$. Equivalently, by (IA15)*,*

$$\widehat{\zeta} = \zeta \times \frac{\operatorname{var}(w_n) + \operatorname{cov}(w_n, e_n)}{\operatorname{var}(w_n) + \operatorname{var}(e_n) + 2\operatorname{cov}(w_n, e_n)}. \tag{IA18}$$

Since $\operatorname{var}(\widehat{w}_n) > 0$, $\widehat{\zeta}$ has the same sign as ζ if and only if $\operatorname{cov}(w_n, \widehat{w}_n) > 0$.

Proof. By the Frisch–Waugh–Lovell theorem, $\widehat{\zeta}$ is the coefficient from the univariate population regression of $g_{i,n}$ on the residual of $a_i\widehat{w}_n$ after projecting out $(1, a_i, \widehat{w}_n)$. Because $a_i \perp \widehat{w}_n$ by Assumption IA3(i), that residual is the product of the demeaned regressors, $\tilde{a}_i\tilde{v}_n$ where $\tilde{a}_i \equiv a_i - \mathbb{E}[a_i]$ and $\tilde{v}_n \equiv \widehat{w}_n - \mathbb{E}[\widehat{w}_n]$, so $\widehat{\zeta} = \operatorname{cov}(g_{i,n}, \tilde{a}_i\tilde{v}_n)/\operatorname{var}(\tilde{a}_i\tilde{v}_n)$.[32] Substituting (IA14): by firm–industry independence and $\mathbb{E}[\tilde{a}_i] = \mathbb{E}[\tilde{v}_n] = 0$, the intercept and both main effects vanish and $\varepsilon_{i,n}$ drops by Assumption IA3(ii); only the interaction survives, giving $\operatorname{cov}(g_{i,n}, \tilde{a}_i\tilde{v}_n) = \zeta \operatorname{var}(a_i) \operatorname{cov}(w_n, \widehat{w}_n)$. Dividing by the denominator $\operatorname{var}(a_i) \operatorname{var}(\widehat{w}_n)$ gives (IA17), and (IA15) gives (IA18). □

Implications. Result IA5 has two implications: (i) If the construction error is classical, $\operatorname{cov}(w_n, e_n) = 0$, then (IA18) reduces to $\widehat{\zeta} = \zeta \operatorname{var}(w_n)/(\operatorname{var}(w_n) + \operatorname{var}(e_n))$, biased toward zero by the proxy's reliability ratio but never across it. The test of $\zeta \neq 0$ is therefore conservative. This is the classical errors-in-variables attenuation, here for the interaction term. (ii) Without the classical assumption, the sign in (IA17) is preserved if and only if $\operatorname{cov}(w_n, \widehat{w}_n) > 0$; by (IA18) this fails only if $\operatorname{cov}(w_n, e_n) \leq -\operatorname{var}(w_n)$, which requires the construction error to be both larger in variance than the signal and strongly negatively correlated with it (a sign flip requires the strict inequality). This condition involves the unobserved w_n—the true expected workstyle change—so it cannot be rejected directly in the data. The first stage nonetheless narrows it down: prediction regression (37) loads with the expected signs and is strongly significant, and the resulting proxy is strongly positively correlated with *realized* workstyle changes ($\rho \approx 0.55$ across industry-years). A sign flip would therefore require the true expected change to covary negatively with our measure while realized changes covary positively with it—a correlation structure that is implausible in any model in which the technology shocks underlying (37) are an important driver of workstyle change.

Discussion of Assumption IA3. Two potential violations of Assumption IA3(i) are worth discussing. First, e_n could inherit a firm-level component correlated with a_i if the first stage used firm-level information. It does not, because $\widehat{w}_n$ is built with no firm identifiers—thus e_n is an industry-level error. Second, a

[32] The normal equations are diagonal and the calculation is identical to (IA5)–(IA6) in the proof of Result IA3 with $\widehat{w}_n$ in place of w_n; the centering is derived from the projection, not assumed.

link between w_n and an industry's *average* age is harmless for ζ. It shifts ρ_a, not the interaction, which is identified from *within*-industry covariation of age with the industry-common regressor. What would threaten ζ is e_n correlated with within-industry age *deviations*, which an industry-level first stage cannot generate.

Assumption IA3(ii) is the usual exogeneity of $\varepsilon_{i,n}$; it would fail if $\widehat{w}_n$ proxied for an omitted industry shock affecting old and young firms differentially—an omitted-variable concern distinct from the generated-regressor question and addressed by the controls in (IA24).

Remark. The bootstrap used throughout the empirical work (resampling industries and re-estimating (37)–(39) in each draw) carries the sampling variability that the generated regressor injects into the standard errors; Result IA5 is the complementary statement about *bias*, showing that this variability shrinks the point estimate toward zero and preserves its sign under positive validity.

IA6.5 Regression with controls

We generalize Result IA3 to an extended regression that includes an additional industry-level control Z_n and an additional firm-level control x_i, along with all pairwise interactions of industry-level and firm-level variables. Adding controls does not change the interpretation of the key interaction coefficient ζ_{wa}, but replaces the bivariate moments in (IA2) with their multivariate (conditional) counterparts.

Notation We maintain the notation from the proof of Result IA3: $w_n = \frac{1}{2}\Delta_n$, $\rho_i = 1 - e^{-\kappa a_i}$, and $\Lambda_n = (\Xi_n \phi f / L_n F^*)\Delta_n$. Let Z_n denote an additional industry-level control and x_i an additional firm-level control. Define the demeaned variables

$$v_n \equiv w_n - \mathbb{E}(w_n), \quad u_n \equiv Z_n - \mathbb{E}(Z_n), \quad \tilde{a}_i \equiv a_i - \mathbb{E}(a_i), \quad \tilde{x}_i \equiv x_i - \mathbb{E}(x_i),$$

and the variance–covariance matrices of the level industry- and firm-level regressors:

$$\Sigma_I = \begin{pmatrix} \operatorname{var}(w_n) & \operatorname{cov}(w_n, Z_n) \\ \operatorname{cov}(w_n, Z_n) & \operatorname{var}(Z_n) \end{pmatrix}, \qquad \Sigma_F = \begin{pmatrix} \operatorname{var}(a_i) & \operatorname{cov}(a_i, x_i) \\ \operatorname{cov}(a_i, x_i) & \operatorname{var}(x_i) \end{pmatrix}. \tag{IA19}$$

Consider two population multivariate regressions. First, the regression of the industry-level loading $\Xi_n \Delta_n / L_n$ on (w_n, Z_n):

$$\frac{\Xi_n \Delta_n}{L_n} = \alpha_\Phi + \Phi_{w|Z}\, w_n + \Phi_{Z|w}\, Z_n + u_n^{(\Phi)}. \tag{IA20}$$

