Who Sets the Price? The Vertical Origins of Uniform Pricing

Alvin Chen[†], Leandro Sanz[‡], and Michael D. Wittry[§]

March 27, 2026

Abstract

We provide evidence that the uniformity of retail prices across vastly different locations originates primarily upstream, within manufacturers and the manufacturer-retailer interactions. Using manufacturer-product scanner data spanning 1.83 billion store-month observations, we decompose retail price variation into manufacturer and retailer components. Manufacturer identity accounts for the large majority of explained variation in price levels, dispersion, and changes—ranging from 73 to 83 percent across decompositions—substantially more than retailer identity. Private label products, which attenuate the manufacturer-retailer information wedge, exhibit significantly greater price dispersion than national brands. We provide quasi-experimental evidence by exploiting differences in manufacturers' exposure to generative AI. Following the introduction of generative AI, products sold by AI-exposed manufacturers exhibit 5 percent lower average prices, greater geographic price dispersion, more frequent price changes, and increased responsiveness to local economic conditions—effects that survive retailer-by-time fixed effects and are concentrated in information-intensive tasks. These findings suggest that retail pricing rigidities reflect upstream informational frictions that technology can relax, rather than immutable features of retail markets.

JEL Codes: L11, L81, D22, D83, O33, E31

Keywords: Uniform pricing, artificial intelligence, retail markets, price dispersion, informational frictions, vertical relationships, price rigidity, technology adoption

For their helpful comments and suggestions, we thank René Stulz and Andrea Eisfeldt. We thank Paul Décaire for assistance with the data. Researcher(s)' own analyses calculated (or derived) based in part on data from The Nielsen Company (US), LLC and marketing databases provided through the Nielsen Datasets at the Kilts Center for Marketing Data Center at The University of Chicago Booth School of Business. The conclusions drawn from the Nielsen data are those of the researcher(s) and do not reflect the views of Nielsen. Nielsen is not responsible for, had no role in, and was not involved in analyzing and preparing the results reported herein.

[†]Stockholm School of Economics, Email: alvin.chen@hhs.se

[‡]Mendoza College of Business, University of Notre Dame, Email: lsanz@nd.edu

[§]Fisher College of Business, Ohio State University, and European Corporate Governance Institute, Email: wittry.2@osu.edu.

1 Introduction

Few corporate decisions are more important than price-setting. Yet, paradoxically, retail prices are remarkably uniform. DellaVigna and Gentzkow (2019) (DG) document that stores within the same retail chain charge nearly identical prices for identical products, despite serving communities with vastly different clienteles and competitive environments. Such uniformity generates significant losses—DG estimate that the median chain sacrifices $16 million in profit annually—suggesting that the frictions preventing localized optimal pricing are substantial. Yet, a fundamental question remains unanswered: who sets that price? That is, where in the vertical chain does the binding friction on price localization originate? The answer determines the nature of the frictions themselves and has deep implications for price rigidity, consumer welfare, and market efficiency.

Final posted retail prices are the joint product of manufacturers upstream—who design, produce, and often provide pricing recommendations for their products—and retailers downstream—who price and place them on the shelf. Each layer possesses distinct complementary information that is costly to communicate across thousands of relationships: retailers observe local foot traffic, demographics, and competitive conditions, while manufacturers hold product-level knowledge of demand and cross-market patterns. If the binding friction to localized pricing is downstream, the natural explanation for uniformity is high implementation and menu costs—the physical and organizational expense of maintaining location-specific prices across hundreds or thousands of stores (Levy, Bergen, Dutta, and Venable, 1997; Zbaracki, Ritson, Levy, Dutta, and Bergen, 2004). Retailers may find it too costly to manage complex pricing strategies or to constantly update shelf tags and point-of-sale systems across many locations. If, however, the binding friction is upstream, the explanation is fundamentally different: the cost of generating and communicating market-specific pricing recommendations through the vertical chain is too high for manufacturers to do so, and absent that information, uniform pricing is the rational default (Mankiw and Reis, 2002; Alvarez, Lippi, and Paciello, 2011). Under this view, the friction is not about the cost of adjusting prices but about the cost of knowing when and how to adjust them optimally.

We provide evidence that the systematic component of uniform pricing in retail markets originates primarily upstream. That is, the frictions generating price uniformity appear to arise before

the point of retail execution—within manufacturers or in the communication and coordination between manufacturers and retailers—rather than in retailers' unilateral pricing decisions. We establish this finding through three complementary empirical strategies. First, using a bridge linking over 300 publicly traded manufacturers to product-level Nielsen Retail Scanner data spanning 1.83 billion store-month observations, we decompose retail pricing variation into its vertical components using variance decomposition methods adapted from the labor literature (e.g., see Abowd, Kramarz, and Margolis, 1999; Card, Heining, and Kline, 2013).[1] This approach decomposes pricing variation into manufacturer and retailer components without imposing strong assumptions about the manufacturer-retailer relationship or contract. Across every model specification, manufacturer identity accounts for substantially more of the explained variation than retailer identity. For example, in a direct head-to-head comparison, manufacturer identity accounts for over 80 percent of the explained variation in prices, compared to just 17 percent for the retailer. This pattern extends to price dispersion—the object at the core of DG's uniform pricing puzzle—and to price changes, with manufacturer shares of 73 and 83 percent, respectively.

Second, we exploit variation in the vertical structure across different products sold within retailers. In particular, we benchmark national brands, which comprise the majority of products in the scanner data and operate through the traditional manufacturer-retailer-consumer chain, against private-label brands, which are vertically integrated and controlled by the same retailer (Hoch and Banerji, 1993; Chintagunta, Bonfrer, and Song, 2002). The private-label structure substantially attenuates the vertical information asymmetry—pricing decisions are made by the same entity that observes all demand information (Hoch, Kim, Montgomery, and Rossi, 1995)—while much of the retailer's pricing infrastructure and implementation costs associated with managing prices across locations (e.g., shelf tags, scanner systems, promotional calendars) remain unchanged.

If retailer-side implementation frictions were the dominant driver of uniform pricing, private-label and national-brand products should exhibit similar uniformity—and several features of private-label markets reinforce this prediction: private labels carry the retailer's own brand, strengthening

[1] Throughout the paper, we use "manufacturer" to refer to the brand owner—the firm that designs, markets, and sets wholesale terms for the product—rather than necessarily the firm that physically produces it. In many cases these are the same entity, but where production is outsourced, the pricing decisions we study are made by the brand owner.

price-image incentives for uniformity; they are often priced as a fixed discount off the category-leading national brand, inheriting any uniformity in that benchmark; they are lower-margin products that may receive less pricing attention; and they face direct competitive pressure from other value-tier brands. If, however, manufacturer-side informational frictions are an important source of uniformity, private labels should display more localized pricing. We find that private-label products exhibit significantly greater price dispersion, even within retailer-by-product-category-by-month. This pattern is consistent with manufacturer-side informational frictions playing an important role in generating uniform pricing.

Third, we provide quasi-experimental evidence on whether reductions in upstream information-processing frictions shift pricing away from uniformity by leveraging manufacturers' differential exposure to the introduction of OpenAI's InstructGPT API in January 2022—the first generative artificial intelligence (AI) model available for firm-level deployment at scale. AI fundamentally reduces information acquisition and processing costs—precisely the type of informational friction highlighted by our earlier decomposition and private-label results (Agrawal, Gans, and Goldfarb, 2019; Goldfarb and Tucker, 2019). We measure firms' AI exposure using occupation-level scores following Eisfeldt, Schubert, Taska, and Zhang (2025), which capture the extent to which generative AI can reduce the time required to complete the tasks associated with each occupation, and aggregate them to the firm level using employment shares.[2] Critically, our treatment variable is defined as above-median AI exposure measured in January 2020—two years before the API introduction—which reduces concerns that the classification reflects endogenous responses to AI availability rather than pre-existing workforce composition.

We compare products sold by high-exposure firms to a matched control group with similar pre-treatment price levels and sales volumes using a difference-in-differences event-study design. Event-study estimates show little evidence of meaningful differential pre-trends in pricing outcomes prior to the introduction of the InstructGPT API. Our central finding is that AI-exposed firms reduce their average CPI-deflated product prices by approximately 5 percent relative to matched control

[2] Our measure captures firms' *exposure* to generative AI based on pre-existing workforce composition, not direct observation of AI implementation. In Section 8, we validate this measure by showing that highly exposed firms significantly increase AI-related hiring and implementation-oriented disclosures following the API introduction.

products following the API introduction. However, the results are not simply a story of uniform cost pass-through. AI-exposed firms simultaneously increase the geographic dispersion of their prices: the cross-store standard deviation and interquartile range of log prices both rise significantly, even within county and retailer, consistent with a shift away from uniform pricing toward greater localization as informational frictions decline (DellaVigna and Gentzkow, 2019; Bergemann and Bonatti, 2019).

AI-exposed firms also reprice more frequently, with the increase driven mostly by price cuts, and their prices become more responsive to local economic conditions. These effects are amplified in more competitive markets and survive retailer-by-time fixed effects, strengthening the interpretation that the pricing changes reflect manufacturer-side rather than retailer-side forces. The estimated treatment effect is negative for virtually every retailer in the data, consistent with a common upstream component, albeit with heterogeneous pass-through across retailers. The effects are concentrated in AI-augmentation tasks and in settings where the returns to improved demand information are likely to be highest, consistent with AI augmenting firms' ability to gather, process, and act on local demand signals rather than merely automating repricing (Zbaracki et al., 2004; Stroebel and Vavra, 2019; Brynjolfsson, Li, and Raymond, 2023; Mackowiak and Wiederholt, 2009).

To formalize this interpretation, we develop a parsimonious vertical pricing framework in which the manufacturer generates market-specific pricing recommendations at a cost, and the retailer implements them subject to operational constraints. The framework identifies two channels through which AI can affect pricing: a production channel, in which lower marginal costs reduce average prices uniformly, and an information channel, in which lower costs of generating pricing recommendations increase geographic dispersion, repricing frequency, and local demand sensitivity. Neither channel alone accounts for the full set of empirical patterns; together they rationalize the joint finding of lower average prices and greater price localization.

We provide several pieces of supporting evidence for this interpretation. High-exposure firms increase AI-related hiring and implementation-oriented disclosures following the API introduction, consistent with the exposure measure capturing genuine deployment rather than passive exposure. The main results are also robust to controls for concurrent macroeconomic shocks, including

interest-rate exposure and COVID-related disruptions, as well as to tighter fixed-effects specifications, placebo event dates, and randomization-inference exercises.

Our findings characterize the pricing of fast-moving consumer goods by large publicly traded manufacturers in brick-and-mortar retail. Whether similar patterns emerge for services, e-commerce, or smaller firms—settings where both AI adoption and pricing frictions differ markedly—remains an open question.

Our study connects to three strands of the literature. First, and most directly, we contribute to a large body of research on the sources of price rigidity and geographic uniformity in retail markets. Canonical models attribute infrequent price adjustment to fixed menu costs (Golosov and Lucas, 2007; Midrigan, 2011), while a growing body of work emphasizes that the informational and organizational costs of pricing decisions dwarf the physical costs of changing price tags (Zbaracki et al., 2004; Sims, 2003; Mackowiak and Wiederholt, 2009; Alvarez, Lippi, and Paciello, 2011). DellaVigna and Gentzkow (2019) document the striking result that retail prices are nearly uniform across locations, even when local demand and competitive conditions differ substantially, and attribute this to the organizational costs of implementing market-specific prices. We provide the first decomposition of retail price variation into its vertical components, showing that manufacturer identity accounts for substantially more of the explained variation in both price levels and price dispersion than retailer identity. The private label comparison further sharpens the mechanism: when the vertical information asymmetry is attenuated, pricing becomes more localized even though implementation costs remain. The AI natural experiment provides quasi-experimental evidence consistent with this interpretation, showing that reductions in manufacturer-side information costs shift pricing away from uniformity and toward localization. Together, these findings suggest that the pricing rigidities documented in the literature are not immutable features of retail markets but rather reflect, in large part, upstream informational frictions that technology can reduce.

Second, we contribute to the rapidly growing literature on the real effects of artificial intelligence. Much of this research has focused on AI's impact on labor markets (Autor, 2015; Acemoglu and Restrepo, 2022; Webb, 2020), firm-level productivity (Brynjolfsson, Rock, and Syverson, 2021), and corporate investment and growth (Babina, Fedyk, He, and Hodson, 2024; Bena, Garlappi,

Ma, and Tanguay, 2024). Eisfeldt et al. (2025) develop the task-level exposure framework that we build on and show that AI exposure is priced in equity markets. Our paper extends this literature to the product market by documenting how AI exposure translates into changes in the prices firms actually charge. The augmentation-automation decomposition aligns with Acemoglu and Restrepo (2022) and Brynjolfsson, Li, and Raymond (2023): pricing effects load primarily on augmentation exposure, consistent with AI enhancing employees' capacity to gather and act on demand information.

Third, our findings complement a growing literature on algorithmic and automated pricing (Calvano, Calzolari, Denicolo, and Pastorello, 2020; Assad, Clark, Ershov, and Xu, 2024; Dubé and Misra, 2023). Our setting differs in a fundamental way: the pricing effects we document are driven not by the deployment of pricing algorithms per se, but by AI's effect on firms' information environments. The increase in geographic dispersion and local demand sensitivity we find is consistent with AI enabling firms to acquire and act on demand information that was previously too costly to obtain (Bergemann and Bonatti, 2019), rather than with algorithmic optimization of existing pricing rules. This distinction matters for welfare: while algorithmic collusion raises consumer prices, the information channel we identify lowers them.

Taken together, our results offer a unified account of how AI reshapes a core corporate decision. The pricing literature asks why prices are so rigid and uniform; the AI literature asks what AI does to firms. We provide evidence on both: AI relaxes the informational frictions that sustain uniform pricing—doing so through task-level augmentation rather than automation—and reduces consumer prices on net. By identifying upstream informational frictions as a central source of uniform pricing, our findings reframe the uniform pricing puzzle from a question about retail operations to one about the flow of information through the vertical chain.

2 Data

This section describes the sample construction, the firm retail pricing data, the AI implementation measure, and other data used in the empirical analysis.

2.1 Sample construction

Our sample starts with publicly listed U.S. firms in Compustat for which we can identify products (Universal Product Code, UPCs) in the Nielsen Retail Scanner (RMS) data from the Kilts Center at the University of Chicago, using a manufacturer-product bridge constructed from the GS1 barcode registry (Décaire and Sosyura, 2024). The unit of observation in the Nielsen RMS data is a firm-UPC-store-week, but we aggregate it to the firm-UPC-month level. For these firms, we obtain generative AI exposure measures from Eisfeldt et al. (2025) and construct augmentation and automation decompositions of AI exposure using our own GPT-based prompts. We also obtain quarterly financial data from Compustat. Our empirical tests focus on the six-month window $t \in \{-6, \ldots, +6\}$ around the introduction of the InstructGPT API on January 27, 2022. The baseline sample contains 90,211 firm-UPC-month observations across 91 firms, spanning 8,750 unique UPCs and 452 Nielsen product modules. In addition to grocery products, we observe pricing data for apparel, consumer electronics, household goods, and health and beauty care products. We exclude observations with product module code 9999, which corresponds to unclassified or miscellaneous products.

2.2 Pricing outcomes

We aggregate the RMS data into a firm-UPC-month panel in two steps. First, within each firm-UPC-week, we average prices and sum units sold across all stores. Second, we average weekly prices and units sold within each calendar month. We deflate all prices to constant dollars using the monthly Consumer Price Index from the Bureau of Labor Statistics (BLS) and winsorize the resulting real prices at the 1st and 99th percentiles to mitigate the influence of extreme observations.

Using this data, we construct the natural log of CPI-deflated average price, $\ln(\textit{Average price})$.[3] Panel C of Table 1 reports summary statistics on pricing outcomes. The median of $\ln(\textit{Average price})$ is 1.19, corresponding to a median real price of approximately \$3.29. The median of the natural log of units sold, $\ln(\textit{Units sold})$, is 3.97, which corresponds to approximately 53 units sold across stores

[3] The price level in results in Section 6.1 are robust to deflating prices using category-specific CPIs rather than the aggregate CPI. These results are reported in Table IA.B.2.

in a given month. The large standard deviation in log prices and units sold reflects the diversity of product categories in the Nielsen data.

We also compute two measures of the dispersion of prices within firms, firm-counties and firm-stores: the standard deviation and interquartile range of log prices. Panel C of Table 1 shows that the median standard deviation of log prices is 0.16 and the median interquartile range is 0.09. The substantial mass near zero is consistent with the widespread uniform pricing documented in the literature (e.g., DellaVigna and Gentzkow, 2019).

Finally, we measure the frequency of price adjustments using an indicator variable, $\mathbb{1}(|\Delta \ln P_{ipt}| > 0.05)$, that equals one if the absolute log price change from month $t-1$ to t exceeds 5 percent.[4] The probability of a price change exceeding 5 percent in a given month is 0.37. Notably, the median of the indicator is zero, indicating that most firm-UPC-months exhibit no meaningful price change. This pattern is also consistent with price rigidity.

2.3 Price variance decomposition

Our first empirical strategy focuses on the decomposition of price variation across manufacturers, retailers, and counties. These decompositions use the broader RMS sample and consist of the universe of store×UPC×month observations for products matched to publicly listed manufacturers. This sample spans 2010 to 2023 and contains approximately 1.83 billion observations across 100,250 unique UPCs, 350 manufacturers, 225 retailers, 3,021 counties, and 886 product modules. Panel A of Table 1 reports summary statistics for the three dependent variables used in the price variance decomposition tests, log prices, squared deviations from UPC×month means, and month-over-month log price changes.

[4] The literature documents that price adjustments are infrequent but tend to be large when they occur, typically 10 to 15 percent in U.S. micro data (e.g., Klenow and Kryvtsov, 2008; Klenow and Malin, 2010). At the same time, very small price changes may reflect measurement error (Nakamura and Steinsson, 2008; Eichenbaum, Jaimovich, and Rebelo, 2011). A 5% threshold is conservative relative to the typical magnitude of price adjustments documented in the literature, while being large enough to exclude small fluctuations that may reflect measurement error. However, our results are robust to using a 10% threshold instead.

2.4 Private labels

In our second empirical strategy, we compare the price dispersion of private label products to national brands. The private label sample uses the full Nielsen scanner universe without requiring a manufacturer match, since private label products are by definition not linked to publicly traded manufacturers. This sample covers 2021 to 2023 and contains approximately 215 million UPC×retailer×month observations, of which 22% are private label products identified by the "CTL BR" brand designation in the Nielsen product attributes file. For each UPC×retailer×month, we compute the standard deviation of log prices across stores within that retailer, which measures the degree of geographic price dispersion. Panel B of Table 1 reports the distribution of this measure separately for private label and national brand products. The mean within-retailer price dispersion is 0.114 for private labels compared to 0.097 for national brands, with medians of 0.050 and 0.041, respectively. This difference is present throughout the distribution. At the 75th percentile, private label dispersion is 0.124 versus 0.114 for national brands.

2.5 AI exposure

We measure firm-level exposure to generative AI using the Eisfeldt et al. (2025) approach. Their approach begins with the universe of task descriptions from the Occupational Information Network (O*NET, version 27.2), which characterizes the work activities associated with each occupation in the U.S. economy. Each task is evaluated by Open AI's GPT 3.5 Turbo model for the extent to which generative AI is expected to reduce the time required to perform it.[5] Tasks are classified into three exposure categories: directly exposed (scored as 1), partially exposed (scored as 0.5), and not exposed (scored as 0). The occupation-level AI exposure score X_o is then computed as the simple average of these scores across all tasks within occupation o.

To construct firm-level AI exposure, we combine the occupation-level scores with workforce composition data from Revelio Labs, which provides individual-level employment information with

[5] We develop a similar task scoring approach where we specify the capabilities of a target model and use GPT 5-mini as the scoring model. We use this approach to obtain an AI exposure measure based on the capabilities of InstructGPT as of January 2022. We show in Internet Appendix B that all three sets of baseline results—price levels, price dispersion, and repricing frequency—are robust to using this alternative exposure measure (Table IA.B.4, Table IA.B.5, and Table IA.B.6).

Standard Occupational Classification (SOC) codes, start and end dates, and firm identifiers. Following Eisfeldt et al. (2025), we compute firm-level exposure as the employment-share-weighted average of occupation-level scores:

$$\text{AI Exposure}_{it} = \sum_{o} S_{iot} \cdot X_o, \tag{1}$$

where S_{iot} is the share of firm i's employees in occupation o at time t and X_o is the fixed occupation-level AI score. Figure 1 plots the cross-sectional distribution of firm-level AI exposure measured in January 2020. The distribution is approximately bell-shaped and concentrated in a narrow range. The interquartile range spans 0.35 to 0.42, and the standard deviation is 0.06 (Panel D of Table 1). This tight clustering reflects the large, diversified firms in our Nielsen sample. Since these firms employ workers across many occupations, the firm-level scores are averaged over a broad mix of tasks and converge toward the cross-firm mean. Despite this relatively tight distribution, there is meaningful variation. A firm at the 25th percentile has an AI exposure score of 0.35, compared to 0.42 at the 75th percentile, a difference of roughly one standard deviation.

We use AI exposure as a proxy for AI implementation, similar to the use of occupation-based measures as proxies for technology adoption in the broader economics literature (e.g., Autor, Levy, and Murnane, 2003; Acemoglu and Restrepo, 2022). The logic is that firms whose workforces are concentrated in occupations where generative AI can meaningfully reduce task completion times are more likely to benefit from adopting these tools and, consequently, more likely to deploy them. We validate this assumption directly in Section 8 by showing that high-exposure firms significantly increase their use of AI-related language in job postings and SEC 10-K filings after the InstructGPT API introduction, consistent with increased AI adoption.

2.6 Augmentation vs. automation decomposition

The AI exposure measure from Eisfeldt et al. (2025) combines two conceptually distinct channels through which AI adoption may affect firm outcomes: augmentation and automation (Autor, 2015; Acemoglu and Restrepo, 2022; Autor and Thompson, 2025). Under augmentation, AI enhances worker productivity while requiring human judgment. Under automation, AI enables tasks to

be completed with minimal human involvement. To distinguish these channels empirically, we decompose AI exposure into separate augmentation and automation components.[6]

Each task is evaluated along two dimensions by GPT-5-mini.[7] We first compute an augmentation score, ranging from 0 to 10, that captures the extent to which AI assistance may increase worker productivity while human judgment is still likely to be required for a task to be completed. For example, high augmentation scores (8 to 10) are assigned to tasks such as data analysis, where AI processes information but humans are still likely required to interpret the data and make decisions based on it. The automation score also ranges from 0 to 10 and measures the likelihood that AI could complete the task independently with minimal human involvement while achieving acceptable quality. High automation scores (8 to 10) are assigned to tasks such as structured data entry and standardized report generation, since these tasks tend to have well-defined rules that govern workflow, and human oversight adds little value.

We aggregate the task-level scores to the firm level following the same procedure used for the overall exposure measure. Task scores are first mapped to three categories—high (8 to 10, scored as 1), medium (4 to 7, scored as 0.5), and low (0 to 3, scored as 0)—and averaged across tasks within each O*NET occupation code. Occupation-level scores are then aggregated to the firm level using employment shares from Revelio Labs. Internet Appendix B provides the complete system prompt submitted to the scoring model and the full scoring rubric.

