NBER WORKING PAPER SERIES

AI AND THE QUANTITY AND QUALITY OF CREATIVE PRODUCTS: HAVE LLMS BOOSTED CREATION OF VALUABLE BOOKS?

Imke Reimers
Joel Waldfogel

Working Paper 34777
http://www.nber.org/papers/w34777

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
January 2026, revised May 2026

We are grateful to seminar participants at the NBER Winter 2026 Digitization meetings, New York University, the University of Illinois, the University of South Carolina, the 2026 IIOC, Netflix, and the University of Minnesota. We thank Shane Greenstein and Ginger Jin for helpful comments, and we thank Pangram Labs for support with the use of their AI detection tool. The views expressed herein are those of the authors and do not necessarily reflect the views of the National Bureau of Economic Research.

AI and the Quantity and Quality of Creative Products: Have LLMs Boosted Creation of Valuable Books?

Imke Reimers and Joel Waldfogel

NBER Working Paper No. 34777
January 2026, revised May 2026
JEL No. L16, L82

ABSTRACT

The diffusion of LLMs from 2022 to 2025 tripled new book releases. While average book quality, measured by usage, declined, the surge in releases raised the number of modest-quality books. Direct evidence using AI detection shows that AI-containing books have lower quality, and their rising share – topping half of 2025 releases – drives the overall decline. A nested logit calibration shows that AI books raised consumer surplus by seven percent in 2025. Author selection accounts for most of the AI quality differential, and the AI-human differential shrinks over time. Finally, AI has not displaced authors active prior to LLMs.

Imke Reimers
Cornell University
SC Johnson College of Business
imke.reimers@cornell.edu

Joel Waldfogel
University of Minnesota
Carlson School of Management
and NBER
jwaldfog@umn.edu

Introduction

Artificial intelligence, in the form of large language models (LLMs), promises to revolutionize many tasks, occupations, and industries, including – and perhaps especially – the creative industries. Because of LLMs' ability to create text, book publishing may be particularly susceptible; and during the period 2022-2025, LLMs became capable of producing sustained text passages. LLMs reduce the cost of producing books; and reductions in entry costs raise the number of new products brought to market. Additional entry can be especially beneficial when product quality is unpredictable at entry, so that some new products have realized quality in, or near, the right tail of the quality distribution.

The appearance of LLMs would not simply facilitate more "draws from the urn." LLMs may also change the distribution of quality realizations, for example by granting writing ability to those who lack it, while either complementing, or substituting for, the ability of skilled writers. For example, if LLMs allow the creation of books with low average quality and little quality variance, they will deliver no books in the right tail of the quality distribution and little unpredictable benefit to consumers. By contrast, if LLMs facilitated the creation of more books drawn from the pre-LLM quality distribution, then the number of high-quality new releases would rise (e.g. the n^{th}-best release in a month would be better than before), with potentially substantial welfare benefits. Hence, the welfare effect of AI in the book market depends on what it does to both the *number* of new books and their *quality distribution*.[1] The eventual effect of LLMs also depends on whether their adoption displaces activity by incumbent authors.

These considerations motivate this paper's questions. First, how do LLMs affect the number of new books created? Second, are AI books better or worse; and how does LLM-enhanced book production affect the resulting distributions of usage/quality? Third, how

[1] By "quality," we mean whatever induces people to buy books, resulting in the "mean utilities" that enter welfare calculations.

do these effects on creation and quality together affect consumer welfare? Finally, which authors use AI, and how does its diffusion affect releases by incumbent authors? Answering these questions requires data on the supply of, and demand for, new books.[2] Using data on new books offered for sale at Amazon, we document that the number of new titles appearing each month nearly tripled between 2022 and late 2025 and rose by a factor of nearly ten in some categories; and the increase in new titles 2022-2025 coincides with both the diffusion of LLMs and the incidence of detected AI in books.

We document the impact of LLMs on the book quality distribution in two broad ways. First, we employ an indirect approach that does not rely on AI detection, asking how the quality of books in a category evolves as the number of book releases in that category increases. To this end, we assemble a stratified random sample of over 330,000 titles that are representative of the 10 million ebooks released at Amazon between 2020 and 2025. We observe the number of ratings for each book as of April 2026 – our main usage/quality measure – as well as each book's April-2026 sales rank and average star rating. With population weighting, the stratified random sample allows us to draw inferences about the effect of the LLM influx on the distribution of new book qualities. The LLM-affected vintages have lower average quality and lower quality at particular percentiles of the distribution. At the same time, given the dispersion in realized quality, the large increase in releases raises the absolute number of moderately high-quality releases.

Second, we take a direct approach employing AI detection on over 50,000 randomly selected titles. We document the growth of AI usage and compare the quality of "AI books" and "non-AI books" (books with and without detected AI). We have four findings. First, the timing of AI growth tracks the growth in releases: Detected AI use is roughly zero through 2022, rises to 30 percent in 2023, to 45 percent in 2024, and surpasses 60 percent during

[2]These data are not readily available. Neither copyright registrations nor repositories of international standard book numbers (ISBNs) show the patterns we document in this study.

2025. Second, AI books garner substantially less usage per title. Most have very little usage, and a modest share is somewhat used, both when measured by the number of ratings and eventual sales ranks. AI books are also worse in the sense of having lower star ratings. Third, the human-AI usage gap narrows substantially between 2023 and 2025. Finally, the number and quality of human-authored books has remained stable. The AI influx thus raises welfare modestly.

Using a simple calibrated nested logit demand model for comparing the value of all releases to just the human-written releases, we find that AI books raise consumer surplus (CS) by less than one percent in 2023, by over three percent in 2024, and by about seven percent in 2025, when over half of new books employ AI and better authors use it. The rise in CS due to AI is small in comparison to the rapid increase in new releases.

We also explore questions about author careers and the human-AI gap using a separate "census" dataset on all 479,000 books released in eight subcategories between 2008 and 2025, along with an AI detection subsample covering all of the books released by over 6,000 authors. We find that author selection explains most of the human-AI gap: Inexperienced and lower-quality authors – whose pre-LLM books had lower average usage – are more likely to adopt AI; and just over a third of the raw human-AI gap remains when we identify the gap using within-author variation. Finally, the influx of new authors – many of whom use AI – does not displace activity by authors active prior to the LLM influx, whose releases rise after 2023.

The paper proceeds in seven sections after the introduction. Section 1 discusses the background, including the development of LLMs, author use of LLMs, the evolution of the book market, and the related academic literatures. Section 2 provides a simple theoretical framework showing how the welfare effect of an influx of new products depends on both the number of new products and the quality distribution. Section 3 describes the data used in the study, including information on the volume of books released, as well as book-level data

on books' success measures and, for a subset, book-level information on detected AI use. Section 4 presents results on the effects of LLMs on new book entry and on the distribution of book quality, based on an "indirect" approach that uses the whole sample without AI detection. Section 5 uses the AI detection subsample to directly measure the AI quality differential as well as its effect on consumer welfare. Section 6 uses the subcategory census data to explore author AI adoption, the sources of the human-AI quality gap, and incumbent author displacement. A brief conclusion follows in Section 7.

1 Background

1.1 AI and authors

The public deployment of ChatGPT in late 2022 brought LLMs to prominence. According to Liang et al. (2025), "LLM usage surged following the release of ChatGPT in November 2022." In book publishing, surveys indicate that LLMs are both attracting new writers and being used by existing writers. An April 2025 survey of 1,229 authors – most of whom were exclusively self-published and had started publishing in 2023 or earlier – found authors divided in their views. Nearly half (45 percent) were using AI to "assist with their work" while 48 percent were neither using nor planning to use AI. The remaining 7 percent indicated they might use AI in the future.[3] Among those not using AI, more than three quarters thought its use "unethical." Among authors using AI, ChatGPT was the most popular LLM, used by 85 percent of users.[4]

A cottage industry now provides LLM advice to would-be first-time authors. For example, `bookautoai.com` offers support so that "you can publish a fully formatted, human-level book

[3] See `https://insights.bookbub.com/how-authors-are-thinking-about-ai-survey/`.

[4] In another survey from November 2025, 61 percent of writers – and 42 percent of fiction authors – reported using AI tools. See `https://www.publishersweekly.com/pw/by-topic/digital/copyright/article/99019-new-report-examines-writers-attitudes-toward-ai.html`.

in under 30 minutes." Some authors have apparently achieved success using AI. Tim Boucher has written hundreds of books using AI and reports having earned thousands of dollars.[5] At the same time, there is a great deal of low-quality work, which critics derisively term "AI-slop."[6]

Participants in the book publishing industry regard AI with concern and suspicion. The Authors Guild argues that publishers and sellers should be compelled to disclose the use of AI in book creation.[7] An open letter from authors, including Colleen Hoover and Dennis Lehane, exhorted publishers to "pledge that they will never release books that were created by machines" and not to "replace their human staff with AI tools or degrade their positions into AI monitors."[8]

Intellectual property protection for AI-enabled writing is limited. The US Copyright Office has determined that "the outputs of generative AI can be protected by copyright only where a human author has determined sufficient expressive elements."[9] Amazon also appears to be cautious in its embrace of AI. Amazon requires authors to disclose AI-generated, but not AI-assisted, content; but Amazon does not disclose the information to consumers.[10]

[5] See https://www.newsweek.com/ai-books-art-money-artificial-intelligence-1799923.

[6] Goodreads has many user-created lists of works identified as "AI slop" by Goodreads users. See https://www.goodreads.com/list/tag/ai-slop, https://www.goodreads.com/list/tag/ai-book, and https://www.goodreads.com/list/show/219041.Books_I_Have_No_Doubt_Are_Written_By_A_I_.

[7] The Authors Guild argues for rules requiring "authors, publishers, platforms, and marketplaces to label AI-generated works and otherwise identify when a significant portion of a written work has been generated by AI." See https://authorsguild.org/advocacy/artificial-intelligence/faq/.

[8] The initial list, from June 2025, included 70 authors in total. By April 16, 2026, the list included 448 authors and other professionals in the book publishing industry. See https://lithub.com/against-ai-an-open-letter-from-writers-to-publishers/ and https://docs.google.com/document/d/1TmtwfEvRAQkR35lg4xqBNilU3TUIYNWN1xNxx9hrM14/edit?tab=t.0.

