Hiding From Generative AI*

Sijie Lin†

Version: August 2024

Abstract

How does generative Artificial Intelligence (AI) change the behaviors of content creators? I investigate the effect of an AI image generator on artists' incentives to publish artworks using data from an online art platform, DeviantArt. On November 11 2022, DeviantArt introduced a generative AI image generator into the platform and artworks on this platform entered training data by default. Using a difference-in-differences estimation with artists who do not use AI, I show that digital artists publish 21% fewer artworks following AI's introduction on this platform, in contrast to artisan crafts artists. This reduction could potentially hinder knowledge spillovers to other artists and AI training data availability. By matching the artworks of artists who publish both on DeviantArt and Instagram, I find that despite artists publishing fewer artworks on DeviantArt, the quality of published artworks for a given artist remains the same after the introduction of AI.

*I am grateful to my advisors, Matthew Mitchell, Heski Bar-Isaac, Avi Goldfarb and Ambarish Chandra for their guidance and support. I also thank Siyuan Liu, Irisa Zhou, Tommaso Alba, Joshua Gans, Christian Peukert, Koleman Strumpf, and participants in North American Summer Meeting of Econometric Society, ESIF Economics and AI+ML Meeting of Econometric Society, American Law and Economics Association Annual Meeting, Canadian Economics Association Annual Meeting, student workshop and Rotman IO group at University of Toronto for their helpful comments. This research is funded by TD MDAL Research Grant at Rotman School of Management in University of Toronto. Previously circulated as "From Creation to Caution: The Effect of Generative AI on Online Art Market". All errors are my own.

†Rotman School of Management, University of Toronto (sijie.lin@mail.utoronto.ca).

1 Introduction

The power of generative Artificial Intelligence (AI) lies in its extensive training on a substantial volume of data. This data largely consists of copyrighted material, raising concerns among creators. Lack of consent and compensation for the use of their original creations brings pushback against generative AI and multiple lawsuits between AI companies and creators around the world. Artists have initiated a class-action lawsuit against AI companies for utilizing their artwork in training datasets without permission (Andersen v. Stability AI Ltd.), and programmers have litigated against GitHub for incorporating their publicly available code to develop the AI code-writing assistant, Copilot (Doe v. GitHub, Inc.). Authors including George Martin sued OpenAI for using their copyrighted materials to train large language model (Authors Guild v. OpenAI). There are also anti-AI protests on online art platforms like DeviantArt and LOFTER following the introduction of generative AI to the websites. Yet empirical evidence on the effects of copyright concerns of AI training data on incentives to share knowledge and on innovation is scarce. This paper presents an empirical context to study whether such copyright concerns disincentivize creators from disclosing more and higher-quality knowledge.

In particular, this paper exploits how the copyright concerns associated with generative AI impact the decision of content creators. To answer this question, I look into the introduction of DreamUp[1], an AI image generator, to an influential online art platform DeviantArt[2]. When DreamUp was launched on DeviantArt on November 11, 2022, all artwork on this website was automatically "opted-in" to be part of the training data of DreamUp, a decision that quickly sparked widespread discontent among the platform's artists.

My research questions are the following. First, how do copyright concerns undermine artist's incentives of disclosing their artwork? To address this, I collect the historical publishing data of 6835 artists from daily featured section on DeviantArt from January 2020

[1] www.deviantart.com/dreamup
[2] www.deviantart.com

to December 2023. Among these artists, digital artists, who make artwork using Adobe Photoshop or Procreate on drawing tables or iPads, are more exposed to AI image generators compared to artisan crafts artists, who usually produce hand-made jewelry, dolls and woodworks, etc. I focus on artists who do not use AI and employ a difference-in-differences approach and show that following the introduction of DreamUp on DeviantArt, digital artists reduced their monthly artwork postings by 21%. This result indicates that artists who are more exposed to AI publish fewer of their artworks on this platform in response to the introduction of generative AI.

Another piece of evidence of copyright concerns disincentivizing artists from disclosing more artworks comes from data of non-AI artists who are also publishing on Instagram. I obtain art publication records on Instagram of 1,467 multi-homing artists in the treatment group (digital artists) and control group (artisan crafts artists) and show that digital artists only reduce the number of artworks published on DeviantArt, not on Instagram, compared to artisan crafts artists.

