The Consequences of Generative AI for UGC and Online Community Engagement

Gordon Burtch, Dokyun "DK" Lee, Zhichen Chen
Boston University, {gburtch, dokyun,zhichenc}@bu.edu

Generative artificial intelligence (AI) technologies, such as ChatGPT and Midjourney, have recently revolutionized the way individuals create, share, and consume content across various domains. While these technologies hold the potential to democratize access to information and streamline content production, they also hold the potential to inadvertently undermine the health and sustainability of online communities, and the open exchange of knowledge more generally. In this work, we investigate the impact of generative AI on user content production, considering the context of online Q&A at StackOverflow, attending to potential shifts in the demand for content (shifts in audience size, attention, and question volumes) that may be induced by ChatGPT, as well as shifts in the quality of answers users may produce (potentially with ChatGPTs assistance). Leveraging a rich dataset from StackOverflow spanning October of 2021 and March of 2023, we estimate that ChatGPT has led to a systematic decline in the demand for information at StackOverflow. Further, we estimate that ChatGPT has also led to a systematic rise in the prevalence of low-quality answers, as reflected by an increase in the rate with which answers are down-voted. Our work contributes to a recent, growing body of literature on the social and economic implications of AI and generative technologies. Further, our work sheds light on the potential challenges and opportunities that online communities may face in the coming months and years. More generally, our findings offer useful insights that may inform the operation and design of online communities going forward, and they bear implications for the open exchange of knowledge in society going forward.

Key words: Generative AI, Platform Economics, User-Generated Content, Difference-in-differences

1. Introduction

The advancement and adoption of generative artificial intelligence (AI) technologies have transformed the way individuals create, share, and consume content across various domains, with significant implications for management scholars (Berg et al. 2023). Since its release in November of 2022, ChatGPT, one of the most advanced iterations of large language models, has set a record, reaching 100 million active users in just two months.[1] Midjourney, a text-to-image generative model that can create stunning visual outputs from textual inputs, has been used to win an art contest, demonstrating the powerful capabilities of Generative AI in content creation.[2]

[1] Reuters: ChatGPT Sets Record for Fastest-growing User Base.

[2] New York Times: AI Generated Art Won a Prize.

While these technologies hold the potential to democratize access to information and streamline content generation, they may inadvertently undermine the health and sustainability of online communities, which center around user-generated content (UGC). This paper aims to empirically investigate the impact of generative AI on user contributions to online communities, any resulting shifts in content quality, with an eye toward the long-term implications for the health and sustainability of online communities, and the open exchange of knowledge more generally.

Generative AI is likely to alter the dynamics of content consumption, creation, and exchange, for several reasons; as generative AI becomes an increasingly prevalent source of information or content for individuals, the latter's *participation and demand for content* in online communities may wane. Intuitively, in the presence of generative AI, individuals may increasingly turn to that technology for answers or solutions, in lieu of other humans, e.g., those participating in online communities. Thus, rather than visit Stack Overflow, Quora, or Reddit to pose a question, to solicit creative ideas, or to engage in dialog or idea exchange, individuals may increasingly turn to generative AI tools, such as ChatGPT, Bard, GitHub Co-pilot, or Midjourney. Were this shift to occur *en masse*, it may have a very significant impact on the long-term sustainability of those online communities.

Generative AI technologies may also influence the *supply side* of user-generated creative and knowledge content. First, users who were previously core, active contributors may reduce their engagement in the community and cease contributions as their audience shrinks (Zhang and Zhu 2011), or if they perceive that AI-generated answers, e.g., from ChatGPT, are of sufficient or superior quality. Second, generative AI technologies can be viewed as lowering the fixed cost of UGC production, making it easier for users with lower expertise or knowledge to participate in content creation. If the quality of content that low-expertise individuals are able to produce with the help of generative AI tools falls short of experts, this might lead to an increase in the volume of low-quality content, and thus an average decline in content quality. When it comes to UGC that is informational in nature, content generated with the assistance of AI is very likely to fall short of that provided by individual human experts in many cases, because low-expertise users will often lack the ability to assess the veracity, accuracy, or quality of information that generative AI tools supply. This point is notable, given many well-known LLMs are known to engage in 'hallucination' (Bang et al. 2023), i.e., providing factually incorrect information with an air of confidence.

The above dynamic, were it to play out as anticipated, would imply a vicious long-term cycle, wherein users in search of reliable content may become increasingly reliant on generative AI tools (provided by private firms), rather than open communities. Ultimately, this would translate to a lack of open exchange of ideas and knowledge among individuals in society, and a shift away from crowd-sourced models of content production. More formally, our work thus seeks to address the following questions. **How does the availability of generative AI technologies affect the volume of user contributions? Further, how do these technologies affect the average quality of content that users contribute?**

To answer these questions, we leverage a rich dataset capturing user contributions to the Stack Overflow community between October of 2022 and March of 2023, as well as a 'paired' (control) sample from the same window of time, between October 2021 and March of 2022. We exploit a sharp inter-temporal shock, namely the introduction and rapid adoption of ChatGPT, to causally identify the impact of generative AI tools on user content generation in the community. Further, we will examine heterogeneity in those effects across different users and content domains.

Our study contributes to the growing body of literature on the social and economic implications of AI and generative technologies. By providing empirical evidence on the consequences of generative AI adoption in the Stack Overflow community, we shed light on the potential challenges and opportunities that similar online communities may face. Furthermore, our findings have important practical implications for the design and management of online communities, particularly in terms of preserving user engagement, fostering content exchange, and maintaining content quality.

In sum, this paper aims to provide a comprehensive understanding of the impact of generative AI on online communities centered around UGC, such as Stack Overflow. We hope that our findings will inform policymakers, platform designers, and users alike as they navigate the complex interplay between technology, human creativity, and the concentration of knowledge in the era of AI.

2. Literature Review

Prior research has identified a variety of motivating factors behind individuals' engagement in online communities and their production of user-generated content (UGC). The various motives generally operate to differing degrees across community contexts, depending on a community's design, composition, and focus. For example, many sub-communities on Reddit are highly creative and social in nature, playing host to conversations and discussion (Burtch et al. 2022). By contrast, other venues, namely knowledge forums like Quora or Stack Overflow, are relatively less social, and more focused on the efficient production and exchange of objective knowledge and information.