By definition of population OLS, $(\Phi_{w|Z}, \Phi_{Z|w})'$ satisfies

$$\Sigma_I \begin{pmatrix} \Phi_{w|Z} \\ \Phi_{Z|w} \end{pmatrix} \;=\; \begin{pmatrix} \mathrm{cov}(\Xi_n \Delta_n / L_n,\; w_n) \\ \mathrm{cov}(\Xi_n \Delta_n / L_n,\; Z_n) \end{pmatrix}. \tag{IA21}$$

Second, the regression of the firm-level rule stock ρ_i on (a_i, x_i):

$$\rho_i \;=\; \alpha_\rho \;+\; \beta_{a|x}\, a_i \;+\; \beta_{x|a}\, x_i \;+\; u_i^{(\rho)}. \tag{IA22}$$

The coefficient vector $(\beta_{a|x}, \beta_{x|a})'$ satisfies

$$\Sigma_F \begin{pmatrix} \beta_{a|x} \\ \beta_{x|a} \end{pmatrix} \;=\; \begin{pmatrix} \mathrm{cov}(\rho_i,\; a_i) \\ \mathrm{cov}(\rho_i,\; x_i) \end{pmatrix}. \tag{IA23}$$

Assumption IA4. *The joint distribution of (a_i, ξ_i, x_i) is identical across industries n and is independent of all industry-level variables (including $Q_{A,n}$, $Q_{B,n}$, Δ_n, w_n, Z_n, L_n, and Ξ_n). No restriction is imposed on $\mathrm{cov}(w_n, Z_n)$ or on $\mathrm{cov}(a_i, x_i)$.*

Result IA6. *Under Assumption IA4, the population OLS coefficient on $w_n \times a_i$ in the extended regression*

$$\begin{aligned} g_{i,n} \;=\; \nu \;+\; \kappa_w\, w_n \;+\; \kappa_Z\, Z_n \;+\; \rho_a\, a_i \;+\; \rho_x\, x_i \\ +\; \zeta_{wa}\,(w_n \times a_i) \;+\; \zeta_{wx}\,(w_n \times x_i) \;+\; \zeta_{Za}\,(Z_n \times a_i) \;+\; \zeta_{Zx}\,(Z_n \times x_i) \;+\; \varepsilon_{i,n} \end{aligned} \tag{IA24}$$

is

$$\zeta_{wa} \;=\; -\,\frac{\phi\, f}{F^*}\, \Phi_{w|Z}\, \beta_{a|x}, \tag{IA25}$$

where $\Phi_{w|Z}$ is the partial coefficient on w_n in (IA20) and $\beta_{a|x}$ is the partial coefficient on a_i in (IA22). More generally, the full 4×1 vector of interaction coefficients satisfies

$$\begin{pmatrix} \zeta_{wa} \\ \zeta_{wx} \\ \zeta_{Za} \\ \zeta_{Zx} \end{pmatrix} \;=\; -\,\frac{\phi\, f}{F^*} \begin{pmatrix} \Phi_{w|Z} \\ \Phi_{Z|w} \end{pmatrix} \otimes \begin{pmatrix} \beta_{a|x} \\ \beta_{x|a} \end{pmatrix}, \tag{IA26}$$

where $\otimes$ denotes the Kronecker product.

Relative to Result IA3, the key changes are: (i) the industry-level scalar Φ is replaced by $(\phi f / F^*)\, \Phi_{w|Z}$, the conditional regression coefficient of $\Xi_n \Delta_n / L_n$ on w_n given Z_n, scaled by $\phi f / F^*$; and (ii) the ratio $\mathrm{cov}(\rho_i, a_i)/\mathrm{var}(a_i)$ is replaced by $\beta_{a|x}$, the conditional regression coefficient of ρ_i on a_i given x_i. In the bivariate special case $Z_n \equiv 0, x_i \equiv 0$, equation (IA25) reduces to (IA2): $\Phi_w = \mathrm{cov}(\Xi_n \Delta_n / L_n, w_n)/\mathrm{var}(w_n) = 2\,\mathrm{cov}(\Xi_n \Delta_n / L_n, \Delta_n)/\mathrm{var}(\Delta_n)$ (using $w_n = \frac{1}{2}\Delta_n$), so $(\phi f / F^*)\, \Phi_w = \Phi$.

Proof. We apply the FWL theorem in three steps: (i) project each of the four interaction terms onto the level regressors $(1, w_n, Z_n, a_i, x_i)$ and collect residuals; (ii) project $g_{i,n}$ onto the same level regressors and collect the residual; (iii) run a multivariate regression of the residual from (ii) on the four residuals from (i).

Regression (i): projecting interaction terms onto level regressors.

The variance matrix of the demeaned level regressors $(v_n, u_n, \tilde{a}_i, \tilde{x}_i)$ is block-diagonal by firm–industry independence:

$$\operatorname{var}\begin{pmatrix} v_n \\ u_n \\ \tilde{a}_i \\ \tilde{x}_i \end{pmatrix} = \begin{pmatrix} \Sigma_I & 0 \\ 0 & \Sigma_F \end{pmatrix}.$$

For each interaction (industry variable) $\times$ (firm variable), the right-hand side of the normal equations factors by firm–industry independence. Take $w_n \times a_i$ as an example. Its covariances with the level regressors are

$$\operatorname{cov}(w_n \times a_i,\, w_n) = \mathbb{E}(a_i)\operatorname{var}(w_n), \quad \operatorname{cov}(w_n \times a_i,\, Z_n) = \mathbb{E}(a_i)\operatorname{cov}(w_n, Z_n),$$

$$\operatorname{cov}(w_n \times a_i,\, a_i) = \mathbb{E}(w_n)\operatorname{var}(a_i), \quad \operatorname{cov}(w_n \times a_i,\, x_i) = \mathbb{E}(w_n)\operatorname{cov}(a_i, x_i).$$

The industry block of the right-hand side is $\mathbb{E}(a_i)$ times the first column of Σ_I, giving coefficients $(\mathbb{E}(a_i), 0)$ on (w_n, Z_n). The firm block is $\mathbb{E}(w_n)$ times the first column of Σ_F, giving coefficients $(\mathbb{E}(w_n), 0)$ on (a_i, x_i). The intercept is $\alpha = -\mathbb{E}(w_n)\mathbb{E}(a_i)$, and the residual is

$$r_{i,n}^{(w,a)} \;=\; v_n\,\tilde{a}_i. \tag{IA27}$$

The same block-diagonal argument applies to each of the remaining three interactions:

$$r_{i,n}^{(w,x)} \;=\; v_n\,\tilde{x}_i, \qquad r_{i,n}^{(Z,a)} \;=\; u_n\,\tilde{a}_i, \qquad r_{i,n}^{(Z,x)} \;=\; u_n\,\tilde{x}_i. \tag{IA28}$$

Regression (ii): projecting $g_{i,n}$ onto level regressors.

