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

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

November 2026

Abstract

Since at least Schumpeter (1942), new firms have been widely viewed as the driving force for implementing new technologies. Yet there is little economics research on when and why new technologies require new firms. We examine the view that incumbents struggle when new technologies create organizational conflict by shifting employment toward occupations whose approaches to solving problems differ from those of 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 their industry induce larger shifts in workstyle. By contrast, the sheer volume of new technologies (e.g., the number of all or important patents) is not significantly related to the outperformance 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, Matthew Denes, Ed Glaeser, Oliver Hart, Greg Kaplan, Steve Kaplan, Josh Lerner, Gustavo Manso, Raffaella Sadun, Bryan Seegmiller, Andrei Shleifer, Scott Stern, Chad Syverson, Rob Vishny, Ting Xu, Hanbin Yang, Ben Zhang, and seminar participants at Blue Collar Working Group Meeting, Boston College, Boston University, Carnegie Mellon, Columbia, e61, ESSFM Gerzensee, FIRS, FOM Conference, Harvard, IWH, Kellogg, Labor & Finance Group, Macquarie, Michigan Ross, NBER Summer Institute Entrepreneurship meeting, Princeton Macro Finance Conference, RAPS/RCFS Europe Conference, Red Rock Finance Conference, SED, Stanford, Stanford SITE conference, Tsinghua, UCLA, University of Kentucky Finance Conference, UT Austin, VGSF Vienna, and Washington Foster 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 self-evident.

When—and why—do new technologies require new firms to be implemented? 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?" But 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, 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"; therefore 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 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 that implementing different technologies requires different occupations, which are associated with different approaches to solving problems. We measure these traits using O*NET data on occupations' workstyles, and use this term throughout the paper. 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 relatively easy to integrate into existing systems.[1]

Motivated by these observations, we build a model that lays out how new technologies associated with greater 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 of 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 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.

We assume that 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 must justify their approach in light of existing rules laid down by business model A that codify how to solve problems based on past experiences in A. Doing so takes time and reduces the productivity of business model B, shrinking its optimal size. Because older firms have more existing A rules, these losses are greater. Older firms therefore operate business model B at a smaller scale, and grow more slowly, when new technologies require larger changes in workstyle. To validate that older firms indeed have more rules and more time-wasting meetings, we process employee reviews from Revelio, computing the fraction of reviews that mention the presence of rules and the fraction that mention an excess of meetings. Across firms at a given point in time, both fractions are positively correlated with firm age.

We build on the model to develop a measure of workstyle changes in an industry induced by new

[1] Numerous news reports attest to the difficulty for legacy car makers to adapt to the different workstyles associated with software (George, 2022; Davis, 2023).

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 the text of breakthrough patents (Kelly et al., 2021). 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 compare the associated (employment-weighted average) future workstyle of the industry with its current one (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.

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 relative to incumbent firm valuations (the log ratio of VC investment to total Compustat market capitalization in an industry-year). This captures the forward-looking valuation of startups relative to incumbents, reflecting the differential growth potential of young vs old firms. We find that a one standard deviation increase in the predicted industry-level workstyle change is associated with a roughly 0.4 log point higher ratio of VC investment to market capitalization. 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—is not significantly related to VC investment and does not 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, breakthrough, creative, or rapidly evolving patents in the industry is not associated with significantly different outcomes for young vs old firms—that is, old firms are not invariably incapable when new technologies are more abundant, important, or creative. Further, the effects

of firm age are not due to age being correlated with size,[2] or due to older firms being profitable so they suffer more from having existing businesses 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 dataset may reflect subsistence rather than innovative entrepreneurship).[3] We find that when the technology-induced workstyle change is larger, employment growth is lower for older age groups, with the oldest group growing significantly more slowly than the youngest. The number of patents shows no comparable pattern: if anything, older age groups grow faster when patenting in the industry is high, the opposite of the workstyle pattern.

A classic argument for incumbents lagging behind is Arrow's replacement effect: an incumbent gains less from an innovation that cannibalizes its own profits (Arrow, 1962; Reinganum, 1983; Grossman and Helpman, 1991; Aghion and Howitt, 1992). This force differs from ours in two ways. First, the two forces work on distinct margins: cannibalization weighs on incumbents with large profits to protect, whereas our frictions weigh on firms whose new technologies require a different workstyle, whatever their profits. Second, cannibalization à la Arrow (1962) should be stronger when new products are closer substitutes for existing ones, whereas our frictions are stronger when new technologies are most different in organizational terms. Finally, in the data, we do not find that current profitability is a key impediment to adaptation.

Finally, both our model and our empirical analyses focus on which companies can *implement* new technologies. We take as given both the emergence of these technologies and the spillover of the knowledge they embody: artificial intelligence and information technology, for example, each emerged from decades of public and private research, and once a technology exists, in principle any company can work to adopt it. New technologies affect growth and business dynamism only once companies put them to use, and our results indicate that neither young nor old firms hold a uniform advantage in doing so: old firms can implement new technologies that fit the workstyle of their existing business, but they struggle, and young firms rise, when the required workstyle differs. Interestingly, in the data, we do not find that old firms patent less in those settings, which indicates that the obstacles they face are not simply "technological challenges" or an inability to learn new technical knowledge, as highlighted by Nadella.

[2] If anything, we find an opposite effect of firm size: larger firms experience faster growth when the industry faces a greater change in workstyle. Since size is positively correlated with age, it cannot account for the age effect we document. We discuss the potential mechanism behind this size effect in Section V.B.

[3] 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).

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; Braguinsky et al., 2024; Cohen, Gurun, and Nguyen, 2026). 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. Ma and Yang (2026) show how frictions in integrating new and legacy technologies shape the allocation of ideas between startups and incumbents, and the implications for market concentration. 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 from incomplete contracts and property rights,[4] 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 à 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, 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 evolve endogenously as the firm operates. In the appendix, we show that our key prediction continues

[4]Strictly speaking, the core takeaway of 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 that studies firms' organization structure based on the framework of property rights, showing that flat hierarchies feature more distinctive technologies than steep ones.

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.[5]

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. Relatedly, Bresnahan, Brynjolfsson, and Hitt (2002) show that the productivity gains from information technology depend on complementary organizational change. 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, a logic formalized in the multitask 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.

The rules that firms accumulate in our model connect to the literature on organizational capital (Prescott and Visscher, 1980; Atkeson and Kehoe, 2005; Lev and Radhakrishnan, 2005; Eisfeldt and Papanikolaou, 2013, 2014). A firm's stock of rules is a form of organizational capital, with two differences from standard formulations. First, it accumulates as a byproduct of operations (rules codify solutions to problems the firm has encountered) rather than through deliberate investment. Second, it can become a net liability. On this dimension, our setup is closer to the putty-clay literature (Johansen, 1959; Atkeson and Kehoe, 1999; Gilchrist and Williams, 2000): rules are putty when devised and clay once codified, when a new business model that they do not fit arrives. Yet putty-clay capital is at worst abandoned, whereas the old rule stock in our framework actively obstructs the new business model, and the firm cannot readily disinvest it. Such negative spillovers from old onto new capital are largely absent from both literatures. Another set of

[5] 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.

models studies coordination and adaptation in organizations, emphasizing how firms use local information in different settings (Milgrom and Roberts, 1992);[6] 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.

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. Section V shows the main empirical results. Section VI concludes.

II New Technologies and New Firms: Some Examples

In this section, we present stylized facts and examples to show that significant technological inventions are not invariably implemented by young firms. Some new technologies favor young firms, while others are implemented most successfully 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 in which 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 use the lists of significant technological innovations assembled by Kelly et al. (2021) and Kalyani et al. (2025). For each technology on the two lists, we use large language models (LLMs) to classify the firm that successfully implemented and commercialized it early on as either young (founded within the previous ten years) or old. The procedure is described in detail in Appendix Section IA2.

Figure 1 shows, for each set of technologies, the share that was most successfully implemented by young firms vs old firms (as well as cases that were unclear or never commercially implemented). Overall, we see a mix: both young and old firms play a substantial role in implementing new technologies. The share implemented by young firms is sizable and, in both samples, exceeds the employment share of firms less

[6] 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).

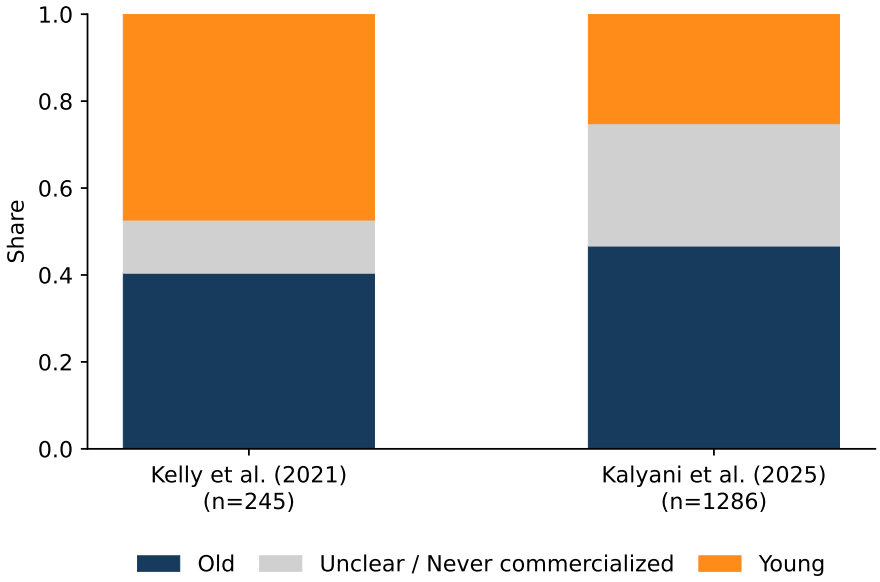


Figure 1. New Technologies Implemented by Young vs Old Firms

Notes: This figure compares the share of technologies most successfully implemented by young firms, old firms, and unclear or never-commercialized implementers across two samples. The first bar uses the full sample of important patents from Kelly et al. (2021). The second bar uses the full sample of technologies from Kalyani et al. (2025). Young firms are firms less than ten years old at implementation; old firms are firms at least ten years old at implementation. The unclear or never-commercialized category includes cases where the implementing firm age cannot be determined or where the technology was not commercially implemented by a firm. The shares are normalized to one within each sample. Colors correspond to implementation categories.

than ten years old in the Census Business Dynamics Statistics (BDS)—that is, young firms implement new technologies out of proportion to their size. Nonetheless, old firms are far from incapable of implementing new technologies, as they would be if, for example, fear of cannibalization were universal.

II.B Classic Examples and the Role of Workstyle

Business school classes have analyzed 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 over 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 followed 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 can we 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 car manufacturing comes to rely more heavily on software, organizational processes must change, given the different attributes of software vs hardware engineering that we later formalize as workstyle. This change is reflected in occupation composition, as software engineers replace hardware engineers. Changes in the composition of occupations with different workstyles thus provide a window into 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 producing 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 problems it has previously encountered. Fourth, a new business model, B, arrives at a single, industry-wide moment, and its different occupation composition exposes workers to a different mix of problems. As a result, 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 from the introduction, builds on the stylized facts of Section II, and motivates the empirical specifications 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 $\psi_{X,j} > 0$ 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 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

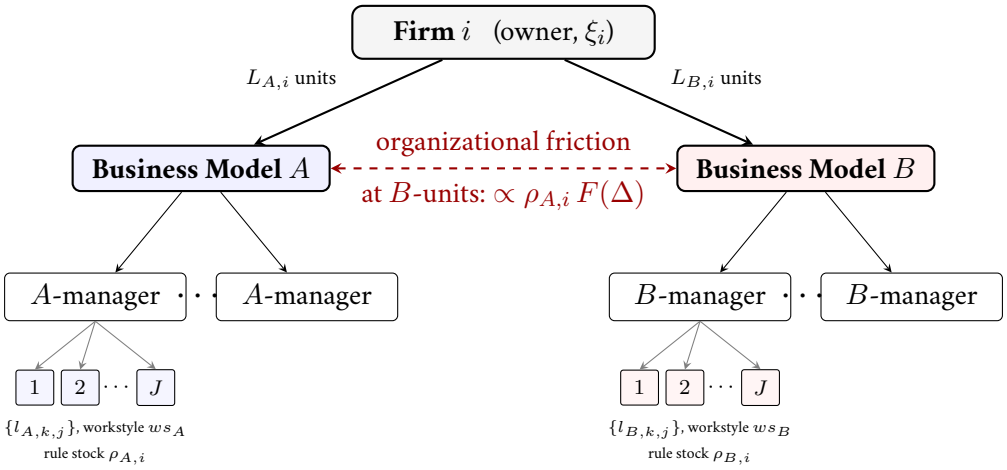


Figure 2. Model Structure

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.

a normalization.

Turning to managers, the firm owner chooses the number of units—that is, the number of managers—of each business model to operate, given the endogenous manager wage in business model X, denoted by $w_{X,t}$. 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 determined by market clearing 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 constant-elasticity-of-substitution (CES) output,

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

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:

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

Note, this is a static optimization with no intertemporal link, so the outcome in (3) is the same whoever is in charge of the unit, its manager or the owner.