[9] The Copyright Office further clarifies: "This can include situations where a human-authored work is perceptible in an AI output, or a human makes creative arrangements or modifications of the output, but not the mere provision of prompts." See https://www.copyright.gov/newsnet/2025/1060.html and https://www.copyright.gov/ai/Copyright-and-Artificial-Intelligence-Part-2-Copyrightability-Report.pdf. See also Lutes et al. (2025).

[10] "We define AI-generated content as text, images, or translations created by an AI-based tool. If you used an AI-based tool to create the actual content (whether text, images, or translations), it is considered "AI-generated," even if you applied substantial edits afterwards." See https://kdp.amazon.com/en_US/help/topic/G200672390.

Moreover, Amazon limits the number of books that authors may upload per day.[11]

1.2 The evolution of the book market

Prior to digitization, books – and the book market – were easy to define. Books were physical products, initially released as hardbound editions and later as paperbacks, for sale at retail establishments. In sales terms, the market was dominated by a handful of New York publishing houses. Digitization changed the market substantially. With Amazon's creation of the Kindle environment in 2008, including the Kindle device and the Kindle Direct Publishing platform, many books were released primarily as electronic products. Moreover, books came to market without the curation of the traditional gatekeepers. As Waldfogel and Reimers (2015) document, self-published works became a significant part of the market in the years after the release of the Kindle. Many works initially appearing as self-published books became best sellers (Peukert and Reimers, 2022). Self-publishing, the main vehicle for LLM-aided books post-2022, was already an established avenue to commercial success.

Measuring the number of new books is not easy in the digital era. Traditional publishers sought copyright registration and international standard book numbers (ISBNs). Individuals publishing outside of traditional avenues often do neither. Between 2020 and 2024, annual nondramatic literary copyright registrations show no evidence of an LLM-induced influx.[12] The number of ISBNs issued per year, from Bowker's Books in Print, also shows no evidence of an LLM-induced increase in releases; it falls from 2.74 million new books in 2023 to 2.37 million in 2024. Hence, the influx we document below is largely a digital – and an Amazon-centered – phenomenon. While a single company, Amazon has over two thirds of the global ebook market.[13]

[11] See `https://www.theguardian.com/books/2023/sep/20/amazon-restricts-authors-from-self-publishing-more-than-three-books-a-day-after-ai-concerns`.

[12] New registrations went from 152 thousand to 156 thousand, albeit with fluctuation between 129 and 189 thousand. See, for example, `https://www.copyright.gov/reports/annual/2020/ar2020.pdf`.

[13] "Amazon holds a dominant 68% market share in global e-book sales, which rises to 83% when Kindle

1.3 Relevant literatures

This paper connects three strands of prior research: (i) the economics of digitization and creative-goods supply, (ii) empirical and theoretical work on generative AI in creative and knowledge-intensive tasks, and (iii) recent analysis of AI-driven content markets.

A first strand examines how digitization has affected the production and quality of creative goods. Brynjolfsson et al. (2003) document the benefit of access to a long tail of extant products. Other work explores supply impacts. In the book market, Waldfogel and Reimers (2015) show that digital distribution and self-publishing substantially increased the number of book titles released, while also lowering entry barriers for new authors. Related work in music finds similar effects: digitization expanded variety and may have increased realized quality by bringing forth many products of unpredictable quality, some of which turn out to be good (Aguiar and Waldfogel, 2018). These studies show that technological changes that reduce production and distribution costs can simultaneously raise both the volume and the realized quality of new creative outputs in contexts where the new technology did not directly affect the quality of the underlying work.

A second strand evaluates how generative AI affects cognitive and creative work. Experiments show large productivity gains from access to large-language models. Noy and Zhang (2023) report that GPT-based writing assistance reduces task completion times by roughly 40 percent and disproportionately benefits lower-skilled workers. Brynjolfsson et al. (2025) find that an AI-based assistant in a large call-center improves worker productivity by 14 percent on average, with the largest gains for novice workers. In creative tasks, recent work documents improvements in fluency and coherence but mixed effects on novelty and diversity (Doshi and Hauser, 2024; Padmakumar et al., 2024). On the other hand, while time saved in some tasks may be reallocated to higher-level activities, the need to verify can offset

Unlimited is included." See `https://goodereader.com/blog/electronic-readers/amazon-is-the-dominant-player-in-ebook-sales`.

gains when error costs are high or outputs are difficult to evaluate (Dell'Acqua et al., 2023; Vaithilingam et al., 2022). Nevertheless, Kusumegi et al. (2025) document that academics adopting LLMs produce substantially more working papers. Together, these studies raise the possibility that AI may significantly increase new book production, but they leave open the exact nature of effects on quality. AI need not raise the quality of incumbent creators: Chen et al. (2025) conduct an experiment in which incumbents facing competition from AI sacrifice quality.

A third strand considers the effect of generative-AI on creative markets, thus far with a focus on images. Goldberg and Lam (2025) show that generative-AI tools increase entry and product variety while sharply displacing output of non-AI images in a large stock image marketplace. Peukert et al. (2024) find that stock image creators reduce their output when they learn that their work is used for AI training. Finally, Kim et al. (2026) document that the launch of text-to-image generative AI depresses output among incumbent creators of anime- and manga-style artwork.

2 Theory

Potential authors create a book if the expected benefit of doing so exceeds its cost. AI tools have reduced the cost of book creation, and cost reduction stimulates additional creation. Moreover, the quality of creative products is understood to be unpredictable at the time of entry (Caves, 2000), and with unpredictability, general cost reduction can raise welfare by delivering additional products from a quality distribution similar to that for existing products, as digitization did with media products (Aguiar and Waldfogel, 2018). Yet, the arrival of AI is not simply a reduction in the cost of entry. AI may allow people to write better books, or more books of similar quality. It might also change investment incentives for incumbents. Hence, LLMs can also bring changes throughout the realized product qual-

ity distribution. These concerns motivate the paper's main question: What happens to the quality distribution – and aggregate value – of entering books when AI tools become available?

What matters for consumer welfare is the distribution of realized product qualities and, in particular, the number of products attracting substantial usage. A simple model of the realized distributions illustrates how AI tools might affect the distribution of product qualities and the value of the choice set. Define δ_j as the mean utility, or "quality" of product j.[14] Define μ_r and σ_r as the mean and standard deviations of normally distributed realized qualities in regime r, where r is either pre- or post-LLM. Define n_r as the number of products introduced in each regime. With $\{\delta_j\}$ as the set of available products, per capita CS in the plain logit is proportional to $\ln(1 + \Sigma e^{\delta_j})$.

Some cases are helpful for organizing thinking. First, suppose that n rises while the distribution from which qualities are drawn remains constant. Then the average quality would remain the same, and the quality of a product at any particular percentile would remain constant. If the number of entering products doubled from 1,000 to 2,000, then the 90^{th}-percentile products – the 100^{th}-best before and the 200^{th}-best after – would be of equal quality. Conversely, the 100^{th}-best product would be better when 2,000 products entered than when 1,000 products entered; and there would now be more products above the former 100^{th}-best book threshold. The increase in entry would, of course, raise the value of the choice set.

For a second case, suppose that μ declines while n and σ remain constant. Then – comparing the new vs the old cohort – the average quality would fall, as would the quality at any percentile of the distribution. Moreover, the k^{th}-best new product's quality would also fall. A choice set produced in this way, rather than with the prior distribution, would

[14] In the plain logit model, $\delta_j = \ln(q_j/M) - \ln((M - \Sigma q_j)/M)$, where q_j is the quantity sold for book j and M is market size. We return to this approach in Section 5.3.

have lower value.

Finally, consider a third – and potentially more relevant – case, where n rises while μ declines. Then effects are slightly more nuanced: The average quality would, of course, decline, as would the quality at any particular percentile of the distribution; but the quality of the k^{th}-best product might rise or fall, depending on the magnitude of the increase in entry relative to the fall in average appeal.

We illustrate the third case in Figure 1. Panel (a) compares the pre-LLM distribution (in blue) with the larger post-LLM distribution (in red), where the post-LLM distribution is illustrated with five times more products and a lower mean.[15] Panel (b) compares the quality of products at the same *percentile* of the pre- and post-LLM distributions, and the quality at any percentile of the distribution is lower in the post-LLM distribution. Panel (c) compares products at the same *rank position* (e.g. 200$^{\text{th}}$-best product) in the pre and post-LLM distributions, and the product at the k^{th}-best position is better under the post-LLM distribution. An increase in the number of new products in, or near, the right tail of the quality distribution is important, as it can substantially raise the value of the choice set to consumers.

The framework in this section helps to focus our attention on this paper's main objects of study: the number of new AI and non-AI products, the average quality of entering books, the quality at particular percentiles of the distribution, the qualities at absolute rank positions, and the distributions of AI and non-AI book quality.[16]

[15] We characterize the pre-LLM period with 200,000 draws from a standard normal distribution, and we represent the LLM era with 1 million observations from a distribution with the same variance and a mean of -0.5.

[16] The LLM-era distribution may be a mixture of a human-generated distribution resembling the pre-LLM era and the machine-enabled distribution. This distinction does not affect our empirical tests, but it does shape the interpretation of the LLM welfare impact in Section 5.3.

3 Data

We are interested in characterizing volumes of new books entering before and after the influx of AI tools, as well as the distributions of those books' usage and therefore the quality that we infer. Ideally, we would have a full list of books published over time, along with measures of those books' sales in each time period (and perhaps other measures of perceived quality, such as star ratings). While ideal data are not available, we have two promising ways to measure the phenomena of interest. First, we observe the number of new books by category and month, as well as product-level information on usage and satisfaction as of early 2026. We use these data for an indirect approach in which we compare quality distributions across categories and over time. Second, we use AI detection on a random sample of these books to compare usage measures for books that do, and do not, contain detected AI.