This contrast in community characteristics is important to consider, because they bear implications for the role that generative AI may play in content production and consumption, as thus the effects that generative AI tools may have. This is because generative AI technologies may exhibit differing potentials to complement or substitute for the traditional element of online communities when it comes to meeting the needs of content producers and consumers. Forums that are heavily focused on creativity may thrive in the presence of generative AI, if the available tools enable humans to focus their attention on non-rote tasks, e.g., exploring creative solutions or generating new or more elaborate ideas, while substituting for more inane or rote tasks, e.g., correcting typos. Further, generative AI tools may offer a relatively inferior social experience (at least for the time being). This story would reflect an oft repeated dynamic of technology's effects on the labor market and innovation; rote tasks are replaced, thereby enabling workers or individuals to focus on more value-adding activities (e.g., Barrett et al. 2012).

By contrast, in community contexts that are highly centered on the production and exchange of objective information (and relatively less so on creativity or socialization), the dynamics may be quite different. Most obviously, if users who would previously have posted questions begin to turn to generative AI to for solutions, attention and eyeballs will decline, which in turn may reduce the incentive for individuals to contribute. When it comes to posting questions in Q&A forums, contributors' central motive will generally be a desire to learn and accrue information. To the extent a generative model can provide that information, the demand for information is likely to fall.

By contrast, the motives for contributing answers in Q&A forums are more varied. For example, users are known to be driven by a desire for status and recognition (Wasko and Faraj 2005, Levina and Arriaga 2014), as well as potential tangible benefits, such as career advancement (Xu et al. 2020, Roberts et al. 2006, Hann et al. 2013). For both of these incentives, the size of one's potential audience will be a major determinant of the of their influence (Zhang and Zhu 2011). If the demand for content declines and consumers cease visiting the community, answering experts may have lesser potential to earn status and recognition, and they may also perceive weaker career benefits from prospective employers in turn.

That said, there is an inherent tension here. Many of these users may actually increase their contributions in the presence of generative AI, if they perceive that these tools effectively reduce the cost of content production. This may imply an increase in content production, helping to sustain the community and offset any negative consequences. Indeed, some early studies have documented that generative AI tools can have strong, productivity-enhancing benefits for workers. Peng et al. (2023) examined the effect of GitHub Co-Pilot on the productivity of programmers, undertaking a controlled experiment wherein programmers were recruited to complete a specific programming task. Subjects treated with access to Co-Pilot were able to complete the task 55% faster than control subjects. Noy and Zhang (2023) report on a similar experiment, involving ChatGPT and writing-intensive tasks. Those authors recruit several hundred college-educated professionals and assign them to complete a variety of "mid-level professional writing tasks." Half of the participants were randomly assigned access to ChatGPT. The authors observed that subjects with access to ChatGPT exhibited a significant, substantial increase in the volume of tasks they were able to complete, and the quality of their output. The authors found that the tool primarily substituted for subjects' effort on rote tasks, enabling a focus on more value-adding exercises, namely ideation and editing. Considered in our context, users may thus leverage these tools to respond to more questions, more quickly, and they may begin to respond to questions they would previously have chosen to ignore, due to constraints on their time, or because the anticipated benefits of responding would previously not have outweighed the cost of generating an answer. In sum, the potential exists that generative AI tools may enhance productivity, creating a countervailing mechanism by which the cannibalization of attention by generative AI tools may be offset by a commensurate rise in the volume and quality of content supplied in the community.

That said, the net effect of the above mechanisms will depend a great deal on which types of content are cannibalized or complemented, as generative AI tools are likely to vary in the quality of content they produce across domains (e.g., depending on the volume of training data available in a subject area). Further, the productivity enhancing benefits of these tools is likely to vary across individuals as well. Here, it is worth noting that both Noy and Zhang (2023) and Peng et al. (2023) observed heterogeneity in the treatment effects, such that systematically greater productivity benefits were observed among lesser-skilled or lower-quality workers. This observation is secondary cause for possible concern in information-oriented online communities, because it suggests the potential that the cost of contributing answers may fall to a systematically greater degree among lower-quality (e.g., non-expert) contributors. As these users will be systematically less capable of verifying the quality or accuracy of informational content produced by generative AI tools, their relatively greater participation, making use of such tools to aid their responses, may lead to a systematic decline in the quality of answer content. Importantly, this dynamic would be less of a concern in other community settings, e.g., those focused on creativity or social exchange, where quality (e.g., beauty or interest) is more in the eye of the beholder.

2.1. Descriptive Evidence of ChatGPT's Effects

We begin by providing a descriptive evaluation of the influence that ChatGPT has had on Stack Overflow since its introduction on November 30th 2022. First, to understand potential shifts in the demand for Stack Overflow content, we turn to descriptive data on web traffic volumes from Simple Web. Figure 1 depicts the time series of visitors to StackOverflow.com, by month, between January of 2021 and February of 2023, as reported by SimilarWeb. The red vertical dashed line indicates the release of ChatGPT. We observe a discontinuous decline in traffic to StackOverflow, consistent with a decline in the demand for content. Indeed, similar stats have been reported directly by SimilarWeb, which noted that Stack Overflow web traffic was done by approximately 14% year-over-year in March 2023.[3]

Next, we attempt to obtain some evidence of whether ChatGPT is influencing the content of answers provided on StackOverflow. Achieving this is a more complicated exercise, as labels of AI-generated content are obviously not available. We achieve this via a rather involved process, consider 500 answers posted to Stack Overflow in the days prior to ChatGPTs release, along with the associated questions. We then prompt ChatGPT to provide answers to these same questions, and store them. This sample thus provides us with a labeled dataset of answers to Stack Overflow questions that are known to be human generated versus AI generated.

[3] SimilarWeb: StackOverflow is ChatGPT Casualty

Figure 1 Web Traffic Volumes at Stack Overflow by Month (source: SimilarWeb)

Notes

We applied AI text detectors to the labeled answers; we considered multiple such detectors, but ultimately settled on the GPT-2 Output Detector, as it exhibited best performance (as we describe below). The yields a 'fake score' for any input text, which can be loosely interpreted as the probability that text is AI-generated. The scores are continuous values that ranges between 0 and 1. Figure 2 depicts the distribution of fake scores returned by the GPT2 output detector for our labeled sample. As can be seen, although the detector is far from perfect, applying extreme thresholds to its prediction output can yield an informative signal.