Workstyle of business model A Recall ws_j denotes the workstyle of occupation j. We 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 for firm i.

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 $r \in [0, R]$, with $R \in \mathbb{R}^{+}$. The firm encounters problems at a constant rate $\mu^{p} > 0$, with the type r drawn uniformly from $[0, R]$ each time.

A rule for problem type r 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 we have used the fact that the firm has no rules $\rho_{A,i,t^i_A} = 0$ when born at t^i_A, and firm i's age is $a_{i,t} \equiv t - t^i_A$ at time t. Rule coverage rises monotonically with age, with the rate of accumulation slowing as the firm has already learned more.[7] Section III.D returns to this prediction empirically. The rule stock can be read as a form of organizational capital (Prescott and Visscher, 1980; Eisfeldt and Papanikolaou, 2013), one that accumulates as a byproduct of operations rather than through deliberate investment. In Section III.C, after showing that existing rules can turn into a liability once a new business model arrives, we ask what would change if the owner can constrain the length of the rulebook in anticipation.

Per-A-unit profit The expected per-unit downtime cost in business model A is $\mu^p\,\delta_p\,(1-\rho_{A,i,t})$, which is the product of the problem arrival rate, the probability the problem is uncovered, and the per-incident cost. Writing $\phi \equiv \mu^p\,\delta_p > 0$, the expected per-$A$-unit flow profit is

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

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 operational problems without improvisation. For simplicity, we assume $\pi_{A,t} > 0$ for all $\rho_{A,i,t} \in [0,1]$ and at every date.[8]

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:

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

[7] To understand (5), over $[t, t+dt]$ the firm experiences $\mu^p dt$ incidents, among which a fraction of $(1-\rho_{A,i,t})$ are new; so the growth of the rule-coverage fraction is $\frac{\mu^p}{R}(1-\rho_{A,i,t})dt$. Note that 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.

[8] In (6), the time subscript of $\pi_{A,t}$ captures the equilibrium A-manager wage $w_{A,t}$ embedded in $z_{A,t}$. In Appendix Section IA7.1 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 $L/\Xi > \phi$.

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,

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

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 derivation of business model A by clearing the industry-level A-manager market: $\int_i L^*_{A,i,t}\,di = L$, where L is the exogenous supply of A-managers.[9] Aggregating (8) across firms and using the fact that the spanning capacity ξ_i and firm age are independent, the equilibrium per-A-unit baseline profit (that is, before costs related to solving operational problems) is:

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

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, depressing 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.[10]

Substituting (9) back into (8), one can express firm i's 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 greater 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.

[9] 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.

[10] One can solve for the industry average rule-coverage fraction $\bar{\rho}_{A,t}$ under entry and exit dynamics for firms within an industry, but we do not pursue this extension here as it does not affect our subsequent mapping from the model's predictions to the data.

III.C Business Model B

A new business model B becomes available at a single, industry-wide Poisson time t_B. From this moment on, each firm may add B-units alongside its existing A-units. In what follows, firms take the arrival time t_B as given, and we normalize $t_B = 0$.

Manager's problem Each B-unit's manager solves the same within-unit problem as under A given by (2), now with the B-productivities $\{\psi_{B,j}\}$ in place of the A-productivities. The solution is again common to all units 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 being 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, increasingly so as the distance Δ grows, leading to productivity losses at type-B units. The microfoundation is as follows.

After business model B becomes available, B-units encounter operational problems at the same Poisson rate μ^p, drawn from the same set $[0, R]$ as under A, so that problems encountered under the two business models may overlap. As before, the firm builds a rulebook for its type-B units. Let $\rho_{B,i,t} \in [0, 1]$ denote the fraction of problem types 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) With exogenous probability $1 - F(\Delta)$ the A-rule does not conflict with the workstyle of business model B; we discuss the economics 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.[11]

(ii.b) With probability $F(\Delta)$ the A-rule conflicts with the workstyle of business model B. In this case, B-workers must complete two steps before they can operate. First, the entrenched A-rule must be set aside, which requires B-managers to escalate the conflict and spend time persuading A-managers that the existing rule should not apply. This incurs an *escalation cost* $\delta_c \equiv \chi\,\delta_p$, where $\chi > 0$ scales the cost relative to 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.

We highlight two aspects of the model. First, 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 $F(0) = 0$.

Second, the escalation cost breaks the property that A-rules always have net positive spillovers for B-units: a conflicting A-rule is not merely useless but, when $\chi > 0$, actively *costly*, because overriding it consumes managerial time. Conflicts are resolved at every B-unit, so per-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.

We are now ready to calculate the expected per-B-unit flow profit $\pi_{B,t}(\rho_A,\, \rho_B;\, \Delta)$. The expected downtime cost 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, we have:[12]

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

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

[11] 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.

[12] Just like $\pi_{A,t}$, the profit $\pi_{B,t}$ is time-varying, because i) the lengths of rulebooks (ρ_A and ρ_B) are time-varying, and ii) the equilibrium B-manager wage $w_{B,t}$ embedded in $z_{B,t}$ changes over time.

$\rho_A = 1$ effectively starts B with a complete rulebook. Rules can be translated from A to B without friction, and the productivity gap between B units of high-ρ_A and low-ρ_A firms is largest.

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

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

We can then write the marginal value of an additional unit of ρ_A for B-profit as

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

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 $F(\Delta) < 1$. The rule stock is thus putty ex ante and clay ex post (Johansen, 1959; Gilchrist and Williams, 2000): rules are molded to the problems the firm encounters under A, but cannot be reshaped when B arrives. Unlike installed capital, however, a conflicting rulebook is not merely obsolete: it lowers the productivity of the new business model, and the firm cannot readily dispose of it (Section III.D discusses whether insulating B from A can achieve this).

Should the owner cap the rulebook? Rules accumulate as a by-product of operations. An owner who foresees that B may arrive at a distance with $F(\Delta) > F^*$ could stop codification once the A-rule length reaches a cap $\bar{\rho}_A \in (0,1]$, trading a higher downtime cost $\phi(1-\rho_A)$ under A for a smaller liability $\phi(1-\rho_B)\rho_A\,[F(\Delta)/F^* - 1]$ once B arrives.[13] When rules are a net asset, $F(\Delta) < F^*$, the cap is $\bar{\rho}_A = 1$; when they are a net liability, it is interior and tighter the larger $F(\Delta)$ and the sooner B is expected. Firm i's A-rule length becomes $\min\{1 - e^{-\kappa a_i}, \bar{\rho}_A\}$, still weakly increasing in age, so the age gradient behind our predictions survives, flattened among the oldest firms. The owner has no reason to cap B's own rulebook: ρ_B enters π_B only through $(1-\rho_B)$, so every B-rule raises B-profit, and nothing exists yet for it to conflict with.

[13] If the arrival of B is an unanticipated "MIT shock," there is nothing to foresee: $\bar{\rho}_A = 1$ and the owner does not intervene. The cap matters only when B is anticipated with positive probability.

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

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

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 firm owner 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:

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

As a result, the size of business model A is given by (8) and that of business model B is given by

$$L^*_{B,i,t} \;=\; \pi_{B,t}(\rho_{A,i,t},\,\rho_{B,i,t};\Delta)\,\xi_i. \tag{18}$$

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 $d\rho_{B,i,t} = \kappa(1-\rho_{B,i,t})dt$. All firms start from no rules under B and the arrival of B is an industry-wide event, yielding the closed form $\rho_{B,i,t} = 1 - e^{-\kappa t}$.

Equilibrium wage of B-managers As before we derive $w_{B,t}$ by invoking the industry-level B-manager market-clearing condition $\int_i L^*_{B,i,t}\,di \;=\; L$. Using the firm-level first-order condition for $L^*_{B,i,t}$ and aggregating across firms, this yields the following equilibrium per-B-unit baseline profit:

$$z_{B,t} \;=\; \frac{L}{\Xi} \;+\; \phi\,(1-\bar{\rho}_{B,t})\left[1 \;-\; \left(1 - \frac{F(\Delta)}{F^*}\right)\,\bar{\rho}_{A,t}\right], \tag{19}$$

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 substituting this expression into the first-order condition for $L_{B,i,t}$ given in (18), with $\pi_{B,t}(\rho_{A,i,t},\,\rho_{B,i,t};\Delta)$ given in (13).

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; see (5). To support this observation, we collect information through employee reviews from Revelio (Revelio Labs, 2025*b*). 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 Section IA5 (which is also used in empirical analyses in Section V), and we calculate the intensity of rules in a GVKEY-year as:[14]

$$\text{Rule Index} = \frac{(\#\text{ of reviews mentioning the presence of rules - }\#\text{ of reviews mentioning the lack of rules})}{\#\text{ of total reviews}}. \tag{20}$$

Similarly, we use employee reviews to check whether old firms have more meetings, committees, and layers of approval that take time, with the detailed prompt given in Appendix Section IA3.2. First, we use an LLM to screen for reviews discussing meetings, committees, and layers of approval. Second, we use it to classify the flagged reviews as either complaining about having too many such activities or praising the company's efficiency. We calculate the intensity of meetings that waste time in a GVKEY-year as:

$$\text{Meeting Index} = \frac{(\#\text{ of reviews complaining too many - }\#\text{ of reviews praising not too many})}{\#\text{ of total reviews}}. \tag{21}$$

Figure 3 shows binscatter plots of the Rule Index and the Meeting Index against firm age. Both are strongly increasing with 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 (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 the B-division.

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

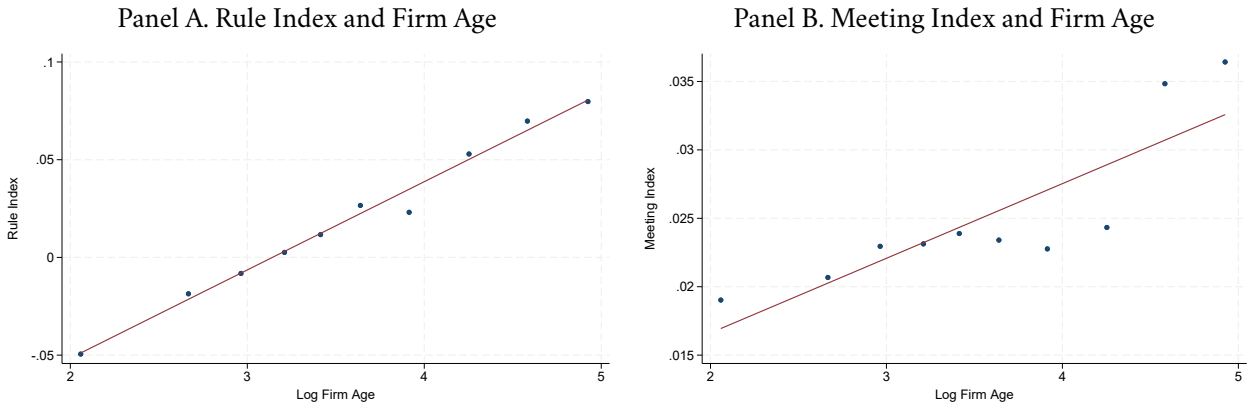


Figure 3. Firm Age, Rules, and Meetings that Waste Time

Notes: This figure shows binscatter plots of the Rule Index in (20) in Panel A and the Meeting Index in (21) 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.

The division-level conflict, often modeled as a game between headquarters and managers, lies at the heart of the incomplete contracts literature (Stein, 1997; Scharfstein and Stein, 2000; Rajan, Servaes, and Zingales, 2000; Hart and Moore, 2008), and is one we abstract from in our baseline analysis. 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; see Appendix Section IA7.2 for details. The per-B-unit profit under internal separation is the rule-free profit, net of this shading cost:

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

with the shading parameter $\theta > 0$ capturing the weakness of headquarters governance.

An incumbent that wants to implement B internally compares separation with integration, and adopts the former only when the conflict probability exceeds a threshold $F^{**} \geq F^{*}$, derived in Appendix Section IA7.2. Importantly, whichever structure it chooses, the division-level friction inherits the comparative statics of the baseline. That is, the friction worsens with workstyle distance and, because shading is more effective the larger the A-division's rule stock, bears down disproportionately on older firms:

$$\frac{\partial^2 \pi_B^{\text{sep}}}{\partial \rho_A\,\partial \Delta} \;=\; -\,\theta\,\phi\,(1-\rho_B)\,F'(\Delta) \;<\; 0, \tag{23}$$

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 predictions 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 $\pi_B(0, \rho_B)$. In positive terms, this sheds light on observed organizational design. Because operationally and financially separate businesses avoid both the worker-level conflict we model above and the headquarters shading of Appendix Section IA7.2, our theory 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).