3.1 Aggregate book release data

We obtain the number of books published over time using queries to Amazon's advanced book search function at `https://www.amazon.com/advanced-search/books`. We request new, English-language ebooks, and we specify their publication dates by month and by category for all 30 Amazon categories between 2020 and 2025.[17] A query delivers a list whose first page indicates "1-16 of x results." When there are fewer than 1,000 results, x is the exact number of releases in the category and release month. Up to 10,000, the number of search results is reported to the nearest thousand. From 10,000 on, the result is reported to the nearest 10,000. The search result also includes the number of search-result pages, with a maximum of 400, indicating 400 $\times$ 16, or 6,400 titles in the result. Hence, the number of

[17]The categories are arts & photography; biographies & memoirs, business & money, children's books, comics & graphic novels, computers & technology; cookbooks, food & wine; crafts, hobbies & home; education & teaching; engineering & transportation; health, fitness & dieting; history; humor & entertainment; law; LGBTQ+ books; literature & fiction; medical books; mystery, thriller & suspense; parenting & relationships; politics & social sciences; reference; religion & spirituality; romance; science & math; science fiction & fantasy; self-help; sports & outdoors; teen & young adult; test preparation; and travel.

books in a search is exact for 1-999, to the nearest 16 for 1,000-6,400, to the nearest 1,000 up to 10,000, and to the nearest 10,000 for higher numbers.[18]

3.2 Book-level analysis datasets

We use the search results pages to obtain book-level information. Although Amazon reports the number of pages in a search result up to 400, it is only possible to directly access the first and last 75 pages of search results and therefore only 2,400 ($= 150 \times 16$) books. Sorting by publication date, we sample book Amazon Standard Identification Numbers (ASINs) from these pages, and we obtain updated information on the numbers of ratings, sales ranks, and average star ratings per book as of February 6 and April 5, 2026, from ASIN Data API (`https://app.asindataapi.com`).

We create two main book-level datasets, along with corresponding subsets with AI detection: a random sample of books released 2020-2025, and the "census" of all books released 2008-2025 in eight narrow categories. These census data allow us to follow author careers and AI adoption. We employ an additional dataset for validating our usage measure based on the correlation of sales and the number of ratings that books receive.

Categories random sample: The first dataset – which we employ for our indirect approach – is a random sample of books from all categories and publication months, 2020-2025. The aggregate series show ten million releases between the start of 2020 and the end of 2025. We assemble a list of all possible 16-title search pages using the information above on the number of titles released per category and month. We choose 10 random pages per

[18]Seven of the 30 categories (children's books; health, fitness & dieting; literature & fiction; religion & spirituality; romance; science fiction & fantasy; and self-help) become sufficiently large during the sample period that the monthly category totals are reported only to the nearest 10,000. For those categories, we obtain their 148 subcategory totals by month to create the aggregate time series of new books released by category.

month and category.[19]

Our random sampling approach obtains 160 random titles within each category and month, or a total of 331,360 observations. That is, our sample is stratified, so that it includes equal numbers of books in each category and month, even though the number of books released differs across categories and is changing over time. We use this random sample, suitably weighted, to explore the effects of the AI influx on the distribution of book qualities. The top panel of Table 1 summarizes the categories random sample. Weighted to reflect the population, the average number of ratings is 112.4, and the average star rating (among the books with ratings) is 4.44. The average price is $9.75, and the average publication year is 2023.1.

Categories AI detection sample: We have a direct measure of AI use for a random sample of 48,193 books among those in the categories random sample. This subsample includes 9,923 random titles from 2020-2022 and 38,270 for 2023-2025.[20] We refer to this as the "categories AI detection sample." We obtained 1,000-word text previews of books at Amazon and ran them through Pangram Labs' AI detector.[21] The quality of AI detection varies across tools, and Pangram's tool has been shown to perform well.[22] The AI detector returns that a text passage is fully human, fully AI, or it returns an AI share. These data allow us to directly compare the quality of AI and non-AI books. We classify books that are not fully human as AI.[23]

[19]For the largest seven categories, we choose ten random pages from their subcategories to avoid oversampling of books published near the beginning or end of the month.

[20]We sampled more heavily from 2023 to 2025.

[21]The text previews are obtained from `https://read.amazon.com/sample/[ASIN]`. Pangram Labs AI detection is available at `https://www.pangram.com/`.

[22]Jabarian and Imas (2025) provide a thorough comparison of AI detectors and conclude that "a clear ranking in detectors emerges. Pangram maintains near-perfect accuracy across long and medium length texts. It achieves very low error rates even on shorter passages and 'stubs.' These results are robust to exogenous changes in the threshold and the use of 'humanizers' such as StealthGPT."

[23]The vast majority of text samples are classified as 0 or 100 percent AI. Just 4.94 percent of the classified books are partially AI; of the partial AI books, the median AI share is 0.65.

Table 1 shows that books in the AI detection sample have an average of 59.4 ratings, an average star rating of 4.42, and an average price of $13.97. The human-created books have much higher usage, 81.9 vs 5.9 average ratings, and higher star ratings as well, 4.46 vs 4.22. Finally, the AI books have lower average prices, $7.93 vs $16.53. The last column reports log sales ranks as of April 5, 2026, which average 13.8 in the categories AI detection sample. Log sales ranks are worse – 14.1 vs 13.7 – for AI vs human-created books, even though AI books tend to be more recent.

Subcategories census sample: For some questions, such as whether incumbent authors change their productivity in the face of AI, it is useful to have all of the titles released in a genre. Hence we collect a complete "census" of new titles in eight subcategories that are sufficiently small to allow their full collection each month, from January 2008 to December 2025.[24] For each title, we observe the author, the publication date, the number of persons rating the book, and the average star rating, as of early January 2026. These data are described in the bottom panel of Table 1. The subcategories census sample contains 479,650 books in total. These books have more ratings, lower average star ratings, and higher prices than those in the categories random sample.

Subcategories AI detection sample: In order to study AI use over time within author, we create an LLM-era sample that contains all of the books (within these categories) by two subsets of authors. First, we search for text samples on all of the works by 1,436 randomly chosen authors with no works before 2023, resulting in a total of 2,000 works. We seek text samples for 8,000 additional LLM-era books, which are all of the LLM-era works by 4,957 randomly selected authors active prior to 2023.[25]

[24]These subcategories are romantasy (a subcategory of romance), economics (a subcategory of business & money), fantasy (science fiction & fantasy), women sleuths (mystery), alternative history (literature & fiction), world history (history), US travel (travel), and sports biography (sports & outdoors).

[25]We excluded authors with more than ten books total.

Because some books are unavailable for preview at Amazon, the AI detection sample contains 9,120 books released between 2023 and 2025, 7,269 by incumbents and 1,851 by authors debuting in 2023 or later. As in the categories AI detection sample, AI-created works have less usage, lower average star ratings, and lower prices. However, the share of AI books in this sample – which disproportionately includes books by incumbent authors – is lower than in the categories AI detection sample.

3.3 Usage measure validity

Our basic usage measure has two potential shortcomings that we address in detail in the Appendix. First, we use information on the number of ratings that books have received rather than actual sales. We show in Appendix A.1 that the number of ratings is highly correlated with sales in a large sample of Amazon ebooks from Bookstat. We also undertake analyses using the April 2026 Amazon sales rank as a usage measure to account for the possibility that the propensity to leave a review varies between human and AI books.

Second, the number of ratings that books have received depends not only on consumers' level of interest but also how long the books have been available. Some of our measurement approaches depend on comparability of the usage measure across release months. We use information on the number of ratings collected at two points in time, February 6 and April 5, 2026. These data show how the number of ratings received – and the probability that books have any ratings – grows over time for books of different ages. We use the resulting growth rates to calculate the adjusted number of ratings as of 72 months after release, which we term $\tilde{r}_j$. We use the February-to-April transitions to adjust zero ratings to the unconditional averages (the share changing to positive $\times$ the average ratings among now-positive), so that all adjusted ratings are positive. See Appendix A.2.

4 Entry and quality distribution effects

This section explores two broad questions. Section 4.1 documents new book entry and the evolving AI-detected share. Section 4.2 uses the random category sample – and the indirect approach – to explore how the distribution of usage/quality evolves as AI usage rises from 2023 to 2025. We revisit the quality question using the direct approach, and AI detection data, in Section 5.

4.1 Aggregate book release patterns and AI

The left panel of Figure 2 shows aggregate new titles by month for all Amazon categories between January 2020 and December 2025. The number of new titles is roughly stable at 100,000 per month between 2020 and mid-2022. It rises during 2023 and 2024 and then rises more sharply in 2025, peaking at over 300,000 per month in late 2025. Over ten million titles are released as ebooks during the six-year period.

The solid line in panel (b) of Figure 2 shows the categories AI-detected share of titles in the AI-detection sample, weighted to the population. The share remains close to zero between 2020 and 2022. The share then rises to 30 percent during 2023, to 45 percent during 2024, and tops 60 percent during 2025. Panel (b) also includes the implied AI share of releases under the assumption that the growth in releases since 2022 is attributable to AI.[26] The implied and detected AI share are close, suggesting that our AI detection is accurate.

The growth rate of new releases is large across most categories. The left panel of Figure 3 shows the ratios of the highest monthly totals during 2025 and 2022, by category. Categories with particularly large proportionate growth are mostly nonfiction and include travel (9.1 times higher), sports & outdoors (6.9), self-help (5.3), and computers & technology (4.5). The right panel of Figure 3 shows the category share of titles published between 2023 and

[26] If n_t is releases in month t, and n_{2022} is average monthly releases during 2022, then the implied AI share is $(n_t - n_{2022})/n_t$.

2025 in which AI is detected. The share is lowest in comics & graphic novels and high in travel and computers & technology. Categories with fast growth tend to have higher AI shares for 2023-2025.

4.2 Effects of LLMs on the quality distribution

The data in the categories random sample allow three indirect approaches to measuring the impact of the LLM influx on the overall distribution of product quality. These include a temporal approach, a high-low approach, and a continuous approach.

Temporal approach: First, using the adjusted measure of the number of ratings $\tilde{r}_j$, which is comparable across vintages, we simply ask whether average quality falls after 2022. We implement this by regressing log adjusted ratings $\tilde{r}_j$ on release month and category fixed effects:

$$\ln(\tilde{r}_j) = \mu_c + \psi_t + \varepsilon_j, \tag{1}$$

where μ_c is a category fixed effect, ψ_t is a time effect, and ε_j is an error term. We weight observations in a category and publication month to reflect the total number of books, and we cluster standard errors by category $\times$ release month. Panel (a.1) of Figure 4 shows the resulting estimates of month effects ψ_t. Between 2020 and early 2022, the average rating measure is stable, then it falls noticeably after 2022.

We measure effects on quality at particular percentiles of the distribution by augmenting Equation (1) to include fixed effects for 20 within-category-and-month rank percentile quantiles. Panel (b.1) of Figure 4 shows the within-percentile results. Like the overall average, the quality conditional on rank percentiles is stable prior to late 2022, then falls.

To measure effects on quality at particular rank positions, we augment Equation (1) to include fixed effects for within-category-and-month rank ranges 1-100, 101-200, etc. Panel

(c.1) shows the resulting time fixed effects. In contrast to the quality at the mean and at percentiles, the quality at particular rank positions rises. The temporal results together indicate that average quality falls but that the number of high-quality books rises.