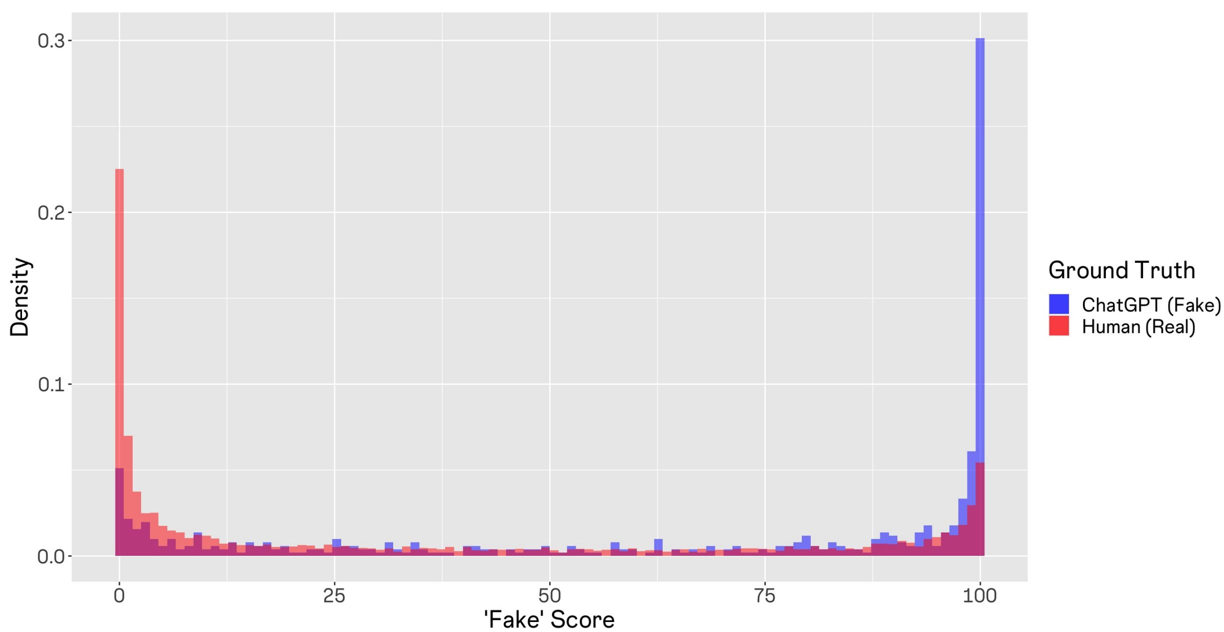

Figure 2 'Fake Score' Distributions of Human- vs. ChatGPT-generated Stack Overflow Answers

For example, the precision associated with the GPT2 Output Detector employing a classification threshold of 99.97% is 70%; that is, when the detector labels content AI-generated with 99.97% confidence, it is correct 70% of the time. The precision rises to nearly 80% employing a threshold of 99.98%. Having some confidence that the resulting labels can be informative of shifts in the prevalence of AI-generated content, we proceed to obtain fake score predictions for a larger sample of answers posted to Stack Overflow, arriving over the days surrounding the release of ChatGPT. We then calculate the proportion of answers labeled as AI-generated, over time. The result employing a threshold of 99.9% is reported in Figure 3.

Unsurprisingly, as the label is noisy, we find a non-zero (1-2%) prevalence of AI-generated content labels in the days prior to ChatGPT's release. What is notable, however, is that we observe a rise in the label prevalence in the days following ChatGPT's release, to 3.5%. AI-generated content declines on December 5^{th}, when Stack Overflow announced a ban on the use of ChatGPT in answering questions.[4] However, as the ban is effectively unenforceable (as there is no way to ensure accurate labeling of AI-generated content), the volume of AI-generated content appears to rebound.

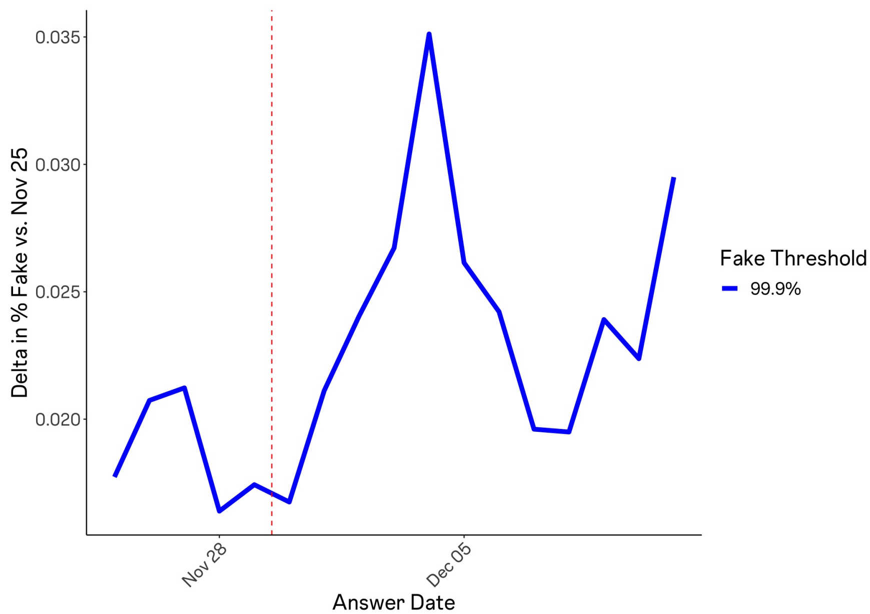

Figure 3 **Shifts in 'AI-Generated' Probability by Day**

3. Methods

Our analyses focus on contributions to Stack Overflow, both questions and answers. We collect data from the Stack Exchange Data Explorer, related to all answers posted between October 1^{st} of 2022 and March 19^{th} of 2023, a period that brackets the public release of ChatGPT (November 30, 2022).