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. We therefore stress that the boundary of the firm in our setting 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. 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:

$$g_{i,n} \equiv \frac{L^*_{A,i,0^+} + L^*_{B,i,0^+}}{L^*_{A,i,0^-}} - 1 = \frac{L^*_{B,i,0^+}}{L^*_{A,i,0^-}}, \tag{24}$$

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

Evaluating (19) at $t = 0^-$, and substituting into the firm's first order condition (18) for the number of B units at time $t = 0^+$, we obtain:

$$L^*_{B,i,0^+} \;=\; \left[\frac{L_n}{\Xi_n} \;+\; \phi\left(1 - \frac{F(\Delta_n)}{F^*}\right)\left(\rho_{A,i,0^-} \;-\; \bar{\rho}_{A,n,0^-}\right)\right]\xi_i. \tag{25}$$

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 in boosting B-unit productivity), the coefficient on $\rho_{A,i,0^-}$ is positive, implying that older firms open *more* B-units than younger firms in absolute terms. When $F(\Delta) > F^*$ (existing rules are a net liability as the escalation cost dominates), the coefficient turns negative, and older firms open *fewer* B-units.

However, older firms always *grow* more slowly in proportional terms once business model B requires a different workstyle. Young firms are small in both business models, having no rules to carry over. Without the friction, old firms are large in both A and B (their rules carry over perfectly), so proportional growth is the same regardless of firm age. The friction shrinks an old firm's B-side without touching its A-side: part of its knowledge conflicts with the B-workstyle and is lost, creating an asymmetry between its small B-units and large A ones. The resulting differential is thus an effect of the friction, scales with the workstyle distance, and vanishes when workstyles coincide. The following result makes this precise.

Result 1. *Consider the linear case in which the workstyle friction is $F(\Delta) = f\,\Delta$, where $f > 0$ is a primitive parameter capturing the degree of friction.*[15] *Then the 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.[16]

[15] 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.

[16] The factor $1/2$ is the employment share of B-units at $t = 0^+$, which equals $1/2$ under the normalization that the supply

Proof. See Appendix Section IA7. □

Three features of (26)–(27) matter for the empirical analysis. First, the spanning capacity ξ_i does not appear in $g_{i,n}$: in this baseline, impulse growth is independent of firm size. Second, the proportionality (27) justifies using the observable 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:

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

Equation (28), which follows from Equation (16), says that the gap in growth between firms with long and short rule books (equivalently, between old and young firms) widens with the workstyle distance. This prediction holds regardless of whether existing rules are a net asset ($F(\Delta) < F^*$) or a net liability ($F(\Delta) > F^*$). Intuitively, the marginal benefit of additional A-rules at B-units declines with Δ_n. As workstyle distance grows, the likelihood of conflict between A-rules and the workstyle of B workers erodes the substitutability between old and new rules that would otherwise benefit older firms.[17]

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 (24).[18] 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 Section IA7.

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.

of managers of types A and B is the same in each industry, $L_{A,n} = L_{B,n} \equiv L$.

[17] In the baseline, owners delegate occupation choice to unit managers, who maximize per-unit output and do not internalize the cross-unit friction. Appendix Section IA7.4 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.

[18] 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.

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

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

Proof. See Appendix Section IA7. □

The coefficient $\hat{\zeta}$ in (30) is the product of two terms. The first is the industry-level scalar $\Phi > 0$, which captures the strength of the rule–workstyle friction. The second is the firm-level covariance $\mathrm{cov}(\rho_{A,i,0^-}, a_i)$ between rule coverage and age. This covariance is also positive because older firms have systematically more A-rules. As a result, $\hat{\zeta} < 0$: older firms grow less in industries undergoing larger workstyle change. Hence, regression (29) provides a direct test of our core hypothesis: as workstyle distance grows, the value of legacy A-rules for B-units declines—whether or not they remain a net asset overall—so firms with longer rule histories grow relatively less.[19]

The contribution of young vs old firms to the implementation of new technologies In Section II.A, we highlighted how the empirical evidence on whether young vs old firms are more likely to bring new technologies to market is mixed: both contribute, though the young do so out of proportion to their share of sales or employment. We now discuss how the model speaks to this motivating fact.

In the model, let firm i's share of total B-output in industry n at the time of the arrival of B be $\sigma_{B,i,0^+}$. Since per-B-unit output is firm-invariant, the output share is equal to the managerial employment share. Appendix Section IA7.6 then shows that this share is given by:

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

Here, $\omega_B(\rho_A;\Delta)$ is a function given in Appendix Section IA7.6, with the following properties: it is increasing in ρ_A when $F(\Delta) < F^*$ and decreasing when $F(\Delta) > F^*$.

This result has two implications. First, conditional on age $(\rho_{A,i,0^-})$, larger firms (those with higher ξ_i) implement more of the new technology B in absolute terms. Second, conditional on size (ξ_i), the effect of age on market share could be positive or negative depending on whether existing rules are a net asset (when $F(\Delta) < F^*$) or a net liability ($F(\Delta) > F^*$). Thus in principle either old firms or young firms could be market leaders in the implementation of B, depending on how severe the workstyle conflict is (or

[19] Appendix Section IA7 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. Appendix Section IA7 also generalizes (29) with additional industry- and firm-level controls, leaving this interpretation of the interaction coefficient unchanged.

equivalently, depending on whether existing rules are a net asset or a net liability).

The pre-arrival A market-share admits a similar solution, reported in Appendix Section IA7.6. The ratio of firm i's B-share to its pre-arrival A-share is independent of ξ_i, and equal to:

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

Unlike the absolute market share ω_B, the relative, within-firm market share is *always strictly decreasing* in $\rho_{A,i,0^-}$ whenever $F(\Delta_n) > 0$, regardless of whether existing rules are a net asset or a net liability.[20] 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. Moreover, in industries with larger workstyle changes, young firms are more over-represented in their share of B output, relative to their share of A. This second 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.

IV Data and Measurement

Our empirical work builds an industry-year counterpart to the workstyle change $\left|d\overline{ws}\right|_n$ in the model, capturing 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 the industry-level predicted workstyle change as the Euclidean distance between the industry-level workstyle in year t and the one implied by the occupation composition that year-t technologies predict for year $t+5$. 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

[20] The proof is reported in Appendix Section IA7.6.

workstyle.[21] 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 from 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. O*NET classifies occupations using a slight variant of the Bureau of Labor Statistics' Standard Occupational Classification (SOC) codes.

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.[22] 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. These workstyle characteristics are also used by Guenzel et al. (2026) to measure the traits that each occupation demands, and their alignments with workers' personality. As we explain in detail below, we treat these workstyle characteristics at the *occupation level* as fixed. What changes over time is that new technologies may require different occupations to implement, which in turn shifts the overall workstyle in production.

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,

[21] 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.

[22] 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 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 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 distance between the two is the predicted workstyle change, our key variable 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, following 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. We then query LLM to label them as routine tasks vs nonroutine tasks.[23] 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 nonroutine for each occupation. The median occupation has 16 tasks and 33.3% are routine.[24]

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

[23] 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"}"

[24] 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.

embeddings. Specifically, the embedding of document i, denoted $\mathcal{X}_i$, 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,[25] and word embeddings x_k are obtained using the *GloVe* model.[26] 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.

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.

Sum over patents in industry n and year t to get occupation j's routine and nonroutine exposures. To obtain occupation j's total exposures (for both routine and nonroutine) in industry n at year t, we sum the similarities between occupation j and all patents granted 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).[27]

[25] 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.

[26] 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.

[27] 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.

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, in line with Kogan et al. (2024).

IV.C Predicted Workstyle Change

Finally, we combine the ingredients in Sections 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 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 that old firms 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 this dataset may reflect subsistence 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).[28]

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. An industry-level measure is also more plausibly exogenous to the growth of any individual firm: a firm-level measure of workstyle change would mechanically reflect the firm's own hiring, and hence its growth. Finally, measuring predicted workstyle change at the industry level is convenient, because shifts in employment by occupation due to new technologies are easiest to construct 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. Appendix Table IA1 reports summary statistics for the three samples.

V.A Venture Capital Investment

We start by 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

[28] 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.

retrieve a list of deals from Refinitiv's Private Equity database (Refinitiv, 2025*a*), and restrict to deals whose fund investor type is 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 VC investment value by three-digit NAICS industry 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.[29] 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 (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. A one standard deviation increase in $|d\overline{ws}|^e_{n,t,5}$ is associated with a roughly 0.4 log point higher ratio of VC investment to market capitalization. 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 $|d\overline{ws}|^e_{n,t,5}$. In column (2), we control for the log number of total patents in industry n and year t, showing that the sheer quantity of new technologies is not significantly related to the strength of

[29] The bootstrap addresses the inference problem that the generated regressor creates, and Appendix Section IA7.7 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.

Table 1 – Venture Capital Investment

	Forward 1-Year Log(VC Investment/Compustat Market Cap)				
	(1)	(2)	(3)	(4)	(5)
Predicted Workstyle Change in Industry	0.431***	0.427***	0.407***	0.418***	0.410***
	(0.094)	(0.097)	(0.092)	(0.094)	(0.095)
Log Patents in Industry		-0.035	0.068	0.031	0.062
		(0.074)	(0.085)	(0.098)	(0.101)
Patent Type	-	All	Breakthrough	Creative	RETech
Year FE	X	X	X	X	X
Observations	748	748	748	748	748
Within R^2	0.045	0.047	0.051	0.046	0.049

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 ($|\overline{dws}|^e_{n,t,5}$), 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 (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. 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 prior patents 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 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; 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 firm (based on its 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.[30]

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 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 $t+1$ (so that year-t technologies are within investors' information sets, as in 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).[31] We include year fixed effects, and bootstrap standard errors as explained earlier.

[30] 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.

[31] Appendix Section IA7.8 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.

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.A) 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 convert the coefficient on log firm age into the implied coefficient at each age for ease of illustration.

A natural question is whether the results on firm age reflect a correlation between age and size: perhaps 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 $|d\overline{ws}|^e_{n,t,5}$ and firm size tends to be positive. It could be possible that, all else equal, larger firms can have more financial resources that help them with adjustment and adaptation.[32]

[32]The baseline model does not generate an interaction between size and workstyle change, since the growth rate in (26) is independent of spanning capacity. There are two ways to deliver a positive interaction by relaxing some of the model's assumptions. The first is to introduce heterogeneous downtime costs δ_p, which measure the difficulty of solving problems but are uncorrelated with spanning capacity. The smaller the downtime cost, the larger the firm's business model A. But because the downtime cost drives our organizational friction, when B arrives firms with smaller δ_p suffer less rule–workstyle

Table 2 – Tobin's Q among Compustat Firms

	Tobin's Q				
	(1)	(2)	(3)	(4)	(5)
Predicted Workstyle Change in Industry × Log Age	-0.117***	-0.119***	-0.109***	-0.112***	-0.119***
	(0.045)	(0.039)	(0.033)	(0.035)	(0.037)
Predicted Workstyle Change in Industry × Log Size	0.033	0.035*	0.036**	0.036*	0.037*
	(0.019)	(0.018)	(0.017)	(0.017)	(0.019)
Predicted Workstyle Change in Industry	0.212	0.179	0.139	0.156	0.165
	(0.133)	(0.133)	(0.131)	(0.131)	(0.128)
Log Patents in Industry × Log Age		-0.010	-0.018	-0.022	-0.013
		(0.020)	(0.023)	(0.025)	(0.026)
Log Patents in Industry × Log Size		-0.005	-0.005	-0.006	-0.010
		(0.010)	(0.013)	(0.014)	(0.015)
Log Patents in Industry		0.096	0.131*	0.166*	0.169*
		(0.105)	(0.101)	(0.121)	(0.114)
Log Age	-0.266***	-0.206	-0.197**	-0.188**	-0.217***
	(0.060)	(0.119)	(0.086)	(0.087)	(0.075)
Log Size	-0.194***	-0.155**	-0.164***	-0.159***	-0.148***
	(0.031)	(0.055)	(0.045)	(0.045)	(0.040)
Patent Type	-	All	Breakthrough	Creative	RETech
Year FE	X	X	X	X	X
Observations	27,977	27,977	27,977	27,977	27,977
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.

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 the slower growth that we see in Tables 3 and

conflict—implying faster growth. Together, these effects generate a positive interaction between size and workstyle change. The second route is to assume 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; larger firms (with higher spanning capacities) therefore respond more positively to the shock, conditional on age.

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	-0.027**	-0.026**	-0.026***	-0.027***	-0.027***
	(0.009)	(0.009)	(0.008)	(0.008)	(0.008)
Predicted Workstyle Change in Industry × Log Size	0.011**	0.011**	0.009**	0.010**	0.010**
	(0.005)	(0.005)	(0.005)	(0.005)	(0.005)
Predicted Workstyle Change in Industry	0.021	0.021	0.028	0.025	0.026
	(0.031)	(0.033)	(0.031)	(0.034)	(0.034)
Log Patents in Industry × Log Age		0.004	0.002	0.004	0.004
		(0.006)	(0.007)	(0.008)	(0.007)
Log Patents in Industry × Log Size		-0.001	0.001	-0.000	-0.000
		(0.003)	(0.003)	(0.004)	(0.004)
Log Patents in Industry		-0.010	-0.017	-0.014	-0.010
		(0.019)	(0.020)	(0.024)	(0.022)
Log Age	-0.085***	-0.108***	-0.094***	-0.099***	-0.096***
	(0.014)	(0.032)	(0.026)	(0.026)	(0.022)
Log Size	-0.031***	-0.023*	-0.033***	-0.028***	-0.029***
	(0.010)	(0.012)	(0.011)	(0.010)	(0.008)
Patent Type	-	All	Breakthrough	Creative	RETech
Year FE	X	X	X	X	X
Observations	27,977	27,977	27,977	27,977	27,977
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.