High-low approach: Our second approach asks whether success measures evolve differently after the LLM influx in categories with rapid release growth ($\mathbb{1}_j^h = 1$) relative to others. We divide the 30 book categories into two groups, those with top quartile weighted release growth and those less exposed.[27] We then ask how the usage measure changes across release dates for the highly exposed groups relative to the less exposed groups via regressions of the form:

$$\ln(\tilde{r}_j) = \mu_c + \mu_t + \lambda_t \times \mathbb{1}_j^h + \varepsilon_j, \tag{2}$$

where μ_c and μ_t are category and publication month fixed effects. As above, we weight for frequencies to reflect total releases; and we cluster standard errors by category $\times$ release month. The coefficients λ_t show how the outcome measures for books in the highly exposed groups evolve over time, relative to the time pattern for the less exposed groups. The right-side graphs in Figure 4 report results on the average usage (a.2), as well as usage conditional on percentile (b.2) and rank position (c.2). Results are similar to the results from the temporal approach: Quality falls on average and conditional on percentiles, but quality rises conditional on rank position.

Continuous approach: Finally, we ask how the usage measure evolves with the number of releases by category by regressing log ratings on release month and category fixed effects as well as the log of the number of releases by category and month, or

[27] We obtain similar results dividing at the weighted median.

$$\ln(\tilde{r}_j) = \mu_c + \mu_t + \phi \ln(n_{ct}) + \varepsilon_j, \tag{3}$$

where n_{ct} is the number of books entering in category c in month t and μ_c and μ_t are category and publication-month fixed effects. The regressions are weighted as above, with standard errors clustered at the category $\times$ release month level; and we employ both the adjusted and raw usage measures. We also add rank percentile and rank range fixed effects to examine effects at different points in the distribution.

Table 2 shows the results, first – in columns (1) and (2) – for average quality. Using both adjusted and unadjusted usage measures, categories with faster growth in entry experience larger decreases in average quality. The raw and adjusted samples differ in that books with zero raw ratings are excluded.[28] Columns (3) and (4) explore impacts within percentile ranges. The negative coefficients for both usage measures indicate that categories with more growth in releases experience larger reductions in quality, conditional on percentile. Columns (5) and (6) explore impacts within rank position ranges. Using both usage measures, quality rises more – within rank ranges – in categories with faster growth in entry.

While we have constrained ϕ to be constant across rank positions in Table 2, in Figure 5 we relax this. The figure reports the rank-position-specific ϕ coefficients from a regression of $\ln(\tilde{r}_j)$ with month, category, and rank-position fixed effects; and results are similar for unadjusted ratings. The coefficient on $\ln(n_{ct})$ is not statistically distinguishable from zero for books ranked 1-200 among the titles released in the category and month. The coefficient then grows as ranks worsen, reaching roughly 1.2 for ranks beyond 600. Thus, the large, LLM-aided growth in books does not raise the quality of the top 200 books released per category-month; but it does raise the quality of books outside of the top 200.

The results above paint a consistent picture of the relationship between LLM-induced entry and the distribution of quality, and the framework in Section 2 is useful for guiding

[28] We obtain virtually identical results using both measures estimated on the raw subsample.

interpretation of the results. The number of books released rises substantially, and despite a decline in average quality, the increase in the quality of books at particular rank positions for all but the top 200 titles released per category and month suggests modest welfare gains.

5 Direct effects on quality and welfare

This section uses the categories AI detection sample to explore three questions. Section 5.1 uses a direct approach to measuring the quality differential between LLM- and human-generated books. Section 5.2 explores the evolution of the human-AI gap, and Section 5.3 presents estimates of the welfare impact of the LLM book influx.

5.1 Quality evidence from the AI detection sample

We measure the difference in usage between AI- and human-generated books in two ways, with the adjusted log number of ratings and with a sales rankings measure. The first two columns of Table 3 present weighted regressions of adjusted usage on the book-level AI indicator.[29] Column (1) includes no fixed effects and shows that the AI-human differential is -1.147 (se = 0.034). Accounting for category and month fixed effects, in column (2), the gap remains large, at -0.945 (0.034), indicating a 61.1 percent differential in the number of ratings received after six years. The books created with AI during 2023-2025 attract substantially less usage than their human-generated counterparts.

It is possible that the tendency to leave ratings differs between AI and non-AI books. We address this with an alternative usage measure that does not depend on whether users leave ratings, the sales rank as of April 5, 2026. Despite its advantage, this measure has the shortcoming that it is the sales rank at a date with differing delays after publication, and ranks weight recent sales more heavily. The measure has the further feature that books

[29] The weights here inflate the AI detection subset of the categories random sample, rather than the overall sample, to the population.

without recent sales have an undefined sales rank, even if they have sold substantially in the past and have large numbers of ratings.[30]

We regress the log sales rank for book j on an indicator for whether it contains AI, along with category and publication month fixed effects. We estimate this by least squares using only the observations with an observed sales rank, and we weight and cluster as above. We also estimate a censored normal model in which books without observed sales ranks are assumed to have a rank worse (higher) than the highest-observed sales rank. Columns (3) to (5) of Table 3 report the coefficients from these regressions. Column (3) reports a specification without fixed effects, yielding a coefficient of 0.321 (se = 0.037), while column (4) adds the fixed effects, giving a coefficient of 0.303 (se = 0.040). The coefficient in the censored model (column (5)) is 0.773 (se = 0.044). Hence, this approach corroborates the main approach in showing that books containing AI are substantially less appealing to consumers, and the difference is statistically significant.

5.2 The human-AI gap over time

The categories AI detection sample also allows calculation of usage measures over time. The solid line in the left panel of Figure 6 shows the evolution of the human book usage measure over time, and it remains roughly constant. The quality of AI books – the dotted line – is below the human quality but is rising. Because the share of AI works grows over time, the quality of average releases – the weighted average of human and AI books during the LLM era – falls after 2022.

Star ratings, too, are lower for AI-containing books. A weighted regression of star ratings on the AI indicator yields a coefficient of -0.146 (0.055); with category and month fixed effects the coefficient is -0.200 (0.036). The differential is also visible in the right panel of Figure

[30] Of the 134,873 books lacking a sales rank on April 5, 2026, 19.4 percent have a positive number of raw ratings.

6, which repeats the time series usage exercise of the left panel using star ratings. The average star ratings for human-authored books are relatively stable, AI books are worse, and the human-AI gap shrinks over time. This has two implications. First, it provides further evidence that AI books are, on average, worse. Second, given evidence elsewhere on star ratings as welfare-enhancing pre-purchase information (Reimers and Waldfogel, 2021), the star ratings may help consumers navigate the AI-enlarged product choice set.

5.3 Consumer welfare effects

Given the addition of AI books after 2022, what happens to the number and share of AI books throughout the quality distribution? Figure 7 presents histograms showing the number of releases, by year, in each adjusted ratings interval.[31] The bar at each interval shows the number of releases at that usage level; and the dark components of the bars – starting in 2023 – show the AI books. In 2023, AI books are concentrated in the lower intervals and absent from the highest intervals. The dark areas become larger in 2024 but remain concentrated at the lower intervals. In 2025, AI books appear in the higher intervals to greater extents. AI books add to the choice set in all years; by 2025 the AI influx delivers books with higher usage than in prior years.

The area of the light portions of the bars in Figure 7 is similar across years, indicating that the volume and quality distribution for human-authored books remains similar as LLMs appear and find wide use.[32] Hence, the growing dark portions of the bars in Figure 7 foreshadow the effects of AI on welfare, and the AI detection sample allows a simple calculation of the welfare effects of AI. For each year 2023-2025 we compare the CS from the entire choice set to the subset that would arise if only the non-AI books had been available.

[31] We exclude books whose ln adjusted ratings are below 2, as these books have little effect on total usage or welfare. Changes in CS calculated excluding these books are nearly identical to welfare effects reported below.

[32] The constancy of the human-authored book quality in Figure 6 provides additional evidence of the stability of the human-authored quality distributions.

We measure the magnitude of the change in consumer surplus using a nested logit calibration in the spirit of Berry (1994). Consumer i chooses among books $j \in J$ and the outside good according to utility given by

$$u_{ij} = \delta_j + \zeta_{ig} + (1-\sigma)\epsilon_{ij},$$

where ϵ_{ij} follows an extreme value distribution, and the idiosyncratic inside-good preference ζ_{ig} follows a distribution such that $\zeta_{ig} + (1-\sigma)\epsilon_{ij}$ also follows an extreme value distribution. The outside good has utility zero, or $u_{i0} = 0$.

The mean utility δ_j of product j is given by the observed market shares and a calibrated taste correlation parameter σ with $\delta_j = \ln(s_j) - \ln(s_0) - \sigma \ln(s_j/(1-s_0))$, where $s_j = q_j/M$. The resulting consumer surplus CS associated with a choice set is proportional to

$$\ln\left(1 + \left(\sum_{j \in J} e^{\delta_j/(1-\sigma)}\right)^{1-\sigma}\right).$$

We treat our adjusted usage measures $\tilde{r}_j$ as quantities, and we estimate the market size M for each year as the US population times ten. We use an estimate of σ from Reimers and Waldfogel (2021) of 0.443 for the baseline nested logit implementation, and we explore robustness to other values of σ. We focus on the percent change in CS across environments.

Table 4 shows the ensuing welfare calculations. The column headed "$\sigma = 0.443$" reports baseline estimates. Using the 2023 releases in the categories AI detection sample weighted to the population of releases, all releases (including those with AI) deliver 0.44 percent more CS than just the non-AI releases. For 2024, the AI books raise CS by 3.26 percent. Finally, the AI books in 2025 – when total output reaches triple the pre-LLM output of 2022 – raise CS by 7.23 percent. The last two columns of Table 4 report welfare estimates for alternative substitution parameters. If $\sigma = 0.25$, then the increase in CS ranges from 0.6 percent in 2023

to 9.8 percent in 2025. If $\sigma = 0.75$, the analogous figures are 0.2 and 3.2, respectively.[33] We conclude that the influx of AI-assisted books creates positive consumer benefits, and these benefits are small relative to the increase in the number of books.

6 Author AI adoption, quality, and displacement

This section explores two questions related to AI adoption and possible incumbent author displacement. Section 6.1 discusses how author selection and the quality of the LLM determine the gap between AI and non-AI book quality. Section 6.2 explores whether the influx of AI displaces incumbent authors.