We consider two levels of analyses. We begin with estimations that consider question-level activity, to understand effects on the supply of questions to Stack Overflow before versus after the introduction of ChatGPT. Subsequently, we examine effects at the answer-level, to understand shifts in answer characteristics. Our question-level

[4] Stack Overflow: Temporary policy: ChatGPT is banned.

sample is constructed by topic-tag. Thus, we construct separate panels for each tag topic, e.g., Python, Java, containing the count of questions posted mentioning each tag over time, by week. We also collect answer-level information, which includes each answer's date of posting, the text comprising each answer, the net of up- and down-votes, i.e., the answer score, an ID associated with the posting user, and the answering user's reputation score within the Stack Overflow community. We associate each answer to its parent question.

Our question-level panel includes a count of questions by topic tag and time. Our answer-level sample includes one observation per answer. We undertake two sets of analyses to understand the consequences that ChatGPT has had on content production at Stack Overflow. First, we consider the supply of questions arriving at the platform. Second, we consider shifts in average answer characteristics, proxying quality based on peer voting behavior.

We employ a difference-in-differences estimator to estimate the impact on content (questions or answers) volume and characteristics from the introduction of ChatGPT. As ChatGPT's introduction is a cross-sectional shock that applies to all users and content on the platform, we construct a control group employing a strategy similar to that of (Eichenbaum et al. 2020). That is, we obtain data from the same calendar period that brackets a point in time one year prior to ChatGPTs introduction (i.e., the control period brackets November 30^{th} of 2021). Our design exploits the sharp inter-temporal shock to content production from ChatGPT's release. This shock is likely to be sharp, given the speed with which ChatGPT has been adopted around the world is, simply put, unprecedented. In just two months post-launch, the app reached 100 million monthly users in January, becoming the fastest-growing consumer application in history.[5]

If patterns of contribution are seasonally stable, e.g., subject merely to a level-shift from one year to the next, then activity observed one year prior, conditional on appropriate combinations of fixed effects, e.g., by calendar time, topic, or user, can serve as a plausible counterfactual. Our control group thus identifies any shifts in question or answer characteristics that arose in the second half of the latter period of observation, which are presumably attributable to the presence of ChatGPT. Observations associated with the earlier time window are labeled as our control group, i.e., $Treat = 0$, whereas observations associated with the later time window are labeled as our treatment group, i.e., $Treat = 1$.

3.1. Question Volume Analysis

As noted above, based on the topic tags associated with each question, we aggregate counts of question volumes by time. To achieve this, we multi one-hot encode the 50 most popular tags, resulting in a series of dummy variables capturing whether each of the 50 most popular tags was mentioned for a given question, e.g., Python, R, Java, and so on. Aggregating question counts by tag yields a panel of tag-time question volumes, reflecting question volumes

[5] Reuters: ChatGPT Sets Record for Fastest-growing User Base.

by topic domain. Again, this panel captures question posting volumes over time in the period surrounding ChatGPT's release, as well as a window of time surrounding the same point one year prior. The former sample of data serves as our set of "treatment" observations, whereas the latter serves as "control." Using this panel, we estimate a regression of the form expressed in Equation 1. In this equation, i indexes topic tags, and t indexes weeks. Our estimation focuses on the interaction between $Treat$ and the $Post$ dummy, which indicates whether an observation takes place within its respective sample at a time after (vs. before) the calendar time of ChatGPT's introduction. Subsequently, as expressed in Equation 3, we estimate a variant of this model, exploding the $Post$ dummy into a set of $RelWeek$ dummies, omitting the week just prior to the calendar timing of ChatGPT's introduction, as reference. In each model, we incorporate a topic tag fixed effect, and we cluster our standard errors by tag. Further, we consider log-linear models as well, to recover estimates of percentage effects on question supply. Lastly, we turn our attention to heterogeneity, re-estimating the simple difference-in-differences model by topic tag, considering both linear and log-linear effects, to understand which topics among the tag set were most heavily affected by ChatGPT's introduction.

$$QuestionVolume_{i,t} = Treat_i + Post_t + Treat_i \cdot Post_t + \alpha_i + \mu_{i,t} \tag{1}$$

$$QuestionVolume_{i,t} = Treat_i + RelWeek_t + Treat_i \cdot RelWeek_t + \alpha_i + \mu_{i,t} \tag{2}$$

3.2. Answer-level Analysis

Next, we turn our attention to the answer-level data. Employing a similar identification strategy, we examine, shifts in vote activity associated with posted answers. Recognizing that a decline in web visitors can be expected to drive a decline in total votes, given fewer individuals are present to cast their votes, we do not model the shift in average net votes an answer receives. Instead, we focus on answers that receive at least one vote, and we model the probability that the net score is negative, rather than positive. We estimate a regression of the form expressed in Equation **??**. Here, q indexes questions and u indexes responding users. [6] Our outcome of interest is a binary indicator of whether the net vote score is negative, $NetNegative$.

Our initial estimation once again consider the $Treat$ and $Post$ dummies. Subsequently, we consider heterogeneity tag topics, and again conclude by examining the dynamics of treatment effects, exploding the $Post$ dummy into a vector of $RelWeek$ dummies, omitting the week prior to treatment as reference. Additionally, our estimation includes a question fixed effect. By doing so, our estimation isolates variation in down-vote propensity conditional on the question being answered, contrasting answers that emerge before versus after ChatGPT is released. We cluster standard errors in this estimation by responding user and question.

$$NetNegative_{q,u} = Treat_q + Post_u + Treat_q \cdot Post_u + \delta_q + \mu_{q,u} \tag{3}$$

[6] Note that, because a given user typically does not answer the same question twice, any given answer will be uniquely identified by a question-user ID pair.