Second, how do copyright concerns related to generative AI change the quality of published artwork? If artists publish fewer high-quality artworks due to copyright concerns, it could imply poorer training data for future AI algorithms, potentially diminishing the efficacy of subsequent generative AI models. For non-AI artists, the artworks they choose to publish also serve as a reference for other non-AI artists. Lower quality published artworks could also negatively affect the knowledge spillover for non-AI artists.

To address the question of whether non-AI artists are withholding high-quality or low-quality artworks from DeviantArt after the introduction of DreamUp, I focus on multi-homing non-AI artists, and examine the performance of artworks that are exclusively posted on Instagram versus those shared on both Instagram and DeviantArt. Artworks posted on both platforms generally receive more likes and comments on Instagram compared to those only posted on Instagram. However, such difference has no significant change after the introduction of DreamUp. These findings jointly show that despite the introduction of AI raising copyright concerns and leading to a decrease in the number of artworks published, the quality of artworks published does not change for a given digital artist.

These findings shed light on knowledge spillover. A decline in the number of artworks published is likely to negatively impact knowledge spillover. Artists sharing less of their work publicly means fewer opportunities for others to learn, get inspired, or build upon existing ideas, which could hinder future innovation.

Third, how does AI change the behaviors among AI adopters? Unlike many other settings, this paper has a direct measure of which artists use generative AI to create their artworks. My analysis reveals that approximately 11% of artists have adopted AI. Among these adopters, incumbent artists experience a significant increase in the number of artworks published, ranging between 55% and 60%. Additionally, AI artists who entered the platform after the shock publish significantly more artwork each month. Despite this increased output, the new artworks by these entrants receive fewer downloads on average than those created by incumbent artists. It's worth noticing that using AI image generators is not a simple task. Some artists are skilled at formulating with better prompts and selecting high-quality artworks with their professional taste. There are even prompts markets where individuals sell prompts that other artists can then refine and use. Although the cost of creating art has significantly decreased with the availability of AI image generators, it is still not negligible.

Overall, this paper presents evidence that while generative AI increase output among adopters, it simultaneously raises copyright concerns among non-AI artists, disincentivizing them from sharing their work. This implies decreased future knowledge spillovers and highlights the need for careful consideration of optimal copyright policies in the context of AI training data.

Related Literature

This paper is closely aligned with recent studies addressing copyright concerns associated with generative AI. Theoretical models highlight the potential harm AI models can cause to original content creators by demonstrating a decrease in their monopoly profits (Gans, 2024; Yang and Zhang, 2024). Gans (2024) examines how the bargaining power of AI companies affects social welfare under different copyright protection regimes. Yang and Zhang (2024) argues that stronger copyright protection can have different welfare implications depending on the availability of training data. When training data is abundant, strict copyright protection can hinder AI development and reduce social welfare.

Recent empirical studies also suggest that copyright concerns limit the availability of AI training data. For instance, Huang, Fu, and Ghose (2023) finds that artists published 40% to 50% fewer artworks after the online art platform LOFTER introduced an AI image generator. This reduction persisted even after the AI image generator was removed from the platform. In addition, the study shows that more prolific and multi-homing users reduced more than the others. My paper further shows that despite artists reducing the number of artworks published on the treated platform, their behavior on Instagram remains unchanged, suggesting that artists do not reduce their overall production. Peukert, Abeillon, Haese, Kaiser, and Staub (2024) shows that after a stock image website released a keyword-labeled dataset of 25,000 images available for AI research, users whose photographs were included in the dataset reduced the number of new uploads by 40% and were more likely to leave the platform. Professional and more popular users showed a greater reduction in uploads than others. Treated users also began uploading images that were more similar to pre-existing ones, as measured by keywords. My paper further explores the behavior of AI adopters and shows that artists who use AI increase the number of artworks they publish. AI artists who enter the market after DreamUp became available are also significantly more productive than previous entrants, highlighting the positive effect of AI on online art platform.

This paper also relates to the literature on the piracy of information products. Some studies focus on the effect of piracy on the incentives of innovation (e.g., sales, profits, revenues). Hui and Png (2003), Rob and Waldfogel (2006, 2007), and Zentner (2006) use instrumental variables for consumption of pirated music and find that piracy has a significantly negative impact on music sales, suggesting that piracy undermines the monetary incentives for innovation. In contrast, other studies suggest that piracy can positively impact sales by increasing consumer awareness of creative products through word-of-mouth effects or by enhancing the benefits of purchasing software with a positive network externality (Aguiar and Martens, 2016; Givon, Mahajan, and Muller, 1995; Peukert, Claussen, and Kretschmer, 2017). Oberholzer-Gee and Strumpf (2007) and Blackburn (2004) argue that the availability of file sharing websites has a negligibly negative impact on sales, as the displacement effect and word-of-mouth effect cancel each other out. Some papers study the impact of piracy on subsequent innovations. Waldfogel (2012a, 2012b) find that the quality of new recorded music did not change significantly when the file-sharing website Napster became available.