6.1 Why are AI books less appealing?

The quality of AI and non-AI books might differ for two reasons, the selection of authors using AI and the effect of AI on book quality for authors adopting it. We decompose the AI quality differential using the subcategories AI detection sample, which allows us to follow author output and success across multiple books.

We have two pieces of evidence about the selection of authors into AI use. First, as the left panel of Figure 8 shows, authors whose first book appears after 2022 are more likely than returning incumbents to use AI. More than 40 percent of the 2023-2025 releases of those authors with no pre-LLM-era books – the "entrants" – contain AI. The AI share is substantially lower for returning incumbents, averaging about 10 percent. The right panel of Figure 8 shows the relationship between an author's pre-2023 success and the AI share. Among returning incumbents, the tendency for their 2023-2025 releases to contain AI is lower for those with more pre-AI success. Hence, some part of the low quality of AI books may reflect low-quality authors using AI.

[33] AI books tend to have lower prices than non-AI books, so the CS benefit arises not just from more books of modest quality but rather from more books of modest quality and low prices.

We can directly decompose the AI quality differential into between-author and within-author components. To this end, we estimate variants of

$$\ln(\tilde{r}_j) = a_{it} + \lambda_t \times \mathbb{1}_{ij}^{AI} + \mu_i + \mu_t + \varepsilon_{ij}, \tag{4}$$

where the dependent variable is the log of adjusted usage for book j, a_{it} is the author's age (calendar time t less the month in which the author's first book appeared), $\mathbb{1}_{ij}^{AI}$ is an indicator for whether book j by author i contains AI, μ_i is an author fixed effect, μ_t is a publication calendar-month fixed effect, and ε_{ij} is an error term.

Column (1) of Table 5 reports a variant of Equation (4) that excludes the author fixed effects and includes a single AI differential for the entire period. As in the categories AI detection sample, AI books have lower usage (-1.74, se = 0.08). Column (2) reports year-specific AI differentials, again without an author fixed effect, and the AI usage differential declines in absolute value from -2.20 in 2023 to -1.50 in 2025. Column (3) adds author fixed effects, so that the AI coefficients show the within-author differentials. The differential remains significant: When a given author uses AI, the resulting book is less appealing. However, the coefficients shrink substantially in column (3) relative to column (2): The within-author component of the differentials is between 35 and 40 percent of total yearly differentials.[34] Hence, the majority of the AI quality differential stems from selection, or that low-quality authors use AI.

The decline in column (3) AI coefficients over time shows that the within-author AI penalty falls, which could reflect author learning or improving technology. Column (4) explores this by adding a rolling count of authors' prior AI-containing books. The count variable obtains a positive but insignificant coefficient. The declining within-author AI differentials in columns (3) and (4) therefore appear to reflect improvement in the technology.

[34]For example, in year 2023, the ratio is $e^{-0.438}-1/e^{-2.197}-1 = 0.399$. These ratios remain similar if we estimate the column (2) regression on the column (3) sample.

6.2 Has AI displaced incumbent authors?

The appearance of a large number of new books by entering – and in many cases AI-using – authors might displace continued activity by incumbent authors (active prior to LLMs). The previous subsection relied on the subcategories AI detection subsample, which was selected to disproportionately include returning authors. Here, we employ the full subcategories census dataset to examine author productivity regardless of AI use.

We document the productivity of pre-LLM authors with a cohort approach that accounts for author activity patterns relative to the time of their first release. Define n_{tv} as the number of books released in time t by authors debuting at time (vintage) v. Productivity of a cohort evolves over time as it ages, where "age" $a = t - v$. We estimate the productivity of pre-LLM authors over time from these age effects. In particular, we estimate the following model on author cohorts entering between 2010 and 2022:

$$\ln(n_{tv}) = \omega_t + \mu_a + \varepsilon_{tv}, \tag{5}$$

where ω_t and μ_a are time and age fixed effects, and ε_{tv} is an idiosyncratic error. We use robust standard errors.

Figure 9 shows the time pattern of the time coefficients ω_t. They are stable between 2020 and the middle of 2023. The coefficients then rise gradually beyond pre-LLM levels starting in 2024. This shows that incumbent authors increase their output during the LLM era.[35]

[35]These results, which are robust to the inclusion of different pre-LLM time spans, contrast with the existing literature (Goldberg and Lam, 2025; Peukert et al., 2024; Kim et al., 2026) that finds substantial crowding out of human-generated creative work by AI.

7 Conclusion

LLMs, which have diffused rapidly between 2022 and 2025, have become increasingly capable of writing long passages of text. Accordingly, they have seen rapid adoption by book authors. We document a tripling in the number of new books coming to market between late 2022 and late 2025 that mirrors the use of AI that we detect in new books. The effects of this influx on consumer welfare depend on the quality of the additional books. The average quality of new books has fallen with the LLM-induced influx, and books with detected AI are substantially worse than human-authored books, so that much of the new work is of little value to consumers. Still, the LLM influx has delivered some books in the middle range of the usage/quality distribution, and the LLM-era entry process delivered seven percent more consumer surplus from books than the pre-LLM process in 2025.

Much of the AI-human quality gap stems from low-quality authors' adoption of AI, but the AI-human gap shrinks between 2023 and 2025. Moreover, the arrival of LLMs does not appear to have displaced activity by incumbent authors. Despite the controversy surrounding LLMs, their effect on book consumers – like other cost-reducing technological changes in the cultural industries – is positive. However, because the new books are mostly of low quality, the effects are modest.

While we document that AI facilitates the creation of content that is valuable to consumers, there is a separate question of whether AI facilitates the creation of content that would be well-regarded by critics or other cultural elites. Commercial success and cultural importance are not the same thing, but the absence of an effect at the top suggests a limit to how much LLMs can help with elite cultural production. Definitive answers to these questions depend on hard-to-predict developments in the capabilities of LLMs.

References

Aguiar, L. and J. Waldfogel (2018): "Quality predictability and the welfare benefits from new products: Evidence from the digitization of recorded music," *Journal of Political Economy*, 126, 492–524.

Berry, S. T. (1994): "Estimating discrete-choice models of product differentiation," *The RAND Journal of Economics*, 242–262.

Brynjolfsson, E., Y. Hu, and M. D. Smith (2003): "Consumer surplus in the digital economy: Estimating the value of increased product variety at online booksellers," *Management science*, 49, 1580–1596.

Brynjolfsson, E., D. Li, and L. Raymond (2025): "Generative AI at work," *The Quarterly Journal of Economics*, 140, 889–942.

Caves, R. E. (2000): *Creative industries: Contracts between art and commerce*, 20, Harvard university press.

Chen, Y. J., J. Gong, J. Li, and Z. Zhao (2025): "Better Technology, Worse Motivation: GenAI's Mediocrity Trap," *Available at SSRN 5208163*.

Dell'Acqua, F., S. Rajendran, E. I. McFowland, E. Mollick, H. Lifshitz-Assaf, L. Krayer, F. Candelon, K. R. Lakhani, and K. C. Kellogg (2023): "Navigating the Jagged Technological Frontier: Field Experimental Evidence of the Effects of AI on Knowledge Worker Productivity and Quality," Tech. Rep. 24-013, Harvard Business School Working Paper.

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

Goldberg, S. and H. T. Lam (2025): "Generative AI in Equilibrium: Evidence from a Creative Goods Marketplace," *Working Paper*, available at SSRN.

Jabarian, B. and A. Imas (2025): "Artificial writing and automated detection," Tech. rep., National Bureau of Economic Research.

Kim, S., G. Z. Jin, and E. Lee (2026): "Does Generative AI Crowd Out Human Creators? Evidence from Pixiv," Tech. rep., National Bureau of Economic Research.

Kusumegi, K., X. Yang, P. Ginsparg, M. de Vaan, T. Stuart, and Y. Yin (2025): "Scientific production in the era of large language models," *Science*, 390, 1240–1243.

Liang, W., Y. Zhang, M. Codreanu, J. Wang, H. Cao, and J. Zou (2025): "The widespread adoption of large language model-assisted writing across society," *arXiv preprint arXiv:2502.09747*.

Lutes, B., J. Gans, S. Greenstein, A. B. Jaffe, A. Nagaraj, I. Reimers, M. D. Smith, R. Telang, C. E. Tucker, and J. Waldfogel (2025): "Identifying the Economic Implications of Artificial Intelligence for Copyright Policy," *First published by the US Copyright Office*.

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

Padmakumar, V., B. Dhingra, W. Ammar, F. Metze, D. Downey, H. Hajishirzi, E. Hovy, and N. A. Smith (2024): "Does Writing with Language Models Reduce Content Diversity?" in *International Conference on Learning Representations (ICLR)*.

Peukert, C., F. Abeillon, J. Haese, F. Kaiser, and A. Staub (2024): "AI and the Dynamic Supply of Training Data," *arXiv preprint arXiv:2404.18445*.

Peukert, C. and I. Reimers (2022): "Digitization, prediction, and market efficiency: Evidence from book publishing deals," *Management Science*, 68, 6907–6924.

Reimers, I. and J. Waldfogel (2021): "Digitization and pre-purchase information: the causal and welfare impacts of reviews and crowd ratings," *American Economic Review*, 111, 1944–1971.

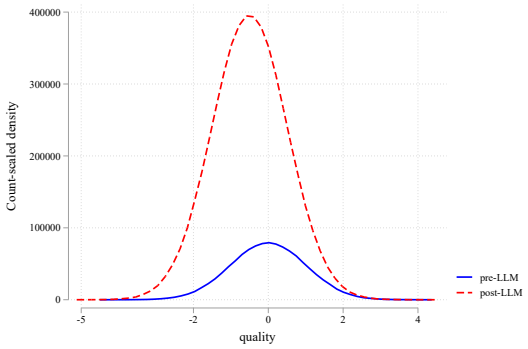
Vaithilingam, P., T. Zhang, and E. L. Glassman (2022): "Expectation vs. Experience: Evaluating the Usability of Code Generation Tools Powered by Large Language Models," *Proceedings of the 2022 CHI Conference on Human Factors in Computing Systems*.

Waldfogel, J. and I. Reimers (2015): "Storming the gatekeepers: Digital disintermediation in the market for books," *Information economics and policy*, 31, 47–58.