4, 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

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	-0.025**	-0.021***	-0.021***	-0.021***	-0.021**
	(0.007)	(0.007)	(0.007)	(0.007)	(0.007)
Predicted Workstyle Change in Industry × Log Size	0.010***	0.008***	0.007***	0.007***	0.007***
	(0.003)	(0.002)	(0.002)	(0.002)	(0.002)
Predicted Workstyle Change in Industry	0.087**	0.078***	0.078***	0.078***	0.076***
	(0.024)	(0.024)	(0.024)	(0.024)	(0.025)
Log Patents in Industry × Log Age		0.000	-0.001	-0.001	-0.002
		(0.006)	(0.007)	(0.008)	(0.007)
Log Patents in Industry × Log Size		0.002	0.003	0.003	0.003
		(0.002)	(0.002)	(0.002)	(0.002)
Log Patents in Industry		-0.005	0.001	0.001	0.006
		(0.021)	(0.025)	(0.029)	(0.026)
Log Age	-0.086***	-0.088***	-0.085***	-0.084***	-0.082***
	(0.013)	(0.031)	(0.023)	(0.024)	(0.020)
Log Size	-0.011**	-0.020**	-0.021***	-0.020***	-0.020***
	(0.005)	(0.009)	(0.007)	(0.007)	(0.006)
Patent Type	-	All	Breakthrough	Creative	RETech
Year FE	X	X	X	X	X
Observations	27,977	27,977	27,977	27,977	27,977
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 ($|\overline{dws}|^e_{n,t,5}$), 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.

the following test on this hypothesis. Table IA2 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 acquisition 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 in which an observation is an industry-age group-year cell, indexed by (n, g, t):

$$\Delta \text{emp}_{n,g,t\to t+5} = \alpha_t + \sum_{g'\neq g_0} \zeta_{g'} \left(|d\overline{ws}|^e_{n,t,5} \times \mathbf{1}_{g=g'} \right) + \eta_w \ |d\overline{ws}|^e_{n,t,5} + \sum_{g'\neq g_0} \eta_{a,g'} \, \mathbf{1}_{g=g'} + \gamma \, z_{n,t} + \epsilon_{n,g,t}. \tag{43}$$

The outcome variable is the growth rate of employment among firms in age group g 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 ($|d\overline{ws}|^e_{n,t,5}$), and its interaction with age group indicators $\mathbf{1}\{g = g'\}$, where g' runs over the age groups 6-10, 11-15, and 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. In Table 5, we see that when $|d\overline{ws}|^e_{n,t,5}$ is high, employment growth is lower for older age groups, with the oldest group (16-20) growing significantly more slowly than the youngest (1-5).

V.D Additional Tests

We perform several additional analyses to address alternative mechanisms.

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.[33] Conceptually, this force differs from ours in two

[33]The same replacement force runs through the Schumpeterian quality-ladder models: an incumbent monopolist gains less from the next innovation than an outsider does, because that innovation would destroy its own monopoly rents, so in Grossman and Helpman (1991) and Aghion and Howitt (1992) the leading innovations are made by entrants (see also Reinganum, 1983), unless the incumbent can preempt them (Gilbert and Newbery, 1982). As Acemoglu and Cao (2015) put it, because of this replacement effect the baseline models predict that all innovation is undertaken by new firms; their extension lets incumbents improve their own products while entrants pursue more radical innovations.

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	0.008	0.008	0.007	0.007	0.007
	(0.010)	(0.010)	(0.010)	(0.010)	(0.010)
Predicted Workstyle Change in Industry × Age 11–15	-0.023	-0.023	-0.023	-0.023	-0.023
	(0.016)	(0.016)	(0.016)	(0.016)	(0.016)
Predicted Workstyle Change in Industry × Age 16–20	-0.024**	-0.023**	-0.024**	-0.023**	-0.024**
	(0.010)	(0.010)	(0.010)	(0.010)	(0.010)
Predicted Workstyle Change in Industry	0.035**	0.035**	0.035**	0.035**	0.034**
	(0.017)	(0.017)	(0.017)	(0.017)	(0.017)
Log Patents in Industry × Age 6–10		0.004	0.005	0.005	0.007
		(0.005)	(0.005)	(0.006)	(0.007)
Log Patents in Industry × Age 11–15		0.001	0.005	0.001	-0.002
		(0.006)	(0.008)	(0.009)	(0.009)
Log Patents in Industry × Age 16–20		0.002	0.005	0.001	0.003
		(0.006)	(0.008)	(0.009)	(0.009)
Log Patents in Industry		-0.005	-0.005	-0.002	0.003
		(0.009)	(0.011)	(0.012)	(0.012)
Age 6–10	0.032**	0.014	0.020	0.021	0.021
	(0.014)	(0.023)	(0.020)	(0.019)	(0.017)
Age 11–15	0.035**	0.030	0.025	0.034	0.040*
	(0.017)	(0.028)	(0.026)	(0.023)	(0.021)
Age 16–20	0.058***	0.048	0.047**	0.057**	0.053**
	(0.017)	(0.030)	(0.024)	(0.024)	(0.021)
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 $\Delta Y_{n,g,t\to t+5}$ is employment growth of age group g in industry n between year t and $t+5$, calculated following Davis, Haltiwanger, and Schuh (1992) as $\Delta Y_{n,g,t\to t+5} = (Y_{n,g,t+5} - Y_{n,g,t})/(0.5 \times Y_{n,g,t+5} + 0.5 \times Y_{n,g,t})$. 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 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 (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.

ways. First, the two forces operate on distinct margins when a new technology arrives: cannibalization weighs on incumbents with large current profits to protect, whereas the frictions we study weigh on firms whose new technologies require a different workstyle, whatever their profits. We check this separation empirically shortly. Second, cannibalization à la Arrow (1962) is stronger when new products are similar to, and hence closer substitutes for, existing ones, whereas our frictions are stronger when new technologies require a workstyle further from the old one; to the extent that product closeness and workstyle closeness go together, the two forces point in opposite directions.

Empirically, if cannibalization à la Arrow (1962) drove our results, firm age would matter only as a proxy for current profits: older firms have more profit to lose, so the effect of predicted workstyle change should depend on profitability, and the role of firm age should weaken once we allow for this dependence. In Appendix Table IA4, 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 the 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.

Knowledge versus implementation A common question is whether old firms are simply slower at learning new technologies, even without organizational frictions. Does larger predicted workstyle change simply capture technologies that are harder for old firms to learn? We offer two responses. First, slower learning by old firms would itself require some friction. Second, we test whether our findings can be explained by old firms falling behind young firms in invention: if so, the share of an industry's patents granted to young firms should rise with the predicted workstyle change. We obtain the founding year of each patent assignee from Ewens and Marx (2024), and calculate the share of patents granted to young firms (ten years old or younger). Appendix Table IA6 shows that the share of an industry's patents granted to young firms does not vary with the predicted workstyle change: the estimates are close to zero and statistically insignificant for all patents, as well as for breakthrough, creative, and RETech patents. In other words, when predicted workstyle change is larger, old firms are not necessarily worse at inventing and patenting. This pattern resembles the well-known example of Xerox's Palo Alto Research Center (PARC), which developed the graphical user interface as well as the Ethernet but was not able to scale or commercialize them (except for the laser printer).

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 Appendix Table IA3, we check that our results on predicted workstyle change are not driven by this variable being correlated with uncertainty. We measure

uncertainty using the World Uncertainty Index, as in 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.

CEO age and employee age Finally, a common question is whether the implementation of new technologies depends on CEO or employee age, not just firm age. For example, Acemoglu, Akcigit, and Celik (2022) find that younger CEOs and inventors are more likely to produce creative innovations, and relate these patterns to the creative destruction process among firms. To test the role of firm age vs CEO and employee age, we obtain CEO age from BoardEx (BoardEx, 2025), which records executives' birth dates, and we estimate the weighted average age of employees using Revelio profiles (Revelio Labs, 2025*a*), as explained in more detail in Appendix Section IA6. In the data, log firm age is mildly positively correlated with log CEO age and slightly negatively correlated with log average employee age. When we include predicted workstyle change interacted with log CEO age and with log average employee age in Appendix Table IA5, we find that firm age appears to matter the most. These results suggest that frictions in implementing new technologies arise more from rigidity that builds up in the firm over time, rather than older people being outdated or incapable of catching up with new technologies.

VI 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 uniformly incapable of implementing new technologies, as they would be if slowness of learning or fear of cannibalization were the driving force. However, they are especially vulnerable when new technologies require changes in organizational processes and priorities, which favor new organizations. With advances in data and measurement—in particular new ways to extract information from text—future research can make further progress in capturing organizational frictions and investigating their role in innovation, productivity, and growth.

References

Acemoglu, Daron, and Dan Cao. 2015. "Innovation by entrants and incumbents." *Journal of Economic Theory*, 157: 255–294.

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, and Murat Alp Celik. 2022. "Radical and incremental innovation: The roles of firms, managers, and innovators." *American Economic Journal: Macroeconomics*, 14(3): 199–249.

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*, 60(2): 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. 1962. "Economic welfare and the allocation of resources for invention." In *The Rate and Direction of Inventive Activity: Economic and Social Factors*. 609–626. Princeton University Press.

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

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

Atkeson, Andrew, and Patrick J Kehoe. 1999. "Models of energy use: Putty-putty versus putty-clay." *American Economic Review*, 89(4): 1028–1043.

Atkeson, Andrew, and Patrick J Kehoe. 2005. "Modeling and measuring organization capital." *Journal of Political Economy*, 113(5): 1026–1053.

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.

BoardEx. 2025. "BoardEx database of board and executive profiles." https://wrds-www.wharton.upenn.edu/pages/get-data/boardex/ Accessed November 6, 2025.

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. 2026. "The ESG-innovation disconnect: Evidence from green patenting." *Management Science*. Forthcoming.

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.

Eisfeldt, Andrea L, and Dimitris Papanikolaou. 2013. "Organization capital and the cross-section of expected returns." *The Journal of Finance*, 68(4): 1365–1406.

Eisfeldt, Andrea L, and Dimitris Papanikolaou. 2014. "The value and ownership of intangible capital." *American Economic Review*, 104(5): 189–194.

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.

Gilbert, Richard J, and David M G Newbery. 1982. "Preemptive patenting and the persistence of monopoly." *American Economic Review*, 72(3): 514–526.

Gilchrist, Simon, and John C Williams. 2000. "Putty-clay and investment: A business cycle analysis." *Journal of Political Economy*, 108(5): 928–960.

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, Gene M, and Elhanan Helpman. 1991. "Quality ladders in the theory of growth." *Review of Economic Studies*, 58(1): 43–61.

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.

Guenzel, Marius, Shimon Kogan, Marina Niessner, and Kelly Shue. 2026. "AI personality extraction from faces: Labor market implications." Working Paper.

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.

Johansen, Leif. 1959. "Substitution versus fixed production coefficients in the theory of economic growth: A synthesis." *Econometrica*, 27(2): 157–176.

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.

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.

Lev, Baruch, and Suresh Radhakrishnan. 2005. "The valuation of organization capital." In *Measuring Capital in the New Economy*. , ed. Carol Corrado, John Haltiwanger and Daniel Sichel, 73–110. University of Chicago Press.

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.

Ma, Yueyuan, and Shaoshuang Yang. 2026. "Technology-driven market concentration through idea allocation." Working Paper.

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.

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.

Prescott, Edward C, and Michael Visscher. 1980. "Organization capital." *Journal of Political Economy*, 88(3): 446–461.

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, 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.

Reinganum, Jennifer F. 1983. "Uncertain innovation and the persistence of monopoly." *American Economic Review*, 73(4): 741–748.

Revelio Labs. 2025*a*. "Individual profiles dataset." https://wrds-www.wharton.upenn.edu/pages/get-data/revelio-labs/revelio-individual/ Accessed November 6, 2025.