2.7 Other data

We obtain quarterly financial data from Compustat and data on the total number of employees from Revelio Labs. All firm-level controls are measured in the quarter before the introduction of the InstructGPT API in 2022 Q1. Panel E of Table 1 reports summary statistics. Relative to the broader Compustat sample, sample firms tend to be larger, older, and more capital-intensive. For

[6] Eisfeldt et al. (2025) also distinguish between these channels, but do so by classifying each occupation's tasks as either core or supplementary using O*NET task categories. With their approach, exposure through core tasks proxies for automation, while exposure through supplementary tasks proxies for augmentation. By contrast, we score every task along both dimensions simultaneously, allowing the same task to exhibit varying degrees of augmentation and automation rather than assigning each task to a single channel based on its role within the occupation.
[7] Internet Appendix IA.A.1 provides the exact prompt and notes that the categorization was conducted in December 2025 with temperature zero.

example, the median sample firm has total assets of approximately $8 billion compared to $1.1 billion in the unrestricted Compustat sample, and a median age of 35 years compared to 16 years. Sample firms also hold less cash (median cash-to-assets of 7% compared to 12%) and have a higher share of fixed assets (17% compared to 8%). This is not surprising given that the Nielsen sample consists of large firms that manufacture physical consumer products. Profitability is similar across samples, with a median return on assets (ROA) near zero in both groups.

We also obtain data on county economic conditions and firm- and county-level pandemic exposure. First, we use monthly county-level unemployment rates from the Bureau of Labor Statistics (BLS) Local Area Unemployment Statistics. Second, we construct a county-level measure of interest rate (IR) exposure using the BLS Quarterly Census of Employment and Wages (QCEW). Specifically, for each county-month, we compute the share of total private employment in interest-rate-sensitive industries, including construction (NAICS 23), real estate (NAICS 531), and finance and insurance (NAICS 52).

For firm-level pandemic exposure, we follow Barry, Campello, Graham, and Ma (2022) and construct occupation-based measures that capture the extent to which a firm's workforce relies on physical proximity. Specifically, we use the work-from-home feasibility index of Dingel and Neiman (2020) and an analogous measure constructed from the American Time Use Survey. Both measures are aggregated to the firm level using pre-pandemic employment shares and are fixed prior to the onset of COVID-19. At the county level, we measure local pandemic intensity using cumulative COVID-19 cases and deaths from the *New York Times* county-level COVID-19 database.

Finally, to assess whether AI exposure translated into actual implementation, we collect data on AI-related hiring and AI discussion in corporate disclosures. We obtain job posting data from LinkUp and textual data from annual 10-K filings from the Securities and Exchange Commission (SEC) EDGAR database.

3 Empirical Strategy

This section describes the empirical strategies used to study the vertical origins of uniform pricing. We begin with variance decomposition and private label comparisons, then turn to a quasi-

experimental design exploiting the January 2022 introduction of OpenAI's InstructGPT API.

3.1 Price variance decomposition

We begin by decomposing the variance of retail prices to establish where in the vertical chain—manufacturers, retailers, or geography—the bottleneck in localized pricing originates. We apply two complementary approaches to the full Nielsen scanner dataset at the store×UPC×month level.

First, we use Shapley value decompositions of R^2 across nested fixed-effects specifications. We estimate models with progressively richer sets of fixed effects—UPC×month, retailer, county, and their interactions—and attribute each factor's contribution to explained variance using the Shapley formula, which averages marginal contributions over all orderings. We report three variants: a product model (baseline: UPC×month), a manufacturer model (baseline: month, factors: manufacturer, retailer, county), and a bilateral model following DellaVigna and Gentzkow (2019) that includes a retailer×UPC interaction to capture product-specific bilateral pricing relationships.

Second, we estimate AKM variance decompositions in the spirit of Abowd, Kramarz, and Margolis (1999) and Card, Heining, and Kline (2013). We regress the dependent variable on additive fixed effects for product (or manufacturer), retailer, county, and month, extract the estimated fixed effects via iterative demeaning, and decompose the variance of the dependent variable into variance and covariance components. We apply both approaches to three dependent variables: log prices (levels), squared deviations from UPC×month means (dispersion), and month-over-month log price changes (adjustments).

3.2 Private label test

Our second empirical strategy exploits the vertical structure of private label pricing. Private labels are manufactured and priced by retailers themselves, which substantially attenuates the manufacturer-retailer information asymmetry that characterizes national brand pricing while holding many implementation costs—shelf tags, scanner systems, promotional calendars—roughly fixed. If upstream informational frictions are a first-order driver of uniform pricing, private label products should exhibit greater geographic price dispersion than national brands sold by the same retailer

in the same category and month. We estimate regressions of within-UPC price dispersion on a private label indicator with progressively tighter fixed-effects structures, from sub-category×month and retailer effects up to retailer×category×month effects.

3.3 AI experiment

Our third empirical strategy provides quasi-experimental evidence on the upstream information channel. We exploit the January 2022 introduction of OpenAI's InstructGPT API, the first generative AI model available for firm-level deployment at scale, and compare pricing outcomes for manufacturers with high versus low pre-existing AI exposure. Importantly, the InstructGPT API was a firm-facing technology event. As shown in Panel A of Figure 2, Google Trends search interest for "artificial intelligence" remained flat throughout the estimation window, spiking only after the consumer-facing ChatGPT release in November 2022. The pricing effects we study thus predate the period in which AI became salient to the general public. If upstream informational frictions are a binding constraint on price localization, as the decomposition and private label results suggest, then a reduction in the cost of acquiring and processing demand information should shift pricing away from uniformity.

Our baseline difference-in-differences specification is as follows:

$$Y_{ipt} = \alpha_{ip} + \alpha_t + \beta(\mathit{Treated}_i \times \mathit{Post}_t) + \delta \ln(\mathit{Units\ sold})_{ipt} + \gamma' (\mathbf{X}_{i0} \times \mathit{Post}_t) + \varepsilon_{ipt}, \tag{2}$$

where i indexes firms, p products (UPCs), and t calendar months. The dependent variable Y_{ipt} is one of the following pricing outcomes: the log real average price, the standard deviation or interquartile range of log prices within a firm-UPC-month, or an indicator for whether the month-over-month price change exceeds a given threshold. $\mathit{Treated}_i$ equals one for firms above the median of AI exposure in January 2020. Post_t equals one from January 2022 onward. The vector $\mathbf{X}_{i0}$ includes pre-event firm characteristics: $\ln(\mathit{Employees})$, $\ln(\mathit{Assets})$, Tobin's q, ROA, cash-to-total assets, fixed assets-to-total assets, R&D-to-total assets, market-to-book, and $\ln(\mathit{Age})$. Because these controls are fixed in the pre-period and interacted with Post_t, they allow for differential changes in pricing outcomes based on pre-treatment characteristics, so that β is not confounded

by the possibility that, for example, larger or more profitable firms would have changed prices differently in the post-period regardless of AI exposure.

The specification includes firm-UPC fixed effects α_{ip} and calendar-month fixed effects α_t. Standard errors are clustered at the firm level. The coefficient of interest, β, captures the differential change in pricing outcomes for high-exposure firms after the InstructGPT introduction. We also estimate variations of this specification at the firm-UPC-county-month and firm-UPC-retailer-month levels, which allow us to exploit within-product geographic and retailer variation in pricing outcomes and to include finer fixed effects. The identifying assumption is that, conditional on the controls, treated and control firms would have followed parallel trends in the outcomes of interest absent the introduction.

We also balance the products of treated and control firms using propensity score matching on pre-event average log prices and log units sold. This ensures that the products entering the comparison have similar pre-treatment pricing levels and sales volumes, so that pre-existing differences in these characteristics do not drive the results.

3.4 Event-Study specification

To test for differential pre-trends and plot dynamic treatment effects, we estimate:

$$\begin{aligned} Y_{ipt} = \alpha_{ip} + \alpha_t + \sum_{k \in \{-6,\ldots,+6\},\, k \neq -3} \beta_k \left(\mathbb{1}[Rel_t = k] \times Treated_i \right) \\ + \delta \ln(Units\ sold)_{ipt} + \gamma' \left(\mathbf{X}_{i0} \times Post_t \right) + \varepsilon_{ipt}, \end{aligned} \tag{3}$$

where Rel_t is event time in months relative to January 2022. We omit $k = -3$ (October 2021) as the reference period so that the baseline precedes the November 18, 2021 expansion of GPT-3 API access, which could have allowed early experimentation before the January 27, 2022 InstructGPT introduction (see Table 2). This event-study design is estimated on the same window used in the baseline analysis. The pre-event coefficients provide a direct test of the parallel-trends assumption, while post-event coefficients characterize the timing and persistence of treatment effects.

4 Price Variance Decomposition

Table 3 decomposes the variance of retail prices using the full Nielsen scanner dataset. We report three complementary decompositions—of price levels, price dispersion, and price changes—each using both Shapley value and AKM variance decomposition methods.

Panel A decomposes the variance of log prices across 1.83 billion store-UPC-month observations. In the product Shapley decomposition (column 1), UPC×month fixed effects account for 93.8% of explained variance, with retailer contributing 5.6% and county only 0.6%. Most of the cross-store price variation is thus attributable to which product is sold and when, not where or by whom it is sold. In the manufacturer Shapley decomposition (column 2), which replaces product with manufacturer identity and uses month as a baseline, manufacturers account for 81.1% of the marginal R^2, retailers 17.1%, and counties 1.7%. The bilateral Shapley decomposition (column 3) following DellaVigna and Gentzkow (2019) introduces a retailer×UPC interaction term, which absorbs 48.3% of explained variance—indicating that a substantial portion of retailer-level pricing reflects product-specific bilateral relationships rather than a uniform retailer markup. The AKM decompositions in columns (4) and (5) confirm the Shapley results. In the product model (column 4), UPC fixed effects explain 87.8% of total variance, with retailer contributing just 1.3%. In the manufacturer model (column 5), manufacturer fixed effects explain 27.9%, retailer 3.4%, and the covariance term is positive at 3.6%, indicating mild positive assortative matching between manufacturers and retailers.

Panel B decomposes the variance of squared deviations from UPC×month means, which isolates cross-store price dispersion conditional on product and time. Despite the much lower overall R^2 (0.049 in column 1 versus 0.907 in Panel A), the manufacturer share remains dominant: 73.1% in the manufacturer Shapley (column 2), compared to 24.7% for retailers. The low R^2 is itself informative—it confirms that conditional on product and time, cross-store price variation is largely unsystematic, consistent with the uniform pricing phenomenon documented by DellaVigna and Gentzkow (2019). What systematic dispersion exists is overwhelmingly attributable to manufacturer identity.

Panel C decomposes the variance of month-over-month log price changes ($\Delta \ln p$), which captures

who determines when and how much prices move. This decomposition shifts the unit of analysis from price levels to price adjustments. In the product Shapley (column 1), UPC×month accounts for 99.9% of explained variance, with retailer contributing just 0.1%—price change timing is almost entirely product-specific. In the manufacturer Shapley (column 2), manufacturer identity accounts for 82.6% of the explained variation, with the retailer share at 13.3% and the county share at 4.1%. The bilateral Shapley (column 3) reinforces this pattern: the retailer×UPC match share falls to 11.8%, compared with 55.2% for dispersion in Panel B, indicating that retailer-specific pricing relationships matter far less for the dynamics of price adjustment than for the cross-sectional pattern of price dispersion. The stability of the manufacturer share across all three decompositions—81.1% for price levels, 73.1% for dispersion, and 82.6% for price changes—indicates that manufacturer identity accounts for the large majority of explained variation not only in price levels but also in how dispersed prices are across stores and when and how much they adjust. The county share remains small in absolute terms but is roughly twice as large for price changes as for price levels (4.1% versus 1.7%), suggesting that the magnitude of price adjustments is somewhat more sensitive to local conditions than the level at which prices are set. This geographic sensitivity, even within a decomposition largely driven by upstream identity, provides a natural channel through which AI-enabled demand sensing could affect pricing: tools that improve local information processing may have their largest effects on how firms calibrate price adjustments across markets.

5 Private Labels vs. National Brands

Table 4 compares the geographic price dispersion of private label products with that of national brands. Private labels are manufactured and priced by retailers themselves, substantially attenuating the manufacturer-retailer information asymmetry that characterizes national brand pricing while holding many implementation costs—shelf tags, scanner systems, promotional calendars—roughly fixed. If upstream informational frictions are a first-order driver of uniform pricing, private label products should exhibit greater geographic price dispersion than national brands sold by the same retailer in the same category and month. Several features of private label pricing work against this prediction: private labels are often priced as a fixed discount off the category-leading national

brand, inheriting any uniformity in that benchmark; they are lower-margin products on which retailers may invest less in price optimization; and they face direct competitive pressure from other value-tier brands, which may compress price variation.

Despite these forces, we find that private labels are substantially more dispersed. As reported in Panel B of Table 1, the mean within-retailer standard deviation of log prices is 0.114 for private labels versus 0.097 for national brands, a difference of approximately 18 percent relative to the national brand mean. The gap is present throughout the distribution: at the median, private label dispersion is 0.050 versus 0.041 for national brands, and at the 25th percentile, private labels exhibit positive dispersion (0.009) while national brands exhibit none. Table 4 shows that this difference survives progressively tighter fixed-effects specifications. Column (1) includes sub-category×month and retailer fixed effects; column (4) includes retailer×category×month fixed effects, which absorb all retailer-category-time variation and isolate the private label effect within the same retailer, category, and month. The coefficient on the private label indicator ranges from 0.007 to 0.008 and is significant at the one-percent level across all four specifications. The stability of the estimate as we tighten the comparison set suggests the result reflects a genuine difference in pricing behavior rather than compositional differences across retailers or categories.

6 AI Experiment

The decomposition and private label results suggest that manufacturer identity is the largest source of systematic pricing variation and that attenuating the upstream information asymmetry increases price localization. We now examine whether these patterns respond to a shock to upstream information costs by exploiting the introduction of generative AI tools that reduce the cost of acquiring and processing demand information. If upstream informational frictions are a binding constraint, manufacturers with greater exposure to AI should exhibit pricing patterns consistent with relaxed information costs: lower average prices, greater geographic dispersion, and more frequent repricing.

To guide interpretation, we distinguish between two channels through which AI can affect pricing. Under a production cost channel, AI lowers marginal costs, leading to uniform reductions in prices across markets. Under an information channel, AI reduces the cost of generating and

processing market-specific demand signals, allowing firms to tailor prices more precisely to local conditions. These channels generate distinct empirical predictions for price dispersion, repricing frequency, and sensitivity to local demand conditions. In Section 9, we formalize this distinction in a parsimonious vertical pricing framework in which manufacturers optimally choose whether to generate market-specific pricing recommendations and transmit them through the retail chain.

6.1 Average price level

Table 5 presents the results of difference-in-differences regressions of Equation 2 with $\ln(\textit{Average price})$ as the dependent variable. The estimated baseline treatment effect, including UPC and calendar month fixed effects and no controls, is -0.059, which is statistically significant at the 1% level, suggesting that high-AI exposure firms reduced prices by approximately 5.9% relative to low-AI exposure firms after the InstructGPT API introduction. Column (2) adds $\ln(\textit{Units sold})$ as a control, column (3) adds pre-treatment $\ln(\textit{Employees})$ interacted with post, and column (4) adds the full set of pre-treatment firm financial controls interacted with post. The estimated treatment effect is negative and strongly significant across all specifications. In the most saturated specification (column (4)), the estimate is -0.048, suggesting a price decline of approximately 4.8%.

Figure 3 presents the corresponding event-study estimates from Equation 3. The pre-treatment coefficients are close to zero and statistically insignificant, which supports the parallel-trends assumption. The treatment effect begins to emerge in the month before the introduction, consistent with some early-adopting firms integrating AI tools following the expansion of access to the GPT-3 API in November 2021. The treatment effect persists through the end of the estimation window.

The decline in average prices is consistent with the production cost channel. Under that channel, a decrease in marginal production costs lowers optimal prices uniformly across markets. However, because the price-level result alone does not distinguish among channels, we turn to additional pricing outcomes whose predictions differ more sharply across mechanisms.

6.2 Price dispersion

We now estimate the effect of AI implementation on the dispersion of firms' product prices. We expect that if AI reduces the cost of gathering and processing local demand information, firms could tailor prices more precisely to local conditions, increasing the dispersion of firms' product prices across stores and markets. Table 6 presents the results. The dependent variable in columns (1) to (3) is the standard deviation of log prices within a firm-product-month, and in columns (4) to (6) is the interquartile range. For each measure, we compute dispersion across all stores (columns (1) and (4)), within counties (columns (2) and (5)), and within retailers (columns (3) and (6)).

The treatment effect is positive and statistically significant across all six specifications. In column (1), the standard deviation of log prices increases by 0.021 after the event for high-exposure firms relative to low-exposure firms, corresponding to approximately a 10% increase relative to the mean of 0.21. We find a larger treatment effect when dispersion is measured across all stores (columns (1) and (4)) than within counties or within retailers (columns (2) to (3) and (5) to (6)), suggesting that much of the increase in price dispersion occurs across geographic areas and retail chains. That said, the within-county and within-retailer effects are also statistically significant, suggesting that AI implementation also enables price changes within local markets. Figure 4 plots the dynamic treatment effects over the estimation window. The pre-treatment coefficients are close to zero for both the standard deviation of log prices (Panel A) and the interquartile range of log prices (Panel B), consistent with no differential pre-trends in price dispersion.

Of the two channels in our pricing framework, the information channel predicts an increase in geographic price dispersion, while the production channel does not. The information channel operates through two margins: lower costs of generating pricing recommendations expand the set of markets for which the manufacturer transmits market-specific guidance (extensive margin), and higher signal precision increases the weight placed on local information relative to priors (intensive margin). Both margins increase geographic price dispersion. In the language of DellaVigna and Gentzkow (2019), our results are consistent with AI eroding the informational constraints that give rise to uniform pricing across heterogeneous markets.

6.3 Price changes

We next examine whether AI implementation affects the frequency with which firms adjust prices. Table 7 presents the results. In columns (1) to (3), the dependent variable is an indicator for whether the month-over-month absolute log price change exceeds 5%. Columns (4) to (6) and (7) to (9) isolate price increases and decreases, respectively. For each dependent variable, we report results at the firm-product, firm-product-county, and firm-product-retailer levels.

Column (1) shows that the probability of an absolute price change exceeding 5% increases by 6.0 percentage points for high-AI exposure firms relative to low-AI exposure firms after the introduction of the InstructGPT API. Given a pre-treatment mean of 0.37, the estimated treatment effect represents approximately a 16% increase in repricing frequency. Columns (2) and (3) show consistent results when we examine changes in repricing frequency within counties and retailers. In columns (4) to (6), we decompose the effect by direction, and show that the frequency of price increases rises significantly at the county level, but the effect is small and insignificant at the firm-product and firm-product-retailer levels. By contrast, the increase in the frequency of price decreases is significant across all three levels of aggregation, with the probability of a price decrease exceeding 5% increasing by 4.9 percentage points at the firm-product level (column (7)). The overall increase in repricing frequency is thus predominantly, though not exclusively, driven by price decreases.

Figure 5 presents the estimated dynamic treatment effects. Panel A shows that the pre-treatment coefficients for the probability of an absolute price change exceeding 5% are close to zero, with a clear increase after the event. The directional decomposition reinforces the regression results. Panel B shows an increase in the frequency of price increases at the county level, though the effect is more muted at the firm-product and retailer levels. Panel C shows a broader post-event increase in the frequency of price decreases across all three levels of aggregation.

These results help distinguish between the channels in our framework. The production channel predicts lower average prices but no change in repricing frequency, and is therefore silent on the patterns in Table 7. The information channel predicts more frequent repricing, consistent with columns (1) through (3). Moreover, the increase in repricing is predominantly driven by price

decreases, which are significant across all three levels of aggregation. This asymmetry is consistent with better demand signals helping identify markets where firms were previously overcharging relative to local willingness to pay: improved information reveals downward pricing opportunities that were previously too costly for the manufacturer to identify.

In Section 7, we probe this interpretation further. As discussed in Section 9, only the information channel predicts that prices should become more sensitive to local economic conditions after the event, because higher signal precision causes the optimal price to load more heavily on the local demand signal and less on a common prior. The pricing effects should also be largest where the returns to better demand information are highest, for example, for products with more volatile or more elastic demand, where the cost of mispricing is potentially higher.

6.4 Robustness tests

We subject these results to a battery of robustness tests addressing two classes of concern. The first is whether the effects reflect manufacturer-level responses to AI or retailer-level pricing strategies. The second is that the InstructGPT API introduction in January 2022 coincided with other macroeconomic developments that could affect retail pricing. In particular, the Federal Reserve began raising interest rates in March 2022, and the economy was still adjusting to the effects of the COVID-19 pandemic. If AI exposure is correlated with firms' sensitivity to these shocks, the estimated treatment effects could reflect confounding macroeconomic forces rather than AI implementation. We address the retailer concern first and then each macroeconomic concern in turn. We also conduct robustness tests using alternative AI exposure measures, a continuous treatment measure, and several placebo tests.

Retailer-specific effects. A natural concern is that our results could reflect retailer-level pricing strategies rather than manufacturer-level responses to AI. Table 8 addresses this by progressively adding retailer×UPC and retailer×calendar month fixed effects, which absorb time-invariant retailer-product pricing differences and retailer-specific time trends such as chain-wide promotional cycles. The estimated treatment effects remain strongly significant across all three outcomes and are similar in magnitude to the baseline. Figure 6 decomposes the price-level treatment effect by

retailer: the estimate is negative for virtually every retailer in the data, consistent with a common upstream component, albeit with heterogeneous pass-through across retail chains.

COVID-19 exposure. As shown in Figure 2, the InstructGPT API introduction coincides with the tail end of the Omicron wave, and COVID-19 cases decline sharply through estimation window. If firms with high AI exposure also had workforces that were more (or less) exposed to pandemic-related workplace disruptions, COVID-19 effects could confound our estimates. We address this concern by including controls for COVID-19 exposure at the firm- and county-levels and present the results in Table IA.B.7. Panel A includes occupation-level measures of workplace physical proximity aggregated to the firm-level and measured in the month before treatment, including the Dingel and Neiman (2020) work-from-home index and an analogous measure from the American Time Use Survey, interacted with post. Panel B estimates the same specification using the firm-product-county-month panel and controls for log cumulative county-level COVID-19 cases and deaths. Both panels report results for all three main outcomes: log average prices, price dispersion, and repricing frequency. The estimated treatment effects remain statistically significant and consistent with the baseline results across all specifications. The COVID-19 controls are themselves small and statistically insignificant in all columns.

Interest rate exposure. The Federal Reserve began raising interest rates in March 2022, which increased borrowing costs and likely dampened demand in interest rate-sensitive sectors such as construction, real estate, and finance during our sample period. If products sold by high-AI exposure firms are disproportionately sold in counties with high employment shares in these sectors, the price decline we document could be driven by weaker local demand. To address this concern, we estimate Equation 2 with controls for county-level interest rate (IR) exposure. We measure IR exposure as the share of private employment in the construction, real estate, and finance industries (NAICS 23, 531, and 52), which we classify as interest-rate-sensitive, using BLS data. We control for both a time-varying and an ex-ante measure of IR exposure to account for potential endogenous changes in local industry composition. Table IA.B.8 presents the results. Across all specifications, the estimated treatment effects remain highly statistically significant and consistent with the base-

line estimates. The IR exposure controls are economically small, though statistically significant in some of the columns. The stability of the treatment effects provides reassurance that the results are not driven by differential exposure to the interest rate tightening cycle.

Alternative AI exposure measure. Finally, we test the robustness of the results to an alternative measure of AI exposure. As discussed in Section 3, we develop a target-judge model framework where GPT-4o-mini evaluates whether the text-davinci-003 model—whose capabilities more closely match those available at the time of the InstructGPT API introduction—can perform each O*NET task. Table IA.B.4, Table IA.B.5, and Table IA.B.6 replicate the baseline results for price levels, price dispersion, and repricing frequency using this alternative exposure measure. Across all specifications, the results are consistent with the baseline, and the corresponding event-study estimates display similar pre-trends and post-event dynamics.