This paper contributes to the literature by showing that even if AI art does not divert consumers' attention away from non-AI artists, non-AI artists still withhold their artworks from DeviantArt. This withholding may be caused by concerns that their artworks will be imitated by AI in the future or as an expression of frustration over their ideas being stolen by AI models. Additionally, I show that non-AI artists are not withholding high-quality artworks following the introduction of DreamUp.

Although this paper finds no evidence supporting that AI-generated artworks divert consumer attention away from non-AI artists, the withholding of artworks still affects knowledge spillovers. Restricted access to existing content can negatively impact subsequent innovations. Murray and Stern (2007) and Williams (2013) demonstrate that after an innovation obtains a patent, knowledge diffusion, as measured by citations, decreases. Galasso and Schankerman (2015) uses exogenous allocations of judges to construct an instrumental variable for patent invalidation and shows that patent invalidation leads to an increase in citations to the patent. Nagaraj (2018) finds that out-of-copyright magazine issues of baseball are more likely to be cited on Wikipedia. Biasi and Moser (2021) shows that when US removed restrictions on publishing scientific books from its rival country during WWII, citations of these books increased, implying a rise in follow-up science works. If artists withhold artworks on DeviantArt due to copyright concerns, future artists will have less human-generated content to learn from, and future AI models will also have smaller training datasets.

Lastly, this paper relates to literature about the impact of copyright protection on innovation and knowledge diffusion. Reimers (2019) shows that stronger copyright protection can lead to increased prices of books and fewer editions per title, which leads to decreased knowledge diffusion. Giorcelli and Moser (2020) shows that when Napoleon's military campaign brought copyright laws to parts of Italy, both the quantity and quality of new operas increased. This paper contributes to this literature by showing that when copyright protection in the domain of generative AI is ambiguous, the overall amount of innovation does not change after DreamUp's introduction. Instead, artists only reduce the number of artworks published on DeviantArt without decreasing the number of artworks produced.

The paper proceeds as follows. Section 2 lays out the empirical setting of DeviantArt and the introduction of the AI image generator DreamUp. Section 3 describes the artwork-level data collected from DeviantArt and Instagram. Section 4 presents the difference-in-differences approach employed in this paper to measure the decline in the number of artworks published after the shock. It also presents the measure of artwork's quality, and shows that quality has not changed after the introduction of DreamUp. Section 5 lays out the estimation results. Section 6 shows the behavior change of AI adopters. Section 7 concludes.

2 Empirical Setting

2.1 Empirical Setting: DeviantArt

DeviantArt (www.deviantart.com) is one of the earliest and largest online art platforms that allow artists to display their artworks and interact with viewers. Viewers are allowed to "favorite" and comment on any artworks and send a direct message to artists.

Artists can also sell artworks in different ways, including premium downloads, adoptables, commissions, and subscription fees. Artists may ask for a fixed price if consumers want to download a version with a higher resolution (premium downloads). They may also sell designed characters, and once a consumer purchases it, he could build his original character based on it (adoptables). Sometimes, consumers ask for a customized artwork (commissions). Some artists have art series for monthly subscription fees. Being famous on DeviantArt can be financially rewarding. Well-known artists on the platform have the potential to earn tens of thousands of dollars per month.

Beyond the direct monetization of their artworks, achieving fame on DeviantArt can offer artists additional rewarding opportunities. Many companies, including games and animation companies, search for potential employees or partners on this platform. For many artists, DeviantArt is a valuable tool for gaining recognition and building a reputation within the art community. Art students also use it to construct portfolios and apply for professional art schools.

The market of online art platform is highly concentrated. By June 2023, DeviantArt has over 75 million registered members and over 550 million artworks on the platform[3]. Patreon (www.patreon.com), one of DeviantArt's main competitors, has only 0.2 million registered artists and less than 7 million artworks[4]. ArtStation (www.artstation.com) is another main competitor of DeviantArt and has only 12% of the search frequency of DeviantArt from 2021 to 2023 according to Google Trends[5].