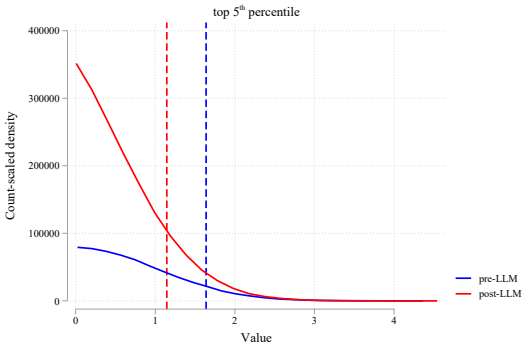
8 Figures and tables

Figure 1: LLM influx and effects on rank positions vs percentiles

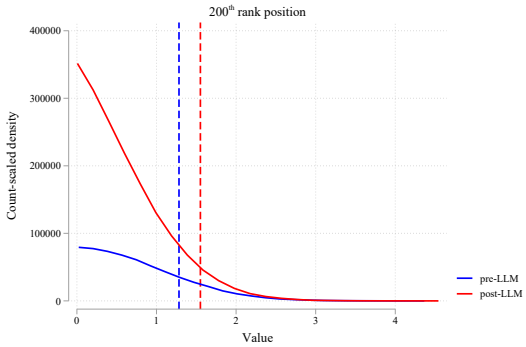
(a) Full distributions



(b) Value at 95$^{\text{th}}$ percentile



(c) Value at 200$^{\text{th}}$ rank position

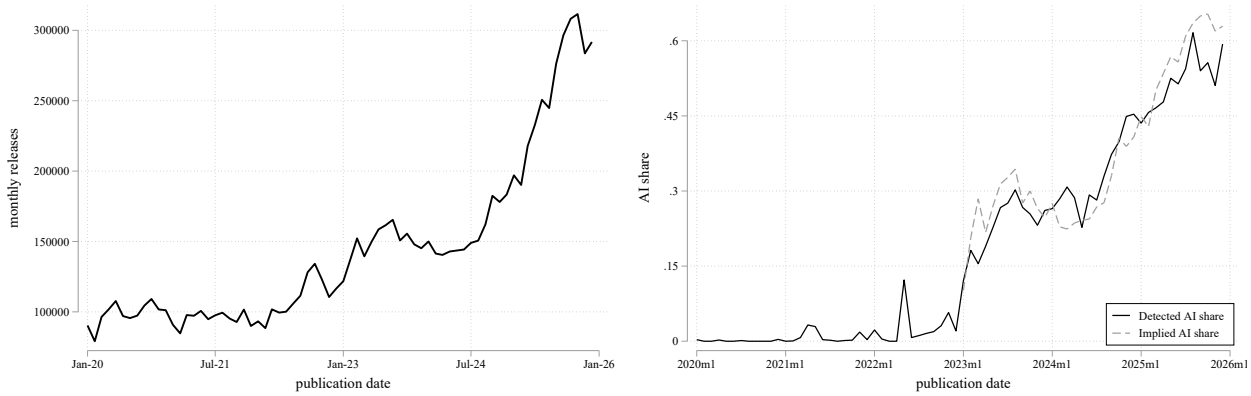


Notes: This figure compares hypothetical pre-LLM (blue) and LLM-era (red) quality distributions, where the LLM-era distribution has a lower mean and more products. Panel (a) shows the overall distributions. Using these distributions, Panel (b) shows that the 95$^{\text{th}}$ (top 5) percentile product has lower quality in the LLM era, while Panel (c) shows that a product at the 200$^{\text{th}}$ rank position has higher quality in the LLM era.

Figure 2: Book entry and detected AI usage

(a) Overall entry

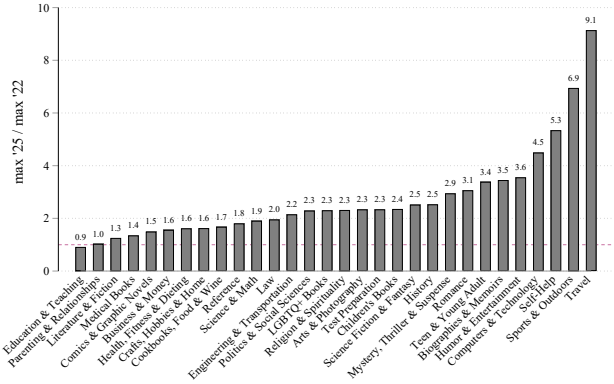
(b) AI share



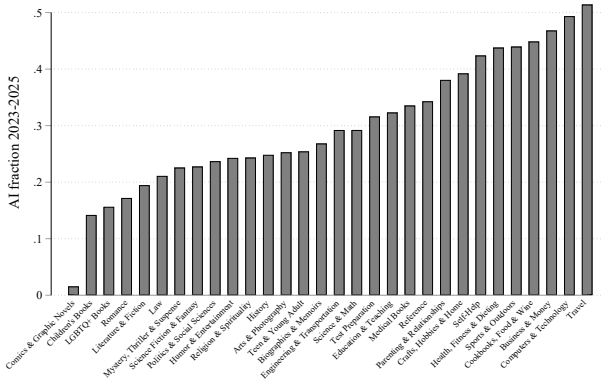
Notes: The left panel shows aggregate book releases across all categories. May 2024 is excluded because queries deliver oddly large results in many categories for that month. The right panel shows the weighted share of books with AI detected, based on 9,923 titles for 2020-2022 and 38,270 for 2023-2025, along with the implied AI share under the assumption that the increase in releases over 2022 is all AI.

Figure 3: Entry and AI use by category

(a) Growth by category



(b) AI share by category

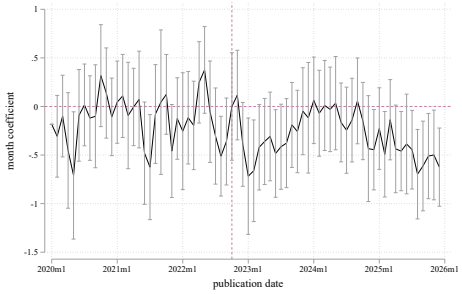


Notes: The left panel shows the ratios of the maximum monthly new publications in 2025 relative to the maximum monthly publications in 2022, for each category. The right panel shows the weighted AI share of the books in the AI-detected sample, 2023-2025.

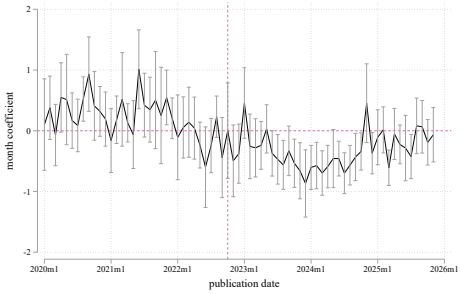
Figure 4: Quality distributions across publication dates

Overall average

(a.1) Temporal approach

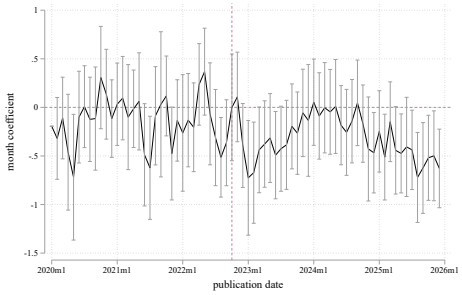


(a.2) High-low approach

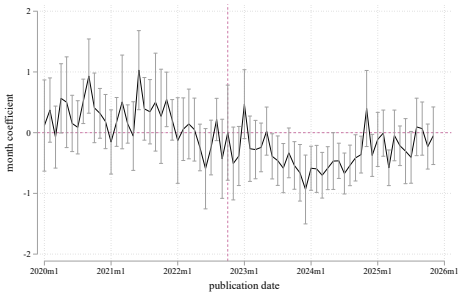


Conditional on rank percentiles

(b.1) Temporal approach

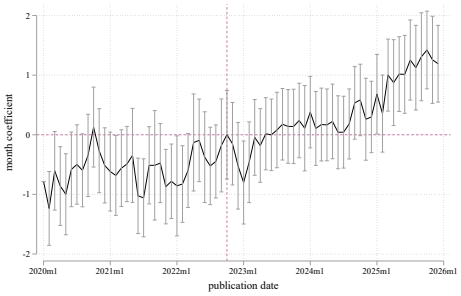


(b.2) High-low approach

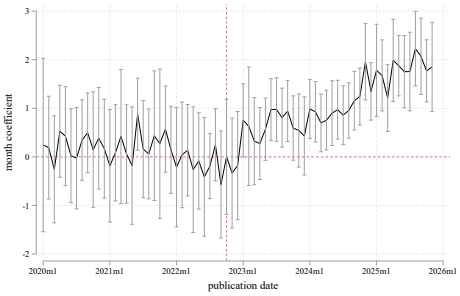


Conditional on rank positions

(c.1) Temporal approach

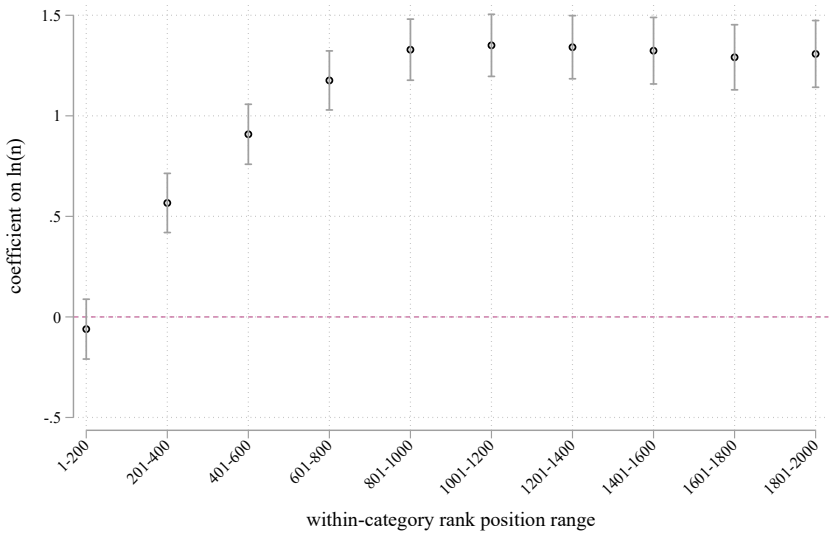


(c.2) High-low approach



Notes: The left figures show month fixed effects from the temporal approach in Equation (1). The right figures show monthly interactions with an indicator for high-growth categories from the high-low approach in Equation (2). Panel (a) is the overall average. Panel (b) includes fixed effects for 20 rank-percentile quantiles. Panel (c) includes fixed effects for rank position ranges within categories by 100, up to 2000.