We also verify that the results are not driven by the propensity score matching procedure or the binary treatment classification. Table IA.B.3 reports results using the full unmatched sample with the binary treatment indicator and the matched sample with continuous AI exposure measured in January 2020, interacted with post. The binary treatment effects in the unmatched sample are similar to those in the matched sample. Similarly, we find that all three effects are large and strongly significant using a continuous measure of AI exposure. Overall, the results confirm that the pricing effects scale with the intensity of AI exposure and are not artifacts of the propensity score matched sample or the binary treatment classification.

Placebo event dates. If our results reflect a genuine treatment effect of AI implementation, they should not appear when the event date is moved to a period with no plausible AI shock. Table IA.B.9 presents placebo tests using January 2019 (36 months before the actual event) and January 2021 (12 months before) as alternative event dates. For each placebo date, we re-estimate the baseline specification for all three main outcomes: the log of average prices, price dispersion, and the probability of an absolute price change exceeding 5%. The estimated treatment effects using these placebo dates are small, have inconsistent signs, and are statistically insignificant across all specifications. The absence of any effect at these placebo dates further supports the interpretation

that our main results are not driven by pre-existing differential trends between high- and low-AI exposure firms.

We further assess the statistical significance of our results using a randomization inference procedure. In each of 5,000 iterations, we randomly reassign treatment across firms, re-estimate the propensity score matching, and re-estimate the baseline specification in Equation 2. Figure IA.B.4 plots the distribution of placebo t-statistics for all three main outcomes. The true t-statistics, indicated by the dashed vertical lines, fall well into the tails of the permutation distributions. The two-sided permutation p-values are 0.011 for log average prices (Panel A), 0.015 for price dispersion (Panel B), and 0.009 for repricing frequency (Panel C). These results suggest that the estimated treatment effects are unlikely to be due to chance treatment assignment.

7 Mechanism: Suggestive Evidence and Heterogeneity

The results in Section 6 are consistent with a reduction in upstream informational frictions, but do not, on their own, isolate the underlying mechanism. In particular, several of the observed patterns—including changes in average price levels—can arise through multiple channels depending on the environment. In this section, we provide additional evidence to help interpret these results by examining how the effects vary with local economic conditions, across product characteristics, and across different dimensions of AI exposure.

These cross-sectional patterns shed some light on the role of information in pricing decisions. If AI primarily affects firms' ability to gather and process demand information, we should expect pricing behavior to become more responsive to local conditions, with effects concentrated in settings where the value of information is highest. We should also expect the effects to operate through the augmentation of worker decision-making rather than through the automation of routine tasks. We evaluate these predictions in turn.

7.1 Price sensitivity to local economic conditions

Table 9 examines whether AI implementation increases the sensitivity of prices to local economic conditions. We augment the specification in Equation 2 with a triple interaction that allows the

treatment effect to vary with the county unemployment rate. Specifically, we interact $Treated \times Post$ with a lagged indicator for whether the county unemployment rate exceeds the cross-county median.

The coefficient on the triple interaction is -0.005 and statistically significant across all four columns. This implies that, after the event, high-AI exposure firms set lower prices in counties with above-median unemployment than in counties with below-median unemployment, relative to control firms. The result holds when we add controls (column 2), replace separate county and calendar month fixed effects with county×calendar month fixed effects (column 3), and further saturate the specification with firm-product×calendar month fixed effects (column 4), which absorb any firm-product-level time trends and isolate within-product geographic variation. The stability of the triple interaction coefficient across these increasingly demanding specifications suggests that AI implementation enables firms to incorporate local economic information at a more granular level into pricing decisions. These results provide evidence consistent with the information channel.

7.2 Product-level heterogeneity

We next investigate whether the treatment effects vary across product categories. Under the information channel, the value of improved demand signals should be highest when demand conditions are more uncertain or when consumers are more price-sensitive. In such cases, the cost of mispricing is higher, and firms have stronger incentives to adjust prices in response to better information. We therefore expect the effects on pricing behavior to be concentrated in product categories where the returns to improved demand information are highest.

Specifically, we interact $Treated \times Post$ with indicator variables for discretionary products, products with high demand volatility, and products with high price elasticity. Using Nielsen product group codes, we classify products as discretionary when they are plausibly more substitutable or non-essential, such as alcohol and consumer electronics. The intuition is that demand for these products is more sensitive to changes in relative prices and local conditions, increasing the returns to improved demand information. We measure demand volatility using the coefficient of variation of pre-treatment units sold within each product module. We estimate price elasticities from pre-

treatment regressions of log quantity changes on log price changes within product modules, and classify products as high-elasticity if their estimated elasticity is above the median.

The estimated coefficients on the triple interaction terms are negative and statistically significant in all specifications, and suggest that the pricing effects are concentrated among discretionary products, products with above-median demand volatility, and products with above-median price elasticity. Overall, these cross-sectional patterns are consistent with the information channel. Under that channel, we should expect the treatment effect to be concentrated in products with more uncertain or more elastic demand, and where the cost of mispricing is higher. This is precisely where improved demand signals from AI adoption should be most valuable.

7.3 Augmentation vs. automation

The previous results are consistent with AI adoption improving firms' ability to process and use demand information. If the mechanism operates through better information rather than task replacement, the pricing effects should be strongest in occupations where AI augments workers' productivity rather than automating their tasks (Autor, 2015; Acemoglu and Restrepo, 2022; Autor and Thompson, 2025). To test this, we decompose the baseline AI exposure measure into augmentation and automation components. As described in Section 2, we score each O*NET task on two dimensions using GPT-5-mini: augmentation, which captures the extent to which AI increases worker productivity for tasks that still require human judgment, and automation, which captures the extent to which AI could perform the task independently. We aggregate these scores to the firm level using employment-share weights and classify firms as having high augmentation or high automation exposure based on the median of each score.

Pricing decisions typically involve interpreting noisy local demand signals and incorporating knowledge of customers and competitors. AI can improve the speed and accuracy with which this information is analyzed, but the final pricing decision may still require human judgment. Therefore, under the information channel, we expect the effects to load on augmentation exposure, where AI enhances employees' ability to gather and interpret market information. By contrast, if the observed price changes were driven by automation of routine pricing tasks, the effects should instead load

on automation exposure.

Table 11 reports estimates of Equation 2 separately for firms with above-median augmentation exposure in columns (1) to (3) and above-median automation exposure in columns (4) to (6), for the three main outcomes. The results are driven entirely by augmentation exposure. Column (1) indicates that prices decline by 4.6% for firms with high augmentation exposure. By contrast, column (4) shows that the corresponding estimate for firms with high automation exposure is small at -0.013 and statistically insignificant.

We find a similar pattern for the price dispersion and repricing frequency results. The estimated treatment effects for augmentation exposure in columns (2) and (3) are positive and consistent with the baseline results, whereas the estimated coefficients for automation exposure in columns (5) and (6) are not statistically different from zero. Taken together, these results are consistent with AI strengthening employees' ability to interpret and act on demand information, rather than simply automating routine pricing activities.

8 Evidence of AI adoption

Our identification strategy exploits cross-sectional variation in pre-existing workforce exposure to generative AI as a proxy for the likelihood of adopting AI following the InstructGPT API introduction on January 27, 2022. It is thus important to examine whether high-AI exposure firms, as identified by our measure, actually exhibit higher AI adoption. Although AI adoption is difficult to measure, we provide evidence from several tests that is consistent with higher AI adoption among treated firms after the event.

We proxy for AI adoption in two complementary ways. First, we match firms in our sample to job posting data from LinkUp, which provides broad coverage of online vacancies. We then classify each job posting into five AI-related categories based on keywords in job titles and descriptions. Specifically, we create indicator variables that equal one for jobs we classify as broadly related to AI, applied AI, which includes jobs in predictive analytics, data science, and postings with O*NET occupational classifications identified as AI-related. We also create an indicator variable that is the

union of these four variables.[8]

Second, using textual data from firms' 10-K annual filings obtained from the SEC's EDGAR database, we measure the frequency of AI-related terms in corporate disclosures. Public filings provide a standardized way for firms to discuss material investments, operational changes, and emerging risks, making them a natural source of data to capture changes in AI strategy and implementation. Specifically, we count the frequency of AI-related terms such as artificial intelligence, machine learning, and automation in the full filing and separately within Item 1 (Business Description), Item 1A (Risk Factors), and Item 7 (Management's Discussion and Analysis). We also distinguish between AI adoption and implementation from discussions of AI risks by separating keywords that may reflect AI adoption and hiring from those that may reflect discussion of AI-related risks and regulation.

Table 12 presents the results. Panel A examines AI-related job postings, and Panel B examines corporate disclosures. In Panel A, we find a highly statistically significant increase in the monthly counts of AI-active job postings across all the different measures. High-exposure firms post approximately 8 more AI positions (column (1)), 4.4 more Applied AI and Data science positions each (columns (2) and (3)), and 22 more positions in the All AI category (column (5)) per month relative to low-exposure firms after the introduction of the InstructGPT API. The coefficients are also economically large. Overall, these findings suggest a reallocation of hiring toward AI-related skills in the first half of 2022, consistent with an increase in AI adoption among treated firms.

In column (1) of Panel B, we find that high-AI exposure firms mention AI-related keywords 1.5 more times per 10-K filing after the event. Columns (2) and (3) show that this increase is concentrated in Items 1 and 1A. The concentration of mentions in these sections suggests that AI is discussed in relation to how the firm operates and the risks it faces. However, we find no significant change in column (4) for Item 7, Management's Discussion and Analysis.

Panel C shows results from difference-in-differences regressions that separately estimate the treatment effects on keywords focused on implementation and hiring versus those focused on risk and regulatory terms. The implementation and hiring category captures keywords related to AI

[8] Appendix A describes the keywords and classification criteria for each indicator variable.

adoption, deployment, and workforce investment, whereas the risk and regulatory category captures keywords related to AI-related threats, compliance, and uncertainty. The increase in AI-related discussion in public disclosures is stronger for implementation and hiring terms compared to risk and regulatory terms. For example, the estimated treatment effect for total implementation keywords is 1.03 in column (1), compared with 0.08 for risk and regulatory keywords in column (4). For example, the estimated treatment effect for total implementation keywords is 1.03 in column (1), compared with 0.08 for risk and regulatory keywords in column (4). The increase in AI-related discussion is thus almost entirely driven by adoption and deployment language, not by discussion of AI-related risks or regulatory concerns. Overall, we find that treated firms discuss AI more often in their public disclosures after the event, particularly in the context of adoption.[9]

Figure 7 presents the dynamic treatment effects corresponding to columns (1) and (4) of Table 12, Panels A and B. Panel A plots the dynamic treatment effects on AI-related job postings from LinkUp at the monthly frequency, and Panel B plots the dynamic treatment effects on 10-K AI keyword counts at the annual frequency. In both panels, the pre-event coefficients are close to zero and statistically insignificant. The post-event coefficients indicate a persistent increase in AI-related job postings and AI-related disclosure following the introduction of the InstructGPT API.

To verify that these patterns reflect AI-specific adoption rather than a broader increase in technology-related discussion, we conduct a placebo test using non-AI technology keywords (e.g., cloud computing, ERP, CRM, cybersecurity, e-commerce infrastructure). Figure IA.B.6 in the Internet Appendix shows no differential change in the use of these terms in earnings calls or 10-K filings following the InstructGPT API rollout, suggesting that the increase in AI-related language is specific to artificial intelligence rather than part of a more general rise in technology-related discussion. Taken together, the results in this section are consistent with our AI exposure measure capturing meaningful differences in AI adoption.

[9] We also examine the use of AI-related keywords in quarterly earnings conference call transcripts. Although conference call discussions are voluntary, they complement the 10-K analysis. Figure IA.B.5 and Table IA.B.11 in the Internet Appendix reports these results. We find that treated firms significantly increase the use of AI keywords in earnings calls after the event.

9 Interpreting the Evidence: A Vertical Pricing Framework

The empirical results point to a common mechanism: pricing decisions are shaped by upstream constraints in generating and transmitting market-specific information. Manufacturer identity accounts for the majority of systematic pricing variation, and pricing becomes more localized when the vertical information asymmetry is attenuated. Reductions in information-processing costs shift pricing away from uniformity and toward localization.

To synthesize these findings, we develop a parsimonious vertical pricing framework that emphasizes two sequential decisions. First, the manufacturer generates pricing guidance based on its information about demand. Second, the retailer implements that guidance, subject to operational constraints. This interaction highlights a central asymmetry: manufacturers observe demand across many geographically dispersed markets and can, in principle, aggregate this information into pricing recommendations, while retailers observe local conditions but rely on upstream guidance to coordinate pricing across products and locations.

Translating this demand intelligence into actionable, market-specific pricing recommendations is nonetheless costly. It requires gathering and processing local demand signals, then transmitting them through organizational or contractual channels to retailers, who must then act on them. The private-label evidence is consistent with this view: retailers are fully capable of implementing localized pricing when they possess the relevant demand information directly, suggesting that the friction lies upstream in generating market-specific pricing recommendations rather than downstream in retail implementation. The model formalizes this upstream information friction and its implications for pricing behavior across markets.

Model Setup. A manufacturer sells a product through retailers operating across N segmented markets. In market $i \in \{1, \ldots, N\}$, demand is $Q_i(P_i) = (\alpha_i - P_i)/\beta_i$, where α_i summarizes local demand conditions and $\beta_i > 0$ parameterizes demand-price sensitivity. For simplicity, we assume that β_i is commonly known, but there is uncertainty about $\alpha_i \sim N(\mu_\alpha, \sigma_\alpha^2)$, which is the source of the information friction. The manufacturer produces at a constant marginal cost c.

Manufacturer's Problem. The manufacturer can generate pricing recommendations by incurring a fixed cost I, which captures broadly the cost of gathering and processing information into pricing recommendations that can be communicated through the vertical chain—via wholesale pricing schedules, trade promotions, minimum advertised price policies, or category captain arrangements.[10] When the manufacturer pays I, the retailer receives a signal $s_i = \alpha_i + \eta_i$, where $\eta_i \overset{i.i.d.}{\sim} N(0, \frac{1}{\phi}\sigma_\alpha^2)$. The retailer's posterior mean becomes

$$\hat{\alpha}_i = \mu_\alpha + \Phi(s_i - \mu_\alpha), \qquad \Phi = \frac{\phi}{1+\phi}, \tag{4}$$

where Φ summarizes the informativeness of the manufacturer's recommendation. When the manufacturer does not pay I, no signal is transmitted and the retailer prices on the prior alone. The manufacturer pays the cost I whenever the resulting expected gains exceed the cost.

Retailer's Problem. The retailer sets prices in two stages to maximize the joint manufacturer-retailer surplus. First, prior to receiving information, the retailer sets a uniform price:

$$\bar{P} = \tfrac{\mu_\alpha + c}{2}, \tag{5}$$

which is the standard monopoly price maximizing the expected joint manufacturer-retailer surplus given the prior. This uniform price serves as the retailer's default across all markets in the absence of communication from the manufacturer.

When the manufacturer transmits signals s_i, the retailer updates its beliefs about α_i and faces a trade-off between the gains from more efficient pricing and the cost of implementing price changes. Deviating from the uniform price $\bar{P}$ incurs a quadratic implementation cost $\frac{k}{2}(P_i - \bar{P})^2$, capturing menu costs, organizational inertia, and the costs of managing price variation across a store's assortment. The retailer's optimal shelf price upon receiving signal s_i is

$$P_i^* = (1 - \lambda_i)\bar{P} + \lambda_i \left(\tfrac{\hat{\alpha}_i + c}{2}\right), \tag{6}$$

[10] This maps onto the "managerial costs" documented by Zbaracki et al. (2004), who find that the informational and deliberation costs of price adjustment are roughly six times larger than the physical costs of changing price tags.

where $\hat{\alpha}_i$ is the posterior mean and $\lambda_i = \frac{2}{2+k\beta_i}$ is the optimal weight placed on the updated information relative to the uniform benchmark. A higher implementation cost (k) lowers the optimal weight (λ). In the extreme ($k \to \infty$), implementation costs drive the optimal shelf price (P_i^*) to $\bar{P}$, regardless of the signal. A lower implementation cost increases the optimal weight. In the extreme ($k \to 0$), the optimal shelf price fully adjusts to reflect the updated posterior.

Upstream Friction. The sensitivity of the shelf price to local demand conditions depends on whether the manufacturer pays the cost of generating pricing recommendations (I). When the manufacturer does not generate pricing recommendations, the retailer leaves prices at $\bar{P}$, meaning that shelf prices are not responsive to local conditions. When the manufacturer pays I, the retailer receives signals s_i, and the sensitivity of shelf prices to local demand becomes

$$\frac{\partial P_i^*}{\partial \alpha_i} = \lambda_i \cdot \frac{\Phi}{2}. \tag{7}$$

The manufacturer's decisions thus affect pricing sensitivity through two margins: the extensive margin, whether to transmit a recommendation at all, and the intensive margin, the informativeness of pricing recommendations. Both are upstream decisions that determine how much local demand information reaches the shelf, and both are affected by AI adoption.

Although the framework is static, we interpret it as the per-period problem of the manufacturer and retailer: each period, market i draws a new α_{it}, and the manufacturer decides whether to transmit a signal, and the retailer decides how much to update prices. Under this interpretation, the within-market price variance $Var(P_i^*) = (\frac{\lambda_i \Phi}{2})^2 \sigma_\alpha^2$ corresponds to time-series price variability, and changes that increase responsiveness to signals also increase repricing frequency.

9.1 AI Adoption and Empirical Predictions

AI can affect pricing through two channels. The following table summarizes the empirical predictions across these channels.

No single channel accounts for the full set of empirical results. A reduction in marginal production costs ($c \downarrow$) shifts the uniform default $\bar{P} = (\mu_\alpha + c)/2$ downward, reducing average prices

	Lower c (Production)	Lower I / Higher Φ (Information)
Average price declines	✓	×
Geographic dispersion increases	×	✓
Repricing frequency increases	×	✓
Local demand sensitivity increases	×	✓

uniformly across markets, but leaving dispersion, repricing frequency, and local price-demand sensitivity unchanged. Lower costs of generating pricing recommendations ($I \downarrow$) expand the set of markets for which the manufacturer transmits a recommendation, increasing dispersion and repricing frequency on the extensive margin, but leave the uniform default $\bar{P}$ unchanged and therefore do not affect average prices. Higher signal precision ($\Phi \uparrow$) increases the weight the retailer places on new information, further raising dispersion and repricing frequency on the intensive margin, and increasing sensitivity to local demand conditions. The full set of empirical patterns—lower average prices, greater geographic dispersion, higher repricing frequency, and increased sensitivity to local economic conditions—is therefore consistent with AI operating through both channels simultaneously.

10 Conclusion

The uniform pricing puzzle has traditionally been framed as a question about retail operations: why don't chains differentiate prices across locations when local conditions differ so markedly? This paper provides evidence that the answer lies upstream. Variance decompositions across 1.83 billion store-product-month observations show that manufacturer identity accounts for 81% of systematic price level variation, 73% of price dispersion variation, and 83% of price change variation. Private label products—which bypass the manufacturer-retailer information channel—exhibit significantly greater geographic price dispersion than national brands within the same retailer, category, and month. And when upstream informational frictions are reduced by the introduction of generative AI tools, prices fall, geographic dispersion increases, repricing becomes more frequent, and prices become more responsive to local economic conditions. These effects load on augmentation rather

than automation exposure and survive retailer-by-time fixed effects, consistent with AI improving the flow of demand information through the vertical chain rather than automating retail execution.

The central implication is that the pricing rigidities documented in the literature are not immutable features of retail markets. They reflect, in substantial part, the cost of generating and communicating market-specific pricing information upstream—a cost that technology can reduce. Future work with richer cost data, retailer-level exposure measures, and structural estimation of the vertical pricing relationship could quantify the welfare consequences of uniform pricing and the gains from relaxing these frictions at scale.

References

Abowd, J. M., F. Kramarz, and D. N. Margolis. 1999. High wage workers and high wage firms. *Econometrica* 67:251–333. doi:10.1111/1468-0262.00020.

Acemoglu, D., and P. Restrepo. 2022. Tasks, automation, and the rise in U.S. wage inequality. *Econometrica* 90:1973–2016.

Agrawal, A., J. S. Gans, and A. Goldfarb. 2019. Artificial intelligence: The ambiguous labor market impact of automating prediction. *Journal of Economic Perspectives* 33:31–50.

Alvarez, F., F. Lippi, and L. Paciello. 2011. Optimal price setting with observation and menu costs. *Quarterly Journal of Economics* 126:1909–60.

Appel, R., M. Massenkoff, P. McCrory, M. McCain, R. Heller, T. Neylon, and A. Tamkin. 2026. Anthropic economic index report: Economic primitives. Working Paper, Anthropic.

Assad, S., R. Clark, D. Ershov, and L. Xu. 2024. Algorithmic pricing and competition: Empirical evidence from the german retail gasoline market. *Journal of Political Economy* 132:723–71.

Autor, D., and N. Thompson. 2025. Expertise. Working Paper, Massachusetts Institute of Technology. Joseph R. Schumpeter Lecture, European Economic Association, August 2024. Revised June 2025.

Autor, D. H. 2015. Why are there still so many jobs? the history and future of workplace automation. *Journal of Economic Perspectives* 29:3–30.

Autor, D. H., F. Levy, and R. J. Murnane. 2003. The skill content of recent technological change: An empirical exploration. *Quarterly Journal of Economics* 118:1279–333.

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

Barry, J., M. Campello, J. R. Graham, and Y. Ma. 2022. Corporate flexibility in a time of crisis. *Journal of Financial Economics* 144:780–806.

Bena, J., L. Garlappi, P. Ma, and D. Tanguay. 2024. Artificial intelligence and corporate investment. Working paper.

Bergemann, D., and A. Bonatti. 2019. Markets for information: An introduction. *Annual Review of Economics* 11:85–107.

Brynjolfsson, E., D. Li, and L. Raymond. 2023. Generative AI at work. Working Paper 31161, NBER.

Brynjolfsson, E., D. Rock, and C. Syverson. 2021. The productivity j-curve: How intangibles complement general purpose technologies. *American Economic Journal: Macroeconomics* 13:333–72.

Calvano, E., G. Calzolari, V. Denicolo, and S. Pastorello. 2020. Artificial intelligence, algorithmic pricing, and collusion. *American Economic Review* 110:3267–97.

Card, D., J. Heining, and P. Kline. 2013. Workplace heterogeneity and the rise of West German wage inequality. *Quarterly Journal of Economics* 128:967–1015. doi:10.1093/qje/qjt006.

Chintagunta, P. K., A. Bonfrer, and I. Song. 2002. Investigating the effects of store-brand introduction on retailer demand and pricing behavior. *Management Science* 48:1242–67.

Décaire, P. H., and D. Sosyura. 2024. On a spending spree: The real effects of heuristics in managerial budgets. Working paper, Arizona State University.

DellaVigna, S., and M. Gentzkow. 2019. Uniform pricing in U.S. retail chains. *Quarterly Journal of Economics* 134:2011–84.

Dingel, J. I., and B. Neiman. 2020. How many jobs can be done at home? *Journal of Public Economics* 189:104235–.

Dubé, J.-P., and S. Misra. 2023. Personalized pricing and consumer welfare. *Journal of Political Economy* 131:131–89.

Eichenbaum, M., N. Jaimovich, and S. Rebelo. 2011. Reference prices, costs, and nominal rigidities. *Quarterly Journal of Economics* 126:1727–72.

Eisfeldt, A., G. Schubert, B. Taska, and M. B. Zhang. 2025. Generative AI and firm values. *Journal of Finance, forthcoming* .

Goldfarb, A., and C. Tucker. 2019. Digital economics. *Journal of Economic Literature* 57:3–43.

Golosov, M., and R. E. Lucas. 2007. Menu costs and phillips curves. *Journal of Political Economy* 115:171–99.

Hoch, S. J., and S. Banerji. 1993. When do private labels succeed? *Sloan Management Review* 34:57–67.

Hoch, S. J., B.-D. Kim, A. L. Montgomery, and P. E. Rossi. 1995. Determinants of store-level price elasticity. *Journal of Marketing Research* 32:17–29.