Revelio Labs. 2025*b*. "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. 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

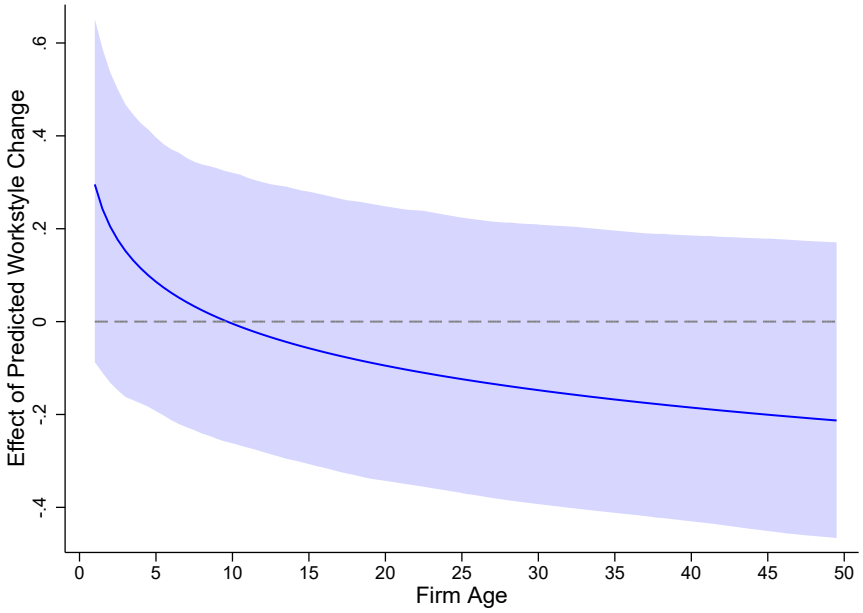


Figure IA1. Tobin's Q among Compustat Firms: Implied Coefficient on $|\overline{dws}|^{e}_{n,t,5}$

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).

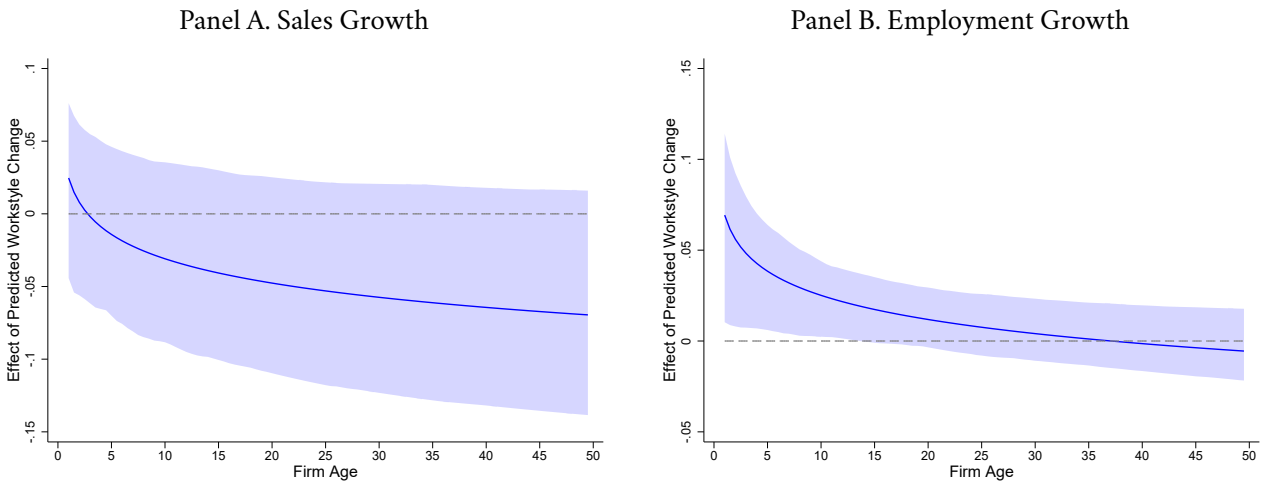


Figure IA2. Sales and Employment Growth among Compustat Firms: Implied Coefficient on $|d\overline{ws}|^{e}_{n,t,5}$

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	-0.15	-0.00	0.30	0.34	1.03	748
Log VC Investment in Industry	16.75	18.15	19.70	18.18	2.29	748
Log Compustat Market Cap in Industry	10.29	11.37	12.60	11.40	1.51	748
Log(VC Investment/Market Cap)	5.60	7.01	8.23	6.78	2.05	748
Log Total Patents in Industry	2.81	5.07	7.16	4.85	2.73	748
Log Breakthrough Patents in Industry	0.37	2.66	4.61	2.73	2.32	748
Log Creative Patents in Industry	0.53	2.35	4.03	2.44	2.01	748
Log RETech Patents in Industry	0.10	1.34	3.44	1.98	2.00	748
Panel B: Compustat Sample						
Predicted Workstyle Change in Industry	-0.07	0.21	0.80	0.56	1.02	27,977
Log Firm Age	2.89	3.47	4.19	3.51	0.87	27,977
Firm Tobin's Q	0.91	1.33	2.15	1.91	2.09	27,977
Sales Growth Next 5 Years	-0.03	0.25	0.58	0.27	0.56	27,977
Employment Growth Next 5 Years	-0.11	0.14	0.44	0.16	0.48	27,977
Log Firm Assets	4.52	6.13	7.65	6.08	2.26	27,977
Log Firm Sales	4.30	6.09	7.56	5.85	2.48	27,977
Log Firm Employment	-1.31	0.45	1.93	0.33	2.25	27,977
Log Total Patents in Industry	3.81	7.60	8.81	6.33	3.21	27,977
Log Breakthrough Patents in Industry	1.15	4.86	6.31	4.25	2.91	27,977
Log Creative Patents in Industry	1.22	4.63	5.76	3.84	2.54	27,977
Log RETech Patents in Industry	0.78	3.63	5.66	3.43	2.60	27,977
Panel C: BDS Sample						
Predicted Workstyle Change in Industry	-0.14	-0.00	0.30	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 – Acquisition Growth among Compustat Firms

	Acquisition Growth (5-Year DHS Rate)				
	(1)	(2)	(3)	(4)	(5)
Predicted Workstyle Change in Industry × Log Age	-0.025	-0.030	-0.032*	-0.031	-0.031*
	(0.018)	(0.018)	(0.019)	(0.018)	(0.019)
Predicted Workstyle Change in Industry × Log Size	-0.000	-0.000	-0.001	-0.001	-0.000
	(0.010)	(0.011)	(0.012)	(0.011)	(0.011)
Predicted Workstyle Change in Industry	0.107	0.115	0.130	0.122	0.123
	(0.094)	(0.101)	(0.100)	(0.099)	(0.099)
Log Patents in Industry × Log Age		0.013	0.012	0.014	0.017
		(0.009)	(0.012)	(0.013)	(0.014)
Log Patents in Industry × Log Size		0.001	0.003	0.002	0.002
		(0.003)	(0.003)	(0.004)	(0.004)
Log Patents in Industry		-0.049	-0.053	-0.057	-0.065
		(0.035)	(0.043)	(0.046)	(0.050)
Log Age	0.072***	-0.010	0.027	0.023	0.021
	(0.026)	(0.055)	(0.044)	(0.046)	(0.040)
Log Size	-0.070***	-0.075***	-0.081***	-0.079***	-0.077***
	(0.012)	(0.018)	(0.014)	(0.015)	(0.013)
Patent Type	-	All	Breakthrough	Creative	RETech
Year FE	X	X	X	X	X
Observations	14,765	14,765	14,765	14,765	14,765
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 IA3 – Robustness Check: Controlling for Alternative Measures of Uncertainty

	Tobin's Q			Sales Growth			Employment Growth		
Uncertainty Measure	WUI	EPU	Return Vol	WUI	EPU	Return Vol	WUI	EPU	Return Vol
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Predicted Workstyle Change in Industry × Log Age	-0.157***	-0.157***	-0.151**	-0.026**	-0.026**	-0.021**	-0.022**	-0.022**	-0.019**
	(0.045)	(0.045)	(0.062)	(0.009)	(0.009)	(0.008)	(0.007)	(0.007)	(0.008)
Predicted Workstyle Change in Industry × Log Size	0.046*	0.043*	0.027*	0.011*	0.011*	0.009	0.009**	0.009**	0.007
	(0.020)	(0.019)	(0.016)	(0.005)	(0.005)	(0.005)	(0.004)	(0.004)	(0.004)
Predicted Workstyle Change in Industry	0.282*	0.293*	0.385*	0.018	0.017	0.007	0.078**	0.078**	0.075***
	(0.174)	(0.175)	(0.229)	(0.032)	(0.031)	(0.035)	(0.026)	(0.026)	(0.026)
Uncertainty × Log Age	-0.050	0.002**	-0.154	0.026	0.001***	0.071**	0.011	0.000	0.053**
	(0.136)	(0.001)	(0.157)	(0.024)	(0.000)	(0.029)	(0.025)	(0.000)	(0.030)
Uncertainty × Log Size	-0.099	-0.001**	-0.141***	0.007	0.000	-0.047**	-0.015	0.000	-0.015
	(0.073)	(0.001)	(0.047)	(0.011)	(0.000)	(0.016)	(0.009)	(0.000)	(0.013)
Uncertainty	1.014**	0.004	0.606	-0.145	-0.003***	-0.256***	-0.050	-0.001	-0.369***
	(0.566)	(0.005)	(0.593)	(0.092)	(0.001)	(0.107)	(0.095)	(0.001)	(0.123)
Log Age	-0.245***	-0.509***	-0.145	-0.102***	-0.167***	-0.127***	-0.093***	-0.141***	-0.110***
	(0.066)	(0.131)	(0.110)	(0.018)	(0.032)	(0.022)	(0.015)	(0.037)	(0.022)
Log Size	-0.174***	-0.039	-0.109**	-0.027**	-0.047**	-0.024*	-0.006	-0.018	-0.015
	(0.034)	(0.064)	(0.043)	(0.012)	(0.017)	(0.013)	(0.006)	(0.014)	(0.010)
Year FE	X	X	X	X	X	X	X	X	X
Observations	27,978	27,978	16,811	27,978	27,978	16,811	27,978	27,978	16,811
Within R^2	0.092	0.092	0.117	0.077	0.077	0.095	0.048	0.048	0.049

Notes: This table presents regressions at the firm-year level in Equations (41) and (42), controlling for alternative measures of uncertainty. The outcome variables are Tobin's Q in columns (1) to (3), sales growth in columns (4) to (6), and employment growth in columns (7) to (9), where sales growth and employment growth between year t and $t+5$ are 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 book assets in year t). We standardize $|\overline{dws}|^{e}_{n,t,5}$ to unit variance and zero median. Columns (1), (4), and (7) control for the World Uncertainty Index (WUI), columns (2), (5), and (8) control for the Economic Policy Uncertainty Index (EPU) from Baker, Bloom, and Davis (2016), and columns (3), (6), and (9) control for firm-level stock return volatility (Return Vol), as well as their 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 IA4 – Robustness Check: Controlling for Alternative Measures of Profitability

	Tobin's Q			Sales Growth			Employment Growth		
Profitability Measure	NI/Assets	NI/Sales	E/P	NI/Assets	NI/Sales	E/P	NI/Assets	NI/Sales	E/P
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Predicted Workstyle Change in Industry × Log Age	-0.122**	-0.134***	-0.160***	-0.025**	-0.024**	-0.023*	-0.024**	-0.023**	-0.019**
	(0.042)	(0.040)	(0.045)	(0.008)	(0.008)	(0.009)	(0.007)	(0.007)	(0.007)
Predicted Workstyle Change in Industry × Log Size	0.047**	0.047*	0.038	0.010*	0.009*	0.010*	0.008*	0.009**	0.008*
	(0.017)	(0.020)	(0.020)	(0.005)	(0.004)	(0.005)	(0.003)	(0.003)	(0.004)
Predicted Workstyle Change in Industry × Profitability	-0.192	-0.033	0.067	0.019	0.003	0.001	0.047	0.000	0.004
	(0.300)	(0.040)	(0.063)	(0.039)	(0.006)	(0.013)	(0.034)	(0.004)	(0.020)
Predicted Workstyle Change in Industry	0.165	0.196	0.342*	0.015	0.030	0.011	0.085***	0.080**	0.070**
	(0.164)	(0.187)	(0.189)	(0.033)	(0.028)	(0.034)	(0.025)	(0.025)	(0.025)
Log Age	-0.190***	-0.234***	-0.275***	-0.092***	-0.099***	-0.099***	-0.089***	-0.090***	-0.094***
	(0.062)	(0.075)	(0.074)	(0.018)	(0.017)	(0.017)	(0.014)	(0.014)	(0.014)
Log Size	-0.114***	-0.165***	-0.196***	-0.019*	-0.006	-0.026**	-0.010*	-0.011*	-0.012*
	(0.024)	(0.034)	(0.036)	(0.010)	(0.007)	(0.013)	(0.006)	(0.006)	(0.006)
Profitability	-1.034***	-0.048***	0.309***	-0.107*	-0.023**	0.094***	0.014	0.003*	0.154***
	(0.248)	(0.069)	(0.055)	(0.038)	(0.007)	(0.023)	(0.059)	(0.008)	(0.024)
Year FE	X	X	X	X	X	X	X	X	X
Observations	27,084	27,204	27,007	27,084	27,204	27,007	27,084	27,204	27,007
Within R^2	0.143	0.108	0.098	0.078	0.075	0.076	0.050	0.050	0.057