There are three primary reasons for selecting this platform as the empirical setting. First, DeviantArt was the pioneer among online art platforms in integrating generative AI, resulting in more unanticipated copyright concerns compared to other platforms. After the introduction of AI on DeviantArt, artists on its competitor ArtStation requested the platform to ban AI-generated content. However, ArtStation declined this request in December 2022,

[3]www.deviantart.com/about#:~:text=We%20have%20over%2075%20million,of%20art%20on%20the%20platform.
[4]https://c5.patreon.com/external/press/resources/fact-sheet.pdf
[5]https://trends.google.com/trends/explore?date=2021-01-01%202023-12-12&q=DeviantArt,ArtStation&hl=en

leading to widespread copyright concerns about ArtStation one month after the situation on DeviantArt. Subsequently, another online art platform, LOFTER, introduced an AI image generator, the Laofuge Drawing Machine, in March 2023, four months after DeviantArt's initial implementation. Secondly, the lawsuit involving artists, DreamUp's parent company Stability AI, another AI firm Midjourney, and DeviantArt, represents one of the initial lawsuits to address copyright issues in the training data of generative AI. It is important for policymakers to understand the effect of copyright concerns on artists' incentives to disclose artworks. Related empirical evidence remains scarce. Thirdly, DeviantArt is one of the most substantial and influential online art platforms in the industry, making it an important site for the study.

2.2 Empirical Setting: Introduction of DreamUp on DeviantArt

AI image generators like DreamUp have some important features. Firstly, creators can specify the style of the artworks. For example, artworks can be generated based on the style of the famous Mexican artist Frida Kahlo. This capability implies that once the AI is sufficiently trained with a particular artist's works, it can effectively replicate or imitate that artist's distinctive style. Second, the cost of using AI image generators is low. For example, on average, it costs less than 10 cents per prompt on DreamUp. Third, they are not time-consuming to use. DreamUp can generate three artworks in 60 seconds.

DreamUp was not the first influential AI image generator. DreamUp was introduced to DeviantArt in November 11, 2022. Four months earlier, on July 12, 2022, the Midjourney image generation platform first entered an open Beta test, and on July 20, 2022, DALL-E 2 entered a beta phase with invitations sent to 1 million waitlisted individuals, and the waitlist requirement was removed on September 28. On August 22, 2022, Stability AI announced the public release of Stable Diffusion.

However, most artists on DeviantArt did not express concerns until the introduction of DreamUp. This is primarily because artists realized that, by default, all artworks on

DeviantArt were included in the AI training data. Many of them deactivated their accounts or stopped posting new artworks on this platform to prevent what they perceived as "stolen by AI," expressing their dissatisfaction on both social media and their personal DeviantArt pages in November 2022. Although artists were provided the option to be "opted-out" of the training data set at the beginning, and DeviantArt announced that it changed the default setting to be "opted-out" after two days, artists continued to show copyright concerns. On January 13, 2023, artists on DeviantArt filed a class action (Andersen v. Stability AI Ltd.) against Stability AI, Midjourney (another AI image generator company) and DeviantArt. The lawsuit alleges copyright infringement by these companies, representing one of the first major legal cases over AI's role in copyright infringement.