Klenow, P. J., and O. Kryvtsov. 2008. State-dependent or time-dependent pricing: Does it matter for recent u.s. inflation? *Quarterly Journal of Economics* 123:863–904.

Klenow, P. J., and B. A. Malin. 2010. Microeconomic evidence on price-setting. In B. M. Friedman and M. Woodford, eds., *Handbook of Monetary Economics*, vol. 3, 231–84. Elsevier.

Levy, D., M. Bergen, S. Dutta, and R. Venable. 1997. The magnitude of menu costs: Direct evidence from large U.S. supermarket chains. *Quarterly Journal of Economics* 112:791–824.

Mackowiak, B., and M. Wiederholt. 2009. Optimal sticky prices under rational inattention. *American Economic Review* 99:769–803.

Mankiw, N. G., and R. Reis. 2002. Sticky information versus sticky prices: A proposal to replace the new keynesian phillips curve. *Quarterly Journal of Economics* 117:1295–328.

Midrigan, V. 2011. Menu costs, multiproduct firms, and aggregate fluctuations. *Econometrica* 79:1139–80.

Nakamura, E., and J. Steinsson. 2008. Five facts about prices: A reevaluation of menu cost models. *Quarterly Journal of Economics* 123:1415–64.

Sims, C. A. 2003. Implications of rational inattention. *Journal of Monetary Economics* 50:665–90.

Stroebel, J., and J. Vavra. 2019. House prices, local demand, and retail prices. *Journal of Political Economy* 127:1391–436.

Webb, M. 2020. The impact of artificial intelligence on the labor market. Working paper, Stanford University.

Zbaracki, M. J., M. Ritson, D. Levy, S. Dutta, and M. Bergen. 2004. Managerial and customer costs of price adjustment: Direct evidence from industrial markets. *Review of Economics and Statistics* 86:514–33.

Figure 1: Distribution of AI exposure

This figure plots the cross-sectional distribution of firm-level generative AI exposure measured in January 2020. AI exposure is computed as the employment-share-weighted average of occupation-level generative AI scores from Eisfeldt et al. (2025), where scores capture the extent to which generative AI is expected to reduce the time required to complete the tasks typically performed by each occupation. Occupation-level scores are aggregated to the firm level using each firm's employment shares across occupations. Higher values indicate greater expected exposure to generative AI.

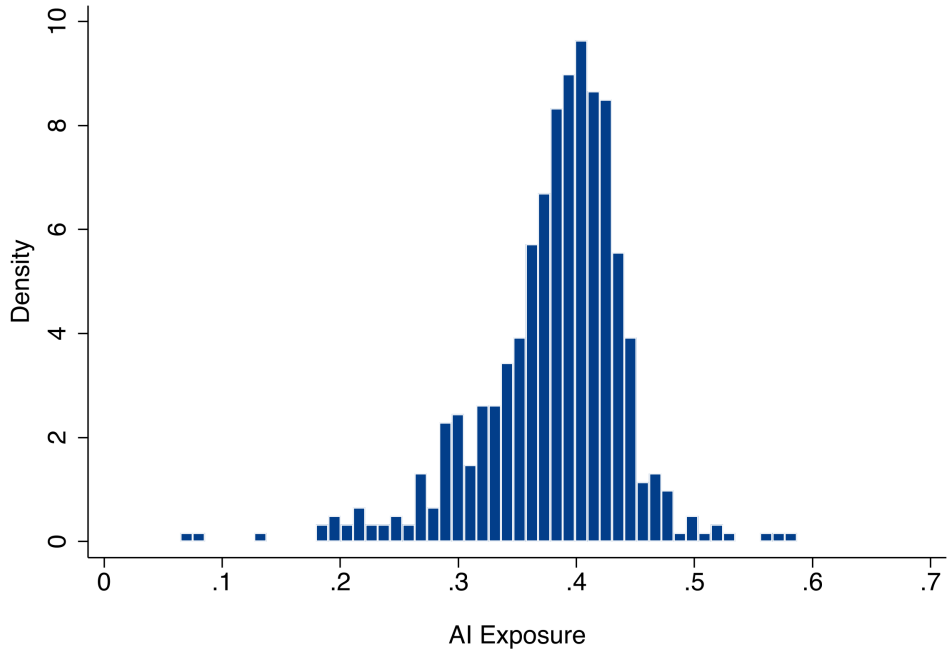
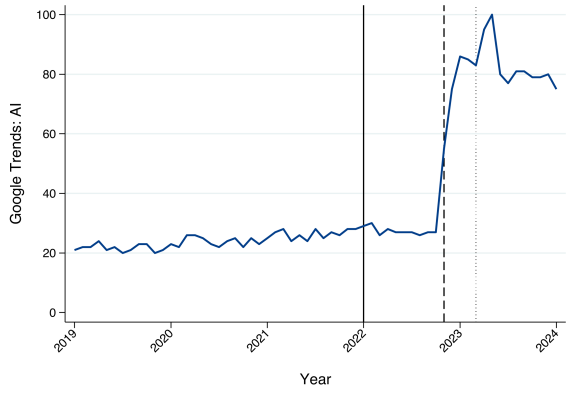


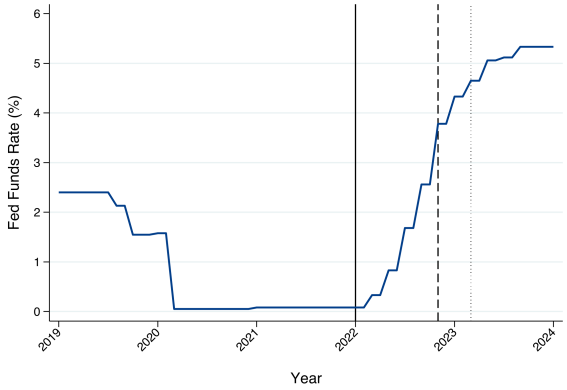
Figure 2: Macro environment around AI implementation events

This figure plots three time series around the InstructGPT API introduction in January 2022. Panel A shows Google Trends search interest for "artificial intelligence" (index, 100 = peak). Search interest remained flat throughout the estimation window, spiking only after the consumer-facing ChatGPT release in November 2022, consistent with the InstructGPT API being a firm-facing event with no immediate effect on public AI awareness. Panel B shows the effective federal funds rate (%). Panel C shows monthly new COVID-19 cases (in millions). In each panel, the solid black vertical line marks the InstructGPT API introduction (January 27, 2022), the dashed gray line marks the ChatGPT public release (November 30, 2022), and the dotted gray line marks the GPT-3.5/4 API introductions (March 2023).

Panel A: General AI search interest



Panel B: Federal Reserve policy



Panel C: COVID-19 waves

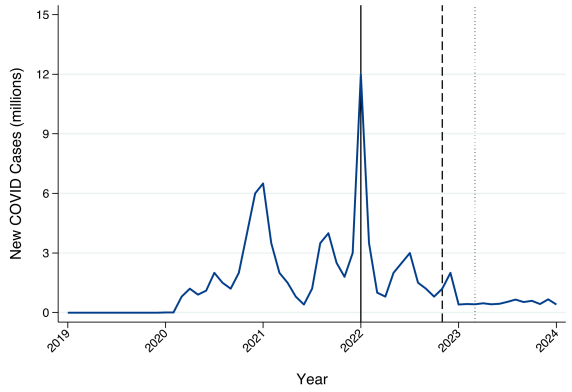


Figure 3: Price level event study

This figure plots the β_k coefficients and 90% confidence intervals from event study regressions of the form: $\ln(\text{Average price})_{ipt} = \alpha_{ip} + \alpha_t + \sum_{k \neq -3} \beta_k \mathbb{1}[\text{Event time} = k] \times \text{Treated}_i + \epsilon_{ipt}$, where i denotes firms, p products (UPCs), and t calendar months. The dependent variable is the log of CPI-deflated average price. AI exposure is computed as the employment-share-weighted average of occupation-level generative AI scores from Eisfeldt et al. (2025), where scores capture the extent to which generative AI is expected to reduce the time required to complete the tasks typically performed by each occupation. *Treated* equals one for firms with above-median overall AI exposure measured in January 2020. Event time runs from −6 to +6 months relative to the InstructGPT API introduction in January 2022, with $k = -3$ as the omitted base period. α_{ip} and α_t denote firm-product and calendar month fixed effects, respectively. Treated and control firm products are balanced on the pre-treatment log average price and log units sold. Standard errors are clustered at the firm level. Appendix A presents variable definitions.

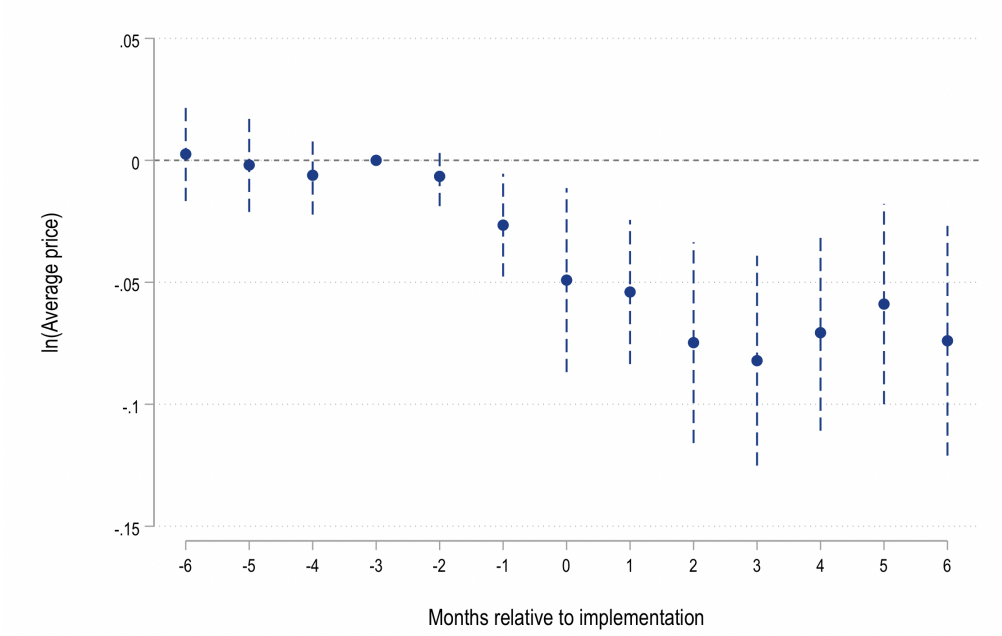
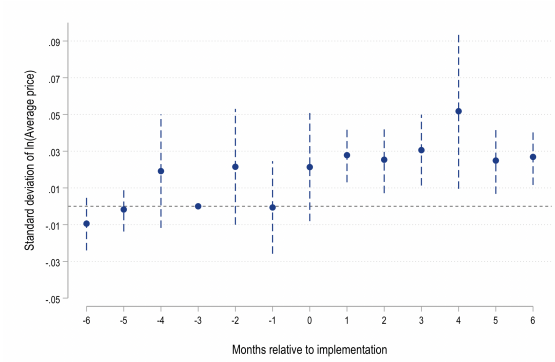


Figure 4: Price dispersion event study

This figure plots the β_k coefficients and 90% confidence intervals from event study regressions of the form: $\text{Dispersion}_{ipt} = \alpha_{ip} + \alpha_t + \sum_{k \neq -3} \beta_k \mathbb{1}[\text{Event time} = k] \times \text{Treated}_i + \epsilon_{ipt}$, where i denotes firms, p products (UPCs), and t calendar months. In Panel A, the dependent variable is the standard deviation of log prices across stores within a firm-product-month. In Panel B, the dependent variable is the interquartile range of log prices across stores within a firm-product-month. AI exposure is computed as the employment-share-weighted average of occupation-level generative AI scores from Eisfeldt et al. (2025), where scores capture the extent to which generative AI is expected to reduce the time required to complete the tasks typically performed by each occupation. *Treated* equals one for firms with above-median AI exposure measured in January 2020. Event time runs from -6 to $+6$ months relative to the InstructGPT API introduction in January 2022, with $k = -3$ as the omitted base period. α_{ip} and α_t denote firm-product and calendar month fixed effects, respectively. Treated and control firm products are balanced on the pre-treatment log average price and log units sold. Standard errors are clustered at the firm level. Appendix A presents variable definitions.

Panel A: Standard deviation



Panel B: Interquartile range

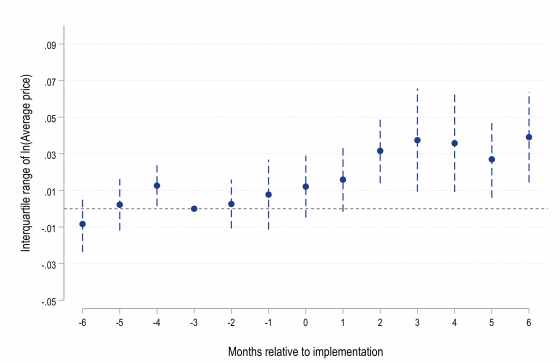
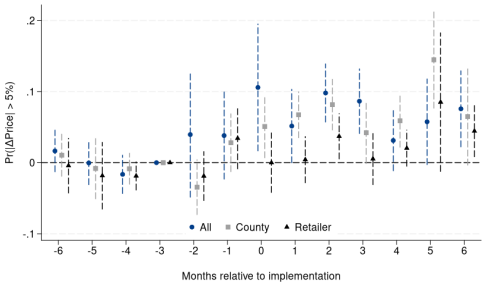


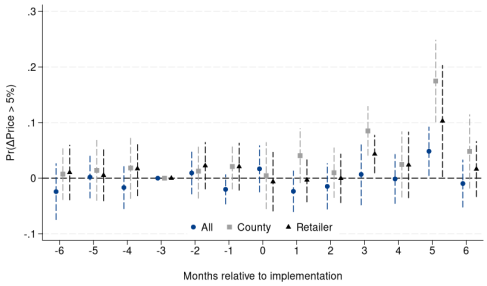
Figure 5: Price changes event study

This figure plots the β_k coefficients and 90% confidence intervals from event study regressions of the form: $\mathbb{P}(|\Delta \ln P| > 5\%)_{ipt} = \alpha_{ip} + \alpha_t + \sum_{k \neq -3} \beta_k \mathbb{1}[\text{Event time} = k] \times \text{Treated}_i + \epsilon_{ipt}$, where i denotes firms, p products (UPCs), and t calendar months. The dependent variable is an indicator for whether the month-over-month log price change exceeds a given threshold. In Panel A, the dependent variable is an indicator for whether the absolute log price change exceeds 5%. Panels B and C focus on price increases and decreases, respectively. AI exposure is computed as the employment-share-weighted average of occupation-level generative AI scores from Eisfeldt et al. (2025), where scores capture the extent to which generative AI is expected to reduce the time required to complete the tasks typically performed by each occupation. *Treated* equals one for firms with above-median AI exposure measured in January 2020. Event time runs from −6 to +6 months relative to the InstructGPT API introduction in January 2022, with $k = -3$ as the omitted base period. α_{ip} and α_t denote firm-product and calendar month fixed effects, respectively. Treated and control firm products are balanced on the pre-treatment log average price and log units sold. Standard errors are clustered at the firm level. Appendix A presents variable definitions.

Panel A: $\mathbb{P}(|\Delta P| > 5\%)$



Panel B: $\mathbb{P}(\Delta P > 5\%)$



Panel C: $\mathbb{P}(\Delta P < -5\%)$

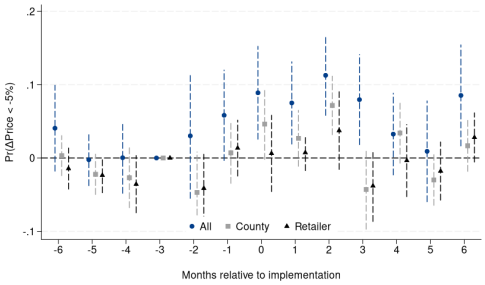


Figure 6: Treatment effect heterogeneity across retailers

This figure plots the estimated β_k coefficients and 95% confidence intervals from a regression that augments the baseline price-level specification with retailer-specific interactions: $\ln(\text{Average price})_{ipt} = \alpha_{ip} + \alpha_t + \sum_k \beta_k (\textit{Treated}_i \times \textit{Post}_t \times \mathbb{1}[\textit{Retailer} = k]) + \mathbf{X}'_{i0}\gamma \times \textit{Post}_t + \epsilon_{ipt}$, where i denotes firms, p products (UPCs), t calendar months, and k indexes retailers. Retailers are ranked along the horizontal axis from the largest (most negative) to the smallest estimated treatment effect. The gray dashed horizontal line indicates the pooled *Treated* × *Post* estimate from the baseline specification. AI exposure is computed as the employment-share-weighted average of occupation-level generative AI scores from Eisfeldt et al. (2025). *Treated* equals one for firms with above-median AI exposure measured in January 2020. *Post* equals one during the six months following the InstructGPT API introduction in January 2022. All specifications include firm-product and calendar month fixed effects and controls for log units sold, log employees interacted with post, and firm financial characteristics interacted with post. Standard errors are clustered at the firm level. Appendix A presents variable definitions.

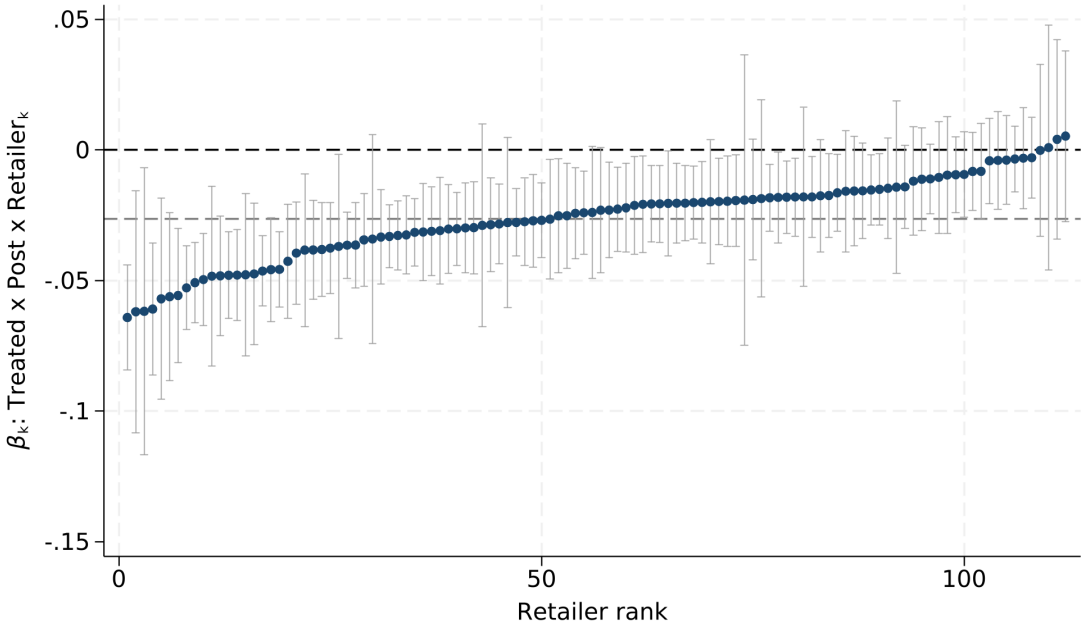
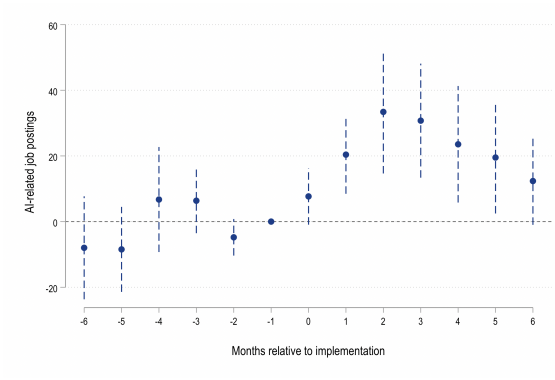


Figure 7: Evidence of AI adoption

This figure plots the β_k coefficients and 90% confidence intervals from event study regressions of the form: $Y_{it} = \alpha_i + \alpha_t + \sum_{k \neq -1} \beta_k \mathbb{1}[\text{Event time} = k] \times \text{Treated}_i + \epsilon_{it}$, where i denotes firms and t denotes calendar periods. In Panel A, the dependent variable is the count of active AI-related job postings from LinkUp, and event time is measured in months relative to the InstructGPT API introduction in January 2022. In Panel B, the dependent variable is the count of AI-related keywords in annual 10-K filings, and event time is measured in years relative to 2022. AI exposure is computed as the employment-share-weighted average of occupation-level generative AI scores from Eisfeldt et al. (2025). *Treated* equals one for firms with above-median AI exposure measured in January 2020. α_i and α_t denote firm and calendar period fixed effects, respectively, with $k = -1$ as the omitted base period. Standard errors are clustered at the firm level. Appendix A presents variable definitions.

Panel A: AI-related job postings (LinkUp)



Panel B: AI keywords in 10-K filings

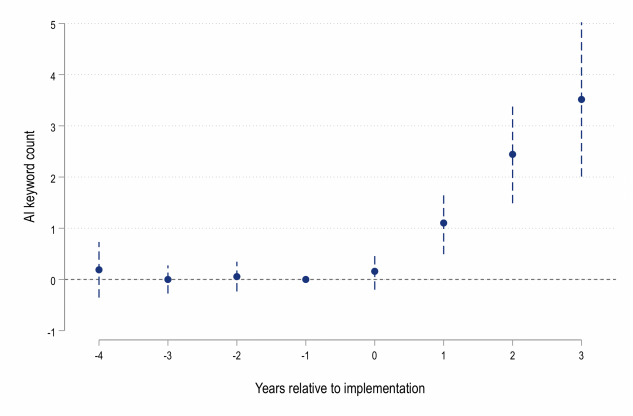


Table 1: Summary Statistics

This table reports summary statistics. Panel A reports the outcome variables used in the price variance decomposition (Table 3), measured at the store×UPC×month level across the full Nielsen Retail Scanner sample. Panel B reports the outcome variable used in the private label tests (Table 4). Panel C reports pricing outcomes for the matched AI experiment sample. Panel D reports the firm-level AI exposure measure. Panel E reports firm characteristics measured in the quarter before the InstructGPT API introduction in January 2022. Panel F reports covariate balance between treated and control firms. Appendix A presents variable definitions.