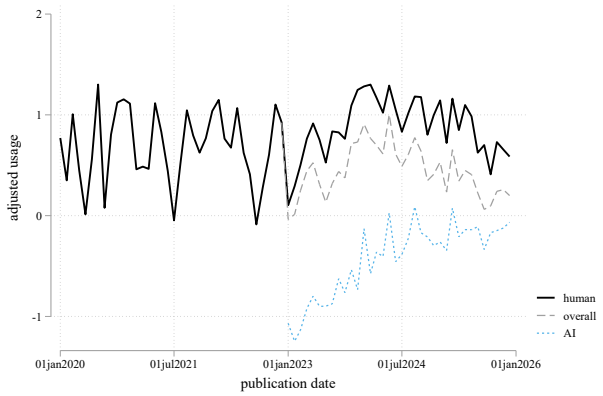
Figure 5: Coefficients on $\ln(n)$ by rank position range



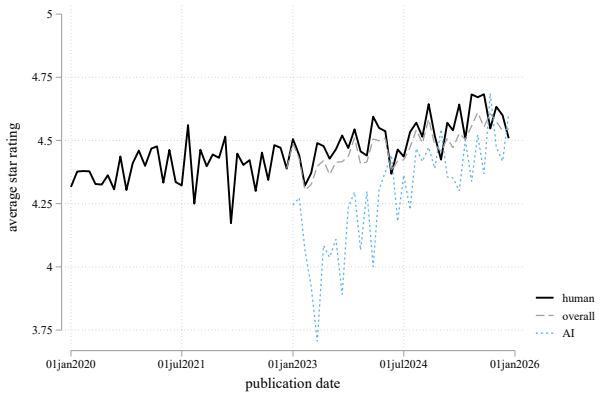
Notes: Rank-position-specific coefficients from regressions of our adjusted usage measure on the log of monthly releases per category. Standard errors are clustered on category × release month. The regression includes month, category, and rank group fixed effects.

Figure 6: AI and non-AI usage and quality time series

(a) ln(adjusted ratings)

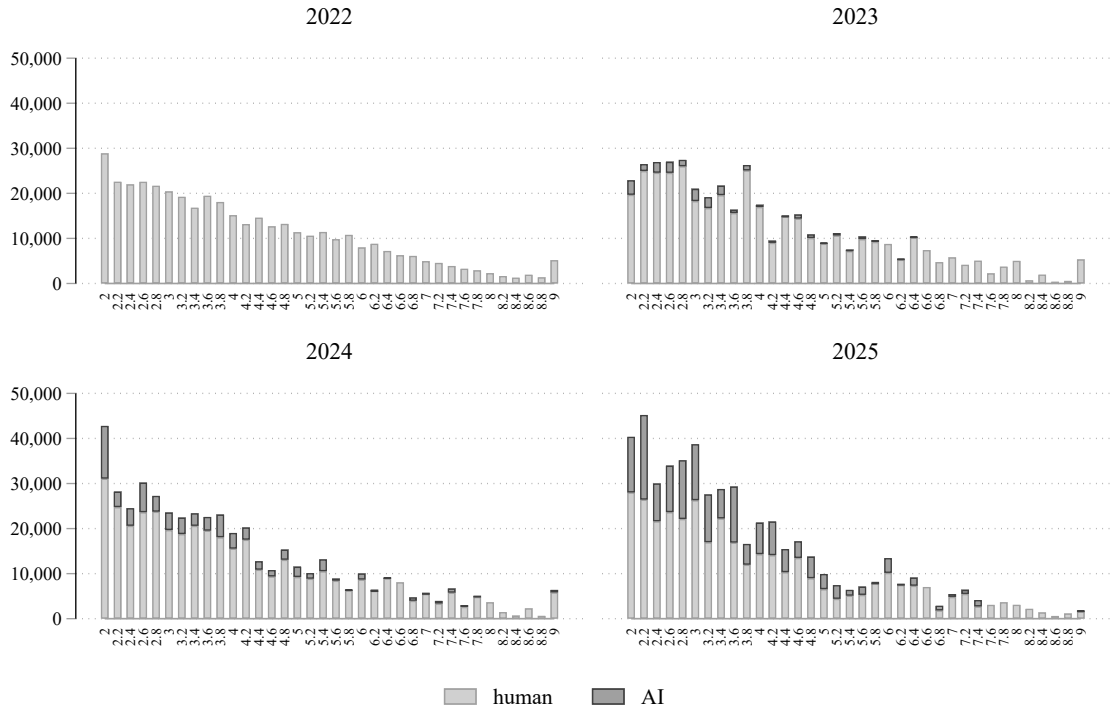


(b) average star ratings



Notes: This figure shows weighted time series of usage (ln adjusted number of ratings), in Panel (a), and average star ratings, in Panel (b), for books of different types. The solid lines show time series for human-authored books (all books released before 2023 and those detected not to contain AI in 2023 and later). For 2023-2025 we also show the average usage and star ratings for AI books (blue, dotted lines), and all (AI and non-AI) books (gray, dashed lines).

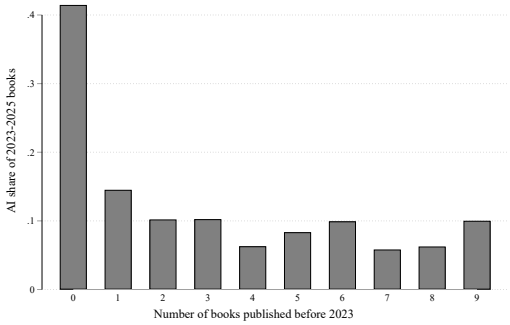
Figure 7: AI and non-AI usage distributions



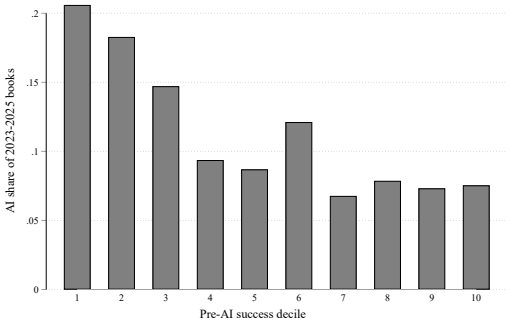
Notes: Distributions of AI and non-AI books by log adjusted rating bins and release years. We clip the data at $\ln(\tilde{r}) = 2$ and group all with more than 9 log adjusted ratings. The figure uses the categories AI detection sample.

Figure 8: AI share by author experience and prior success

(a) Entrants vs incumbents

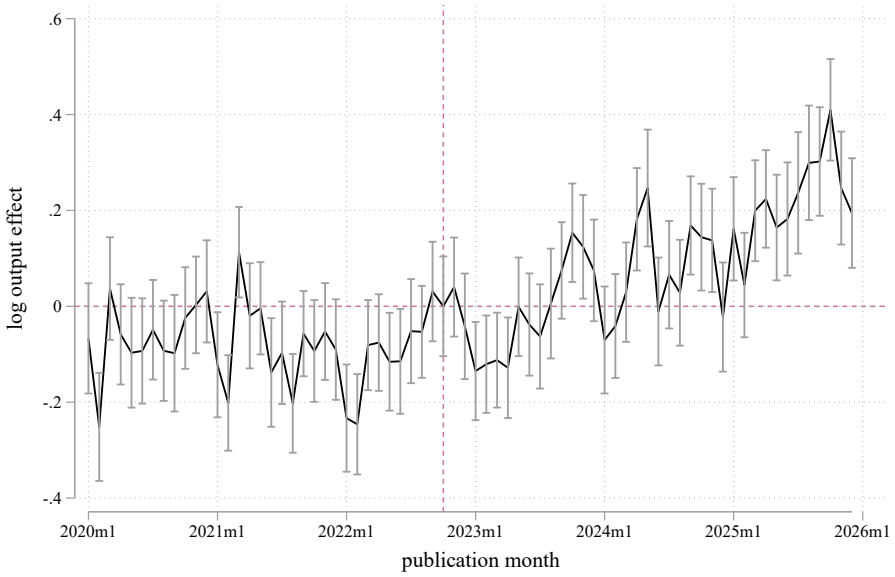


(b) Authors by pre-LLM success



Notes: The left panel shows the share of 2023-2025 books in which AI is detected as a function of the number of books an author released prior to 2023. Authors with 0 are entrants during 2023-2025. The right panel shows the share of 2023-2025 books in which AI is detected as a function of pre-2023 author success, for authors who have released at least one book before 2023. Success is measured as the sum of ratings that books have received, divided by the number of books published in the pre-LLM period. These figures are based on the subcategories AI detection sample.

Figure 9: Productivity of incumbent authors



Notes: This figure shows publication month fixed effects from a regression of log releases aggregated to publication month and author debut month. The regression also includes author age fixed effects. We use the subcategories census sample, and we include authors with debuts between 2010 and 2022. The regression uses robust standard errors.

Table 1: Summary statistics across samples

	# books	# ratings	avg stars	price	pub year	ln rank
Categories dataset:						
Weighted to population	10,233,646	112.40	4.44	9.75	2023.10	
Raw	331,360	83.38	4.40	14.85	2022.52	
AI detection sample	48,193	59.65	4.42	13.90	2023.38	13.82
human-created	33,792	82.53	4.46	16.46	2023.02	13.72
AI-created	14,401	5.96	4.22	7.94	2024.21	14.09
Census dataset:						
Full sample	479,650	166.58	4.33	14.60	2019.47	
AI detection sample	9,120	178.65	4.44	11.23	2024.04	13.59
human-created	7,614	208.63	4.45	12.01	2023.95	13.57
AI-created	1,506	27.14	4.25	7.31	2024.47	13.77

Notes: The table shows average values for variables of interest in the main datasets. The categories random dataset, described in the top panel, is a stratified random sample of books released between 2020 and 2025. The top row reports the averages weighted to the population of all releases over this period. The raw averages are on the second row. We run 1,000-word samples from 48,193 of these books through Pangram Labs AI detection to create the categories AI detection sample. The census dataset, in the bottom panel, contains all 479,650 books released in eight small subcategories between 2008 and 2025. We run 9,120 of these books through AI detection to create the census AI detection sample.