4. Results

4.1. Question-level Effects

For the sake of brevity, we report our results graphically, plotting coefficients from our question volume estimations in Figures 4 and 5. We see clear evidence that question volumes have declined heavily on StackOverflow following the introduction of ChatGPT. We estimate that the arrival rate of questions has fallen by approximately 18% over the 4 months following ChatGPT's release, with roughly 20 fewer questions arriving per topic in any given week. Taking into account the variation of these effects across different topics, our findings indicate that the most substantial decreases in question volumes correspond to topic tags closely related to specific, self-contained coding exercises. These instances, where ChatGPT is anticipated to excel due to the accessible training data on GitHub, exhibit the most pronounced decline in user queries. For example, python, css, flutter, reactjs, django, sql, arrays, pandas, and so on, are all references to programming languages, specific libraries within programming languages, or data types one might encounter while working with a programming language. In contrast, unaffected tags appear most likely to be those reflecting topics involving complex tasks involving not only the syntax of underlying programming languages but also *contextual information* more likely to be outside of the scope of ChatGPT's training data. For example, Spring and Spring-boot are Java-based frameworks for enterprise solutions, often involving back-end (server-side) programming logic with private enterprise knowledge-bases and software services. Questions related to these are intuitively questions for which an automated (i.e., cut-and-paste) solution is less straightforward, and less likely to appear in textual training data that would be readily accessible to ChatGPT. Additional examples here include the tags related to Amazon Web Services, Firebase, Docker, SQL Server, Microsoft Azure, and so on. Considering Figure 5, we see that the effects manifest quickly following ChatGPT's release. Further, the lack of meaningful effects in the pre-treatment period is consistent with the assumption of parallel trends, lending further support to a causal interpretation for our estimates.

4.2. Answer-level Effects

Next, we turn our attention to answer-level effects, particularly in regard to peer voting. Results associated with our answer-level regressions are presented in Figures 6 and 7. Consider the average effect of ChatGPTs release on the probability that an answer receives a net down-vote rather than a net up-vote, we estimate a small, statistically significant increase, of approximately 2 percentage points. When we consider heterogeneity in the effect across topic tags, interestingly, we see an apparent inverse relationship between whether a topic's question volumes were negatively affected by ChatGPT, and whether a topic's answer quality is negatively affected. For example, many of those topics that exhibited the weakest effects for question volumes (Azure, AWS, Excel, Docker) appear to exhibit the largest increases in the probability of down-voted answers. Conversely, many of the topics that were most affected in question volumes (Python, javascript, reactjs, flutter, dart, pandas) generally exhibit null or negative coefficient estimates in the down-vote analysis.

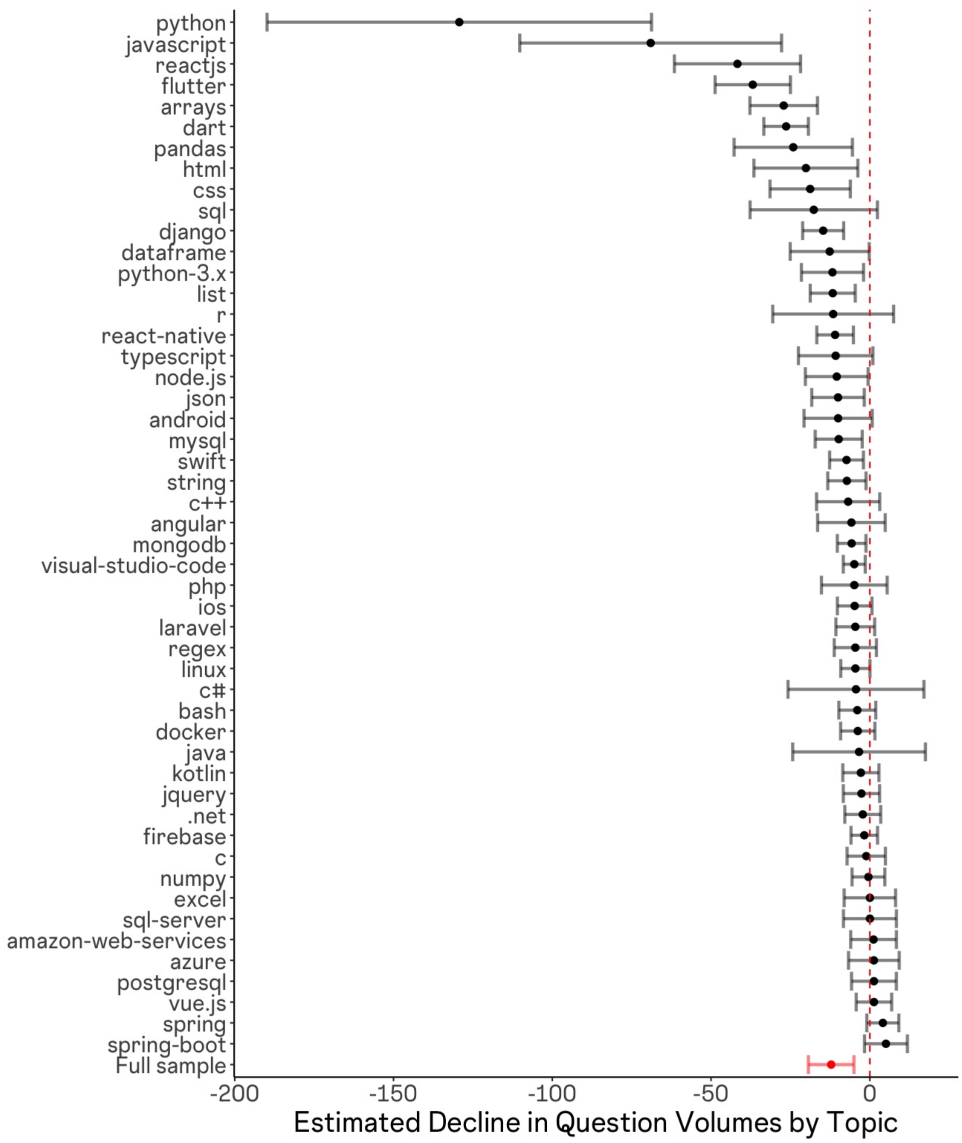

Figure 4 Question Volume Effect by Topic Tag and on Average

A plausible (even likely) explanation for this is that the effects we observe are jointly explained by a single underlying factor: ChatGPT is more effective at answering questions in some domains than others. In those domains where ChatGPT can provide high-quality, effective solutions, we see a systematic decline in questions posed on Stack Overflow (because users can obtain their answers from ChatGPT more quickly and easily). Conversely, for those topics where we observe no decline in question volumes, ChatGPT is presumably less capable. Once we recognize this, it is reasonable to expect that answers users may seek to generate with the help of ChatGPT will be systematically lower quality in those very same domains where ChatGPT is less capable.