	Obs	Mean	Std. Dev.	25th Pct.	Median	75th Pct.
	(1)	(2)	(3)	(4)	(5)	(6)
Panel A: Price variance decomposition						
ln(Average price)	1.83B	1.890	0.810	1.380	1.920	2.380
$(\ln p_{ist} - \overline{\ln p}_{it})^2$	1.83B	0.068	0.920	0.002	0.010	0.032
Δln(Price)	1.60B	0.000	0.489	−0.049	0.000	0.062
Panel B: Private label sample						
SD of ln(Average price), private label	22.7M	0.114	0.237	0.009	0.050	0.124
SD of ln(Average price), national brand	196.0M	0.097	0.206	0.000	0.041	0.114
Panel C: Pricing outcomes						
Ln(Average price)	90,211	1.797	0.924	1.194	1.795	2.407
SD of Ln(Average price)	78,341	0.206	0.189	0.084	0.161	0.264
IQR of Ln(Average price)	90,211	0.144	0.174	0.000	0.088	0.220
P($\lvert\%\Delta P\rvert > 5\%$)	81,461	0.374	0.484	0.000	0.000	1.000
P($\lvert\%\Delta P\rvert > 10\%$)	81,461	0.229	0.420	0.000	0.000	0.000
Ln(Units sold)	90,211	4.374	3.191	1.526	3.970	6.989
Panel D: AI exposure						
AI Exposure	586	0.382	0.060	0.355	0.392	0.418
Panel E: Firm characteristics						
Ln(Employees)	91	7.724	2.201	6.340	7.620	9.627
Ln(Assets)	91	8.671	2.248	7.235	8.987	10.246
Tobin's q	91	2.802	1.861	1.397	2.005	3.582
ROA	91	0.005	0.072	0.005	0.014	0.027
Cash/Assets	91	0.108	0.124	0.031	0.072	0.136
Fixed assets/Assets	91	0.212	0.151	0.102	0.174	0.283
R&D/Assets	91	0.008	0.019	0.000	0.002	0.010
Market-to-book	91	6.287	15.954	2.078	3.368	6.935
Ln(Age)	91	3.439	0.820	3.135	3.555	4.111

Panel F: Covariate balance (pre-treatment, 2017–2021)

	Obs	Treated	Control	Difference	t-stat	Rank-sum
Ln(Average price)	89	2.282	1.762	0.520	2.847	[0.028]
Ln(Units sold)	89	3.826	4.684	−0.858	−1.691	[0.116]
Ln(Employees)	89	8.304	7.433	0.871	1.863	[0.098]
Ln(Assets)	91	9.208	8.071	1.137	2.413	[0.028]
Tobin's q	91	2.496	2.908	−0.412	−0.881	[0.352]
ROA	91	0.012	0.000	0.012	1.120	[0.654]
Cash/Assets	91	0.104	0.120	−0.016	−0.578	[0.328]
Fixed assets/Assets	91	0.121	0.249	−0.128	−4.518	[0.000]
R&D/Assets	91	0.011	0.004	0.007	2.152	[0.000]
Market-to-book	91	0.819	0.552	0.267	0.038	[0.348]
Ln(Age)	91	3.461	3.272	0.189	0.975	[0.051]

Table 2: OpenAI API release schedule
This table lists the OpenAI large language model families and their API availability dates. Each row corresponds to a model version made available through the OpenAI API. The *Event Date* column indicates the date on which the model became available for commercial use via the API.

Family	Model API Name	Event Date
GPT-3	davinci (beta)	2020-06-11
GPT-3	davinci (general release)	2021-11-18
InstructGPT	text-davinci-001	2022-01-27
InstructGPT	text-davinci-002	2022-03-15
InstructGPT	text-davinci-003	2022-11-28
GPT-3.5	gpt-3.5-turbo-0301	2023-03-01
GPT-3.5	gpt-3.5-turbo-0613	2023-06-13
GPT-4	gpt-4	2023-03-14
GPT-4	gpt-4-32k	2023-03-14
GPT-4 Turbo	gpt-4-1106-preview	2023-11-06
GPT-4o	gpt-4o	2024-05-13

Table 3: Price variance decomposition

This table decomposes the variance of log retail prices at the store×UPC×month level using the full Nielsen scanner dataset (2010–2023). Panel A decomposes the variance of log prices. Panel B decomposes the variance of squared deviations from the UPC×month mean, capturing cross-store price dispersion. Panel C decomposes the variance of month-over-month log price changes ($\Delta \ln p$), capturing who determines when and how much prices move. In each panel, columns (1)–(3) report Shapley value decompositions and columns (4)–(5) report AKM variance decompositions (Abowd, Kramarz, and Margolis, 1999; Card, Heining, and Kline, 2013). In the Shapley decompositions, column (1) uses UPC×month as the baseline and attributes the marginal R^2 to retailer and county fixed effects; column (2) uses manufacturer as the baseline and attributes variance to retailer and county; column (3) extends the decomposition following DellaVigna and Gentzkow (2019) to include a bilateral retailer×UPC pricing term. In the AKM decompositions, column (4) decomposes variance into UPC, retailer, covariance, time, and residual components; column (5) replaces UPC with manufacturer fixed effects. Cov(UPC, Retailer) denotes $2 \cdot \text{Cov}(\hat{\alpha}, \hat{\psi})$, where a positive value indicates positive assortative matching between product (or manufacturer) and retailer pricing. All regressions are estimated via iterative demeaning. Appendix A presents variable definitions.

Panel A: Variance of log prices

	Shapley			AKM	
	Product	Manufacturer	Bilateral	Product	Manufacturer
	(1)	(2)	(3)	(4)	(5)
UPC × Month	93.8%		47.7%		
UPC				87.8%	
Manufacturer		81.1%			27.9%
Retailer	5.6%	17.1%	3.6%	1.3%	3.4%
Retailer × UPC			48.3%		
County	0.6%	1.7%	0.4%	0.1%	0.1%
Month				0.9%	1.9%
Cov(UPC, Retailer)				−0.9%	3.6%
R^2	0.907	0.366	0.920	0.898	0.366
Observations (billion)	1.83	1.83	1.83	1.83	1.83

Panel B: Variance of squared deviations from UPC×month mean

	Shapley			AKM	
	Product	Manufacturer	Bilateral	Product	Manufacturer
	(1)	(2)	(3)	(4)	(5)
UPC × Month	97.6%		44.0%		
UPC				2.6%	
Manufacturer		73.1%			0.3%
Retailer	2.2%	24.7%	0.8%	0.1%	0.1%
Retailer × UPC			55.2%		
County	0.2%	2.3%	0.1%	0.0%	0.0%
Month				0.0%	0.0%
Cov(UPC, Retailer)				0.0%	−0.0%
R^2	0.049	0.005	0.079	0.028	0.005
Observations (billion)	1.83	1.83	1.83	1.83	1.83

Panel C: Variance of log price changes

	Shapley			AKM	
	Product	Manufacturer	Bilateral	Product	Manufacturer
	(1)	(2)	(3)	(4)	(5)
UPC × Month	99.9%		88.1%		
UPC				0.6%	
Manufacturer		82.6%			0.0%
Retailer	0.1%	13.3%	0.0%	0.0%	0.0%
Retailer × UPC			11.8%		
County	0.0%	4.1%	0.0%	0.0%	0.0%
Month				0.0%	0.0%
Cov(UPC, Retailer)				−0.0%	0.0%
R^2	0.071	0.001	0.076	0.006	0.001
Observations (billion)	0.92	0.92	0.92	0.92	0.92

Table 4: Private label products and price dispersion

This table presents OLS regressions of the standard deviation of log price on a private label indicator at the UPC-retailer-month level. *Private Label* equals one for store brand products. Column (1) includes sub-category × month and retailer fixed effects. Column (2) replaces the retailer fixed effect with retailer × month. Column (3) uses retailer × category fixed effects. Column (4) uses retailer × category × month fixed effects. Standard errors clustered at the UPC level are reported in parentheses. Statistical significance at the 1, 5, and 10 percent significance levels is denoted by ***, **, and *, respectively.

	Standard deviation of log price			
	(1)	(2)	(3)	(4)
Private Label	0.008***	0.008***	0.007***	0.007***
	(0.000)	(0.001)	(0.000)	(0.000)
Fixed effects				
Month × Product category	✓			
Retailer	✓			
Retailer × Month		✓		
Retailer × Product category			✓	
Retailer × Product category × Month				✓
Observations	215,135,478	218,684,228	215,130,651	213,403,925
Adjusted R^2	0.11	0.06	0.17	0.22

Table 5: Price level

This table presents results from difference-in-differences regressions of prices around the introduction of the InstructGPT API. The sample is a firm-product-month panel of publicly listed U.S. firms. The dependent variable is the log of CPI-deflated average price. *Treated* equals one for firms with above-median AI exposure, measured in January 2020 using employment-share-weighted occupation-level generative AI scores from Eisfeldt et al. (2025). *Post* equals one during the six months following the event. Treated and control firms are balanced on pre-treatment log average price and log units sold. Column (1) includes only the treatment indicator. Column (2) adds log units sold. Column (3) adds log employees measured in the month before the event and interacted with post. Column (4) adds firm financial controls (log assets, Tobin's q, ROA, cash-to-assets, fixed assets-to-assets, R&D-to-assets, market-to-book, and log firm age) measured in the quarter before the event and interacted with post. All specifications include firm-product and calendar month fixed effects. Standard errors clustered at the firm level are reported in parentheses below coefficient estimates. Statistical significance at the 1, 5, and 10 percent significance levels is denoted by ***, **, and *, respectively. Appendix A presents variable definitions.

	ln(Average price)			
	(1)	(2)	(3)	(4)
Treated × Post	−0.059***	−0.061***	−0.063***	−0.048***
	(0.018)	(0.018)	(0.018)	(0.011)
ln(Units sold)		0.016	0.016	0.016*
		(0.010)	(0.010)	(0.009)
ln(Employees) × Post			0.004	−0.026**
			(0.004)	(0.012)
ln(Assets) × Post				0.031**
				(0.013)
Tobin's q × Post				−0.003
				(0.003)
ROA × Post				0.535
				(0.343)
Cash × Post				0.229**
				(0.090)
Fixed assets × Post				0.075
				(0.047)
R&D × Post				−0.227
				(0.484)
Market-to-book × Post				0.000
				(0.000)
ln(Age) × Post				−0.015
				(0.010)
Fixed effects				
UPC	✓	✓	✓	✓
Calendar month	✓	✓	✓	✓
Observations	89,682	89,682	89,682	89,682
Adjusted R^2	0.95	0.95	0.95	0.95

Table 6: Price dispersion

This table presents results from difference-in-differences regressions of price dispersion around the introduction of the InstructGPT API. The sample is a firm-product-month panel of publicly listed U.S. firms. The dependent variable in columns (1) to (3) is the standard deviation of log prices, and in columns (4) to (6) is the interquartile range of log prices, both computed across stores within a firm-product-month. Columns (1) and (4) compute dispersion across all stores; columns (2) and (5) compute dispersion within counties; and columns (3) and (6) compute dispersion within retailers. *Treated* equals one for firms with above-median AI exposure, measured in January 2020 using employment-share-weighted occupation-level generative AI scores from Eisfeldt et al. (2025). *Post* equals one during the six months following the event. Treated and control firms are balanced on pre-treatment log average price and log units sold. All specifications include controls for log units sold, log employees measured in the month before the event and interacted with post, and firm financial controls (log assets, Tobin's q, ROA, cash-to-assets, fixed assets-to-assets, R&D-to-assets, market-to-book, and log firm age) measured in the quarter before the event and interacted with post. All specifications include firm-product and calendar month fixed effects. Standard errors clustered at the firm level are reported in parentheses below coefficient estimates. Statistical significance at the 1, 5, and 10 percent significance levels is denoted by ***, **, and *, respectively. Appendix A presents variable definitions.

	Standard deviation			Interquartile range		
	All	County	Retailer	All	County	Retailer
	(1)	(2)	(3)	(4)	(5)	(6)
Treated × Post	0.021***	0.007**	0.007**	0.017***	0.001**	0.006**
	(0.005)	(0.003)	(0.003)	(0.006)	(0.001)	(0.003)
Controls	✓	✓	✓	✓	✓	✓
Fixed effects						
UPC	✓	✓	✓	✓	✓	✓
Calendar month	✓	✓	✓	✓	✓	✓
Observations	77,972	63,620	75,594	77,972	63,620	75,594
Adjusted R^2	0.46	0.46	0.46	0.50	0.79	0.46

Table 7: Price changes

This table presents results from difference-in-differences regressions of the frequency of price changes around the introduction of the InstructGPT API. The dependent variable is an indicator for whether the month-over-month log price change exceeds 5%. Columns (1) to (3) use the absolute value of the price change; columns (4) to (6) isolate price increases; and columns (7) to (9) isolate price decreases. Within each group, columns are estimated at the firm-product level (All), the firm-product-county level (County), and the firm-product-retailer level (Retail). The All specifications include firm-product and calendar month fixed effects; the County specifications include UPC×county and calendar month fixed effects; and the Retail specifications include UPC×retailer and calendar month fixed effects. *Treated* equals one for firms with above-median AI exposure, measured in January 2020 using employment-share-weighted occupation-level generative AI scores from Eisfeldt et al. (2025). *Post* equals one during the six months following the event. Treated and control firms are balanced on pre-treatment log average price and log units sold. All specifications include controls for log units sold, log employees measured in the month before the event and interacted with post, and firm financial controls (log assets, Tobin's q, ROA, cash-to-assets, fixed assets-to-assets, R&D-to-assets, market-to-book, and log firm age) measured in the quarter before the event and interacted with post. Standard errors clustered at the firm level are reported in parentheses below coefficient estimates. Statistical significance at the 1, 5, and 10 percent significance levels is denoted by ***, **, and *, respectively. Appendix A presents variable definitions.

	$\mathbb{P}(\lvert\%\Delta P\rvert > 5\%)$			$\mathbb{P}(\%\Delta P > 5\%)$			$\mathbb{P}(\%\Delta P < -5\%)$		
	All	County	Retail	All	County	Retail	All	County	Retail
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Treated × Post	0.060***	0.061***	0.031**	0.011	0.027**	0.012	0.049***	0.034**	0.020*
	(0.017)	(0.016)	(0.013)	(0.010)	(0.011)	(0.011)	(0.013)	(0.013)	(0.012)
Controls	✓	✓	✓	✓	✓	✓	✓	✓	✓
Fixed effects									
UPC	✓			✓			✓		
UPC×County		✓			✓			✓	
UPC×Retailer			✓			✓			✓
Calendar month	✓	✓	✓	✓	✓	✓	✓	✓	✓
Observations	89,682	16,580,508	965,916	89,682	16,580,508	965,916	89,682	16,580,508	965,916
Adjusted R^2	0.31	0.20	0.25	0.09	0.05	0.07	0.13	0.06	0.10

Table 8: Controlling for retailer-specific effects

This table presents results from difference-in-differences regressions of prices around the introduction of the InstructGPT API, controlling for retailer-specific fixed effects. *Treated* equals one for firms with above-median AI exposure, measured in January 2020 using employment-share-weighted occupation-level generative AI scores from Eisfeldt et al. (2025). *Post* equals one during the six months following the event. Treated and control firms are balanced on pre-treatment log average price and log units sold. The dependent variables are the log of CPI-deflated average price (columns (1)–(3)), the standard deviation of log prices (columns (4)–(6)), and an indicator for absolute month-over-month price changes exceeding 5% (columns (7)–(9)). Columns (1), (4), and (7) include UPC and calendar month fixed effects (baseline). Columns (2), (5), and (8) replace UPC fixed effects with UPC×retailer fixed effects, which absorb any time-invariant retailer-product pricing differences. Columns (3), (6), and (9) further add retailer×calendar month fixed effects, which absorb retailer-specific time trends such as chain-wide promotional cycles and seasonal pricing strategies. All specifications include controls for log units sold, log employees measured in the month before the event and interacted with post, and firm financial controls (log assets, Tobin's q, ROA, cash-to-assets, fixed assets-to-assets, R&D-to-assets, market-to-book, and log firm age) measured in the quarter before the event and interacted with post. Standard errors clustered at the firm level are reported in parentheses below coefficient estimates. Statistical significance at the 1, 5, and 10 percent significance levels is denoted by ***, **, and *, respectively. Appendix A presents variable definitions.

	ln(Average price)			Standard deviation			$\mathbb{P}(\lvert\%\Delta P\rvert > 5\%)$		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Treated × Post	-0.032^{***}	-0.034^{***}	-0.030^{***}	0.005^{**}	0.006^{**}	0.005^{*}	0.028^{**}	0.031^{**}	0.029^{**}
	(0.007)	(0.007)	(0.006)	(0.003)	(0.002)	(0.002)	(0.012)	(0.013)	(0.012)
Controls	✓	✓	✓	✓	✓	✓	✓	✓	✓
Fixed effects									
UPC	✓			✓			✓		
UPC × Retailer		✓	✓		✓	✓		✓	✓
Retailer × Month			✓			✓			✓
Calendar month	✓	✓		✓	✓		✓	✓	
Observations	969,418	965,916	965,916	859,404	856,610	856,610	969,418	965,916	965,916
Adjusted R^2	0.91	0.96	0.96	0.17	0.41	0.43	0.12	0.25	0.26

Table 9: Price sensitivity to local economic conditions

This table presents results from difference-in-differences regressions of prices around the introduction of the InstructGPT API, including interactions with the county unemployment rate. The sample is a firm-product-county-month panel of publicly listed U.S. firms. The dependent variable is the log of CPI-deflated average price. *Treated* equals one for firms with above-median AI exposure, measured in January 2020 using employment-share-weighted occupation-level generative AI scores from Eisfeldt et al. (2025). *County unemployment rate* is a lagged indicator for whether the county unemployment rate is above the cross-county median. *Post* equals one during the six months following the event. Treated and control firms are balanced on pre-treatment log average price and log units sold. Column (1) includes firm-product, county, and calendar month fixed effects. Column (2) adds controls for log units sold, log employees measured in the month before the event and interacted with post, and firm financial controls (log assets, Tobin's q, ROA, cash-to-assets, fixed assets-to-assets, R&D-to-assets, market-to-book, and log firm age) measured in the quarter before the event and interacted with post. Column (3) replaces separate county and calendar month fixed effects with county×calendar month fixed effects. Column (4) further saturates the specification with firm-product×calendar month and county×calendar month fixed effects. Standard errors clustered at the firm level are reported in parentheses below coefficient estimates. Statistical significance at the 1, 5, and 10 percent significance levels is denoted by ***, **, and *, respectively. Appendix A presents variable definitions.

	ln(Average price)			
	(1)	(2)	(3)	(4)
Treated × Post	−0.019***	−0.019***	−0.019***	
	(0.006)	(0.005)	(0.005)	
Treated × Post × County unemployment rate	−0.005***	−0.005**	−0.005**	−0.005**
	(0.002)	(0.002)	(0.002)	(0.002)
Treated × County unemployment rate	0.006**	0.007**	0.007**	0.008**
	(0.003)	(0.003)	(0.003)	(0.003)
County unemployment rate × Post	0.004*	0.004*		
	(0.002)	(0.002)		
County unemployment rate	−0.005*	−0.006*		
	(0.003)	(0.003)		
Controls		✓		
Fixed effects				
UPC	✓	✓	✓	
County	✓	✓		
Calendar month	✓	✓		
UPC × Calendar month				✓
County × Calendar month			✓	✓
Observations	15,996,249	15,996,249	15,996,249	15,996,249
Adjusted R^2	0.95	0.95	0.95	0.96

Table 10: Product-level heterogeneity

This table presents results from difference-in-differences regressions of prices around the introduction of the InstructGPT API, including triple interactions with product-level characteristics. The sample is a firm-product-month panel of publicly listed U.S. firms. The dependent variable is the log of CPI-deflated average price. *Treated* equals one for firms with above-median AI exposure, measured in January 2020 using employment-share-weighted occupation-level generative AI scores from Eisfeldt et al. (2025). *Product indicator* is a binary variable capturing a product-level characteristic interacted with *Treated* × *Post*. In column (1), the product indicator equals one for discretionary products, classified based on Nielsen product group codes. In column (2), the product indicator equals one for products with above-median demand volatility, measured as the coefficient of variation of pre-treatment units sold within each product module. In column (3), the product indicator equals one for products with above-median price elasticity, estimated from pre-treatment regressions of log quantity changes on log price changes. *Post* equals one during the six months following the event. Treated and control firms are balanced on pre-treatment log average price and log units sold. All specifications include controls for log units sold, log employees measured in the month before the event and interacted with post, and firm financial controls (log assets, Tobin's q, ROA, cash-to-assets, fixed assets-to-assets, R&D-to-assets, market-to-book, and log firm age) measured in the quarter before the event and interacted with post. All specifications include firm-product and product group × calendar month fixed effects. Standard errors clustered at the firm level are reported in parentheses below coefficient estimates. Statistical significance at the 1, 5, and 10 percent significance levels is denoted by ***, **, and *, respectively. Appendix A presents variable definitions.

	ln(Average price)		
	Discretionary	Demand volatility	Price elasticity
	(1)	(2)	(3)
Treated × Post	−0.005	−0.012	−0.008
	(0.009)	(0.010)	(0.011)
Treated × Post × Product indicator	−0.034**	−0.024**	−0.022**
	(0.014)	(0.011)	(0.010)
Product indicator × Post		0.015*	−0.003
		(0.007)	(0.008)
Controls	✓	✓	✓
Fixed effects			
UPC	✓	✓	✓
Product group × Calendar month	✓	✓	✓
Observations	89,572	89,238	89,226
Adjusted R^2	0.95	0.95	0.95

Table 11: Augmentation vs. automation

This table presents results from difference-in-differences regressions of prices around the introduction of the InstructGPT API, where treatment is defined separately using AI augmentation and automation exposure. The sample is a firm-product-month panel of publicly listed U.S. firms. Augmentation scores capture the extent to which AI assistance increases worker productivity while human judgment and oversight remain essential for task completion. Automation scores measure the likelihood that AI could complete the task independently with minimal human involvement while achieving acceptable quality. Both measures are scored at the O*NET task level, aggregated to the occupation level, and then to the firm level using employment shares. In columns (1) to (3), *Treated* equals one for firms with above-median augmentation exposure measured in January 2020. In columns (4) to (6), *Treated* equals one for firms with above-median automation exposure measured in January 2020. *Post* equals one during the six months following the event. Treated and control firms are balanced on pre-treatment log average price and log units sold. The dependent variables are the log of CPI-deflated average price (columns (1) and (4)), the standard deviation of log prices (columns (2) and (5)), and an indicator for absolute month-over-month price changes exceeding 5% (columns (3) and (6)). All specifications include controls for log units sold, log employees measured in the month before the event and interacted with post, and firm financial controls (log assets, Tobin's q, ROA, cash-to-assets, fixed assets-to-assets, R&D-to-assets, market-to-book, and log firm age) measured in the quarter before the event and interacted with post. All specifications include firm-product and calendar month fixed effects. Standard errors clustered at the firm level are reported in parentheses below coefficient estimates. Statistical significance at the 1, 5, and 10 percent significance levels is denoted by ***, **, and *, respectively. Appendix A presents variable definitions.

	Augmentation			Automation		
	ln(Average price)	Standard deviation	$\mathbb{P}(\lvert\%\Delta P\rvert > 5\%)$	ln(Average price)	Standard deviation	$\mathbb{P}(\lvert\%\Delta P\rvert > 5\%)$
	(1)	(2)	(3)	(4)	(5)	(6)
Treated × Post	-0.046^{***}	0.012^{*}	0.041^{**}	-0.013	0.014	0.020
	(0.010)	(0.007)	(0.019)	(0.016)	(0.011)	(0.021)
Controls	✓	✓	✓	✓	✓	✓
Fixed effects						
UPC	✓	✓	✓	✓	✓	✓
Calendar month	✓	✓	✓	✓	✓	✓
Observations	74,050	65,682	74,050	65,240	55,939	65,240
Adjusted R^2	0.94	0.43	0.31	0.95	0.44	0.29

Table 12: Evidence of AI adoption

This table presents results from difference-in-differences regressions validating that firms with high AI exposure increase AI-adoption activity after the introduction of the InstructGPT API. Panel A reports regressions of AI-related job posting counts. The dependent variables are monthly counts of active job postings in five AI categories: AI, Applied AI, Data science, Ø*NET AI, and All AI. Panel B reports regressions of AI-related keyword counts in annual 10-K filings. The dependent variable is the count of AI-related keywords measured across the full filing (Total) and separately for Item 1 (Business Description), Item 1A (Risk Factors), and Item 7 (Management's Discussion and Analysis). Panel C decomposes 10-K AI keywords into implementation and hiring and risk and regulatory categories. *Treated* equals one for firms with above-median AI exposure, measured in January 2020 using employment-share-weighted occupation-level generative AI scores from Eisfeldt et al. (2025). *Post* equals one after the introduction of the InstructGPT API. All specifications include firm and calendar period fixed effects. Standard errors clustered at the firm level are reported in parentheses below coefficient estimates. Statistical significance at the 1, 5, and 10 percent significance levels is denoted by ***, **, and *, respectively. Appendix A presents variable definitions.