Notes: This table presents regressions at the firm-year level in Equations (41) and (42), controlling for alternative measures of current profitability to address concerns that the estimated effects may reflect cannibalization of firms' existing business activities. The outcome variables are Tobin's Q in columns (1) to (3), sales growth in columns (4) to (6), and employment growth in columns (7) to (9), where sales growth and employment growth between year t and $t+5$ are 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 book assets in year t). We standardize $|\overline{dws}|^{e}_{n,t,5}$ to unit variance and zero median. Columns (1), (4), and (7) control for profitability measured as net income divided by assets (NI/Assets), columns (2), (5), and (8) use net income divided by sales (NI/Sales), and columns (3), (6), and (9) use the earnings-to-price ratio (E/P); each specification includes the corresponding profitability measure and its interaction with the predicted workstyle change. 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 IA5 – Robustness Check: Controlling for CEO Age and Employee Age

	Tobin's Q		Sales Growth		Employment Growth	
	(1)	(2)	(3)	(4)	(5)	(6)
Predicted Workstyle Change in Industry × Log Firm Age	-0.147**	-0.086**	-0.021*	-0.017**	-0.016	-0.015**
	(0.055)	(0.033)	(0.009)	(0.007)	(0.008)	(0.006)
Predicted Workstyle Change in Industry × Log CEO Age	0.094		0.006		-0.011	
	(0.184)		(0.056)		(0.034)	
Predicted Workstyle Change in Industry × Log Employee Age		-0.205		-0.179***		-0.149**
		(0.302)		(0.065)		(0.050)
Predicted Workstyle Change in Industry × Log Firm Size	0.033*	0.030	0.011**	0.011**	0.008**	0.009***
	(0.015)	(0.017)	(0.004)	(0.004)	(0.003)	(0.002)
Predicted Workstyle Change in Industry	-0.004	0.886	-0.015	0.620**	0.099	0.588***
	(0.684)	(1.012)	(0.219)	(0.232)	(0.132)	(0.178)
Log Firm Age	-0.245***	-0.208***	-0.086***	-0.055***	-0.083***	-0.051***
	(0.076)	(0.055)	(0.016)	(0.012)	(0.013)	(0.011)
Log CEO Age	-0.735***		-0.294***		-0.240***	
	(0.235)		(0.082)		(0.057)	
Log Employee Age		-0.362		-0.381***		-0.428***
		(0.608)		(0.106)		(0.085)
Log Firm Size	-0.111***	-0.202***	-0.027*	-0.047***	-0.012	-0.032***
	(0.025)	(0.030)	(0.015)	(0.010)	(0.008)	(0.005)
Year FE	X	X	X	X	X	X
Observations	23,739	27,683	23,739	27,683	23,739	27,683
Within R^2	0.073	0.076	0.064	0.065	0.052	0.047

Notes: This table presents regressions at the firm-year level in Equations (41) and (42), controlling for CEO age and average employee age. The outcome variables are Tobin's Q in columns (1) and (2), sales growth in columns (3) and (4), and employment growth in columns (5) and (6), where sales growth and employment growth between year t and $t+5$ are 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. We standardize $|\overline{dws}|^{e}_{n,t,5}$ to unit variance and zero median. Columns (1), (3), and (5) control for log CEO age and its interaction with the predicted workstyle change, whereas columns (2), (4), and (6) control for log average employee age and its interaction with the predicted workstyle change. 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 IA6 – Young Firms' Share of Patenting and Predicted Workstyle Change

	All patents	Breakthrough patents	Creative patents	RETech patents
Predicted Workstyle Change	0.001	-0.003	-0.013	-0.004
	(0.006)	(0.009)	(0.009)	(0.010)
Year FE	X	X	X	X
Observations	1,019	968	986	886
Within R^2	0.000	0.000	0.004	0.000

Notes: This table presents regressions examining the relationship between the share of patents attributed to young firms and predicted industry-level workstyle change. The dependent variable is the young-firm share of patents in each category. Columns (1)–(4) consider all patents, breakthrough patents, creative patents, and RETech patents, respectively. The independent variable is predicted workstyle change in industry n over the next five years due to year t technologies ($|\widetilde{dws}|^{e}_{n,t,5}$), constructed as in Equation (39). All specifications include year fixed effects. Standard errors from 1,000 bootstrap replications are reported in parentheses. Asterisks denote significance levels (***=1%, **=5%, *=10%).

IA2 Technology Implementation and Commercialization

IA2.1 Important Inventions in Kelly et al. (2021)

Kelly et al. (2021) provide a list of 245 significant U.S. inventions using patents granted between 1840 and 2002 (their Appendix Table A1). For each of these inventions in their list, we prompt Gemini-2.5-Flash, with Google Search grounding, to identify the firm that invented it and the firm most successful in its initial implementation and commercialization, and to classify each as a young firm (founded less than ten years earlier) or an old firm. We supply the invention's name together with a short identifier—its named inventor, U.S. patent number, and grant year—and instruct the model to verify firm details through web search. The prompt is reproduced below.

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

Here is a technological invention. Your task is to determine:
  A) The company that invented it.
  B) The company that was most successful in its initial implementation and
  commercialization.
Then classify each as a young firm or an old firm.

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

### A. Inventor

- Return the name of the company that originally invented the technology.
- Always use the name of the parent company and do not use the name of the
division or subsidiary.
- Use the company's name if the technology was invented by specific employees
working at the company.
- If multiple companies were involved, return the primary company most widely
credited.
- If the inventor was a university, government lab, or individual (not a
company), return that name and set inventor_firm_type = 2.
- If unclear, return "Unknown".
```

B. Implementer

- 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 was never commercialized, return "Never Commercialized" with implementing_firm_type = 2.
- If unclear, return "Unknown".

C. Firm type classification

For each company, determine whether it was young or old at the time of the invention or implementation, based on the age of the parent company.

  1 = Young firm (founded <10 years ago at that time)
  -1 = Old firm (founded >=10 years ago at that time)
  0 = Ambiguous / unknown
  2 = Not a company (university, government, individual) / Never commercialized (implementer only)

D. Consistency rules

1. Same entity: If the inventor and implementer are the same company, or are parent and subsidiary (e.g., Google / Alphabet), they MUST receive the same firm name. Set same_entity = true.

2. Same-entity firm type: When same_entity = true, implementing_firm_type should equal inventor_firm_type unless the difference between implementing_year and inventing_year justifies a change (e.g., the firm was young at invention but crossed the 10-year threshold by the time of implementation).

3. Non-company inventors: If the inventor is a university, government agency,

```
   nonprofit, or individual, set inventor_firm_type = 2.

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

Technology: {invention}

Inventor Type:
'''Named inventor: {inventor}. US Patent {patent}, granted {year}.'''

Description:
'''US Patent {patent} ({year}) - {invention}, named inventor {inventor}.'''

------------------------------------------------
OUTPUT FORMAT (JSON only, no other text)
------------------------------------------------

{
  "technology": "{invention}",
  "inventor_name": "Company name or 'Unknown'",
  "inventor_firm_type": 1 or -1 or 0 or 2,
  "inventor_firm_type_reason": "Short explanation (< 30 words)",
  "inventing_year": "YYYY or 'Unknown'",
  "implementing_firm_name": "Company name or 'Never Commercialized'",
  "implementing_firm_type": 1 or -1 or 0 or 2,
  "implementing_firm_type_reason": "Short explanation (< 30 words)",
  "implementing_year": "YYYY or 'Unknown'",
  "same_entity": true or false
}

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

Example 1:
{
  "technology": "iPhone",
```

```
  "inventor_name": "Apple",
  "inventor_firm_type": -1,
  "inventor_firm_type_reason": "Apple founded in 1976; iPhone launched in
  2007.",
  "inventing_year": "2007",
  "implementing_firm_name": "Apple",
  "implementing_firm_type": -1,
  "implementing_firm_type_reason": "Apple founded in 1976; iPhone launched in
  2007.",
  "implementing_year": "2007",
  "same_entity": true
}

Example 2:
{
  "technology": "Dropbox",
  "inventor_name": "Dropbox",
  "inventor_firm_type": 1,
  "inventor_firm_type_reason": "Dropbox founded in 2007; product developed
  immediately.",
  "inventing_year": "2007",
  "implementing_firm_name": "Dropbox",
  "implementing_firm_type": 1,
  "implementing_firm_type_reason": "Dropbox founded in 2007; launched publicly
  in 2008.",
  "implementing_year": "2008",
  "same_entity": true
}

Example 3:
{
  "technology": "Ethernet",
  "inventor_name": "Xerox",
  "inventor_firm_type": -1,
  "inventor_firm_type_reason": "Xerox PARC, part of Xerox (founded in 1906),
  developed Ethernet in 1973",
  "inventing_year": "1973",
  "implementing_firm_name": "3Com",
```

```
  "implementing_firm_type": 1,
  "implementing_firm_type_reason": "3Com founded 1979; sold Ethernet adapters
  by 1981.",
  "implementing_year": "1981",
  "same_entity": false
}

Example 4:
{
  "technology": "World Wide Web",
  "inventor_name": "CERN",
  "inventor_firm_type": 2,
  "inventor_firm_type_reason": "CERN is an intergovernmental research
  organization, not a private firm.",
  "inventing_year": "1989",
  "implementing_firm_name": "Netscape",
  "implementing_firm_type": 1,
  "implementing_firm_type_reason": "Netscape founded in 1994; launched
  Navigator browser in 1994.",
  "implementing_year": "1994",
  "same_entity": false
}

Example 5:
{
  "technology": "Google Glass",
  "inventor_name": "Google",
  "inventor_firm_type": -1,
  "inventor_firm_type_reason": "Google founded in 1998; Glass released in
  2012.",
  "inventing_year": "2012",
  "implementing_firm_name": "Never Commercialized",
  "implementing_firm_type": 2,
  "implementing_firm_type_reason": "Consumer edition cancelled in 2015; never
  achieved mass-market commercialization.",
  "implementing_year": "Unknown",
  "same_entity": false
}
```

```
Example 6:
{
  "technology": "Fusion ignition",
  "inventor_name": "Lawrence Livermore National Laboratory",
  "inventor_firm_type": 2,
  "inventor_firm_type_reason": "LLNL is a US Department of Energy national
  laboratory, not a private firm.",
  "inventing_year": "2021",
  "implementing_firm_name": "Never Commercialized",
  "implementing_firm_type": 2,
  "implementing_firm_type_reason": "Commercial fusion power remains unrealized
  despite achieving ignition.",
  "implementing_year": "Unknown",
  "same_entity": false
}
```

IA2.2 Important New Technologies in Kalyani et al. (2025)

We next apply the same classification procedure to the set of important new technologies in Kalyani et al. (2025). Unlike the Kelly et al. (2021) sample, these technologies are not organized around individual patent records. We therefore first match each technology title to its corresponding Wikipedia article, standardizing capitalization, spaces, underscores, and punctuation in titles. This yields 1,286 matched technology articles.

For each technology, we first construct background information from the Wikipedia article, including when the technology first became workable, where it was first created, and the broad type of inventor.

```
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

```
  ran by engineers from private companies, so it should be classified as
  "Private Company". W3C (World Wide Web Consortium) is ran 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.

You will receive one JSON object containing the technology name and Wikipedia
page text. Return exactly one JSON object matching the provided schema. Do not
return markdown, comments, citations, or extra keys. Preserve `page_id` and
`canonical_title` exactly as supplied.

Set `answer_source` to `wikipedia` or `domain_knowledge`. Use `wikipedia` if
the key
answers are sourced from the supplied Wikipedia text. Use `domain_knowledge`
if
any key answer is approximated from your domain knowledge because the
Wikipedia
text is missing, ambiguous, or incomplete.

Return fields:
{
```

```
  "page_id": input page_id,
  "canonical_title": input canonical_title,
  "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 inventor categories,
  "who_reason": "Brief reason for chosen category",
  "domain_of_application": one of the 15 listed categories,
  "domain_reason": "Brief reason for chosen domain",
  "answer_source": "wikipedia" or "domain_knowledge"
}

Example Output #1:
{
  "page_id": 123,
  "canonical_title": "IPhone",
  "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.",
  "answer_source": "wikipedia"
}

Example Output #2:
{
  "page_id": 456,
  "canonical_title": "CRISPR_gene_editing",
  "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/Non-Profit",
  "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.",
  "answer_source": "wikipedia"
}
```

We then classify the inventor and the firm most successful in the technology's initial implementation and commercialization using Gemini-2.5-Flash with Google Search grounding. The prompt is the same as in the previous subsection, except that the input contains the technology name, the inventor type inferred from the background-information step, and the Wikipedia article text rather than a patent number, grant year, and named inventor.

IA3 Processing Employee Reviews

We collect information through employee reviews from Revelio (Revelio Labs, 2025*b*). 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 the 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 in Section IV.B. The overall approach follows Kogan et al. (2024): we represent each patent's text and each occupation's task descriptions as vectors in embedding space, and compute cosine similarity between them.

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).