Considering Figure 7, we see that for any answer attracting at least one vote from the crowd, there is a systematic rise in the probability that the answer receives a net negative score, versus a net positive score, with the effect becoming significant several weeks after ChatGPTs release. We see that the probability that a submitted answer is

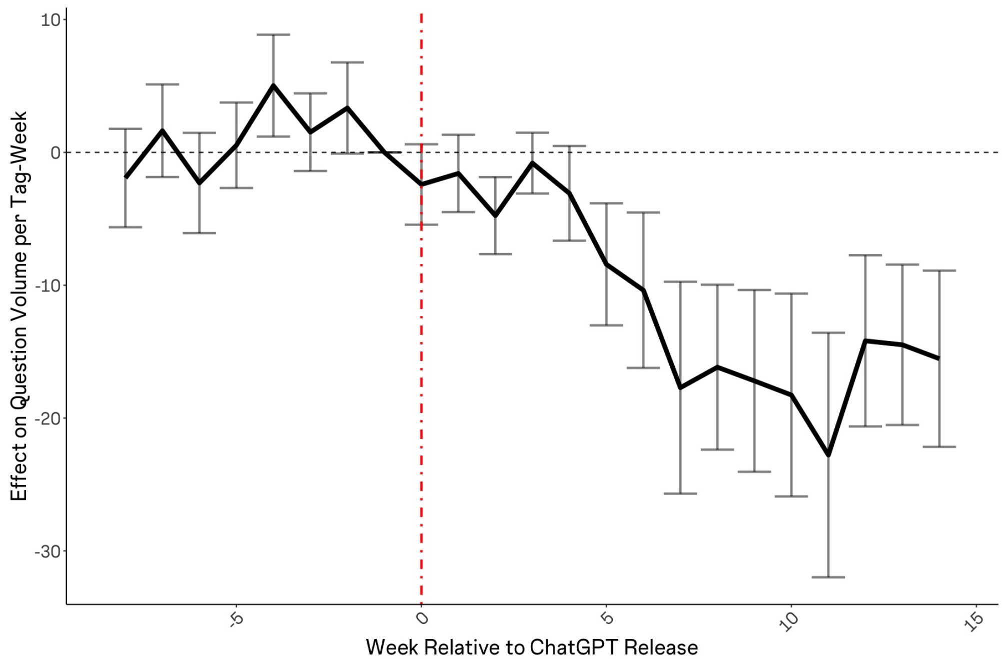

Figure 5 Average Question Volume Effect Over Time

down-voted rises, on average, by approximately 10 percentage points by the end of the 4-month period following ChatGPT's introduction. The lack of significant effects in the pre-treatment period once again implies no evidence to reject the parallel trends assumption, and thus again supports a causal interpretation to our findings.

5. Conclusion and Discussion

Prior studies have demonstrated that technology use in lieu of traditional manual or analog approaches to task execution can have material consequences for human development. As a simple example, work in neuroscience has found that individuals' reliance on navigation devices that provide turn-by-turn directions can lead to reduced spatial awareness and even lesser gray matter in the hippocampus (Alexander and Nitz 2015). This analogy is quite likely to apply to generative AI. A reliance on these tools and technology is likely to hollow out the demand for various skills and knowledge-bases, e.g., artistry, programming.

Specific to the knowledge-sharing online community, The pervasive adoption of generative AI models, such as ChatGPT, may lead to a substantial decrease in human knowledge-sharing activities (e.g., Stack Overflow, Quora, Q&A forums) as illustrated in this paper. This decline may involve the deterioration of high-quality responses and the archival of expert knowledge. Considering that generative AI systems learn from internet data originating from humans, there is a possibility that the depletion and decrease of human-contributed high-quality knowledge sources might eventually exhaust the available AI learning material, resulting in a scenario where generative AI can only marginally improve beyond existing online data.