Panel A: Job postings (LinkUp)

	Job postings				
	AI	Applied AI	Data science	O*NET AI	All AI
	(1)	(2)	(3)	(4)	(5)
Treated × Post	7.952**	4.448**	4.385**	16.014**	22.451**
	(3.381)	(1.771)	(1.773)	(7.247)	(9.200)
Fixed effects					
Firm	✓	✓	✓	✓	✓
Calendar month	✓	✓	✓	✓	✓
Observations	1,586	1,586	1,586	1,586	1,586
Adjusted R^2	0.96	0.93	0.93	0.94	0.95

Panel B: 10-K AI keywords

	AI keyword count			
	Total	Item 1	Item 1A	Item 7
	(1)	(2)	(3)	(4)
Treated × Post	1.516***	0.498**	0.708***	0.038
	(0.392)	(0.215)	(0.218)	(0.058)
Fixed effects				
Firm	✓	✓	✓	✓
Year	✓	✓	✓	✓
Observations	2,384	2,384	2,384	2,384
Adjusted R^2	0.68	0.73	0.43	0.61

Panel C: 10-K implementation vs. risk

	Implementation/hiring			Risk/regulatory		
	Total	Item 1	Item 1A	Total	Item 1A	Item 7
	(1)	(2)	(3)	(4)	(5)	(6)
Treated × Post	1.026***	0.405**	0.433***	0.079**	0.089**	0.002
	(0.283)	(0.174)	(0.135)	(0.037)	(0.035)	(0.002)
Fixed effects						
Firm	✓	✓	✓	✓	✓	✓
Year	✓	✓	✓	✓	✓	✓
Observations	2,384	2,384	2,384	2,384	2,384	2,384
Adjusted R^2	0.69	0.73	0.43	0.28	0.24	-0.01

Appendix A: Variable Definitions

Variable	Definition	Source
Panel A: Dependent variables		
ln(Average price)	The natural log of the CPI-deflated average unit price of a firm-product (UPC) in a given month.	Nielsen Retail Panel; BLS Consumer Price Index (CPI)
SD of ln(Average price)	The standard deviation of log prices across stores within a firm-product-month. Variants are computed across all stores, within counties, and within retailers.	Nielsen Retail Panel
IQR of ln(Average price)	The interquartile range (75th minus 25th percentile) of log prices across stores within a firm-product-month. Variants are computed across all stores, within counties, and within retailers.	Nielsen Retail Panel
$\mathbb{P}(\lvert\%\Delta P\rvert > 5\%)$	An indicator variable equal to one if the absolute month-over-month change in log average price exceeds 5%.	Nielsen Retail Panel
$\mathbb{P}(\%\Delta P > 5\%)$	An indicator variable equal to one if the month-over-month change in log average price exceeds +5%.	Nielsen Retail Panel
$\mathbb{P}(\%\Delta P < -5\%)$	An indicator variable equal to one if the month-over-month change in log average price is less than −5%.	Nielsen Retail Panel
Panel B: Treatment and exposure variables		
AI Exposure	The employment-share-weighted average of occupation-level generative AI scores from Eisfeldt et al. (2025). Scores capture the extent to which generative AI is expected to reduce the time required to complete the tasks typically performed by each occupation.	ONET (v27.2); Eisfeldt et al. (2025); Revelio Labs
Treated	An indicator variable equal to one for firms with above-median AI exposure measured in January 2020.	
Post	An indicator variable equal to one during the six months following the AI implementation event.	
AI Exposure (Davinci-01)	An alternative AI exposure measure derived from a target-judge model framework, where GPT-5-mini evaluates whether the text-davinci-001 model can perform O*NET occupational tasks. Aggregated to the firm level using employment shares.	ONET (v27.2); Eisfeldt et al. (2025); Revelio Labs

Variable	Definition	Source
Augmentation score	A task-level score (0–10) capturing the extent to which AI assistance increases worker productivity while human judgment remains essential, aggregated to the firm level using employment shares.	ONET (v27.2)
Automation score	A task-level score (0–10) capturing the likelihood that AI could complete the task independently with minimal human involvement, aggregated to the firm level using employment shares.	ONET (v27.2)
Panel C: Control variables		
ln(Units sold)	The natural log of total units sold for a firm-product in a given month.	Nielsen Retail Panel
ln(Employees)	The natural log of the number of employees, measured in the month before the event.	Revelio Labs
ln(Assets)	The natural log of total assets (atq), measured in the quarter before the event.	Compustat
Tobin's q	Tobin's q, measured in the quarter before the event.	Compustat
ROA	Return on assets, computed as income before extraordinary items divided by total assets (ibq/atq), measured in the quarter before the event.	Compustat
Cash/Assets	Cash and short-term investments divided by total assets (chq/atq), measured in the quarter before the event.	Compustat
Fixed assets/Assets	Net property, plant, and equipment divided by total assets ($ppentq/atq$), measured in the quarter before the event.	Compustat
R&D/Assets	Research and development expenditures divided by total assets ($xrdq/atq$), measured in the quarter before the event. Missing values set to zero.	Compustat
Market-to-book	Shares outstanding times stock price divided by common equity ($cshoq \times prccq/ceqq$), measured in the quarter before the event.	Compustat
ln(Age)	The natural log of firm age.	CRSP
County unemployment rate	The lagged county-level unemployment rate.	BLS Local Area Unemployment Statistics (LAUS)

Variable	Definition	Source
Workplace Covid exposure (DN)	An occupation-level measure of physical proximity at work from Dingel and Neiman (2020), aggregated to the firm level using employment shares and measured before the event.	Dingel and Neiman (2020); Revelio Labs
Workplace Covid exposure (ATUS)	An occupation-level measure of physical proximity at work from the American Time Use Survey, aggregated to the firm level using employment shares and measured before the event.	American Time Use Survey (ATUS); Revelio Labs
ln(County Covid cases)	The natural log of one plus monthly county-level cumulative COVID-19 cases.	NYT County-Level COVID-19 Database
ln(County Covid deaths)	The natural log of one plus monthly county-level cumulative COVID-19 deaths.	NYT County-Level COVID-19 Database
Panel D: Product-level heterogeneity variables		
Discretionary	An indicator variable equal to one for discretionary products, classified based on Nielsen product group codes.	Nielsen Retail Panel
Demand volatility	An indicator variable equal to one for products with above-median demand volatility, measured as the coefficient of variation (standard deviation divided by mean) of pre-treatment units sold within each product module.	Nielsen Retail Panel
Price elasticity	An indicator variable equal to one for products with above-median price elasticity, estimated from pre-treatment regressions of log quantity changes on log price changes.	Nielsen Retail Panel

Variable	Definition	Source
Panel E: AI adoption variables		
AI job postings	Monthly count of active job postings classified into five AI-related categories based on job title and description keywords. *AI*: postings containing general AI terms (e.g., "artificial intelligence," "machine learning," "deep learning," "natural language processing"). *Applied AI*: postings referencing the deployment of AI tools in business functions (e.g., "predictive analytics," "recommendation system," "chatbot"). *Data science*: postings for data scientists, data engineers, machine learning engineers, and related roles. *O*NET AI*: postings whose occupational classification maps to O*NET occupations identified as AI-related. *All AI*: the union of the preceding four categories.	LinkUp
AI keyword count (10-K)	The count of AI-related keywords (e.g., "artificial intelligence," "machine learning," "automation") in firm i's annual 10-K filing, measured across the full filing (Total) and separately for Item 1 (Business Description), Item 1A (Risk Factors), and Item 7 (Management's Discussion and Analysis).	SEC EDGAR
Implementation/hiring keywords	A subset of 10-K AI keywords capturing references to AI adoption, deployment, and workforce investment.	SEC EDGAR
Risk/regulatory keywords	A subset of 10-K AI keywords capturing references to AI-related threats, compliance, and uncertainty.	SEC EDGAR

Internet Appendix for
Who Sets the Price? The Vertical Origins of Uniform Pricing

Alvin Chen, Leandro Sanz, and Michael Wittry

IA.A Data

IA.A.1 AI implementation exposure decomposition

The original measure of workforce exposure to generative AI developed by Eisfeldt et al. (2025) can manifest through two distinct channels: augmenting worker productivity while preserving human involvement, or automating the completion of tasks with minimal or no human involvement. Therefore, we extend the Eisfeldt et al. (2025) measure by decomposing it into separate measures of augmentation and automation. We conducted this categorization in December 2025, using GPT model 5-mini, and a temperature parameter of zero.

For each occupation, we use the O*NET 30.1 database task descriptions and evaluate each task description along an augmentation and automation dimension. Our augmentation score ranges from 0 to 10 and captures the extent by which AI assistance increases worker productivity while human judgment and oversight remains important for task completion. High scores (8–10) indicate tasks where AI is expected to reduce the time to complete a task by 50% or more, or the use of AI significantly improves quality. Examples of tasks that have high augmentation scores include writing and data analysis, where AI drafts text or processes information but human oversight is important to interpret and make final decisions. Similarly, our automation score ranges from 0 to 10, and measures the likelihood that AI could complete the task independently with minimal human involvement while achieving acceptable quality. Tasks with high automation scores include structured data entry and the creation of standardized reports.

For each occupation, we aggregate task-level augmentation and automation scores by taking an equal-weighted average across tasks within an occupation, and then weighting each occupation score by the share of workers in that occupation within a firm. This decomposition allows us to distinguish firms whose workforces face primarily augmentation effects from those facing automation effects. The system prompt submitted to the GPT model is as follows:

> *Consider the most powerful large language models (LLMs) and AI systems currently available. These models can complete many tasks with text input and output where context can be captured in ˜100,000 words. They can access up-to-date information through web search and can generate and analyze images and code. Assume you are a worker with average expertise in your role trying*

to complete the given task. You have access to AI systems as well as any commonly available software, computer hardware, and technical tools accessible via a laptop. You do not have access to specialized physical tools, materials, or equipment beyond standard office technology.

Your job is to evaluate two separate dimensions for each task:

Augmentation score (0–10): Rate how much AI assistance increases worker productivity while the human remains essential to task completion.

High augmentation (8–10): AI significantly enhances human capability, reducing time by 50% or more, or substantially improving quality, but human judgment, creativity, or oversight remains critical:

- *Complex writing where AI drafts but human provides strategic direction and refinement*
- *Data analysis where AI processes information but human interprets implications and makes decisions*
- *Research where AI gathers and synthesizes information but human evaluates relevance and draws conclusions*
- *Creative work where AI generates options but human provides vision and final selection*
- *Technical work where AI handles routine coding and calculations but human designs architecture*
- *Advisory work where AI provides information and recommendations but human provides contextualized counsel*
- *Planning where AI analyzes scenarios but human makes strategic choices*

Medium augmentation (4–7): AI provides moderate assistance but human expertise drives the majority of value:

- *Editing and proofreading where AI catches errors but human ensures meaning*
- *Routine communication where AI assists with drafting but human handles nuanced interactions*
- *Information organization where AI structures data but human ensures accuracy*
- *Basic analysis where AI highlights patterns but human validates findings*

Low augmentation (1–3): AI provides minimal assistance; task primarily relies on human skills:

- *Tasks requiring real-time human interaction or relationship building*
- *Physical manipulation or in-person demonstration*
- *Work requiring tacit knowledge not easily articulated*
- *Tasks where AI assistance does not meaningfully reduce time or improve quality*

No augmentation (0): AI cannot assist meaningfully with this task.

Automation score (0–10): Rate the likelihood that AI could complete this task independently with minimal human involvement, achieving acceptable quality.

High automation (8–10): AI can complete the task end-to-end with minimal human oversight:

- *Highly structured data entry, transformation, or formatting*
- *Routine documentation following clear templates*
- *Straightforward text translation or summarization*
- *Basic code generation for well-defined specifications*
- *Repetitive communication following standard scripts*
- *Simple scheduling or record-keeping*
- *Standardized report generation from structured inputs*

Medium automation (4–7): AI can complete substantial portions independently but requires human review or input at key decision points:

- *Complex documentation where AI generates drafts requiring human validation*
- *Analysis with clear methodologies but outputs need expert interpretation*
- *Customer service where AI handles routine queries but escalates complex issues*
- *Research where AI gathers information but human validates sources and conclusions*

Low automation (1–3): AI can only handle small, discrete sub-tasks; human drives most of the work:

- *Work requiring situational judgment or contextual expertise*
- *Tasks involving negotiation or persuasion*
- *Creative problem-solving without clear parameters*

- *Processes requiring real-time adaptation based on feedback*

No automation (0): Human must perform this task; AI cannot automate it:

- *Tasks legally requiring human performance*
- *Work requiring physical presence or manipulation*
- *Tasks requiring precise measurements or detailed visual inspection*
- *Decisions directly affecting human welfare (hiring, medical diagnosis, legal judgments)*
- *Work requiring trust, empathy, or human connection*
- *Tasks where accountability must rest with a human*

Relationship between scores:

- *High augmentation + low automation = AI augments human expert (e.g., strategic analysis).*
- *High augmentation + high automation = Transitional state (AI increasingly capable, human adds value but may become optional).*
- *Low augmentation + high automation = Full automation potential (e.g., data entry).*
- *Low augmentation + low automation = AI-resistant work (e.g., skilled trades, in-person services).*

Evaluation guidelines:

- *Augmentation asks: "How much does AI help a human do this better or faster?"*
- *Automation asks: "Could AI do this acceptably without the human?"*
- *Consider both current AI capabilities and easily imaginable near-term applications.*
- *Consider whether quality or judgment trade-offs are acceptable for automation.*
- *Default to lower scores when uncertain.*
- *Account for regulatory, ethical, or practical constraints on AI autonomy.*

Output format: Return valid JSON with:

- *Augmentation: score 0–10*
- *Automation: score 0–10*

- *Augmentation_reason: ≤30 words explaining why AI helps humans perform this task*
- *Automation_reason: ≤30 words explaining whether and why AI could perform this independently*

IA.B Additional results

This section presents additional results and robustness tests that are described and referenced in the manuscript.

IA.B.1 Price variance decomposition by firm size

Table IA.B.1 replicates the price variance decomposition of Table 3 separately for large and small manufacturers, split at the median of total revenue. Panel A decomposes the variance of log price levels and Panel B decomposes the variance of squared deviations from the UPC×month mean. Large manufacturers exhibit price structures similar to the full sample, with UPC×month effects explaining the vast majority of level variance and bilateral retailer×UPC terms dominating the dispersion decomposition. For small manufacturers, the manufacturer fixed effect absorbs a larger share of price-level variance, and retailer and county components are proportionally larger in both panels, consistent with smaller firms facing greater cross-retailer pricing heterogeneity.

Table IA.B.1: Price variance decomposition by firm size

This table replicates Table 3 separately for large manufacturers (columns (1)–(5)) and small manufacturers (columns (6)–(10)), split at the median of total revenue. Panel A decomposes the variance of log prices. Panel B decomposes the variance of squared deviations from the UPC×month mean, capturing cross-store price dispersion. Within each size group, columns (1)–(3) and (6)–(8) report Shapley value decompositions and columns (4)–(5) and (9)–(10) report AKM variance decompositions (Abowd, Kramarz, and Margolis, 1999; Card, Heining, and Kline, 2013). In the Shapley decompositions, the Product column uses UPC×month as the baseline; the Mfr column uses manufacturer as the baseline; and the Bilateral column extends the decomposition following DellaVigna and Gentzkow (2019) to include a bilateral retailer×UPC pricing term. In the AKM decompositions, the UPC column decomposes variance into UPC, retailer, covariance, time, and residual components; the Mfr column replaces UPC with manufacturer fixed effects. Cov(UPC, Retailer) denotes $2 \cdot \text{Cov}(\hat{\alpha}, \hat{\psi})$. All regressions are estimated via iterative demeaning. Appendix A presents variable definitions.

Panel A: Variance of log prices

	Large Manufacturers					Small Manufacturers				
	Shapley			AKM		Shapley			AKM	
	Product	Mfr	Bilateral	Product	Mfr	Product	Mfr	Bilateral	Product	Mfr
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
UPC × Month	93.8%		47.7%			89.6%		46.3%		
UPC				87.8%					85.8%	
Manufacturer		81.1%			27.9%		84.5%			64.3%
Retailer	5.6%	17.2%	3.6%	1.3%	3.4%	8.7%	13.1%	5.7%	2.0%	12.3%
Retailer × UPC			48.3%					46.8%		
County	0.6%	1.7%	0.4%	0.1%	0.1%	1.7%	2.4%	1.2%	0.1%	0.2%
Month				0.9%	1.9%				0.1%	0.4%
Cov(UPC, Retailer)				−0.9%	3.6%				−3.3%	−11.0%
R^2	0.907	0.366	0.920	0.898	0.366	0.863	0.669	0.877	0.851	0.669
Firms	175	175	175	175	175	175	175	175	175	175

Panel B: Variance of squared deviations from UPC×month mean

	Large Manufacturers					Small Manufacturers				
	Shapley			AKM		Shapley			AKM	
	Product	Mfr	Bilateral	Product	Mfr	Product	Mfr	Bilateral	Product	Mfr
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
UPC × Month	97.6%		44.0%			79.0%		22.9%		
UPC				2.6%					4.0%	
Manufacturer		73.0%			0.3%		53.0%			1.5%
Retailer	2.2%	24.7%	0.8%	0.1%	0.1%	14.8%	33.1%	4.1%	1.6%	1.3%
Retailer × UPC			55.1%					70.2%		
County	0.2%	2.3%	0.1%	0.0%	0.0%	6.2%	13.9%	2.8%	0.3%	0.3%
Month				0.0%	0.0%				0.0%	0.0%
Cov(UPC, Retailer)				0.0%	−0.0%				−1.6%	−0.8%
R^2	0.049	0.005	0.079	0.028	0.005	0.048	0.023	0.095	0.043	0.023
Firms	175	175	175	175	175	175	175	175	175	175

IA.B.2 Category-specific CPI deflation

A potential concern is that the baseline price-level result could be driven by sectoral inflation differences: if treated firms disproportionately sell in categories that had different inflation around the event (e.g., food at home versus discretionary goods in 2022), the estimated gap could reflect composition and sectoral price dynamics rather than a within-category relative price response. To address this, we deflate prices using category-specific CPIs from the BLS. We use 26 detailed expenditure categories (e.g., cereals and bakery, dairy, meats, fruits and vegetables, beverages, apparel, personal care, pets) and map each Nielsen product group to the corresponding BLS category. Table IA.B.2 replicates the baseline price-level specifications (Table 5) with the dependent variable defined as the log of average price deflated by the product's category CPI rather than the national CPI. The treatment effect remains negative and statistically significant across all four columns. The point estimates are smaller in magnitude than in the baseline, consistent with part of the raw gap reflecting sectoral inflation differences; the within-category relative price effect remains economically and statistically significant.

Table IA.B.2: Price level with category-specific CPI deflation

This table presents results from difference-in-differences regressions of prices around the introduction of the InstructGPT API, where the dependent variable is the log of average price deflated by category-specific CPIs from the BLS rather than the national CPI. Prices are deflated using 26 detailed expenditure categories (e.g., cereals and bakery, dairy, meats, fruits and vegetables, nonalcoholic beverages, other food at home, apparel, personal care products, pets); each Nielsen product group is mapped to the corresponding BLS category. The sample, treatment definition, event window, matching, and specification are identical to the baseline price-level regressions (Table 5). Column (1) includes only the treatment indicator. Column (2) adds log units sold. Column (3) adds log employees measured in the month before the event and interacted with post. Column (4) adds firm financial controls (log assets, Tobin's q, ROA, cash-to-assets, fixed assets-to-assets, R&D-to-assets, market-to-book, and log firm age) measured in the quarter before the event and interacted with post. All specifications include firm-product and calendar month fixed effects. Standard errors clustered at the firm level are reported in parentheses below coefficient estimates. Statistical significance at the 1, 5, and 10 percent significance levels is denoted by ***, **, and *, respectively. Appendix A presents variable definitions.

	ln(Average price), category-CPI deflated			
	(1)	(2)	(3)	(4)
Treated × Post	−0.038**	−0.039**	−0.041**	−0.033***
	(0.019)	(0.019)	(0.018)	(0.010)
ln(Units sold)		0.016*	0.017*	0.016*
		(0.009)	(0.009)	(0.009)
ln(Employees) × Post			0.006	−0.021*
			(0.004)	(0.012)
ln(Assets) × Post				0.028**
				(0.012)
Tobin's q × Post				−0.002
				(0.003)
ROA × Post				0.547*
				(0.314)
Cash × Post				0.244***
				(0.089)
Fixed assets × Post				0.043
				(0.043)
R&D × Post				−0.218
				(0.472)
Market-to-book × Post				0.000
				(0.000)
ln(Age) × Post				−0.014*
				(0.008)
Fixed effects				
UPC	✓	✓	✓	✓
Calendar month	✓	✓	✓	✓
Observations	89,682	89,682	89,682	89,682
Adjusted R^2	0.95	0.95	0.95	0.95

IA.B.3 Unmatched sample and continuous treatment

Our baseline results use a matched sample constructed via propensity-score matching on pre-treatment price levels and sales volumes. To assess whether the matching procedure drives the results, Table IA.B.3 reports two sets of robustness tests. Odd columns re-estimate the baseline binary treatment specification on the full unmatched sample. The coefficients on *Treated* × *Post* are virtually identical to the matched-sample estimates across all three outcomes: prices decline by approximately 5.1%, within-product price dispersion increases by 0.021, and the probability of a price change exceeding 5% rises by 6.0 percentage points—all significant at the 1% level. Even columns replace the binary treatment indicator with continuous AI exposure, measured as the employment-share-weighted average of occupation-level generative AI scores. The results are consistent across all three outcomes: the continuous exposure measure is negative and significant for price levels (-2.680, $p < 0.01$), positive and significant for price dispersion (1.004, $p < 0.01$), and positive and significant for repricing frequency (3.074, $p < 0.01$). The consistency of the binary and continuous specifications provides evidence that the results are not sensitive to the specific treatment classification or to the matching procedure.

Table IA.B.3: Unmatched sample and continuous treatment

This table presents results from difference-in-differences regressions of the log of CPI-deflated average price the introduction of the InstructGPT API. The sample is a firm-product-month panel of publicly listed U.S. firms. AI exposure is derived from a target-judge model framework, where GPT-5-mini evaluates whether the text-davinci-001 model can perform O*NET occupational tasks. Scores are aggregated to the firm level using employment shares. *Treated* equals one for firms with above-median AI exposure measured in January 2020. *Post* equals one during the six months following the event. Treated and control firms are balanced on pre-treatment log average price and log units sold. All specifications include controls for log units sold, log employees measured in the month before the event and interacted with post, and firm financial controls (log assets, Tobin's q, ROA, cash-to-assets, fixed assets-to-assets, R&D-to-assets, market-to-book, and log firm age) measured in the quarter before the event and interacted with post. All specifications include firm-product and calendar month fixed effects. Standard errors clustered at the firm level are reported in parentheses below coefficient estimates. Statistical significance at the 1, 5, and 10 percent significance levels is denoted by ***, **, and *, respectively. Appendix A presents variable definitions.

	ln(Average price)		Standard deviation		$\Pr(\lvert\Delta P\rvert > 5\%)$	
	(1)	(2)	(3)	(4)	(5)	(6)
Treated × Post	-0.051^{***}		0.021^{***}		0.060^{***}	
	(0.013)		(0.005)		(0.015)	
Exposure × Post		-2.680^{***}		1.004^{***}		3.074^{***}
		(0.857)		(0.278)		(1.010)
Fixed effects						
UPC	✓	✓	✓	✓	✓	✓
Calendar month	✓	✓	✓	✓	✓	✓
Observations	134,341	63,665	114,583	57,511	134,341	63,665
Adjusted R^2	0.96	0.93	0.48	0.45	0.32	0.31

IA.B.4 Specifying the target and judge model

Our baseline AI exposure measure uses the Eisfeldt et al. (2025) scores, which evaluate each O*NET task against the capabilities of GPT 3.5 Turbo. An alternative approach is to specify the capabilities of a particular target model and use a separate judge model to evaluate whether the target model can perform each task. We implement this approach using GPT-5-mini as the judge model and text-davinci-001 as the target model, whose capabilities more closely match those available at the time of the InstructGPT API introduction. Scores are aggregated to the firm level using employment shares, and treatment is defined as above-median exposure, identical to the baseline.