IA5 Measuring the Age of Compustat Firms

For the analyses in Sections III.D and V, 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 Estimating Firm Employee Age from Revelio

This appendix describes how we construct employee birth years and firm-level average employee age measures from the Revelio profiles (Revelio Labs, 2025*a*). The procedure has two steps. First, we infer or impute a user-level birth year for individuals in Revelio associated with Compustat firms. Second, we

merge these user-level birth years into the annual worker-firm panel and aggregate worker ages to the firm-year level.

IA6.1 User-Level Birth Year Construction

We begin with the Revelio position-level data matched to education records. The education data contain, when available, a user's degree type and education start and end dates. We restrict the sample to users who are ever associated with a Compustat-linked firm between 2003 and 2019. Specifically, a user is included if at least one of her positions has an `rcid` that can be matched to a Compustat `gvkey` through the parent-subsidiary bridge in the same calendar year between 2003 and 2019.

For users with non-missing education information, we infer birth year from the education end year and the type of degree. Let e_i denote the education end year for user i. We map degree types into typical graduation ages as follows:

$$a(d_i) = \begin{cases} 18, & \text{high school,} \\ 20, & \text{associate degree,} \\ 22, & \text{bachelor's degree or missing/other degree,} \\ 24, & \text{master's degree or MBA,} \\ 28, & \text{doctoral degree or PhD.} \end{cases}$$

For each education record, we compute an implied birth year

$$\widetilde{b}_{ir} = e_{ir} - a(d_{ir}).$$

We retain education end years between 1940 and 2035. When a user has multiple valid education records, we define the observed user-level birth year as the rounded median of the implied birth years:

$$b_i^{obs} = \text{round}\left(\text{median}_r\{\widetilde{b}_{ir}\}\right).$$

This produces an education-based birth year for users with valid education records.

For users without an observed education-based birth year, we impute birth year using a linear model estimated on users with non-missing b_i^{obs}. The dependent variable is b_i^{obs}. We use a high-dimensional OLS specification with career-history controls and fixed effects. The continuous controls are average seniority, first observed start salary, and first observed end salary. The fixed effects include first and last observed calendar year, binned career span, binned number of positions, binned number of firms, binned number of parent firms, binned number of O*NET occupations, binned number of roles, first and last seniority, maximum seniority bin, first and last O*NET occupation, first and last role, first parent firm, salary bin, and first-position end-status. Rare fixed-effect categories are pooled using a minimum training-cell threshold

of 100.

Formally, the imputation model is

$$b_i^{obs} = X_i'\beta + \alpha_{\text{first year}(i)} + \alpha_{\text{last year}(i)} + \alpha_{\text{career bins}(i)} + \alpha_{\text{seniority}(i)} + \alpha_{\text{occupation}(i)} + \alpha_{\text{role}(i)} + \alpha_{\text{parent firm}(i)} + \varepsilon_i,$$

where X_i contains the continuous salary and seniority controls. The model is estimated on two million users with observed birth years and evaluated on a held-out sample of 500,000 users. In the final specification used for imputation, the out-of-sample R^2 is approximately 0.62, with RMSE around 6.26 years and MAE around 4.33 years.

We then predict birth year for users with missing education-based birth year:

$$\widehat{b}_i = \widehat{X_i'\beta} + \widehat{\alpha}_{\text{first year}(i)} + \widehat{\alpha}_{\text{last year}(i)} + \widehat{\alpha}_{\text{career bins}(i)} + \widehat{\alpha}_{\text{seniority}(i)} + \widehat{\alpha}_{\text{occupation}(i)} + \widehat{\alpha}_{\text{role}(i)} + \widehat{\alpha}_{\text{parent firm}(i)}.$$

The final user-level birth year is

$$b_i = \begin{cases} b_i^{obs}, & \text{if an education-based birth year is available,} \\ \text{round}(\widehat{b}_i), & \text{otherwise.} \end{cases}$$

In the final Compustat 2003–2019 user sample, this procedure fills all missing birth years.

IA6.2 Firm-Level Average Employee Age

We next merge the user-level birth year back into the annual worker-firm panel. For each worker-year observation, we compute age as

$$\text{age}_{it} = t - b_i,$$

where t is the calendar year of the worker-firm observation and b_i is the final user-level birth year defined above.

We map each worker-firm-year observation to a Compustat firm using the parent-subsidiary bridge. Specifically, for each row with Revelio firm identifier `rcid` and year t, we merge to the bridge by `rcid` and year to obtain the contemporaneous Compustat identifier `gvkey`.

Before aggregating, we de-duplicate the data at the `gvkey`-year-user level so that the same worker is counted at most once within a firm-year. Let $\mathcal{I}_{gt}$ denote the set of distinct users linked to firm g in year t. Each user-year observation has a Revelio employment weight w_{it}. The employment-weighted average age is then

$$\text{age_w}_{gt} = \frac{\sum_{i \in \mathcal{I}_{gt}} \text{age}_{it} w_{it}}{\sum_{i \in \mathcal{I}_{gt}} w_{it}},$$

which is the employee age variable in Appendix Table IA5.

IA7 Additional Proofs

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

IA7.1 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

$$\frac{L}{\Xi} \;>\; \max\{1,\chi\}\,\phi \tag{IA1}$$

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 (IA1).

Per-B-unit profit Substituting the equilibrium expression (19) 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 (IA1).

Economic interpretation Condition (IA1) 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.

IA7.2 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

$$\pi_B(\rho_A, \rho_B) \;-\; \pi_B(0, \rho_B) \;=\; \phi\,(1 - \rho_B)\,\rho_A \Big[\, 1 - F(\Delta)/F^* \,\Big], \tag{IA2}$$

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,

$$\pi_B^{\text{sep}} \;=\; \pi_B(0,\rho_B) \;-\; \theta\,\phi\,(1-\rho_B)\,\rho_A\,F(\Delta), \qquad \theta \geq 0, \tag{IA3}$$

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 (Stein, 1997; Scharfstein and Stein, 2000; Rajan, Servaes, and Zingales, 2000): 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

$$\frac{\partial^2 \pi_B^{\text{sep}}}{\partial \rho_A\,\partial \Delta} \;=\; -\,\theta\,\phi\,(1-\rho_B)\,F'(\Delta) \;<\; 0, \tag{IA4}$$

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 IA3 (Integration vs internal separation)**.** *Consider a firm with $\rho_A > 0$ facing workstyle distance Δ, and define*

$$F^{**} \;\equiv\; \frac{1}{1+\chi-\theta}, \tag{IA5}$$

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 (IA3) from $\pi_B(\rho_A,\rho_B)$ and use (IA2):

$$\pi_B(\rho_A,\rho_B) \;-\; \pi_B^{\text{sep}} \;=\; \phi\,(1-\rho_B)\,\rho_A\,\Big[\,1-(1+\chi-\theta)\,F(\Delta)\,\Big], \tag{IA6}$$

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. □

IA7.3 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 (19) 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 $|d\overline{ws}|_n = \frac{1}{2}\,\Delta_n$. □

IA7.4 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

$$\max_{\{l_{A,j}\},\{l_{B,j}\}} \; L_A\, Q_A(\{l_{A,j}\}) \;+\; L_B\, Q_B(\{l_{B,j}\}) \;-\; K\,\Delta \quad \text{s.t.} \sum_j l_{X,j} = 1,\; X \in \{A,B\}, \tag{IA7}$$

with the friction-strength scalar

$$K \;\equiv\; L_B\, \frac{\phi}{F^*}\,(1-\rho_B)\,\rho_A\, f \;\geq\; 0. \tag{IA8}$$

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 (IA7) are

$$L_A\, \frac{\partial Q_A}{\partial l_{A,j}} \;+\; K\,(u \cdot ws_j) \;=\; \mu_A, \quad \forall j, \tag{IA9}$$

$$L_B\, \frac{\partial Q_B}{\partial l_{B,j}} \;-\; K\,(u \cdot ws_j) \;=\; \mu_B, \quad \forall j, \tag{IA10}$$

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 $l_{X,j}^{CES} = \psi_{X,j}/\sum_k \psi_{X,k}$ satisfies the first term in each FOC as a constant across j (Euler's theorem on CES production: $\partial Q_X/\partial l_{X,j}|_{CES} = Q_X^{CES}$), 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 $l_{X,j}^* = l_{X,j}^{CES}(1 + \nu_{X,j})$ with $\sum_j l_{X,j}^{CES}\nu_{X,j} = 0$ 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 $\partial Q_X/\partial l_{X,j}|_{l^*} \approx Q_X^{CES}(1 - \nu_{X,j}/\sigma)$. Substituting into (IA9)–(IA10) and using the

constraint to pin down the multipliers, the first-order solution for the wedges is

$$\nu_{A,j} \;=\; +\frac{\sigma\, K}{L_A\, Q_A^{CES}}\, u \cdot \left(ws_j - ws_A^{CES}\right), \qquad \nu_{B,j} \;=\; -\frac{\sigma\, K}{L_B\, Q_B^{CES}}\, u \cdot \left(ws_j - ws_B^{CES}\right). \tag{IA11}$$

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) $1/(L_X\, Q_X^{CES})$, normalizing by the unit-type's flow output; and (iv) the alignment $u \cdot (ws_j - ws_X^{CES})$ 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

$$\Delta^* \;\equiv\; \|(ws_B + \Delta ws_B) - (ws_A + \Delta ws_A)\| \;\approx\; \Delta \;-\; \sigma\, K\, u^\top \left[\frac{\Sigma_A}{L_A\, Q_A^{CES}} + \frac{\Sigma_B}{L_B\, Q_B^{CES}}\right] u, \tag{IA12}$$

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 (IA12) 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 (IA12). Substituting into the linearized impulse growth equation (IA15), derived in the proof of Result 2 in Section IA7.5 below, 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 (IA11) 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 IA4. *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

$$\Delta^*(t_A, t_B) \;=\; \max\{(t_A\Sigma\psi_A + t_B\Sigma\psi_B - 1)\,\Delta_n,\, 0\}\,. \tag{IA13}$$

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 (IA13) into the owner's problem (17), and evaluating at

the impulse $t = 0^+$:

$$\max_{L_A, L_B \geq 0,\, t_X \in [0, 1/\Sigma\psi_X]} \; L_A\, \pi_{A,0^+}(t_A) + L_B\, \pi_{B,0^+}(t_B, \Delta^*(t_A, t_B)) \; - \; \frac{L_A^2 + L_B^2}{2\xi_i}, \tag{IA14}$$

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 (IA1), $\pi_{X,0^+}^{CES} > 0$ for all firms in the industry (Appendix Section IA7.1), 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 $ws_X = \sum_j l_{X,j}\, ws_j \in \text{conv}(\{ws_j\})$.

Substituting $Q_X = 1$ into the owner's problem (IA7):

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

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.

IA7.5 Proof of Result 2

To make notation lighter, write $w_n \equiv |d\overline{ws}|_n$, $a_i \equiv \mathrm{age}_i$, 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:

$$g_{i,n} \;\approx\; 1 \;-\; \Lambda_n \,(\rho_i - \bar{\rho}_{A,n,0^-}) \;+\; O(\cdot)^2, \qquad \text{where} \quad \Lambda_n \;\equiv\; \frac{\Xi_n \,\phi\, f}{L_n \,F^*}\,\Delta_n. \tag{IA15}$$

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 (i) from the body of Result 2 for convenience.

Assumption IA1 (Assumption (i), 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 IA5 (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*

$$\hat{\zeta} \;=\; -\,\Phi \;\times\; \frac{\mathrm{cov}(\rho_i, a_i)}{\mathrm{var}(a_i)}, \tag{IA16}$$

where

$$\Phi \;\equiv\; \frac{2\,\phi\,f}{F^{*}} \;\times\; \frac{\operatorname{cov}(\Xi_n \Delta_n / L_n,\; \Delta_n)}{\operatorname{var}(\Delta_n)} \tag{IA17}$$

is an industry-level scalar. Under the additional Assumption (ii) of Result 2, that Ξ_n/L_n is independent of Δ_n, this reduces to the body's $\Phi = \frac{2\,\phi\,f}{F^{*}}\,\mathbb{E}[\Xi_n/L_n]$, and further to $\Phi = 2(\Xi/L)\,\phi f/F^{*}$ when $\Xi_n/L_n = \Xi/L$ is industry-invariant. Sufficient conditions for $\Phi > 0$ are given in Section IA7.5.

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

$$g_{i,n} \;=\; \nu \;+\; \kappa_w\, w_n \;+\; \rho_a\, a_i \;+\; \zeta\,(w_n \times a_i) \;+\; \varepsilon_{i,n}. \tag{IA18}$$

The FWL theorem states that the coefficient ζ on $w_n \times a_i$ in (IA18) 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

$$w_n \times a_i \;=\; \alpha_1 \;+\; \beta_2\, w_n \;+\; \beta_1\, a_i \;+\; r^{(1)}_{i,n}, \tag{IA19}$$

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 (IA19), the residual is

$$r^{(1)}_{i,n} \;=\; \Big(w_n - \mathbb{E}(w_n)\Big)\Big(a_i - \mathbb{E}(a_i)\Big) \;\equiv\; v_n\, \tilde{a}_i, \tag{IA20}$$

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

$$g_{i,n} \;=\; \alpha_2 \;+\; \kappa_2\, w_n \;+\; \rho_2\, a_i \;+\; r^{(2)}_{i,n}, \tag{IA21}$$

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, $r_{i,n}^{(2)}$ is uncorrelated with each of the regressors $(1, w_n, a_i)$.