The residual $r_{i,n}^{(g)}$ from the population regression of $g_{i,n}$ on $(1, w_n, Z_n, a_i, x_i)$ is, by construction, uncorrelated with each of the level regressors.

Regression (iii): multivariate regression of $r_{i,n}^{(g)}$ on interaction residuals.

By the FWL theorem, the coefficient vector $(\zeta_{wa}, \zeta_{wx}, \zeta_{Za}, \zeta_{Zx})'$ equals the coefficient from regressing $r_{i,n}^{(g)}$ on $(v_n\tilde{a}_i,\, v_n\tilde{x}_i,\, u_n\tilde{a}_i,\, u_n\tilde{x}_i)$. As in the proof of Result IA3, the interaction residuals are orthogonal to

the level regressors, so $r_{i,n}^{(g)}$ can be replaced by $g_{i,n}$ on the right-hand side. The normal equations are

$$M \begin{pmatrix} \zeta_{wa} \\ \zeta_{wx} \\ \zeta_{Za} \\ \zeta_{Zx} \end{pmatrix} = \begin{pmatrix} \operatorname{cov}(g_{i,n},\, v_n \tilde{a}_i) \\ \operatorname{cov}(g_{i,n},\, v_n \tilde{x}_i) \\ \operatorname{cov}(g_{i,n},\, u_n \tilde{a}_i) \\ \operatorname{cov}(g_{i,n},\, u_n \tilde{x}_i) \end{pmatrix}, \tag{IA29}$$

where M is the 4×4 variance–covariance matrix of $(v_n \tilde{a}_i,\, v_n \tilde{x}_i,\, u_n \tilde{a}_i,\, u_n \tilde{x}_i)$.

Evaluating the left-hand side By firm–industry independence, each entry of M factors:

$$\operatorname{cov}(r_n\, p_i,\, t_n\, q_i) \;=\; \operatorname{cov}(r_n, t_n)\operatorname{cov}(p_i, q_i), \qquad r, t \in \{v, u\},\; p, q \in \{\tilde{a}, \tilde{x}\}.$$

This gives $M = \Sigma_I \otimes \Sigma_F$, with rows and columns ordered as $(v{\cdot}a, v{\cdot}x, u{\cdot}a, u{\cdot}x)$.

Evaluating the right-hand side Substituting the linearized growth equation $g_{i,n} \approx 1 - \Lambda_n(\rho_i - \bar{\rho})$ from (IA1) and using $\operatorname{cov}(1, r_n p_i) = 0$ (since $\mathbb{E}[p_i] = 0$), we have for each $r \in \{v, u\}$ and $p \in \{\tilde{a}, \tilde{x}\}$:

$$\operatorname{cov}(g_{i,n},\, r_n\, p_i) \;=\; -\mathbb{E}[\Lambda_n\, r_n] \operatorname{cov}(\rho_i, p_i),$$

where firm–industry independence is used to factor the expectation. Since $\Lambda_n = (\phi f / F^*) \cdot (\Xi_n \Delta_n / L_n)$, $\mathbb{E}[\Lambda_n v_n] = (\phi f / F^*) \operatorname{cov}(\Xi_n \Delta_n / L_n, w_n)$ and $\mathbb{E}[\Lambda_n u_n] = (\phi f / F^*) \operatorname{cov}(\Xi_n \Delta_n / L_n, Z_n)$. Stacking the four covariances:

$$\begin{pmatrix} \operatorname{cov}(g_{i,n},\, v_n \tilde{a}_i) \\ \operatorname{cov}(g_{i,n},\, v_n \tilde{x}_i) \\ \operatorname{cov}(g_{i,n},\, u_n \tilde{a}_i) \\ \operatorname{cov}(g_{i,n},\, u_n \tilde{x}_i) \end{pmatrix} = -\frac{\phi f}{F^*} \begin{pmatrix} \operatorname{cov}(\Xi_n \Delta_n / L_n,\, w_n) \\ \operatorname{cov}(\Xi_n \Delta_n / L_n,\, Z_n) \end{pmatrix} \otimes \begin{pmatrix} \operatorname{cov}(\rho_i, a_i) \\ \operatorname{cov}(\rho_i, x_i) \end{pmatrix}. \tag{IA30}$$

Solving the system Substituting into (IA29) and using $M^{-1} = \Sigma_I^{-1} \otimes \Sigma_F^{-1}$:

$$\begin{pmatrix} \zeta_{wa} \\ \zeta_{wx} \\ \zeta_{Za} \\ \zeta_{Zx} \end{pmatrix} = -\frac{\phi f}{F^*} \left(\Sigma_I^{-1} \begin{pmatrix} \operatorname{cov}(\Xi_n \Delta_n / L_n,\, w_n) \\ \operatorname{cov}(\Xi_n \Delta_n / L_n,\, Z_n) \end{pmatrix} \right) \otimes \left(\Sigma_F^{-1} \begin{pmatrix} \operatorname{cov}(\rho_i, a_i) \\ \operatorname{cov}(\rho_i, x_i) \end{pmatrix} \right).$$

By (IA21), the first parenthesis equals $(\Phi_{w|Z}, \Phi_{Z|w})'$. By (IA23), the second equals $(\beta_{a|x}, \beta_{x|a})'$. Therefore

$$\begin{pmatrix} \zeta_{wa} \\ \zeta_{wx} \\ \zeta_{Za} \\ \zeta_{Zx} \end{pmatrix} \;=\; -\frac{\phi f}{F^*} \begin{pmatrix} \Phi_{w|Z} \\ \Phi_{Z|w} \end{pmatrix} \otimes \begin{pmatrix} \beta_{a|x} \\ \beta_{x|a} \end{pmatrix},$$

which is (IA26). In particular, $\zeta_{wa} = -\frac{\phi f}{F^*}\,\Phi_{w|Z}\,\beta_{a|x}$, which is (IA25). □

IA6.6 Who implements B?