Table IA.B.4, Table IA.B.5, and Table IA.B.6 replicate the baseline price level, price dispersion, and price changes results using the Davinci-01 exposure measure. Across all specifications, the results are consistent with the baseline results. That is, we find that prices decline, price dispersion increases, and repricing frequency increases. The corresponding event-study estimates (Figure IA.B.1, Figure IA.B.2, and Figure IA.B.3) display parallel pre-trends and post-event treatment effects similar to those in the main analysis.

Figure IA.B.1: Price level event study, using AI exposure based on Davinci-01
This figure plots the β_k coefficients and 90% confidence intervals from event study regressions of the form: $\ln(\text{Average price})_{ipt} = \alpha_{ip} + \alpha_t + \sum_{k \neq -3} \beta_k \mathbb{1}[\text{Event time} = k] \times \text{Treated}_i + \epsilon_{ipt}$, where i denotes firms, p products (UPCs), and t calendar months. The dependent variable is the log of CPI-deflated average price. AI exposure is derived from a target-judge model framework, where GPT-5-mini evaluates whether the text-davinci-001 model can perform O*NET occupational tasks. Scores are aggregated to the firm level using employment shares. *Treated* equals one for firms with above-median AI exposure measured in January 2020. Event time runs from -6 to $+6$ months relative to the event, with $k = -3$ as the omitted base period. α_{ip} and α_t denote firm-product and calendar month fixed effects, respectively. Treated and control firms are balanced on pre-treatment log average price and log units sold. Standard errors are clustered at the firm level. Appendix A presents variable definitions.

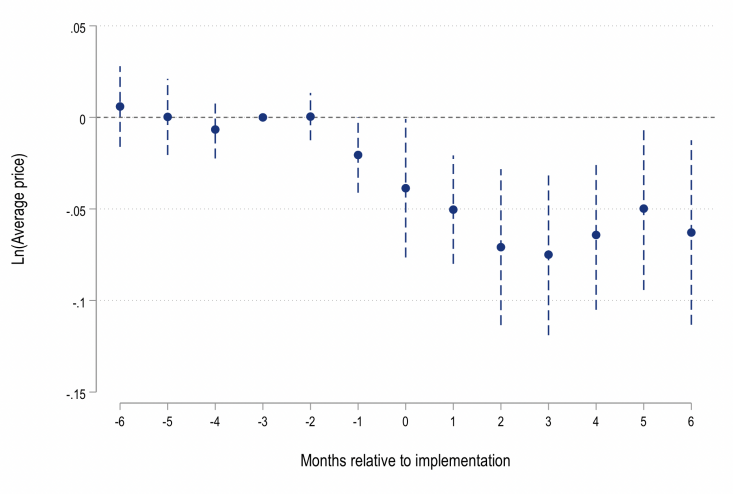
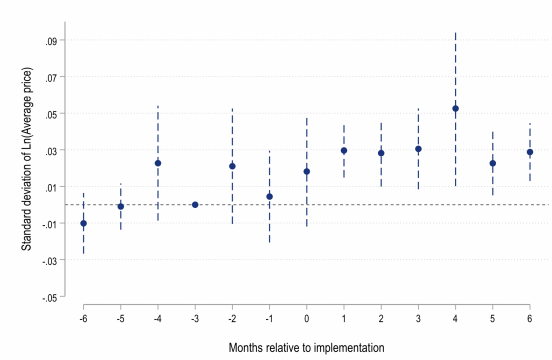


Figure IA.B.2: Price dispersion event study, using AI exposure based on Davinci-01
This figure plots the β_k coefficients and 90% confidence intervals from event study regressions of the form: $\text{Dispersion}_{ipt} = \alpha_{ip} + \alpha_t + \sum_{k \neq -3} \beta_k \mathbb{1}[\text{Event time} = k] \times \text{Treated}_i + \epsilon_{ipt}$, where i denotes firms, p products (UPCs), and t calendar months. In Panel A, the dependent variable is the standard deviation of log prices across stores within a firm-product-month. In Panel B, the dependent variable is the interquartile range of log prices across stores within a firm-product-month. AI exposure is derived from a target-judge model framework, where GPT-5-mini evaluates whether the text-davinci-001 model can perform O*NET occupational tasks. Scores are aggregated to the firm level using employment shares. *Treated* equals one for firms with above-median AI exposure measured in January 2020. Event time runs from -6 to $+6$ months relative to the implementation event, with $k = -3$ as the omitted base period. α_{ip} and α_t denote firm-product and calendar month fixed effects, respectively. Treated and control firms are balanced on pre-treatment log average price and log units sold. Standard errors are clustered at the firm level. Appendix A presents variable definitions.

Panel A: Standard deviation



Panel B: Interquartile range

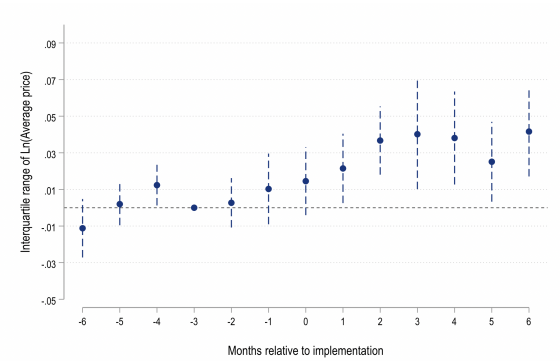
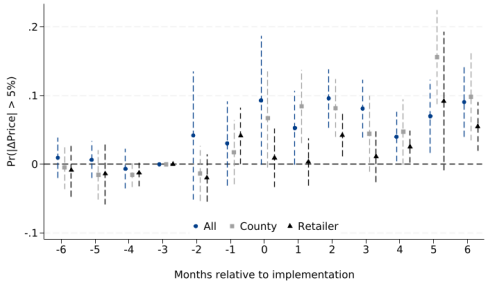


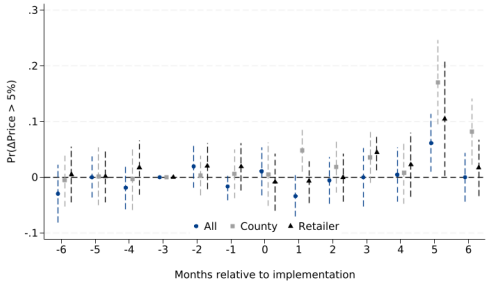
Figure IA.B.3: Price changes event study, using AI exposure based on Davinci-01

This figure plots the β_k coefficients and 90% confidence intervals from event study regressions of the form: $\mathbb{P}(|\Delta \ln P| > 5\%)_{ipt} = \alpha_{ip} + \alpha_t + \sum_{k \neq -3} \beta_k \mathbb{1}[\text{Event time} = k] \times \text{Treated}_i + \epsilon_{ipt}$, where i denotes firms, p products (UPCs), and t calendar months. The dependent variable is an indicator for whether the month-over-month log price change exceeds a given threshold. In Panel A, the dependent variable is an indicator for whether the absolute log price change exceeds 5%. Panels B and C focus on price increases and decreases, respectively. AI exposure is computed as the employment-share-weighted average of occupation-level generative AI scores from Eisfeldt et al. (2025), where scores capture the extent to which generative AI is expected to reduce the time required to complete the tasks typically performed by each occupation. *Treated* equals one for firms with above-median AI exposure measured in January 2020. Event time runs from −6 to +6 months relative to the InstructGPT API introduction in January 2022, with $k = -3$ as the omitted base period. α_{ip} and α_t denote firm-product and calendar month fixed effects, respectively. Treated and control firm products are balanced on the pre-treatment log average price and log units sold. Standard errors are clustered at the firm level. Appendix A presents variable definitions.

Panel A: $\mathbb{P}(|\Delta P| > 5\%)$



Panel B: $\mathbb{P}(\Delta P > 5\%)$



Panel C: $\mathbb{P}(\Delta P < -5\%)$

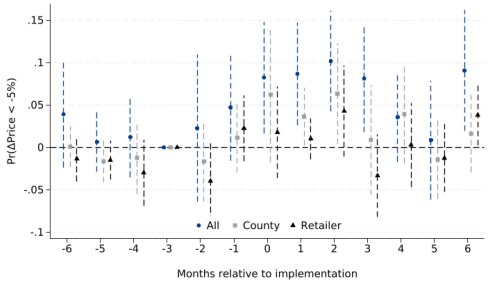


Table IA.B.4: Price level, using AI exposure based on Davinci-01

This table presents results from difference-in-differences regressions of the log of CPI-deflated average price around the introduction of the InstructGPT API. The sample is a firm-product-month panel of publicly listed U.S. firms. AI exposure is derived from a target-judge model framework, where GPT-5-mini evaluates whether the text-davinci-001 model can perform O*NET occupational tasks. Scores are aggregated to the firm level using employment shares. *Treated* equals one for firms with above-median AI exposure measured in January 2020. *Post* equals one during the six months following the event. Treated and control firms are balanced on pre-treatment log average price and log units sold. All specifications include controls for log units sold, log employees measured in the month before the event and interacted with post, and firm financial controls (log assets, Tobin's q, ROA, cash-to-assets, fixed assets-to-assets, R&D-to-assets, market-to-book, and log firm age) measured in the quarter before the event and interacted with post. All specifications include firm-product and calendar month fixed effects. Standard errors clustered at the firm level are reported in parentheses below coefficient estimates. Statistical significance at the 1, 5, and 10 percent significance levels is denoted by ***, **, and *, respectively. Appendix A presents variable definitions.

	ln(Average price)			
	(1)	(2)	(3)	(4)
Treated × Post	−0.055***	−0.056***	−0.060***	−0.061***
	(0.019)	(0.019)	(0.019)	(0.011)
ln(Units sold)		0.015	0.015	0.015
		(0.010)	(0.010)	(0.009)
ln(Employees) × Post			0.007	−0.023**
			(0.005)	(0.011)
ln(Assets) × Post				0.031***
				(0.011)
Tobin's q × Post				−0.002
				(0.004)
ROA × Post				0.692**
				(0.336)
Cash × Post				0.228***
				(0.080)
Fixed assets × Post				0.027
				(0.039)
R&D × Post				−0.141
				(0.445)
Market-to-book × Post				0.000
				(0.000)
ln(Age) × Post				−0.015
				(0.010)
Fixed effects				
UPC	✓	✓	✓	✓
Calendar month	✓	✓	✓	✓
Observations	89,776	89,776	89,776	89,776
Adjusted R^2	0.95	0.95	0.95	0.95

Table IA.B.5: Price dispersion, using AI exposure based on Davinci-01

This table presents results from difference-in-differences regressions of price dispersion around the introduction of the InstructGPT API. The sample is a firm-product-month panel of publicly listed U.S. firms. The dependent variable in columns (1) to (3) is the standard deviation of log prices, and in columns (4) to (6) is the interquartile range of log prices, both computed across stores within a firm-product-month. Columns (1) and (4) compute dispersion across all stores; columns (2) and (5) compute dispersion within counties; and columns (3) and (6) compute dispersion within retailers. AI exposure is derived from a target-judge model framework, where GPT-5-mini evaluates whether the text-davinci-001 model can perform O*NET occupational tasks. Scores are aggregated to the firm level using employment shares. *Treated* equals one for firms with above-median AI exposure measured in January 2020. *Post* equals one during the six months following the event. Treated and control firms are balanced on pre-treatment log average price and log units sold. All specifications include controls for log units sold, log employees measured in the month before the event and interacted with post, and firm financial controls (log assets, Tobin's q, ROA, cash-to-assets, fixed assets-to-assets, R&D-to-assets, market-to-book, and log firm age) measured in the quarter before the event and interacted with post. All specifications include firm-product and calendar month fixed effects. Standard errors clustered at the firm level are reported in parentheses below coefficient estimates. Statistical significance at the 1, 5, and 10 percent significance levels is denoted by ***, **, and *, respectively. Appendix A presents variable definitions.

	Standard deviation			Interquartile range		
	All	County	Retailer	All	County	Retailer
	(1)	(2)	(3)	(4)	(5)	(6)
Treated × Post	0.023***	0.010***	0.009**	0.022***	0.002**	0.007**
	(0.005)	(0.003)	(0.004)	(0.006)	(0.001)	(0.003)
Controls	✓	✓	✓	✓	✓	✓
Fixed effects						
UPC	✓	✓	✓	✓	✓	✓
Calendar month	✓	✓	✓	✓	✓	✓
Observations	77,912	63,354	75,528	77,912	63,354	75,528
Adjusted R^2	0.47	0.46	0.46	0.51	0.79	0.45

Table IA.B.6: Price changes, using AI exposure based on Davinci-01

This table presents results from difference-in-differences regressions of the frequency of price changes around the introduction of the InstructGPT API. The dependent variable is an indicator for whether the month-over-month log price change exceeds 5%. Columns (1) to (3) use the absolute value of the price change; columns (4) to (6) isolate price increases; and columns (7) to (9) isolate price decreases. Within each group, columns are estimated at the firm-product level (All), the firm-product-county level (County), and the firm-product-retailer level (Retail). AI exposure is derived from a target-judge model framework, where GPT-5-mini evaluates whether the text-davinci-001 model can perform O*NET occupational tasks. Scores are aggregated to the firm level using employment shares. *Treated* equals one for firms with above-median AI exposure measured in January 2020. *Post* equals one during the six months following the event. Treated and control firms are balanced on pre-treatment log average price and log units sold. All specifications include controls for log units sold, log employees measured in the month before the event and interacted with post, and firm financial controls (log assets, Tobin's q, ROA, cash-to-assets, fixed assets-to-assets, R&D-to-assets, market-to-book, and log firm age) measured in the quarter before the event and interacted with post. The All specifications include firm-product and calendar month fixed effects; the County specifications include UPC×county and calendar month fixed effects; and the Retail specifications include UPC×retailer and calendar month fixed effects. Standard errors clustered at the firm level are reported in parentheses below coefficient estimates. Statistical significance at the 1, 5, and 10 percent significance levels is denoted by ***, **, and *, respectively. Appendix A presents variable definitions.

	$\mathbb{P}(\lvert\%\Delta P\rvert > 5\%)$			$\mathbb{P}(\%\Delta P > 5\%)$			$\mathbb{P}(\%\Delta P < -5\%)$		
	All	County	Retail	All	County	Retail	All	County	Retail
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Treated × Post	0.061***	0.087***	0.035***	0.013	0.051***	0.013	0.049***	0.036	0.022*
	(0.015)	(0.028)	(0.013)	(0.010)	(0.012)	(0.011)	(0.013)	(0.024)	(0.012)
Controls	✓	✓	✓	✓	✓	✓	✓	✓	✓
Fixed effects									
UPC	✓			✓			✓		
UPC×County		✓			✓			✓	
UPC×Retailer			✓			✓			✓
Calendar month	✓	✓	✓	✓	✓	✓	✓	✓	✓
Observations	89,776	13,296,290	941,982	89,776	13,296,290	941,982	89,776	13,296,290	941,982
Adjusted R^2	0.31	0.20	0.25	0.09	0.05	0.07	0.13	0.06	0.10

IA.B.5 COVID-19 exposure

The InstructGPT API introduction in January 2022 occurred while the economy was still adjusting to the effects of the COVID-19 pandemic. If firms with high AI exposure also had workforces that were more exposed to pandemic-related workplace disruptions, COVID-19 effects could confound the estimated treatment effects. Table IA.B.7 addresses this concern by augmenting the baseline specification with direct controls for pandemic exposure. Panel A adds firm-level measures of workplace physical proximity—the Dingel and Neiman (2020) work-from-home index and an analogous measure from the American Time Use Survey—interacted with post. Panel B moves to the firm-product-county-month panel and controls for cumulative county-level COVID-19 cases and deaths. Across all specifications, the treatment effects remain statistically significant and consistent with the baseline, and the COVID-19 controls are themselves small and statistically insignificant.

Table IA.B.7: Controlling for COVID-19 exposure

This table presents results from difference-in-differences regressions of prices around the introduction of the InstructGPT API, controlling for COVID-19 exposure. *Treated* equals one for firms with above-median AI exposure, measured in January 2020 using employment-share-weighted occupation-level generative AI scores from Eisfeldt et al. (2025). *Post* equals one during the six months following the event. Treated and control firms are balanced on pre-treatment log average price and log units sold. The dependent variables are the log of CPI-deflated average price (columns (1) and (2)), the standard deviation of log prices (columns (3) and (4)), and an indicator for absolute month-over-month price changes exceeding 5% (columns (5) and (6)). Panel A uses the firm-product-month panel with firm-product and calendar month fixed effects. *Workplace Covid exposure (DN)* and *Workplace Covid exposure (ATUS)* are occupation-level measures of physical proximity at work from Dingel and Neiman (2020) and the American Time Use Survey, respectively, aggregated to the firm level using employment shares and interacted with post. Odd columns control for the Dingel and Neiman (2020) measure and even columns control for the ATUS measure. Panel B uses the firm-product-county-month panel with firm-product×county and calendar month fixed effects. Odd columns control for the log of one plus cumulative county-level COVID-19 cases and even columns control for the log of one plus cumulative county-level COVID-19 deaths. All specifications include controls for log units sold, log employees measured in the month before the event and interacted with post, and firm financial controls (log assets, Tobin's q, ROA, cash-to-assets, fixed assets-to-assets, R&D-to-assets, market-to-book, and log firm age) measured in the quarter before the event and interacted with post. Standard errors clustered at the firm level are reported in parentheses below coefficient estimates. Statistical significance at the 1, 5, and 10 percent significance levels is denoted by ***, **, and *, respectively. Appendix A presents variable definitions.

Panel A: Firm-level COVID-19 exposure

	ln(Average price)		Standard deviation		$\mathbb{P}(\lvert\%\Delta P\rvert > 5\%)$	
	(1)	(2)	(3)	(4)	(5)	(6)
Treated × Post	-0.051^{***}	-0.051^{***}	0.021^{**}	0.024^{***}	0.042^{**}	0.059^{***}
	(0.012)	(0.012)	(0.008)	(0.007)	(0.018)	(0.020)
Workplace Covid exposure (DN) × Post	−0.028		0.019		−0.120	
	(0.069)		(0.042)		(0.079)	
Workplace Covid exposure (ATUS)× Post		−0.070		0.110		0.027
		(0.128)		(0.072)		(0.179)
Controls	✓	✓	✓	✓	✓	✓
Fixed effects						
UPC	✓	✓	✓	✓	✓	✓
UPC × County						
Calendar month	✓	✓	✓	✓	✓	✓
Observations	89,515	89,515	77,841	77,841	89,515	89,515
Adjusted R^2	0.95	0.95	0.44	0.44	0.31	0.31

Panel B: County-level COVID-19 exposure

	Ln(Average price)		Standard deviation		$\mathbb{P}(\lvert\%\Delta P\rvert > 5\%)$	
	(1)	(2)	(3)	(4)	(5)	(6)
Treated × Post	-0.022^{***}	-0.022^{***}	0.011^{***}	0.011^{***}	0.030^{***}	0.030^{***}
	(0.005)	(0.005)	(0.003)	(0.003)	(0.009)	(0.009)
Ln(County Covid deaths)		−0.001		0.001		−0.001
		(0.001)		(0.003)		(0.004)
Ln(Units sold)	-0.026^{***}	-0.026^{***}	0.018^{***}	0.018^{***}	-0.014^{***}	-0.014^{***}
	(0.005)	(0.005)	(0.003)	(0.003)	(0.003)	(0.003)
Controls	✓	✓	✓	✓	✓	✓
Fixed effects						
UPC						
UPC × County	✓	✓	✓	✓	✓	✓
Calendar month	✓	✓	✓	✓	✓	✓
Observations	16,482,096	16,482,096	11,068,469	11,068,469	16,482,096	16,482,096
Adjusted R^2	0.97	0.97	0.30	0.30	0.19	0.19

IA.B.6 Interest rate exposure

The Federal Reserve began raising interest rates in March 2022, shortly after the InstructGPT API introduction. If AI exposure is correlated with firms' geographic concentration in interest-rate-sensitive sectors, the price decline could reflect local demand contractions rather than AI implementation. Table IA.B.8 augments the baseline specification with controls for county-level interest rate exposure, measured as the share of private employment in interest-rate-sensitive industries (NAICS 23, 531, and 52). Odd columns control for the time-varying monthly IR exposure measure, and even columns use the baseline (January 2020) value to rule out endogenous changes in local industry composition. The treatment effects are virtually identical to the baseline across all six specifications, indicating that differential exposure to the interest rate tightening cycle does not drive the results.

Table IA.B.8: Controlling for interest rate exposure

This table presents results from difference-in-differences regressions of prices around the introduction of the InstructGPT API, controlling for county-level interest rate exposure. *Treated* equals one for firms with above-median AI exposure, measured in January 2020 using employment-share-weighted occupation-level generative AI scores from Eisfeldt et al. (2025). *Post* equals one during the six months following the implementation event. Treated and control firms are balanced on pre-treatment log average price and log units sold. The dependent variables are the log of CPI-deflated average price (columns (1) and (2)), the standard deviation of log prices (columns (3) and (4)), and an indicator for absolute month-over-month price changes exceeding 5% (columns (5) and (6)). *County IR exposure* is the share of private employment in interest-rate-sensitive industries (NAICS 23: construction, NAICS 531: real estate, and NAICS 52: finance and insurance), constructed using data from the BLS Quarterly Census of Employment and Wages. Odd columns control for the time-varying monthly county IR exposure measure interacted with post. Even columns control for county IR exposure fixed at its pre-treatment (January 2020) value interacted with post. All specifications include UPC×county and calendar month fixed effects. Controls include log units sold, log employees measured in the month before the event and interacted with post, and firm financial controls (log assets, Tobin's q, ROA, cash-to-assets, fixed assets-to-assets, R&D-to-book, market-to-book, and log firm age) measured in the quarter before the event and interacted with post. Standard errors clustered at the firm level are reported in parentheses below coefficient estimates. Statistical significance at the 1, 5, and 10 percent significance levels is denoted by ***, **, and *, respectively. Appendix A presents variable definitions.

	ln(Average price)		Standard deviation		$\mathbb{P}(\lvert\%\Delta P\rvert > 5\%)$	
	(1)	(2)	(3)	(4)	(5)	(6)
Treated × Post	−0.022***	−0.022***	0.011***	0.011***	0.062***	0.062***
	(0.005)	(0.005)	(0.003)	(0.003)	(0.012)	(0.012)
County IR exposure × Post	−0.001***		−0.000		0.003***	
	(0.000)		(0.000)		(0.001)	
County IR exposure (ex-ante) × Post		−0.001***		0.000		0.003***
		(0.000)		(0.000)		(0.001)
County IR exposure	0.000		0.001		−0.001	
	(0.000)		(0.000)		(0.001)	
Controls	✓	✓	✓	✓	✓	✓
Fixed effects						
UPC						
UPC × County	✓	✓	✓	✓	✓	✓
Calendar month	✓	✓	✓	✓	✓	✓
Observations	16,424,020	16,445,069	11,028,225	11,041,353	16,424,020	16,445,069
Adjusted R^2	0.97	0.97	0.30	0.30	0.20	0.20

IA.B.7 Placebo tests

We conduct two types of placebo tests: date-based placebos that move the event window away from the actual API introduction, and randomization-based placebos that reassign treatment status across firms.

Placebo event dates. If the estimated treatment effects reflect a genuine response to the InstructGPT API introduction, they should not appear when the event date is moved to a period with no plausible AI shock. Table IA.B.9 presents placebo tests using January 2019 (36 months before the actual event) and January 2021 (12 months before) as alternative event dates. For each placebo date, we re-estimate the baseline specification for all three main outcomes. Across all six specifications, the placebo coefficients are small, of inconsistent sign, and statistically insignificant, supporting the interpretation that the main results are driven by the actual AI implementation event.