Regression (iii): univariate regression of $r_{i,n}^{(2)}$ on $r_{i,n}^{(1)}$. By the FWL theorem, the coefficient ζ in (IA18) equals the coefficient from the univariate regression of $r_{i,n}^{(2)}$ on $r_{i,n}^{(1)} = v_n \tilde{a}_i$:

$$\zeta \;=\; \frac{\operatorname{cov}(r_{i,n}^{(2)},\, v_n \tilde{a}_i)}{\operatorname{var}(v_n \tilde{a}_i)}. \tag{IA22}$$

We now show that $\operatorname{cov}(r_{i,n}^{(2)}, v_n \tilde{a}_i) = \operatorname{cov}(g_{i,n}, v_n \tilde{a}_i)$, so that $r_{i,n}^{(2)}$ can be replaced by $g_{i,n}$ in the numerator. From (IA21), $r_{i,n}^{(2)} = g_{i,n} - \alpha_2 - \rho_2 a_i - \kappa_2 w_n$, so

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

Both $\operatorname{cov}(a_i, v_n \tilde{a}_i)$ and $\operatorname{cov}(w_n, v_n \tilde{a}_i)$ vanish, because $v_n \tilde{a}_i = r_{i,n}^{(1)}$ is the residual from regression (i) and is therefore orthogonal to $(1, w_n, a_i)$. Hence

$$\zeta \;=\; \frac{\operatorname{cov}(g_{i,n},\, v_n \tilde{a}_i)}{\operatorname{var}(v_n \tilde{a}_i)}. \tag{IA23}$$

It remains to evaluate the numerator and denominator of (IA23) using (IA15) 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 (IA15):

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

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 (IA23) is

$$\operatorname{cov}(g_{i,n},\, v_n\tilde{a}_i) \;=\; -\frac{\phi\, f}{2F^*}\operatorname{cov}(\Xi_n\Delta_n/L_n,\, \Delta_n)\,\operatorname{cov}(\rho_i, a_i). \tag{IA24}$$

Combining (IA24) 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 (IA16) with Φ as defined in (IA17). □

Sufficient conditions for $\Phi > 0$ The covariance form (IA17) 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 IA6. *Under Assumption IA2,*

$$\Phi \;=\; \frac{2\,\phi\, f}{F^*}\,\mathbb{E}\left[\frac{\Xi_n}{L_n}\right] \;>\; 0. \tag{IA25}$$

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 (IA17) 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 (IA17), $\Phi > 0$ if and only if

$$\operatorname{cov}(\Xi_n\Delta_n/L_n,\; \Delta_n) \;>\; 0, \tag{IA26}$$

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 (IA15) and the workstyle distance Δ_n itself. To see that (IA26)

holds even without full independence, define $\widetilde{\Delta}_n \equiv \Delta_n - \mathbb{E}(\Delta_n)$ and decompose

$$\operatorname{cov}\left(\frac{\Xi_n}{L_n}\Delta_n,\, \Delta_n\right) \;=\; \underbrace{\mathbb{E}\left[\frac{\Xi_n}{L_n}\right]\operatorname{var}(\Delta_n)}_{>0} \;+\; \underbrace{\operatorname{cov}\left(\frac{\Xi_n}{L_n},\, \Delta_n\,\widetilde{\Delta}_n\right)}_{\gtrless 0}. \tag{IA27}$$

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.

IA7.6 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 (25) 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 $\lambda^B \equiv (\Xi_n/L_n)\phi\left(1 - F(\Delta_n)/F^*\right)$ 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. □

IA7.7 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

$$g_{i,n} = \nu + \kappa_w w_n + \rho_a a_i + \zeta\,(a_i w_n) + \varepsilon_{i,n}, \qquad \mathbb{E}[\varepsilon_{i,n} \mid a_i, w_n] = 0, \tag{IA28}$$

with proxy

$$\widehat{w}_n = w_n + e_n \tag{IA29}$$

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

$$g_{i,n} = \nu + \kappa_w \widehat{w}_n + \rho_a a_i + \zeta\,(a_i \widehat{w}_n) + \eta_{i,n}, \tag{IA30}$$

where substituting (IA29) into (IA28) gives the error $\eta_{i,n} \equiv \varepsilon_{i,n} - (\kappa_w + \zeta a_i)\, e_n$,[34] 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 (IA28) 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 (IA30) 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 (IA28); 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 (IA30). 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 (IA30) must run through the construction error e_n. Result IA8 formalizes its consequence for the OLS coefficient $\widehat{\zeta}$ in (IA30).

Result IA8. *Under Assumption IA3, the probability limit of the OLS interaction coefficient in (IA30) is*

$$\widehat{\zeta} = \zeta \times \frac{\operatorname{cov}(w_n, \widehat{w}_n)}{\operatorname{var}(\widehat{w}_n)}, \tag{IA31}$$

the population regression coefficient of w_n on the proxy $\widehat{w}_n$. Equivalently, by (IA29),

$$\widehat{\zeta} = \zeta \times \frac{\operatorname{var}(w_n) + \operatorname{cov}(w_n, e_n)}{\operatorname{var}(w_n) + \operatorname{var}(e_n) + 2\operatorname{cov}(w_n, e_n)}. \tag{IA32}$$

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 $\widehat{\zeta} = \operatorname{cov}(g_{i,n}, \tilde{a}_i \tilde{v}_n)/\operatorname{var}(\tilde{a}_i \tilde{v}_n)$.[35] Substituting (IA28): by firm–industry independence and $\mathbb{E}[\tilde{a}_i] = \mathbb{E}[\tilde{v}_n] = 0$,

[34] 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.

[35] The normal equations are diagonal and the calculation is identical to (IA19)–(IA20) in the proof of Result IA5 with $\widehat{w}_n$ in place of w_n; the centering is derived from the projection, not assumed.

the intercept and both main effects vanish and $\varepsilon_{i,n}$ drops by Assumption IA3(ii); only the interaction survives, giving $\operatorname{cov}(g_{i,n}, \tilde{a}_i \widehat{v}_n) = \zeta \operatorname{var}(a_i) \operatorname{cov}(w_n, \widehat{w}_n)$. Dividing by the denominator $\operatorname{var}(a_i) \operatorname{var}(\widehat{w}_n)$ gives (IA31), and (IA29) gives (IA32). □

Implications. Result IA8 has two implications: (i) If the construction error is classical, $\operatorname{cov}(w_n, e_n) = 0$, then (IA32) 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 (IA31) is preserved if and only if $\operatorname{cov}(w_n, \widehat{w}_n) > 0$; by (IA32) 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 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 (IA38).

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 IA8 is the complementary statement about *bias*, showing that this variability shrinks the point estimate toward zero and preserves its sign under positive validity.

IA7.8 Regression with Controls

We generalize Result IA5 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 (IA16) with their multivariate (conditional) counterparts.

Notation We maintain the notation from the proof of Result IA5: $w_n = \frac{1}{2}\Delta_n$, $\rho_i = 1 - e^{-\kappa a_i}$, 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:

$$\Sigma_I = \begin{pmatrix} \mathrm{var}(w_n) & \mathrm{cov}(w_n, Z_n) \\ \mathrm{cov}(w_n, Z_n) & \mathrm{var}(Z_n) \end{pmatrix}, \qquad \Sigma_F = \begin{pmatrix} \mathrm{var}(a_i) & \mathrm{cov}(a_i, x_i) \\ \mathrm{cov}(a_i, x_i) & \mathrm{var}(x_i) \end{pmatrix}. \tag{IA33}$$

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

$$\frac{\Xi_n \Delta_n}{L_n} = \alpha_\Phi + \Phi_{w|Z}\, w_n + \Phi_{Z|w}\, Z_n + u_n^{(\Phi)}. \tag{IA34}$$

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

$$\Sigma_I \begin{pmatrix} \Phi_{w|Z} \\ \Phi_{Z|w} \end{pmatrix} = \begin{pmatrix} \mathrm{cov}(\Xi_n \Delta_n / L_n,\, w_n) \\ \mathrm{cov}(\Xi_n \Delta_n / L_n,\, Z_n) \end{pmatrix}. \tag{IA35}$$

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

$$\rho_i = \alpha_\rho + \beta_{a|x}\, a_i + \beta_{x|a}\, x_i + u_i^{(\rho)}. \tag{IA36}$$

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

$$\Sigma_F \begin{pmatrix} \beta_{a|x} \\ \beta_{x|a} \end{pmatrix} = \begin{pmatrix} \mathrm{cov}(\rho_i,\, a_i) \\ \mathrm{cov}(\rho_i,\, x_i) \end{pmatrix}. \tag{IA37}$$

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 IA9. *Under Assumption IA4, the population OLS coefficient on $w_n \times a_i$ in the extended regression*

$$\begin{aligned} g_{i,n} = \nu &+ \kappa_w\, w_n + \kappa_Z\, Z_n + \rho_a\, a_i + \rho_x\, x_i \\ &+ \zeta_{wa}\,(w_n \times a_i) + \zeta_{wx}\,(w_n \times x_i) + \zeta_{Za}\,(Z_n \times a_i) + \zeta_{Zx}\,(Z_n \times x_i) + \varepsilon_{i,n} \end{aligned} \tag{IA38}$$

is

$$\zeta_{wa} \;=\; -\frac{\phi\, f}{F^*}\,\Phi_{w|Z}\,\beta_{a|x}, \tag{IA39}$$

where $\Phi_{w|Z}$ is the partial coefficient on w_n in (IA34) and $\beta_{a|x}$ is the partial coefficient on a_i in (IA36). More generally, the full 4×1 vector of interaction coefficients satisfies

$$\begin{pmatrix} \zeta_{wa} \\ \zeta_{wx} \\ \zeta_{Za} \\ \zeta_{Zx} \end{pmatrix} \;=\; -\frac{\phi\, f}{F^*} \begin{pmatrix} \Phi_{w|Z} \\ \Phi_{Z|w} \end{pmatrix} \otimes \begin{pmatrix} \beta_{a|x} \\ \beta_{x|a} \end{pmatrix}, \tag{IA40}$$

where $\otimes$ denotes the Kronecker product.

Relative to Result IA5, 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 (IA39) reduces to (IA16): $\Phi_w = \mathrm{cov}(\Xi_n\Delta_n/L_n, w_n)/\mathrm{var}(w_n) = 2\,\mathrm{cov}(\Xi_n\Delta_n/L_n, \Delta_n)/\mathrm{var}(\Delta_n)$ (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:

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

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

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

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

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

$$r_{i,n}^{(w,a)} \;=\; v_n\,\tilde{a}_i. \tag{IA41}$$

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

$$r_{i,n}^{(w,x)} \;=\; v_n\,\tilde{x}_i, \qquad r_{i,n}^{(Z,a)} \;=\; u_n\,\tilde{a}_i, \qquad r_{i,n}^{(Z,x)} \;=\; u_n\,\tilde{x}_i. \tag{IA42}$$

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 IA5, 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

$$M\begin{pmatrix}\zeta_{wa}\\ \zeta_{wx}\\ \zeta_{Za}\\ \zeta_{Zx}\end{pmatrix} \;=\; \begin{pmatrix}\operatorname{cov}(g_{i,n},\ v_n\tilde{a}_i)\\ \operatorname{cov}(g_{i,n},\ v_n\tilde{x}_i)\\ \operatorname{cov}(g_{i,n},\ u_n\tilde{a}_i)\\ \operatorname{cov}(g_{i,n},\ u_n\tilde{x}_i)\end{pmatrix}, \tag{IA43}$$

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 (IA15) 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:

$$\begin{pmatrix} \operatorname{cov}(g_{i,n},\, v_n\tilde{a}_i) \\ \operatorname{cov}(g_{i,n},\, v_n\tilde{x}_i) \\ \operatorname{cov}(g_{i,n},\, u_n\tilde{a}_i) \\ \operatorname{cov}(g_{i,n},\, u_n\tilde{x}_i) \end{pmatrix} = -\frac{\phi f}{F^*} \begin{pmatrix} \operatorname{cov}(\Xi_n\Delta_n/L_n,\, w_n) \\ \operatorname{cov}(\Xi_n\Delta_n/L_n,\, Z_n) \end{pmatrix} \otimes \begin{pmatrix} \operatorname{cov}(\rho_i, a_i) \\ \operatorname{cov}(\rho_i, x_i) \end{pmatrix}. \tag{IA44}$$

Solving the system Substituting into (IA43) 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 (IA35), the first parenthesis equals $(\Phi_{w|Z}, \Phi_{Z|w})'$. By (IA37), 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 (IA40). In particular, $\zeta_{wa} = -\frac{\phi f}{F^*}\, \Phi_{w|Z}\, \beta_{a|x}$, which is (IA39). □