Result IA7. *At impulse, firm i's share of total B-output in industry n, $\sigma_{B,i,0^+} \equiv L^*_{B,i,0^+}/L_n$, admits the multiplicative decomposition*

$$\sigma_{B,i,0^+} \;=\; \frac{\xi_i}{\Xi_n}\,\omega_B(\rho_{A,i,0^-};\Delta_n)\,, \qquad \omega_B(\rho;\Delta) \;\equiv\; 1 \;+\; \frac{\Xi_n}{L_n}\,\phi\left(1 - F(\Delta)/F^*\right)\left(\rho - \bar\rho_{A,n,0^-}\right).$$

The pre-arrival A-share, $\sigma_{A,i,0^-} \equiv L^*_{A,i,0^-}/L_n$, admits the parallel decomposition

$$\sigma_{A,i,0^-} \;=\; \frac{\xi_i}{\Xi_n}\,\omega_A(\rho_{A,i,0^-})\,, \qquad \omega_A(\rho) \;\equiv\; 1 \;+\; \frac{\Xi_n}{L_n}\,\phi\left(\rho - \bar\rho_{A,n,0^-}\right),$$

which is the special case of ω_B at $F(\Delta) = 0$. The ratio of firm i's B-share to its pre-arrival A-share is independent of ξ_i and equal to

$$\frac{\sigma_{B,i,0^+}}{\sigma_{A,i,0^-}} \;=\; \frac{\omega_B(\rho_{A,i,0^-};\Delta_n)}{\omega_A(\rho_{A,i,0^-})},$$

and is strictly decreasing in $\rho_{A,i,0^-}$ whenever $F(\Delta_n) > 0$.

Proof. The per-X-unit output $Q_X = \left(\sum_j \psi_{X,j}\right)^{1/(\sigma-1)}$ is identical across all X-units of all firms in industry n (Sections III.B and III.C), so firm i's total X-output at time t is $Y_{X,i,t} = Q_X\,L^*_{X,i,t}$. Aggregating across firms and using industry-level market clearing $\int_j L^*_{X,j,t}\,dj = L_n$ (with the normalization $L_{A,n} = L_{B,n} \equiv L_n$) gives $\sigma_{X,i,t} \equiv Y_{X,i,t}/Y_{X,n,t} = L^*_{X,i,t}/L_n$: the firm's output share equals its employment share.

Substituting the impulse FOC (24) into $\sigma_{B,i,0^+} = L^*_{B,i,0^+}/L_n$ and factoring out ξ_i/Ξ_n:

$$\begin{aligned} \sigma_{B,i,0^+} \;&=\; \frac{1}{L_n}\left[\frac{L_n}{\Xi_n} \;+\; \phi\left(1 - F(\Delta_n)/F^*\right)\left(\rho_{A,i,0^-} - \bar\rho_{A,n,0^-}\right)\right]\xi_i \\ &=\; \frac{\xi_i}{\Xi_n}\left[1 \;+\; \frac{\Xi_n}{L_n}\,\phi\left(1 - F(\Delta_n)/F^*\right)\left(\rho_{A,i,0^-} - \bar\rho_{A,n,0^-}\right)\right] \;=\; \frac{\xi_i}{\Xi_n}\,\omega_B(\rho_{A,i,0^-};\Delta_n)\,, \end{aligned}$$

which is the B-share decomposition. The pre-arrival A-share follows analogously by substituting $L^*_{A,i,0^-}$ from (10), with Δ_n effectively set to zero in the friction loading.

The ratio $\sigma_{B,i,0^+}/\sigma_{A,i,0^-} = \omega_B/\omega_A$ is independent of ξ_i since the (ξ_i/Ξ_n) factor cancels. Writing $\rho \equiv \rho_{A,i,0^-}$, $\lambda^A \equiv (\Xi_n/L_n)\phi$, and $\lambda^B \equiv (\Xi_n/L_n)\phi\big(1 - F(\Delta_n)/F^*\big)$ for compactness, direct differentiation in ρ gives

$$\frac{\partial}{\partial \rho}\left[\frac{\omega_B(\rho;\Delta_n)}{\omega_A(\rho)}\right] \;=\; \frac{\lambda^B\,\omega_A(\rho) \;-\; \lambda^A\,\omega_B(\rho;\Delta_n)}{\omega_A(\rho)^2}.$$

Substituting the linear expressions for ω_A and ω_B, the cross-products in $\lambda^A\lambda^B(\rho - \bar{\rho}_{A,n,0^-})$ cancel and the numerator collapses to the constant

$$\lambda^B \;-\; \lambda^A \;=\; -\,\frac{1}{F^*}\,\frac{\Xi_n}{L_n}\,\phi\,F(\Delta_n),$$

which is strictly negative whenever $F(\Delta_n) > 0$. Since $\omega_A(\rho)^2 > 0$ (positivity of $L^*_{A,i,0^-}$, Section III.B), the ratio ω_B/ω_A is strictly decreasing in ρ throughout the relevant range. □

IA6.7 Owner's optimal labor composition

In the baseline of Sections III.B–III.C, we decentralize the choice of labor composition $\{l_{X,i,k,j}\}_{j=1}^{J}$ within each operating unit to that unit's manager, who maximizes per-unit output Q_X. This appendix asks what changes if the owner of the firm instead chooses the labor composition directly, internalizing the cross-unit conflict friction at B-units. We derive the resulting deviation from the CES choice (the "composition wedge"), discuss its effect on the cross-sectional growth predictions of Section III.E, and characterize the two edge cases of perfect complements ($\sigma \to 0$) and perfect substitutes ($\sigma \to \infty$). All derivations are carried out under the linear friction $F(\Delta) = f\Delta$ assumed throughout the body of the paper.

The labor-composition choice is purely static The owner's labor-composition choice in our setting solves a per-instant static problem, decoupled from the dynamics of the rule stock. To see this, recall from (5) that the A-rule stock evolves according to $d\rho_{A,i,t} = \kappa(1 - \rho_{A,i,t})\,dt$ with closed-form trajectory $\rho_{A,i,t} = 1 - e^{-\kappa\, a_{i,t}}$, depending only on firm age. The B-rule stock evolves analogously from $t = 0^+$. In neither case does the labor composition $\{l_{X,i,k,j}\}$ enter the law of motion: problem arrival and rule formation are determined by the elapsed time, not by the workstyle of the workers solving the problems. Cast the firm's problem as a continuous-time control problem with state $X(t) \equiv (\rho_{A,i,t}, \rho_{B,i,t})$, dynamic controls $u_L(t) \equiv (L_{A,i,t}, L_{B,i,t})$, and labor-composition controls $u_l(t) \equiv (\{l_{A,i,k,j}\}, \{l_{B,i,k,j}\})$. The drift $f(X, u_L) \equiv (\kappa(1-\rho_{A,i,t}), \kappa(1-\rho_{B,i,t}))$ does not involve u_l. The HJB equation associated with the firm's value function therefore admits the separation