3 Data

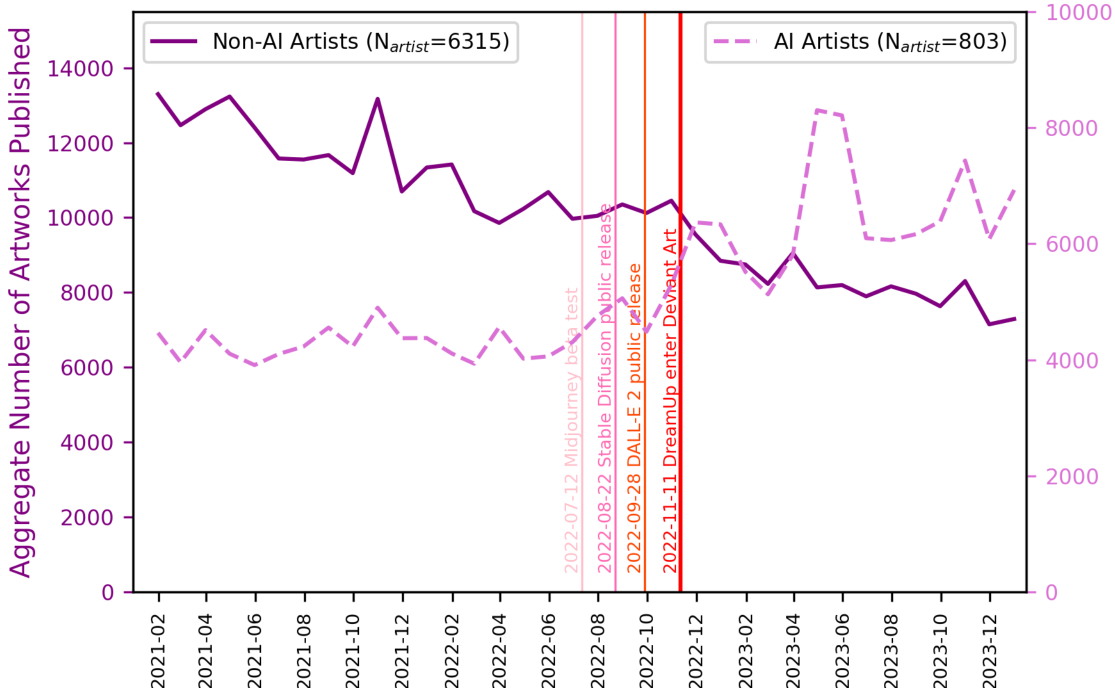
3.1 Data Description

I collect artwork-level data on artists whose artworks were selected for display in the daily featured section on DeviantArt. DeviantArt features approximately 10 to 20 artworks in this section, typically showcasing high-quality pieces. For the main analysis, I select artists who had their work featured in the daily section before January 2021 and have remained active since then to construct a balanced panel. These artists can be considered capable of producing artwork of sufficient quality to be featured before AI image generators became available. The full sample contains 7,118 artists, including information on their user profiles and publication history on DeviantArt. Table 9 in the appendix presents the summary statistics of the artists and artworks included in the main analysis.

Figure 1 shows the time trend of the historical aggregate number of artworks published on DeviantArt by my sample from January 2021 to December 2023.

Artists have heterogeneous reactions over the introduction of Dreamup. Typically, AI-generated artworks are self-identified with AI-related title, descriptions and tags like "ai",

Figure 1: Overall Time Trend



Notes: This figure shows the number of artworks published by all artists in the data. This figure includes both artists entered the platform before and after the shock.

"prompt", "midjourney", etc. I classify an artwork as AI-generated if it has AI-related keywords[6]. Accordingly, artists with at least one AI-tagged artwork are categorized as AI artists. One potential concern is the possibility of artists not being truthful about their use of AI in creating artworks. To address this, I conducted a verification check using "Created with AI" label from DeviantArt, which detects AI-generated artworks using their metadata and labels them accordingly. I cross-referenced these labels with the artists' self-reported information and found that only four out of eight thousand artists who did not disclose their use of AI had the AI-generated label. If some artists are, in fact, using AI image generators but never tagging their work as AI-generated, we might expect to see an increase in the number of artworks published by non-AI artists, which does not exist. Therefore, keywords should be a reliable way to identify AI-generated artworks. Additionally, there is little incentive for artists to falsely claim AI usage if they are not using it, as the platform's audience tends to devalue AI-generated art. If some digital artists secretly use AI without reporting it, the effect observed in the main regression will be underestimated.

There are 11% of AI artists and 88% of non-AI artists in the sample. Figure 1 shows that AI artists increased the number of artworks they published after the shock. Conversely, the publication frequency of non-AI artists decreases after the introduction of Dreamup.

In order to answer the question of whether artists disclose fewer high-quality or low-quality artworks, I also collect the history of posts of multi-homing artists on other platforms. Specifically, I focus on Instagram. Among these 4,947 artists used in the difference-in-differences estimation, 77% of them claim they have an account on other platforms on their DeviantArt profile page, and the the most popular platforms are Instagram, Facebook, Twitter, and YouTube. Table 1 shows the distribution of multi-homing artists across platforms. Among these artists, I find the ones with professional or business accounts on Instagram,

[6]The AI-related keywords are: ai, aiart, artificialintelligence, digitalai, midjourney, midjourneyaiart, midjourneyartwork, midjourneyai, midjourneyart, aiartcommunity, ai_art, aigenerated, aicreated, aiartgenerator. To avoid artworks expressing anti-AI attitude misspecified as AI-generated, I also create a list of anti-AI keywords as follow. If an artwork contains both AI keywords and anti-AI keywords, it will not be classified as AI generated. A list of anti-AI keywords: noai, no_ai, notoai, aiisnotart, stopaiart, DreamUp_is_unethical, aiistheft, ban_ai, notoaigeneratedimage, againstai.

which are accessible through the official Meta Developer Graph API. I am able to construct balanced panel data using 1,458 of the artists.