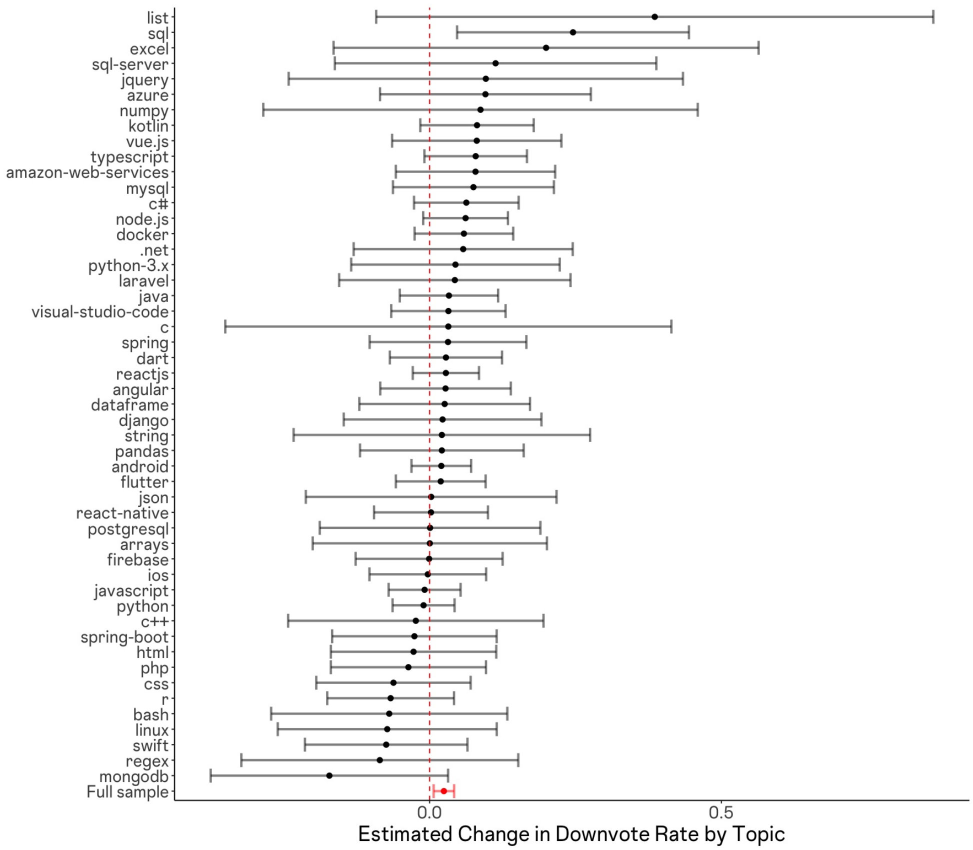

Figure 6 **Estimated Change in Pr(Downvote) by Topic Tag**

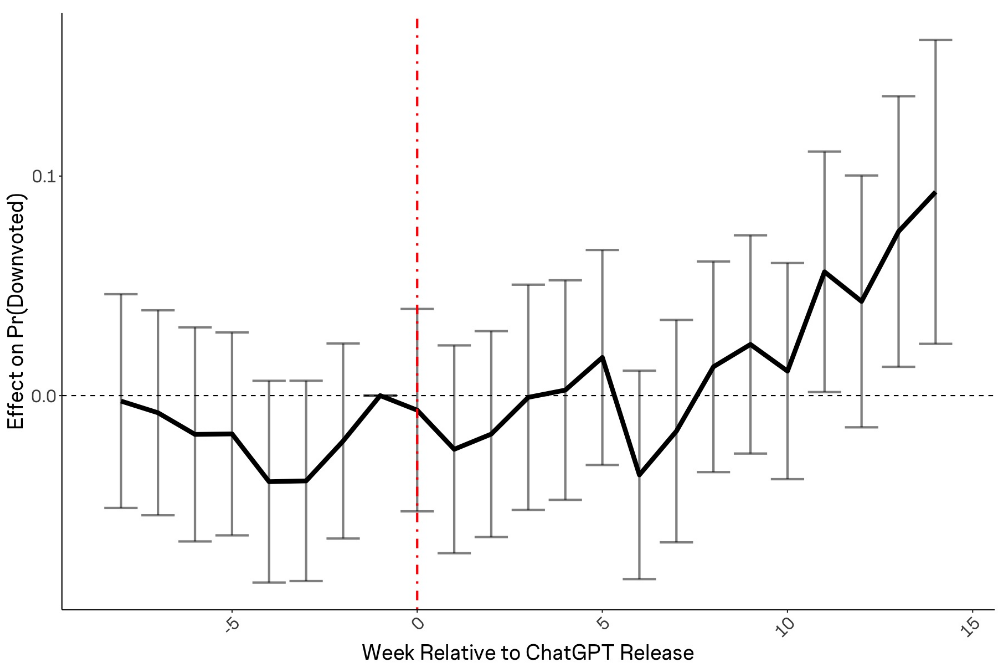

Figure 7 **Estimated Change in Pr(Downvote) Over Time**

Simultaneously, individuals could grow increasingly dependent on AI for problem-solving and knowledge acquisition, which in turn may reduce their inclination to engage in autonomous learning, critical thinking, and exploration. This reliance on AI-generated solutions might contribute to stagnation and an unfavorable equilibrium in the creation of new knowledge, impeding progress across various fields, including programming and other technical domains. In the long run, this could establish a negative feedback loop in which the knowledge repositories within AI systems become outdated, and the absence of innovative human insights further aggravates the issue. Although the future might not necessarily unfold in this manner, it remains essential to remain vigilant about potential unintended adverse consequences, not only for the platforms but also for the broader society that relies on online knowledge-sharing and collaboration.

References

Alexander, Andrew S, Douglas A Nitz. 2015. Retrosplenial cortex maps the conjunction of internal and external spaces. *Nature neuroscience* **18**(8) 1143–1151.

Bang, Yejin, Samuel Cahyawijaya, Nayeon Lee, Wenliang Dai, Dan Su, Bryan Wilie, Holy Lovenia, Ziwei Ji, Tiezheng Yu, Willy Chung, et al. 2023. A multitask, multilingual, multimodal evaluation of chatgpt on reasoning, hallucination, and interactivity. *arXiv preprint arXiv:2302.04023* .

Barrett, Michael, Eivor Oborn, Wanda J Orlikowski, JoAnne Yates. 2012. Reconfiguring boundary relations: Robotic innovations in pharmacy work. *Organization Science* **23**(5) 1448–1466.

Berg, Justin, Manav Raj, Robert Seamans. 2023. Capturing value from artificial intelligence. *Academy of Management Discoveries* (ja).

Burtch, Gordon, Qinglai He, Yili Hong, Dokyun Lee. 2022. How do peer awards motivate creative content? experimental evidence from reddit. *Management Science* **68**(5) 3488–3506.

Eichenbaum, Martin S, Miguel Godinho de Matos, Francisco Lima, Sergio Rebelo, Mathias Trabandt. 2020. Expectations, infections, and economic activity. Tech. rep., National Bureau of Economic Research.

Hann, Il-Horn, Jeffrey A Roberts, Sandra A Slaughter. 2013. All are not equal: An examination of the economic returns to different forms of participation in open source software communities. *Information Systems Research* **24**(3) 520–538.

Levina, Natalia, Manuel Arriaga. 2014. Distinction and status production on user-generated content platforms: Using bourdieu's theory of cultural production to understand social dynamics in online fields. *Information Systems Research* **25**(3) 468–488.

Noy, Shakked, Whitney Zhang. 2023. Experimental evidence on the productivity effects of generative artificial intelligence. *Available at SSRN 4375283* .

Peng, Sida, Eirini Kalliamvakou, Peter Cihon, Mert Demirer. 2023. The impact of ai on developer productivity: Evidence from github copilot. *arXiv preprint arXiv:2302.06590* .

Roberts, Jeffrey A, Il-Horn Hann, Sandra A Slaughter. 2006. Understanding the motivations, participation, and performance of open source software developers: A longitudinal study of the apache projects. *Management science* **52**(7) 984–999.

Wasko, Molly McLure, Samer Faraj. 2005. Why should i share? examining social capital and knowledge contribution in electronic networks of practice. *MIS quarterly* 35–57.

Xu, Lei, Tingting Nian, Luis Cabral. 2020. What makes geeks tick? a study of stack overflow careers. *Management Science* **66**(2) 587–604.