$$\sup_{u_L,\,u_l}\Big[\,\Pi(X,u_L,u_l) + \nabla W \cdot f(X,u_L)\,\Big] \;=\; \sup_{u_L}\Big[\,\Pi^*(X,u_L) + \nabla W \cdot f(X,u_L)\,\Big],$$

where $\Pi(X, u_L, u_l)$ is the per-instant flow profit and $\Pi^*(X, u_L) \equiv \sup_{u_l} \Pi(X, u_L, u_l)$ is the same profit with u_l optimized out at the current (X, u_L). Because f does not depend on u_l, the labor-composition choice never enters the continuation value, and the optimal $u_l^*(t)$ depends only on the current state and dynamic controls. The same separation holds pre-arrival: the post-arrival value $W(\rho_{A,i,t}, 0)$ inherited at the Poisson arrival of B is computed at the post-arrival re-optimized labor composition and so does not bind the pre-arrival choice.

The argument relies on two structural properties of the model: (i) the rule-stock drift does not depend on u_l — problem arrival and rule formation rates are time-driven, not workstyle-driven — and (ii) labor composition is freely re-chosen at every instant, so there is no labor-adjustment cost that would carry the choice forward in time. If either property fails (for instance, if the per-unit problem-arrival rate at a B-unit depended on the workstyle of its workers, or if the firm faced a quadratic cost of changing its labor mix), the term $\nabla W \cdot f$ would acquire a u_l-dependence and the separation above would no longer hold.

Pre-arrival: the wedge is identically zero Pre-arrival, only A-units operate, so the only labor composition controls are $\{l_{A,i,k,j}\}$. The flow profit reduces to $L_{A,i,t}\,\pi_{A,t}(\rho_{A,i,t}) - L_{A,i,t}^2/(2\xi_i)$, with $\pi_{A,t}(\rho_{A,i,t}) = Q_A(\{l_{A,i,k,j}\}) - 1 - w_{A,t} - \phi(1-\rho_{A,i,t})$ from (6). Crucially, the friction term $\phi(1-\rho_{A,i,t})$ depends only on the firm's A-rule stock, not on the workstyle of the A-workers solving the problems — there is no B-side to conflict with — so the only term in the flow profit sensitive to the labor composition is the unit-level output Q_A. The owner's static problem $\sup_{u_l} \Pi^{\text{pre}}(\rho_{A,i,t}, L_{A,i,t}, u_l)$ therefore reduces to $\max_{\{l_{A,i,k,j}\}} Q_A(\{l_{A,i,k,j}\})$ subject to $\sum_j l_{A,i,k,j} = 1$, which is the manager's myopic CES problem of Section III.B. The owner-optimal labor composition pre-arrival coincides with the manager's CES choice, $l^*_{A,i,k,j} = \psi_{A,j} / \sum_k \psi_{A,k}$, and the composition wedge is identically zero. The option value of B-arrival does not feed back into the pre-arrival labor choice: the post-arrival continuation $W(\rho_{A,i,t}, 0)$ is computed at the post-arrival re-optimized labor composition and so does not depend on the pre-arrival labor mix. The composition wedge analyzed below therefore arises entirely post-arrival.

The static post-arrival problem under linear friction Substituting $F(\Delta) = f\Delta$ into the per-B-unit profit (13), the firm's flow profit at $(L_{A,i,t}, L_{B,i,t})$ is, up to terms independent of the labor composition,

$$L_{A,i,t}\,Q_A\Big(\{l_{A,i,k,j}\}\Big) \;+\; L_{B,i,t}\,Q_B\Big(\{l_{B,i,k,j}\}\Big) \;-\; L_{B,i,t}\,\frac{\phi}{F^*}\,(1-\rho_{B,i,t})\,\rho_{A,i,t}\,f\,\Delta_{i,t},$$

where the firm-specific workstyle distance $\Delta_{i,t} = \|ws_{B,i,t} - ws_{A,i,t}\|$ depends on the chosen labor mixes through $ws_{X,i,t} = \sum_j l_{X,i,k,j}\, ws_j$. Suppressing firm and time subscripts and dropping terms independent of the labor composition, the owner solves

$$\max_{\{l_{A,j}\},\{l_{B,j}\}} \; L_A\,Q_A(\{l_{A,j}\}) \;+\; L_B\,Q_B(\{l_{B,j}\}) \;-\; K\,\Delta \quad \text{s.t.} \sum_j l_{X,j} = 1,\; X \in \{A,B\}, \tag{IA31}$$

with the friction-strength scalar

$$K \;\equiv\; L_B \,\frac{\phi}{F^*}\,(1-\rho_B)\,\rho_A\, f \;\geq\; 0. \tag{IA32}$$

Crucially, under linear F, K does not depend on Δ: differentiating $F(\Delta) = f\Delta$ gives $F'(\Delta) = f$ constant, so the marginal cost of widening the workstyle gap is constant in Δ and the FOCs admit a closed-form solution.

First-order conditions and closed-form wedge Letting μ_A, μ_B denote the multipliers on the labor-bundle constraints and $u \equiv (ws_B - ws_A)/\Delta$ the unit vector pointing from ws_A to ws_B, the FOCs of (IA31) are

$$L_A \,\frac{\partial Q_A}{\partial l_{A,j}} \;+\; K\,(u \cdot ws_j) \;=\; \mu_A, \quad \forall j, \tag{IA33}$$

$$L_B \,\frac{\partial Q_B}{\partial l_{B,j}} \;-\; K\,(u \cdot ws_j) \;=\; \mu_B, \quad \forall j, \tag{IA34}$$

where we used $\partial\Delta/\partial l_{A,j} = -u \cdot ws_j$ and $\partial\Delta/\partial l_{B,j} = +u \cdot ws_j$. The myopic CES choice $l^{CES}_{X,j} = \psi_{X,j}/\sum_k \psi_{X,k}$ satisfies the first term in each FOC as a constant across j (Euler's theorem on CES production: $\partial Q_X/\partial l_{X,j}|_{CES} = Q^{CES}_X$), but the wedge term $K(u \cdot ws_j)$ varies with the alignment of occupation j's workstyle and the gap direction u. The owner-optimal choice therefore deviates from CES whenever $K > 0$.