Table 2: Continuous approach usage regressions

	average		percentile		rank position	
	(1)	(2)	(3)	(4)	(5)	(6)
	adjusted	raw	adjusted	raw	adjusted	raw
ln monthly releases	-0.621***	-0.459***	-0.614***	-1.188***	1.580***	0.672***
	(0.0671)	(0.0609)	(0.0677)	(0.0936)	(0.0726)	(0.0921)
Observations	326975	157373	326975	157372	326954	157361

Notes: This table shows weighted regressions of adjusted and raw ln ratings measures on the natural log of the number of monthly releases in the category, using the categories random sample. All regressions include category and month fixed effects; and standard errors are clustered at the category × month level. The percentile regressions also include fixed effects for percentiles of adjusted usage within category-month (0-5,6-10, etc). The rank position regressions also include fixed effects for adjusted usage rank positions 1-100, 101-200, etc. within category-month.

Table 3: AI books, usage, and sales ranks

	log adjusted ratings		log sales rank		
	(1)	(2)	(3)	(4)	(5)
AI-generated	-1.147***	-0.945***	0.321***	0.303***	0.773***
	(0.0338)	(0.0341)	(0.0372)	(0.0404)	(0.0435)
Fixed effects	None	cat. + month	None	cat. + month	cat. + month
N	38270	38270	23628	23628	38270

Notes: This table shows regressions of usage measures on an indicator for whether the book contains AI. Columns (1) and (2) use ln(adjusted ratings) as the dependent variable. Column (2) includes category and month fixed effects. Columns (3)-(5) use ln(sales ranks) as the dependent variable. Columns (3) and (4) include only observations with enough recent sales to have ranks. Column (5) is estimated as a censored normal regression, and books without ranks are assumed to have worse ranks than the worst observed rank. All models are estimated on the categories AI detection sample (2023-2025). Standard errors are clustered at the category $\times$ month level, and regressions are weighted by the numbers of entering books by category and birth month.

Table 4: Consumer surplus effects of LLMs by year

			%Δ CS		
year	# of books	AI share	$\sigma = 0.443$	$\sigma = 0.25$	$\sigma = 0.75$
2022	1,287,514				
2023	1,758,187	23.12%	+0.44%	+0.59%	+0.2%
2024	1,841,872	33.57%	+3.26%	+4.38%	+1.45%
2025	3,075,328	52.7%	+7.23%	+9.82%	+3.19%

Notes: This table shows consumer surplus effects from adding AI books in each year. The second column shows the numbers of books released. The "AI share" column shows the weighted average annual AI detected share. The remaining columns show the percent increase in consumer surplus from the full sample in each year, relative to just the human-authored books of the year, based on our preferred nested logit substitution parameter $\sigma = 0.443$ and two alternative substitution parameters.

Table 5: AI books and usage over time

	(1)	(2)	(3)	(4)
AI-generated	-1.740***			
	(0.0760)			
AI × year 2023		-2.197***	-0.438**	-0.425**
		(0.166)	(0.187)	(0.188)
AI × year 2024		-1.863***	-0.354**	-0.365**
		(0.150)	(0.167)	(0.167)
AI × year 2025		-1.503***	-0.341**	-0.395***
		(0.104)	(0.134)	(0.143)
# prior AI books				0.0601
				(0.0559)
Fixed effects	None	None	author + age	author + age
N	9026	9026	4629	4629

Notes: Regressions of ln(adjusted ratings) on an indicator for whether the book contains AI using the subcategories AI detection sample (2023-2025). Columns (1) and (2) include no author fixed effects. Column (3) includes author fixed effects, and column (4) adds a variable for the author's number of prior AI-containing books.

A Appendix

A.1 Validity of our usage measure

Given our reliance on the number of ratings a book has received (r_j) as a usage measure, it is important to validate our quantity/usage proxy. For this, we use an additional dataset. We have daily estimates of sales, average star ratings, and the number of ratings received, for ebooks sold at Amazon between mid-2017 and mid-2022, from Bookstat. The data include titles appearing among the top few hundred thousand per day in sales. For each of these titles, we have book metadata, including publication date, genre, author name, and publisher. For the validation, define $r_{j,\tau}$ as the number of cumulative ratings received by book j as of τ periods since its release. Using the Bookstat estimates of Amazon sales, along with information on the number of Amazon ratings the ebook has received by that time, we explore how well ratings reflect sales. Starting with the ebooks published in April 2017, and in the Januaries of 2018-2020, we calculate the estimated cumulative sales (q) of each title by month (τ) until mid-2022. Figure A.1 shows how the cross-book correlations of $\ln(q)$ and $\ln(r)$ evolve with time since publication. By 30 days after publication it reaches nearly 0.4, and it rises to 0.6 after about seven months. This indicates that our usage measure is good for books published through the end of the sample period. Because the main usage measure is obtained in April of 2026, even the December 2025 sample books have four months to accumulate ratings.

A.2 Adjusting the number of ratings

The number of ratings a book has received as of the time of data collection depends in part on the amount of time that has elapsed since publication. Some of our estimation approaches require a measure of the number of ratings received that is comparable across books; and that requires a way to adjust the number of ratings for the time elapsed between publication and observation.

We create an adjustment using information on the evolution of books' ratings across time. In particular, we observe the number of ratings for each book at two points in time, roughly 60 days apart.[36] That is, we observe r_j at a time τ (τ periods since j's publication) and at time $\tau + k$. Because our books have a range of ages at each observation time, we can characterize how r_j varies over time with the age of the book. We assume that the number of ratings for a book grows proportionally according to

$$r_{j,\tau+k} = r_{j,\tau} e^{g(\tau)k}, \tag{6}$$

where $g(\tau)$ is the daily ratings accumulation rate at time τ since publication. We calculate $g(\tau)$ for each 30-day age interval.

We want to normalize the ratings measure to the number of ratings that would have accumulated by an age of T, i.e. a time that is T after publication (we use 72 months). We

[36] We collected data on February 6 and, later, on April 5, 2026.

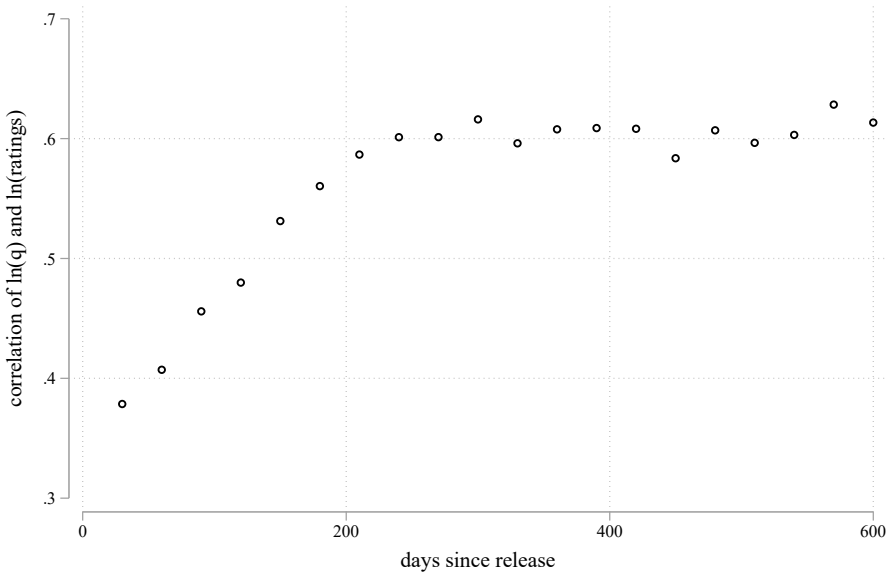
do this as follows: $r_{jT} = r_j e^{\Sigma_{i=\tau}^{T} g(i)}$.

We calculate the daily ratings growth rate from age τ as $\ln(r_{\tau+k}/r_\tau)/k$. Ratings grow quickly immediately after publication and more slowly as books age. Figure A.2 shows the average time pattern of ratings growth, among books with positive ratings, normalized to one at 72 months.

We handle books with zero ratings at the time of observation differently. Using our repeat observation data, we see the share of zero-rating books transitioning to having positive numbers of ratings at each age, as well as the average ratings for those books at time $\tau + k$. We use these empirical transition shares to assign the books obtaining ratings to the average number of ratings that books getting positive ratings obtain, times the probability of transitioning. Adjusted ratings are all then positive, and we adjust the non-zero ratings forward as above. We refer to the adjusted number of ratings as $\tilde{r}_j$.

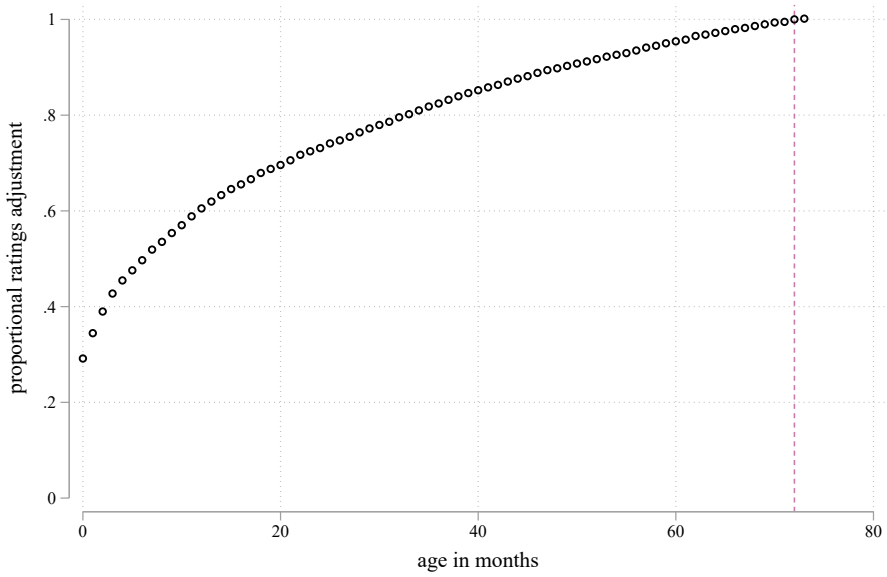
A.3 Appendix figures and tables

Figure A.1: Quantities and the number of ratings



Notes: The Figure shows the correlation between log cumulative sales and the log of ratings as of various numbers of days after publication. Included ebooks were published in April 2017 and Januaries of 2018-2020.

Figure A.2: Ratings adjustments by book age at collection



Notes: The figure shows the proportional continuous adjustment factors for ratings. We divide raw ratings by the adjustment to create our 72-months-old "adjusted rating" ($\tilde{r}_j$).