Zhang, Xiaoquan, Feng Zhu. 2011. Group size and incentives to contribute: A natural experiment at chinese wikipedia. *American Economic Review* **101**(4) 1601–1615.

Appendix A. Log-linear Estimations for Question Volume

Here, we report the result of our log-linear specifications for Question volume effects. Figures A1 and A2 present our estimations. Broadly, we observe results consistent with the linear models reported earlier. The main difference here, of course, is that our estimates bear a percentage interpretation.

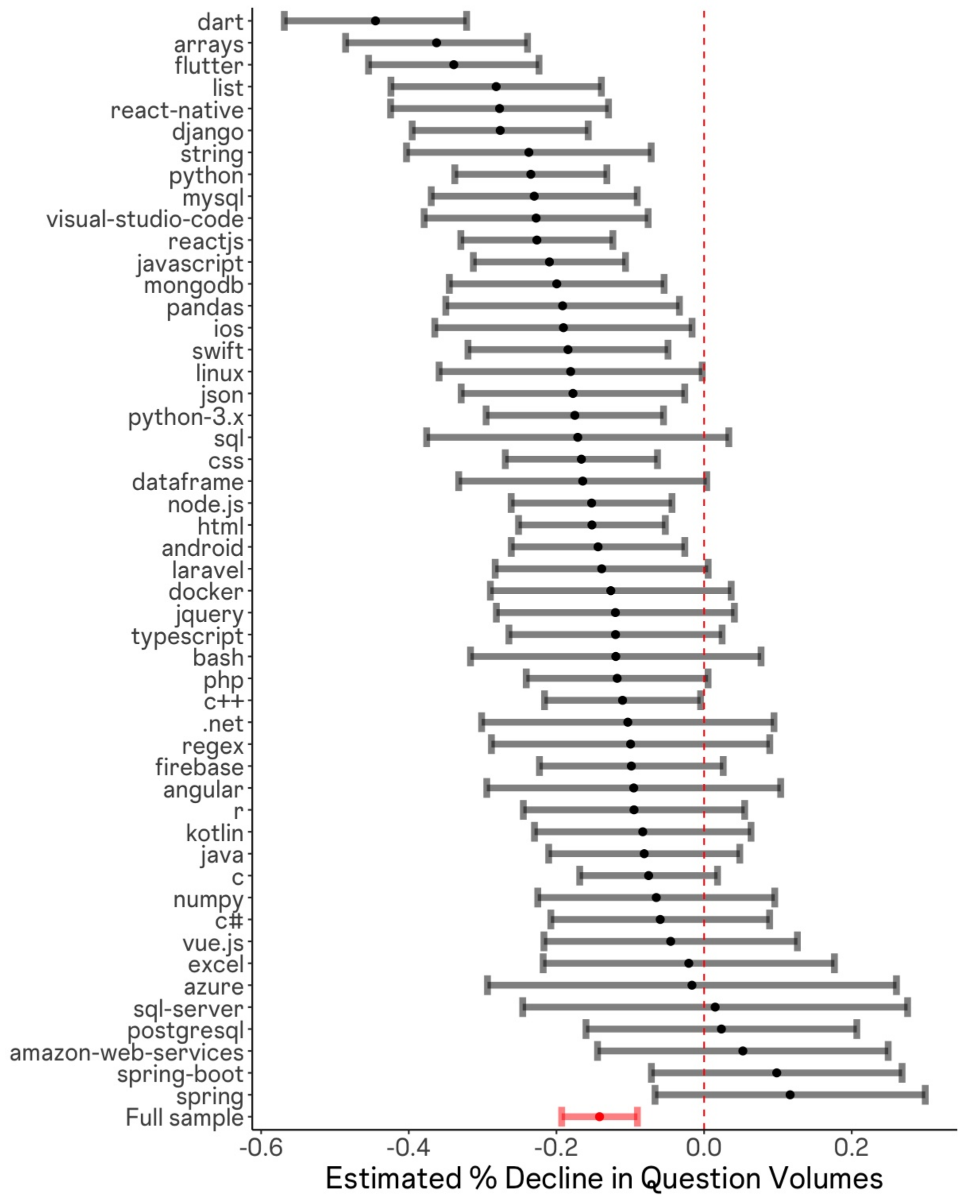

Figure A1 Log-Linear Question Volume Effect by Topic Tag and on Average

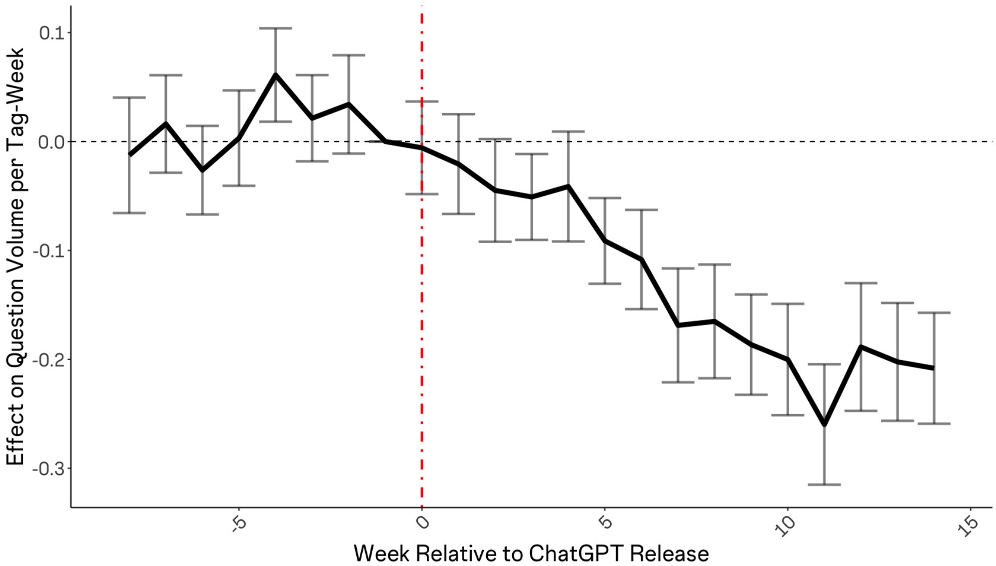

Figure A2 **Log-Linear Question Volume Effect by Topic Tag and on Average**