To first order around the CES choice, write $l^*_{X,j} = l^{CES}_{X,j}(1 + \nu_{X,j})$ with $\sum_j l^{CES}_{X,j}\nu_{X,j} = 0$ to preserve the labor-bundle constraint. Differentiating the CES marginal product $\partial Q_X/\partial l_{X,j} = (Q_X\,\psi_{X,j}/l_{X,j})^{1/\sigma}$ around the CES yields $\partial Q_X/\partial l_{X,j}|_{l^*} \approx Q^{CES}_X(1 - \nu_{X,j}/\sigma)$. Substituting into (IA33)–(IA34) and using the constraint to pin down the multipliers, the first-order solution for the wedges is

$$\nu_{A,j} \;=\; +\frac{\sigma\, K}{L_A\, Q^{CES}_A}\, u \cdot \left(ws_j - ws^{CES}_A\right), \qquad \nu_{B,j} \;=\; -\frac{\sigma\, K}{L_B\, Q^{CES}_B}\, u \cdot \left(ws_j - ws^{CES}_B\right). \tag{IA35}$$

The wedge at A-units shifts labor toward occupations whose workstyle is aligned with the direction u (occupations whose workstyle is "B-like" relative to A's CES mean); the wedge at B-units shifts labor away from B-like occupations and toward A-like ones. Both shifts pull the realized workstyles together. The intensity of each wedge scales with: (i) σ, the elasticity of substitution (more substitutability means a larger feasible deviation); (ii) K, the friction strength (active only post-arrival, with $L_B, \rho_A > 0$ and $\rho_B < 1$); (iii) $1/(L_X\, Q^{CES}_X)$, normalizing by the unit-type's flow output; and (iv) the alignment $u \cdot (ws_j - ws^{CES}_X)$ between occupation j's workstyle deviation from the CES mean and the gap direction u.

Compressed workstyle distance The shift in unit workstyle is $\Delta ws_X = \sum_j l_{X,j}^{CES} \nu_{X,j} ws_j$. Using $\sum_j l_{X,j}^{CES}(ws_j - ws_X^{CES}) = 0$, we obtain

$$\Delta ws_A = +\frac{\sigma K}{L_A Q_A^{CES}} \Sigma_A u, \qquad \Delta ws_B = -\frac{\sigma K}{L_B Q_B^{CES}} \Sigma_B u,$$

where $\Sigma_X \equiv \sum_j l_{X,j}^{CES}(ws_j - ws_X^{CES})(ws_j - ws_X^{CES})^\top$ is the labor-share-weighted occupation-workstyle covariance matrix at unit-type X. To first order in K, the owner-optimal workstyle distance is

$$\Delta^* \equiv \|(ws_B + \Delta ws_B) - (ws_A + \Delta ws_A)\| \approx \Delta - \sigma K u^\top \left[\frac{\Sigma_A}{L_A Q_A^{CES}} + \frac{\Sigma_B}{L_B Q_B^{CES}}\right] u, \quad \text{(IA36)}$$

strictly smaller than Δ whenever $K > 0$ and either Σ_A or Σ_B has a positive component along u (the generic case). The trade-off behind (IA36) is the classical envelope one: deviating from CES costs a second-order Q_X loss (the firm is perturbing around the CES maximum) for a first-order Δ gain, with the gain magnified by the friction strength K.

Implication for the cross-derivative result With the owner's labor-composition wedge, the per-B-unit friction is reduced from $\frac{\phi}{F^*}(1-\rho_B)\rho_A f \Delta$ to $\frac{\phi}{F^*}(1-\rho_B)\rho_A f \Delta^*$, with Δ^* given by (IA36). Substituting into the linearized impulse growth equation (IA1), the firm-level loading $\Lambda_n = (\Xi_n \phi f / L_n F^*)\Delta_n$ that appears in $g_{i,n} \approx 1 - \Lambda_n(\rho_i - \bar{\rho}_{A,n,0^-}) + O(\cdot)^2$ is replaced by an effective loading $\Lambda_{n,i}^* = (\Xi_n \phi f / L_n F^*)\Delta_{n,i}^*$, which is firm-specific (because Δ^* depends on the firm-level K), strictly smaller than Λ_n whenever $\sigma > 0$ and $K > 0$, but still strictly positive and still scaling with Δ_n and $\rho_{A,i,0^-}$. The cross-derivative result $\partial^2 g_{i,n} / \partial \rho_{A,i,0^-} \partial \Delta_n < 0$ of (28) therefore survives the labor-composition margin: it is reduced in magnitude (because the owner can partially offset the friction at the labor-composition margin) but preserved in sign. In the reduced-form regression of Section III.E, the coefficient $\hat{\zeta}$ in Result 2 continues to be negative, with magnitude attenuated by a factor that depends on σ and on the industry-level occupation covariance matrices $\Sigma_{A,n}$ and $\Sigma_{B,n}$.

Edge case: $\sigma = 0$ (Leontief) At $\sigma = 0$, the CES production function in (2) converges to the strict Leontief form $Q_X^{\text{Leontief}}(\{l_{X,j}\}) = \min_j(l_{X,j}/\psi_{X,j})$, in which occupations are perfect complements at the productivity-weighted proportions. Because Q_X^{Leontief} is non-smooth at the manager's CES choice, the first-order expansion (IA35) does not apply at the kink, and we solve the owner's joint problem (over L_X, the unit-level output, and the labor mix) from first principles.

Result IA8. *At $\sigma = 0$ with linear friction $F(\Delta) = f\Delta$, define*

$$\kappa_X \equiv \frac{\phi}{F^*} \rho_{A,i,0^+} f \Sigma\psi_X \Delta_n, \qquad X \in \{A, B\},$$

where $\Sigma\psi_X \equiv \sum_k \psi_{X,k}$ and $\Delta_n \equiv \|ws_{B,n}^{CES} - ws_{A,n}^{CES}\|$ is the industry workstyle distance at the manager's

CES choice. Assuming that $\kappa_A, \kappa_B < 1$, the owner operates each side at the manager's CES proportions:

$$t_X^* \;=\; 1/\Sigma\psi_X, \qquad L_X^* \;=\; \pi_{X,0^+}^{CES}\,\xi_i \;>\; 0 \qquad (X \in \{A,B\}).$$

The realized workstyle distance is $\Delta^* = \Delta_n$ (no labor-composition wedge), the friction is fully borne, and the cross-derivative result of (28) holds in its strongest form.

Proof. The proof proceeds in five steps.