Table 1: Distribution of Multi-Homing Artists in Main Analysis

	Proportion Among Multi-Homing Artists
Instagram	62.9%
Twitter	51.1%
Facebook	37.8%
YouTube	21.8%
Tumblr	20.8%
Fraction of Multi-homing Artists	77.4%

Notes: The platforms are not mutually exclusive. Many artists are multi-homing on more than two platforms.

3.2 Matching Artworks Across DeviantArt and Instagram

Artworks on DeviantArt and Instagram are matched based on their title, description, publication date, and tags. For each pair of artworks across platforms for a given artist, I am able to calculate four similarity scores: title similarity, description similarity, date similarity, and tags similarity. Title similarity and description similarity are calculated using SequenceMatcher in Python's "difflib" module. Date similarity is defined as a decreasing function of the publish date difference between two artworks. Tag similarity is calculated using the Jaccard index, which measures the overlap between the sets of tags used for each artwork, providing a proportion of shared tags to the total number of unique tags across both artworks.

For an artist with m artworks on DeviantArt and n artworks on Instagram, this algorithm produces four $m \times n$ matrices. They are summed up with weight $(w_{title}, w_{description}, w_{date}, w_{tags}) = (\frac{15}{36}, \frac{8}{36}, \frac{12}{36}, \frac{1}{36})$[7]. Then the Hungarian algorithm described in Munkres (1957) is employed to match artworks according to this $m \times n$ matrix. This algorithm has been widely used in

[7] The weights are determined through iterative adjustment based on manual verification of accuracy. An initial weight of $(\frac{1}{4}, \frac{1}{4}, \frac{1}{4}, \frac{1}{4})$ was assigned to match the artworks. And then I manually check the matching accuracy and adjust the weight and then recheck. I repeat the same procedure until accuracy is higher than 85%.

one-to-one matching problems based on similarity in engineering and biology literature to match strings like bus names and genes (e.g., Guo, Jiang, Thornton, and Saunders, 2019; Mahmood et al., 2010). After the matching, I randomly select 150 artworks and manually check the matching, with 85% of them correctly matched or unmatched.

4 Identification Strategy

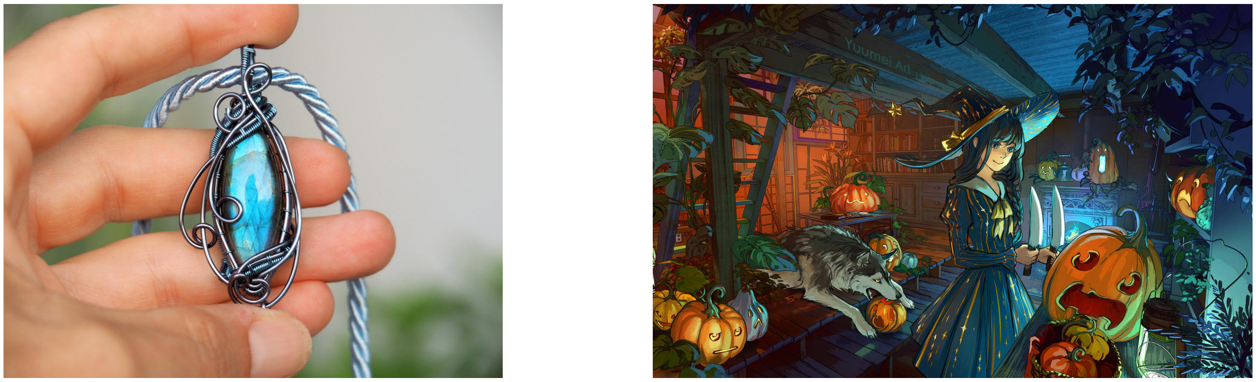
4.1 Impact on the Number of Artworks Published

I use a difference-in-differences approach to identify a decrease in the number of artworks published after DreamUp was introduced on DeviantArt in November 2022. Even though all artists on this platform face the potential threat of being part of the training data, some artists are more exposed to AI than others. The control group is non-AI artists specialized in artisan crafts who are less exposed to AI. These artists typically create physical objects like jewelry, dolls, and woodwork by hand, which are then photographed from various angles. The treatment group is non-AI artists specialized in digital art, who are more exposed to AI. Their work often involves fantasy themes, including dragons and sea monsters, created using software like Adobe Photoshop or Procreate on drawing tablets or iPads. These artworks can be easily generated by AI. The left panel of Figure 2 shows a typical artisan artwork, while the right panel shows a digital artwork.