Randomization inference. As an alternative to conventional asymptotic inference, we conduct a permutation test that directly evaluates the likelihood of observing our treatment effects under random assignment. In each of 5,000 iterations, we randomly reassign the treatment label across firms—preserving the number of treated firms—drop three outlier firms whose product portfolios are sufficiently large to exert disproportionate influence on the permutation distribution, re-match treated and control products using propensity-score matching on pre-treatment price levels and sales volumes, and re-estimate the full baseline specification. For each regression, we store the t-statistic on the treatment interaction. Figure IA.B.4 plots the resulting placebo distributions alongside the true t-statistics from our main estimates (dashed vertical lines). Across all three outcomes, the true t-statistics fall in the extreme tails of the placebo distributions: the two-sided permutation p-values are 0.011 for average prices, 0.015 for price dispersion, and 0.009 for repricing frequency. These results provide nonparametric evidence that the estimated treatment effects are unlikely to arise from any arbitrary partition of firms into treated and control groups.

Table IA.B.9: Placebo tests using January 2019 and January 2021 as event dates

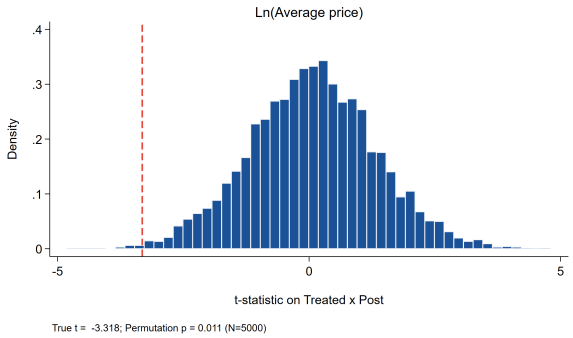
This table presents results from placebo tests that move the event date away from the actual InstructGPT API introduction in January 2022. Columns (1) to (3) use January 2019 (36 months before) and columns (4) to (6) use January 2021 (12 months before) as placebo event dates. The dependent variables are the log of CPI-deflated average price, within-product price standard deviation, and an indicator for whether the absolute month-over-month log price change exceeds 5%. *Treated* equals one for firms with above-median AI exposure, measured in January 2020 using employment-share-weighted occupation-level generative AI scores from Eisfeldt et al. (2025). *Post* equals one during the six months following the placebo event date. All specifications include firm-product and calendar month fixed effects. Standard errors clustered at the firm level are reported in parentheses below coefficient estimates. Statistical significance at the 1, 5, and 10 percent significance levels is denoted by ***, **, and *, respectively. Appendix A presents variable definitions.

	January 2019			January 2021		
	ln(Average price)	Standard deviation	$\mathbb{P}(\lvert\%\Delta P\rvert > 5\%)$	Ln(Average price)	Standard deviation	$\mathbb{P}(\lvert\%\Delta P\rvert > 5\%)$
	(1)	(2)	(3)	(4)	(5)	(6)
Treated × Post	−0.003	0.005	−0.005	−0.024	−0.001	0.012
	(0.010)	(0.008)	(0.012)	(0.017)	(0.008)	(0.014)
Fixed effects						
UPC	✓	✓	✓	✓	✓	✓
Calendar month	✓	✓	✓	✓	✓	✓
Observations	109,472	105,372	109,472	109,645	97,303	109,645
Adjusted R^2	0.97	0.49	0.27	0.95	0.40	0.28

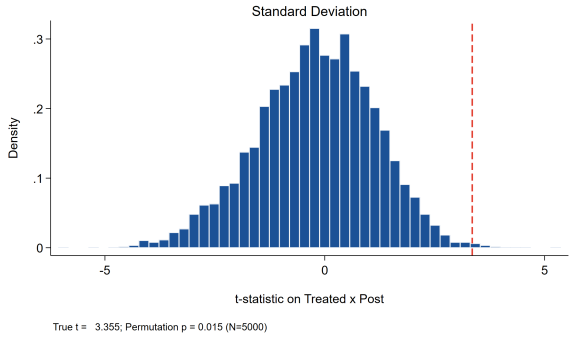
Figure IA.B.4: Placebo tests randomizing treatment status

This figure presents the distribution of placebo t-statistics from 5,000 random permutations of the treatment indicator. In each iteration, we randomly reassign the treated label across firms (preserving the number of treated firms), drop three outlier firms whose product portfolios are sufficiently large to exert disproportionate influence on the permutation distribution, re-match treated and control products using propensity-score matching on pre-treatment price levels and sales volumes, and re-estimate the baseline difference-in-differences specification. Panel A reports results for log average price, Panel B for the standard deviation of log price, and Panel C for the frequency of absolute price changes exceeding 5%. The dashed vertical line indicates the true estimated t-statistic. Two-sided permutation p-values are 0.011, 0.015, and 0.009 for Panels A, B, and C, respectively. All specifications include firm-product and calendar month fixed effects, with controls for log units sold, log employees interacted with post, and firm financial characteristics (log assets, Tobin's q, ROA, cash-to-assets, fixed assets-to-assets, R&D-to-assets, market-to-book, and log firm age) interacted with post. Standard errors are clustered at the firm level. Appendix A presents variable definitions.

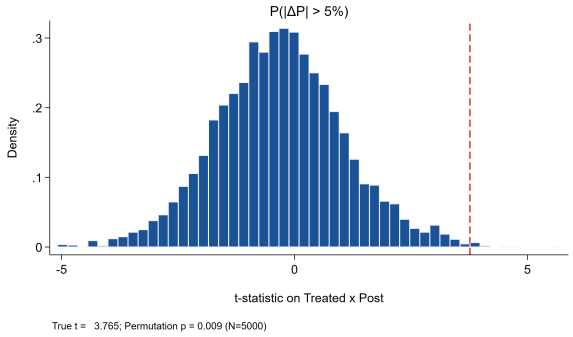
Panel A: ln(Average price)



Panel B: Standard Deviation



Panel C: $\mathbb{P}(|\Delta P| > 5\%)$



IA.B.8 Alternative augmentation vs. automation decomposition

As a complement to the augmentation vs. automation decomposition in Table 11, we examine an alternative decomposition based on the Anthropic Economic Index (Appel, Massenkoff, McCrory, McCain, Heller, Neylon, and Tamkin, 2026), which classifies the mode of human–AI interaction in conversations tagged to O*NET tasks. We focus on two interaction modes: *task iteration*, in which users iteratively refine outputs with the AI model, and *directive*, in which users issue a single instruction and the AI executes independently. Task iteration most closely corresponds to augmentation—the human remains in the loop, using AI to enhance analysis—while directive corresponds to automation.

Table IA.B.10 reports the results. Firms with above-median task iteration exposure exhibit a statistically significant price decline of 4.2% (column 1) and a 4.9 percentage point increase in repricing frequency (column 3). By contrast, the directive split produces smaller and less significant effects: a 1.5% price decline that is not statistically significant (column 4) and a marginally significant increase in repricing frequency (column 6). These results mirror the augmentation vs. automation pattern and provide independent corroboration using an external data source: the pricing effects of AI load on occupations where workers interact iteratively with AI tools, not where AI executes tasks autonomously.

Table IA.B.10: Augmentation vs. automation (Anthropic Economic Index)

This table presents results from difference-in-differences regressions that decompose AI exposure into augmentation and automation components using the Anthropic Economic Index (AEI). The AEI classifies the mode of human–AI interaction in conversations tagged to O*NET occupational tasks. Augmentation is proxied by the *task iteration* mode, in which users iteratively refine outputs with the AI model. Automation is proxied by the *directive* mode, in which users issue single instructions that the AI executes independently. Firm-level scores are computed as employment-share-weighted averages of occupation-level AEI scores from Revelio Labs. Treatment is defined as above-median exposure measured in January 2020, with propensity score matching on pre-event log average price and log units sold. *Post* equals one during the six months following the InstructGPT API introduction in January 2022. All specifications include firm-product and calendar month fixed effects and controls for log units sold, log employees interacted with post, and firm financial characteristics interacted with post. Standard errors clustered at the firm level are reported in parentheses. Statistical significance at the 1, 5, and 10 percent significance levels is denoted by ***, **, and *, respectively. Appendix A presents variable definitions.

	Augmentation (task iteration)			Automation (directive)		
	Ln(Average price)	Standard deviation	$\mathbb{P}(\lvert\%\Delta P\rvert > 5\%)$	Ln(Average price)	Standard deviation	$\mathbb{P}(\lvert\%\Delta P\rvert > 5\%)$
	(1)	(2)	(3)	(4)	(5)	(6)
Treated × Post	-0.042^{**}	0.010	0.049^{***}	−0.015	0.011	0.037^{*}
	(0.018)	(0.008)	(0.018)	(0.017)	(0.009)	(0.021)
Fixed effects						
Firm product	✓	✓	✓	✓	✓	✓
Calendar month	✓	✓	✓	✓	✓	✓
Observations	87,772	75,371	87,772	81,529	70,411	81,529
Adjusted R^2	0.95	0.45	0.30	0.95	0.45	0.29

IA.B.9 AI adoption: Earnings conference calls

As a complement to the job posting and 10-K analyses in the main text, we examine AI-related keywords in quarterly earnings conference call transcripts, which provide more timely information about corporate strategy. Table IA.B.11 reports PPML regressions of AI keyword counts on the *Treated* × *Post* interaction, with firm and calendar quarter fixed effects. High-exposure firms significantly increase AI keyword usage in earnings calls after the event, both in total and across individual sections (prepared remarks, management Q&A, and analyst Q&A). Figure IA.B.5 presents the corresponding event-study estimates, which show parallel pre-trends and a sustained increase in AI mentions beginning at the event date, consistent with the job posting and 10-K evidence.

Figure IA.B.5: AI keyword event study using earnings conference calls
This figure plots the β_k coefficients and 90% confidence intervals from event study regressions of the form: AI Keywords$_{it}$ $= \alpha_i + \alpha_t + \sum_{k \neq -1} \beta_k \mathbb{1}[\text{Event time} = k] \times \text{Treated}_i + \epsilon_{it}$, where i denotes firms and t calendar quarters. The dependent variable is the count of AI-related keywords (artificial intelligence, machine learning, ChatGPT, LLM, neural network, predictive analytics, pricing algorithm) in firm i's earnings call transcript in quarter t. AI exposure is computed as the employment-share-weighted average of occupation-level generative AI scores from Eisfeldt et al. (2025), where scores capture the extent to which generative AI is expected to reduce the time required to complete the tasks typically performed by each occupation. *Treated* equals one for firms with above-median AI exposure measured in January 2020. Event time runs from −4 to +4 quarters relative to the InstructGPT API introduction in 2022 Q1, with $k = -1$ as the omitted base period. α_i and α_t denote firm and calendar quarter fixed effects, respectively. Standard errors are clustered at the firm level. Appendix A presents variable definitions.

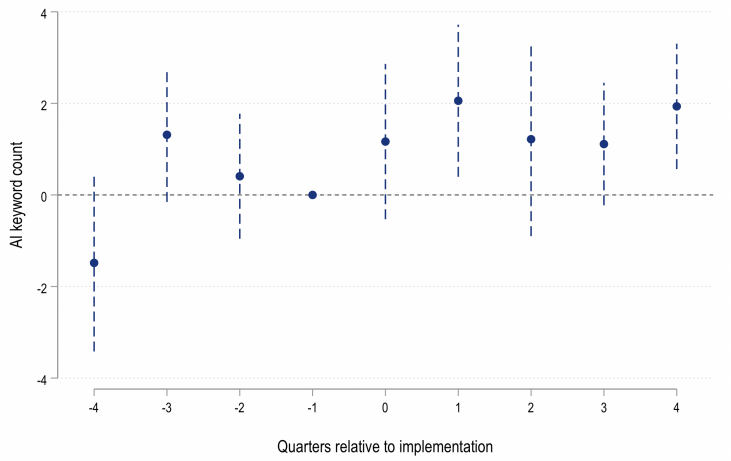


Table IA.B.11: AI keyword counts in earnings conference calls

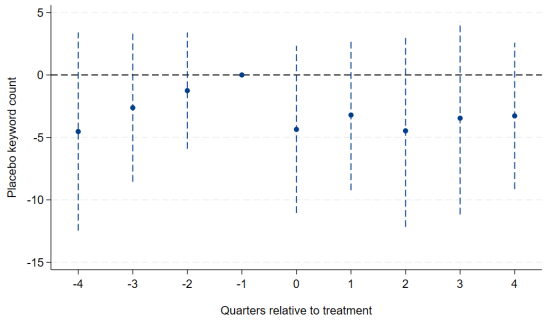
This table presents results from Poisson pseudo-maximum-likelihood (PPML) regressions of AI-related keyword counts in quarterly earnings conference call transcripts around AI implementation events. The dependent variable is the count of AI-related keywords (artificial intelligence, machine learning, ChatGPT, LLM, neural network, predictive analytics, pricing algorithm) in firm i's earnings call transcript in quarter t. Column (1) uses the total keyword count across the full transcript. Columns (2)–(4) decompose the count by transcript section: prepared remarks (column (2)), management Q&A (column (3)), and analyst Q&A (column (4)). *Treated* equals one for firms with above-median AI exposure, measured in January 2020 using employment-share-weighted occupation-level generative AI scores from Eisfeldt et al. (2025). *Post* equals one after the InstructGPT API introduction 2022 Q1. All specifications include firm and calendar quarter fixed effects. Standard errors clustered at the firm level are reported in parentheses below coefficient estimates. Statistical significance at the 1, 5, and 10 percent significance levels is denoted by ***, **, and *, respectively. Appendix A presents variable definitions.

	AI keyword count			
	Total	Prepared remarks	Management Q&A	Analyst Q&A
	(1)	(2)	(3)	(4)
Treated × Post	1.800***	1.629**	2.265***	0.348
	(0.609)	(0.746)	(0.740)	(2.092)
Fixed effects				
Firm	✓	✓	✓	✓
Quarter	✓	✓	✓	✓
Observations	297	238	155	30

Figure IA.B.6: Placebo keyword event studies

This figure plots the β_k coefficients and 90% confidence intervals from event study regressions of the form: Placebo Keywords$_{it} = \alpha_i + \alpha_t + \sum_{k \neq -1} \beta_k \mathbb{1}[\text{Event time} = k] \times \text{Treated}_i + \epsilon_{it}$, where i denotes firms and t denotes calendar periods. The dependent variable is the count of non-AI technology keywords (cloud, ERP, CRM, digital transformation, digitization, cybersecurity, data warehouse, software modernization, omnichannel, e-commerce infrastructure, automation excluding AI contexts, enterprise systems). In Panel A, the dependent variable is measured in quarterly earnings conference call transcripts, and event time runs from -4 to $+4$ quarters relative to the InstructGPT API introduction in 2022 Q1. In Panel B, the dependent variable is measured in annual 10-K filings, and event time runs from -4 to $+3$ years relative to 2022. AI exposure is computed as the employment-share-weighted average of occupation-level generative AI scores from Eisfeldt et al. (2025). *Treated* equals one for firms with above-median AI exposure measured in January 2020. α_i and α_t denote firm and calendar period fixed effects, respectively, with $k = -1$ as the omitted base period. Standard errors are clustered at the firm level. All coefficients are statistically indistinguishable from zero, indicating no differential trend in non-AI technology discussion between treated and control firms. Appendix A presents variable definitions.

Panel A: Earnings conference calls



Panel B: 10-K filings

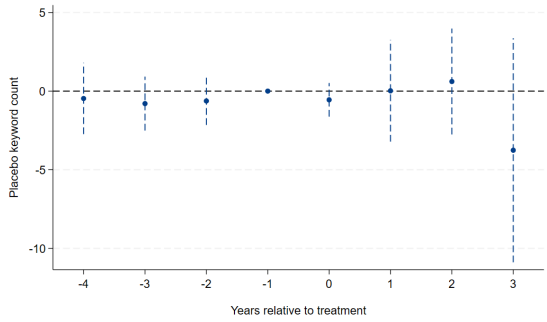
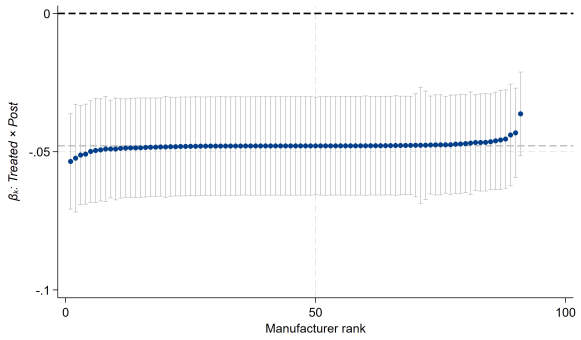


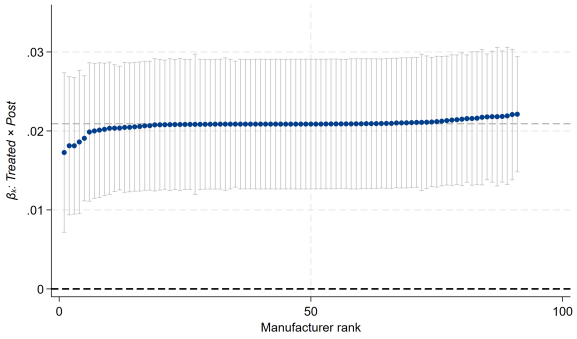
Figure IA.B.7: Leave-one-out manufacturer analysis

This figure presents the estimated difference-in-differences coefficient on Treated × Post when iteratively dropping each manufacturer from the sample. For each of the 91 manufacturers in the matched sample, we exclude that manufacturer and re-estimate the baseline specification. Coefficients are ranked from smallest to largest, plotted with 90% confidence intervals (gray whiskers). The thin dashed horizontal line indicates the full-sample estimate, and the thick dashed horizontal line marks zero. Panel A reports results for log average price (full-sample $\hat{\beta} = -0.048$), Panel B for the standard deviation of log price (full-sample $\hat{\beta} = 0.021$), and Panel C for the frequency of absolute price changes exceeding 5% (full-sample $\hat{\beta} = 0.063$). No single manufacturer drives the results: all leave-one-out estimates remain negative for price levels, and positive for price dispersion and repricing frequency. All specifications include UPC and calendar month fixed effects, with controls for log units sold, log employees interacted with post, and firm financial characteristics (log assets, Tobin's q, ROA, cash-to-assets, fixed assets-to-assets, R&D-to-assets, market-to-book, and log firm age) interacted with post. Standard errors are clustered at the firm level. Appendix A presents variable definitions.

Panel A: ln(Average price)



Panel B: Standard Deviation



Panel C: $\mathbb{P}(|\Delta P| > 5\%)$

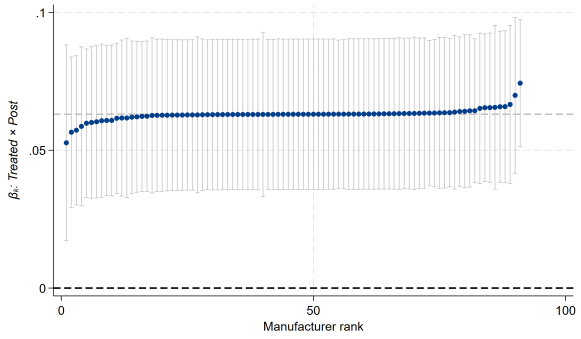
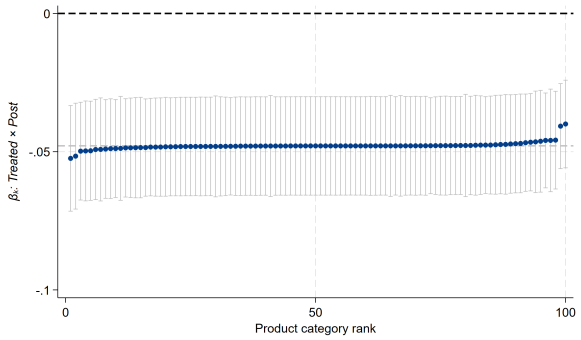


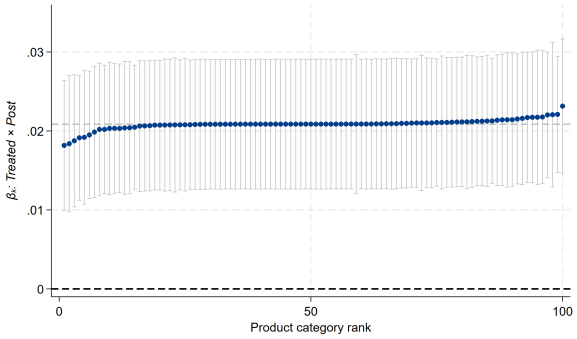
Figure IA.B.8: Leave-one-out product category analysis

This figure presents the estimated difference-in-differences coefficient on Treated × Post when iteratively dropping each product category from the sample. For each of the 100 Nielsen product groups in the matched sample, we exclude all products in that category and re-estimate the baseline specification. Coefficients are ranked from smallest to largest, plotted with 90% confidence intervals (gray whiskers). The thin dashed horizontal line indicates the full-sample estimate, and the thick dashed horizontal line marks zero. Panel A reports results for log average price (full-sample $\hat{\beta} = -0.048$), Panel B for the standard deviation of log price (full-sample $\hat{\beta} = 0.021$), and Panel C for the frequency of absolute price changes exceeding 5% (full-sample $\hat{\beta} = 0.063$). No single product category drives the results: all leave-one-out estimates remain negative for price levels, and positive for price dispersion and repricing frequency. All specifications include UPC and calendar month fixed effects, with controls for log units sold, log employees interacted with post, and firm financial characteristics (log assets, Tobin's q, ROA, cash-to-assets, fixed assets-to-assets, R&D-to-assets, market-to-book, and log firm age) interacted with post. Standard errors are clustered at the firm level. Appendix A presents variable definitions.

Panel A: ln(Average price)



Panel B: Standard Deviation



Panel C: $\mathbb{P}(|\Delta P| > 5\%)$

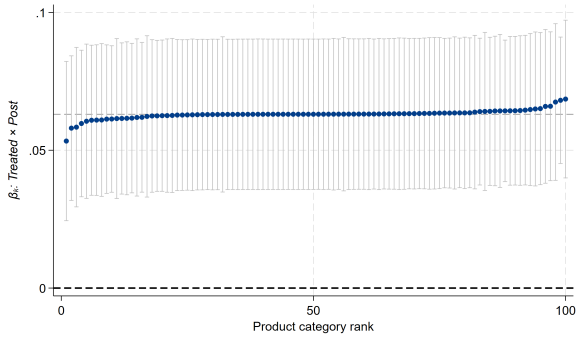


Table IA.B.12: Wild cluster bootstrap inference

This table reports wild cluster bootstrap p-values for the three main difference-in-differences specifications. Column (1) reports the estimated coefficient on Treated × Post, column (2) reports the standard error clustered at the firm level, column (3) reports the analytical p-value from the clustered variance-covariance matrix, and column (4) reports the p-value from a wild cluster bootstrap-t procedure with 9,999 replications (Roodman et al., 2019). The bootstrap imposes the null hypothesis and uses Webb (2023) six-point weights. All specifications include UPC and calendar month fixed effects, with controls for log units sold, log employees interacted with post, and firm financial characteristics (log assets, Tobin's q, ROA, cash-to-assets, fixed assets-to-assets, R&D-to-assets, market-to-book, and log firm age) interacted with post. Appendix A presents variable definitions.

	(1) $\hat{\beta}$	(2) SE	(3) p(cluster)	(4) p(bootstrap)
Ln(Average price)	−0.048	0.011	<0.001	<0.001
SD(ln price)	0.021	0.005	<0.001	0.004
$\mathbb{P}(\lvert\Delta P\rvert > 5\%)$	0.063	0.017	<0.001	0.002
Fixed effects	UPC, Calendar month			
Controls	✓			
Bootstrap replications	9,999			
Clusters (firms)	91			