Step 1: parameterization of feasible labor mixes. Any labor mix on side X satisfying $\sum_j l_{X,j} = 1$ delivers a unit-level output $t_X \equiv Q_X^{\text{Leontief}}(\{l_{X,j}\}) \in [0, 1/\Sigma\psi_X]$. Decompose $l_{X,j} = t_X\,\psi_{X,j} + e_{X,j}$ with $e_{X,j} \geq 0$, and let $E_X \equiv \sum_j e_{X,j} = 1 - t_X\Sigma\psi_X \in [0,1]$ be the "extra" labor budget. The first part of the labor pins Q_X^{Leontief} to t_X; the second part leaves Q_X^{Leontief} unchanged but is freely allocable across occupations and shifts the realized workstyle:

$$ws_X \;=\; (1-E_X)\,ws_X^{CES} + E_X\,\tilde{ws}_X, \qquad \tilde{ws}_X \;\equiv\; \sum_j (e_{X,j}/E_X)\,ws_j \;\in\; \text{conv}(\{ws_j\}) \;\text{ when } E_X > 0.$$

Step 2: minimum achievable workstyle distance. For given (t_A, t_B), the firm chooses $(\tilde{ws}_A, \tilde{ws}_B)$ to minimize $\|ws_B - ws_A\|$. Setting $\tilde{ws}_A = ws_B^{CES}$ and $\tilde{ws}_B = ws_A^{CES}$ (both feasible, since each CES workstyle is a convex combination of $\{ws_j\}_{j=1}^J$) pulls ws_A toward ws_B^{CES} by $E_A\Delta_n$ and ws_B toward ws_A^{CES} by $E_B\Delta_n$, delivering the minimum workstyle distance

$$\Delta^*(t_A, t_B) \;=\; \max\{(t_A\Sigma\psi_A + t_B\Sigma\psi_B - 1)\,\Delta_n,\, 0\}\,. \tag{IA37}$$

At full CES on both sides ($t_X = 1/\Sigma\psi_X$), $E_A + E_B = 0$ and $\Delta^* = \Delta_n$; the gap closes entirely once $t_A\Sigma\psi_A + t_B\Sigma\psi_B \leq 1$.

Step 3: reduced problem and Lagrangian. Substituting (IA37) into the owner's problem (17), and evaluating at the impulse $t = 0^+$:

$$\max_{L_A,L_B\geq 0,\; t_X\in[0,1/\Sigma\psi_X]} \; L_A\,\pi_{A,0^+}(t_A) + L_B\,\pi_{B,0^+}(t_B, \Delta^*(t_A,t_B)) \;-\; \frac{L_A^2 + L_B^2}{2\xi_i}, \tag{IA38}$$

with $\pi_{A,0^+}(t_A) \equiv t_A - 1 - w_{A,n,0^+} - \phi(1 - \rho_{A,i,0^+})$ and $\pi_{B,0^+}(t_B, \Delta) \equiv t_B - 1 - w_{B,n,0^+} - \phi + \phi\rho_{A,i,0^+}(1 - f\Delta/F^*)$. The Lagrangian is

$$\mathcal{L} \;=\; L_A\pi_{A,0^+} + L_B\pi_{B,0^+} - \frac{L_A^2 + L_B^2}{2\xi_i} \;+\; \mu_A^L L_A + \mu_B^L L_B \;+\; \mu_A^t\Big(\tfrac{1}{\Sigma\psi_A} - t_A\Big) + \mu_B^t\Big(\tfrac{1}{\Sigma\psi_B} - t_B\Big)\,,$$

with non-negativity multipliers $\mu_X^L, \mu_X^t \geq 0$ subject to complementary slackness.

Step 4: KKT for L_X. The FOC gives $L_X^* = \pi_{X,0^+}(t_X^*, \Delta^*)\,\xi_i$ when $\pi_{X,0^+} > 0$. Under the standing positivity restriction (A), $\pi_{X,0^+}^{CES} > 0$ for all firms in the industry (Appendix IA6.1), so the corner $L_X^* = 0$ does not

bind at the full-CES point. Substituting L_X^* back, the firm's payoff from operating side X at level t_X is $(\xi_i/2)\pi_{X,0^+}^2$, and the choice of (t_A, t_B) maximizes $(\xi_i/2)[\pi_{A,0^+}^2 + \pi_{B,0^+}^2]$.

Step 5: KKT for t_X. At an interior $\Delta^* > 0$ (so $\partial\Delta^*/\partial t_X = \Sigma\psi_X\Delta_n$), the FOCs in t_X are

$$\begin{aligned}\partial\mathcal{L}/\partial t_A &= \xi_i\Big(\pi_{A,0^+} - \kappa_A\pi_{B,0^+}\Big) - \mu_A^t = 0,\\ \partial\mathcal{L}/\partial t_B &= \xi_i\pi_{B,0^+}(1-\kappa_B) - \mu_B^t = 0.\end{aligned}$$

Under $\kappa_B < 1$, the second equation has no interior solution: the upper-bound multiplier $\mu_B^t > 0$ binds at $t_B^* = 1/\Sigma\psi_B$ (the firm wants the maximum unit-level output and pays full friction). Similarly, under $\kappa_A < 1$ and $\pi_{A,0^+} > \kappa_A\pi_{B,0^+}$ (a generic condition that follows from $\kappa_A < 1$ when $\pi_{A,0^+}$ and $\pi_{B,0^+}$ are of comparable magnitude), the corner $t_A^* = 1/\Sigma\psi_A$ binds. With both corners binding, $\Delta^* = \Delta_n$ and $L_X^* = \pi_{X,0^+}^{CES}\,\xi_i$. □

In the Leontief limit, the wedge between managers' and owner choice of labor composition collapses to zero. The firm owner cannot smoothly trade off unit-level output for workstyle alignment, so it stays at the manager's CES proportions and bears the full friction $\frac{\phi}{F^*}(1-\rho_B)\rho_A f\Delta_n$. The cross-derivative result of (28) therefore holds exactly.

Edge case: $\sigma \to \infty$ (perfect substitutes) At $\sigma \to \infty$, the CES production function in (2) collapses to perfect substitution. From the marginal product $\partial Q_X/\partial l_{X,j} = (Q_X\,\psi_{X,j}/l_{X,j})^{1/\sigma}$, all marginal products converge to 1 as $\sigma \to \infty$, regardless of the realized $l_{X,j}$ or $\psi_{X,j}$. Combined with the labor-bundle constraint $\sum_j l_{X,j} = 1$, this implies $Q_X \to 1$ for any feasible labor mix in the simplex, and the firm faces no production cost whatsoever from its choice of $ws_X = \sum_j l_{X,j}\,ws_j \in \mathrm{conv}(\{ws_j\})$.

Substituting $Q_X = 1$ into the owner's problem (IA31):

$$\max_{ws_A,\,ws_B \in \mathrm{conv}(\{ws_j\})} \; L_A + L_B \;-\; K\,\|ws_B - ws_A\|.$$

The first two terms are independent of (ws_A, ws_B), so the firm maximizes by minimizing the third (friction) term. Provided the convex hull $\mathrm{conv}(\{ws_j\})$ admits a common point $\bar{ws}$ achievable by both A- and B-units (a generic condition when the occupation set is rich enough), the firm chooses

$$ws_A^* = ws_B^* = \bar{ws}, \qquad \Delta^* = 0.$$

With full substitutability, the firm rebalances occupations across A- and B-units to align the realized workstyles exactly, paying no production cost in the limit. The friction is fully neutralized at the labor-composition margin, and the cross-derivative result of Section III.E vanishes: $\partial^2 g_{i,n}/\partial\rho_{A,i,0^-}\partial\Delta_n \to 0$ at $\sigma \to \infty$. For the model's empirical prediction to bite, σ must therefore be bounded away from infinity; put differently, the empirical prediction holds so long as occupations are not perfect substitutes in production.

IA6.8 Insulating B from A: separation, shading, and the boundary of the firm

This appendix formalizes the discussion of Section III.D. We work in terms of the per-B-unit profit $\pi_B(\rho_A, \rho_B; \Delta)$ of (13), suppress firm and industry subscripts. The maintained scope is that the incumbent wishes to implement B *as its own*, so that the relevant choice is between integration and internal separation.

Integration Under integration, the per-B-unit profit is $\pi_B(\rho_A, \rho_B)$ from (13). Write $\pi_B(0, \rho_B) = z_B - \phi(1 - \rho_B)$ for the profit a *rule-free* B-side would earn — one with no A-rulebook to draw on, which improvises every problem the B-rulebook does not yet cover from scratch, neither transferring an A-rule nor paying to override one. The value of the firm's A-rule stock to its B-side is the difference

$$\pi_B(\rho_A, \rho_B) \;-\; \pi_B(0, \rho_B) \;=\; \phi\,(1 - \rho_B)\,\rho_A \left[\,1 - F(\Delta)/F^*\,\right], \tag{IA39}$$

which is positive when $F(\Delta) < F^*$ and negative when $F(\Delta) > F^*$. Existing rules thus raise B's profit when the workstyle distance is small (a net asset) and depress it when it is large (a net liability); equivalently, $\partial \pi_B(\rho_A, \rho_B)/\partial \rho_A = \phi(1 - \rho_B)[1 - F(\Delta)/F^*]$ changes sign at F^*.

Internal separation and a shading microfoundation The within-firm way for an incumbent to shed the conflict, while keeping B as its own, is to run B as a stand-alone division on its own rulebook, so that B-workers no longer consult the A-rulebook. This eliminates the worker-level conflict, but it does not make B independent: the two divisions continue to report to a common headquarters that controls non-contractible shared resources — discretionary budgets, infrastructure, shared services. We show that this residual link reintroduces a cost with the same comparative statics as the original conflict, following the "contract as reference point" mechanism of Hart and Moore (2008).

At the impulse, headquarters has one unit of a shared resource to allocate between the two divisions; let $a \in [0, 1]$ be the share going to A and $1 - a$ to B. The allocation cannot be contracted upon ex ante and is set ex post to favor the division with the higher marginal return to the resource. Because business model B is adopted only when it is productive enough to overcome the organizational friction, and because that productivity is increasing in the workstyle distance, a larger Δ raises B's marginal return and tilts the ex post allocation toward B, so a falls with Δ. The A-division takes the maximal feasible allocation as its reference point and is aggrieved by the shortfall, which is increasing in Δ; write the grievance as $g(F(\Delta))$ with $g' > 0$ and $g(0) = 0$. Following Hart and Moore (2008), an aggrieved division *shades*: it withholds cooperation on shared services and contests B's initiatives at headquarters, reducing B's realized output in proportion to the grievance and to the division's ability to act on it. We take that ability to be increasing in the A-division's rule stock ρ_A, since a longer and more codified body of procedures makes A's objections more legible and more persuasive to headquarters. Normalizing the linear case $g(F(\Delta)) = F(\Delta)$, the per-B-unit profit under internal separation is the rule-free profit net of a shading cost,

$$\pi_B^{\text{sep}} \;=\; \pi_B(0, \rho_B) \;-\; \theta\,\phi\,(1 - \rho_B)\,\rho_A\,F(\Delta), \qquad \theta \geq 0, \tag{IA40}$$

where θ indexes the effectiveness of shading, equivalently the weakness of headquarters governance. An alternative microfoundation with the same reduced form follows the internal-capital-markets literature (Scharfstein and Stein, 2000; Rajan, Servaes, and Zingales, 2000; Stein, 1997): when A and B draw on a common internal budget, the entrenched A-division diverts funds away from B by an amount that increases in the divisions' workstyle distance and in A's influence. Either way, the division-level friction inherits the comparative statics of the worker-level one: it worsens with the workstyle distance and falls disproportionately on older firms, since

$$\frac{\partial^2 \pi_B^{\text{sep}}}{\partial \rho_A \, \partial \Delta} \;=\; -\,\theta\,\phi\,(1-\rho_B)\,F'(\Delta) \;<\; 0, \tag{IA41}$$

the same sign as the integrated cross-derivative $-\frac{\phi}{F^*}(1-\rho_B)F'(\Delta)$. The empirical predictions of Section III.E are therefore unchanged regardless of which organizational form the incumbent adopts.

The incumbent's choice An incumbent that keeps B chooses the more profitable of integration and internal separation. The following result characterizes that choice.

Result IA9 (Integration versus internal separation)**.** *Consider a firm with $\rho_A > 0$ facing workstyle distance Δ, and define*

$$F^{**} \;\equiv\; \frac{1}{1+\chi-\theta}, \tag{IA42}$$

taken to be $+\infty$ when $\theta \geq 1+\chi$. The incumbent integrates when $F(\Delta) < F^{**}$ and runs B as an internally separated division when $F(\Delta) > F^{**}$; in the latter case the residual shading cost $\theta\phi(1-\rho_B)\rho_A F(\Delta)$ remains.

Proof. Subtract (IA40) from $\pi_B(\rho_A, \rho_B)$ and use (IA39):

$$\pi_B(\rho_A,\rho_B) \;-\; \pi_B^{\text{sep}} \;=\; \phi\,(1-\rho_B)\,\rho_A\,\Big[\,1-(1+\chi-\theta)\,F(\Delta)\,\Big], \tag{IA43}$$

which is negative — so internal separation is strictly more profitable — iff $(1+\chi-\theta)F(\Delta) > 1$, i.e. iff $F(\Delta) > F^{**}$; otherwise integration is weakly more profitable. □