Figure 3(a) presents the time trend of the aggregate number of artworks published by both the control group (non-AI artisan crafts artists) and the treatment group (non-AI digital artists). These artists are the subset of the 13,843 non-AI artists displayed in Figure 1. Before DreamUp's introduction, the control group and treatment group exhibits similar patterns, while after the introduction, digital artists' publications decrease compared to artisan crafts artists. There is a slight increase of total number of artisan artworks published after the shock. There are primarily two potential reasons. Artisan artists can be using DreamUp for idea searching even if they are not directly publishing AI-generated artworks.

Figure 2: Illustration of Artisan Crafts and Digital Art

(a) Artisan Crafts (b) Digital Art



Notes: I use artists specialized in artisan crafts as the control group and artists specialized in digital art as the treatment group. There are 559 non-AI artisan artists and 4,388 non-AI digital artists used in Table 2.

And since there is a sharp increase in the number of AI artworks published, artisan crafts artists may be publishing more frequently to attract attention. Similar stories should also apply to digital artists. In fact, it's more likely for digital artists to get inspiration from DreamUp than artisan crafts artists because generative AI is mostly used to generate digital artworks.

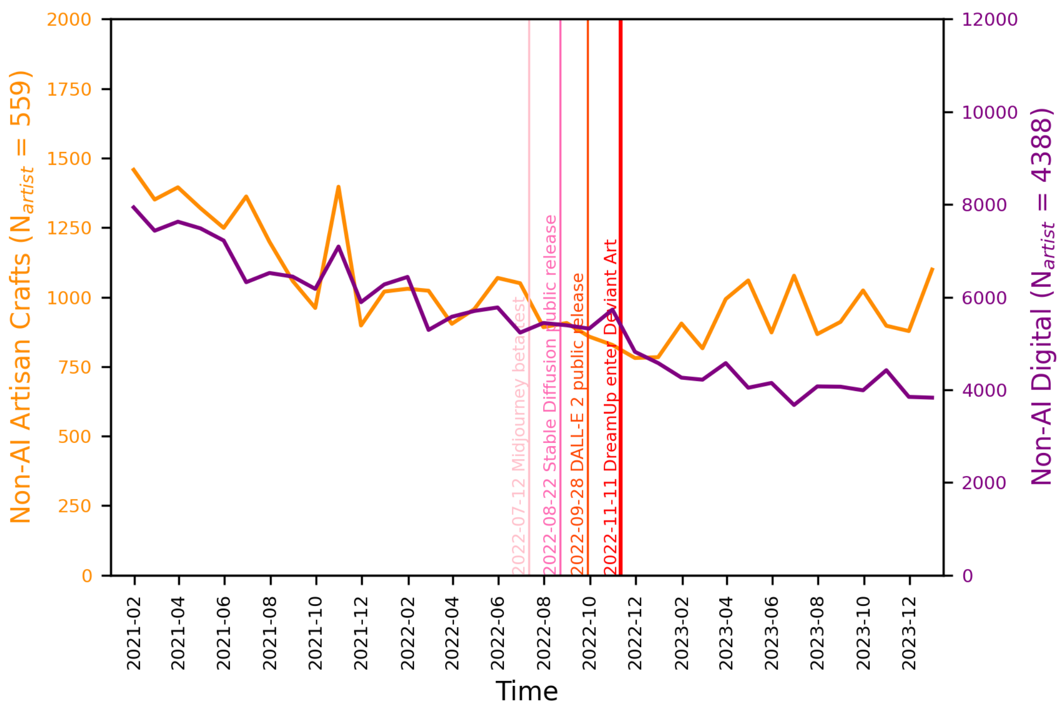
The main specification uses two-way fixed effects:

$$Artwork_{it} = \beta Post_t \times Treated_i + \delta_i + \delta_t + \epsilon_{it} \tag{1}$$

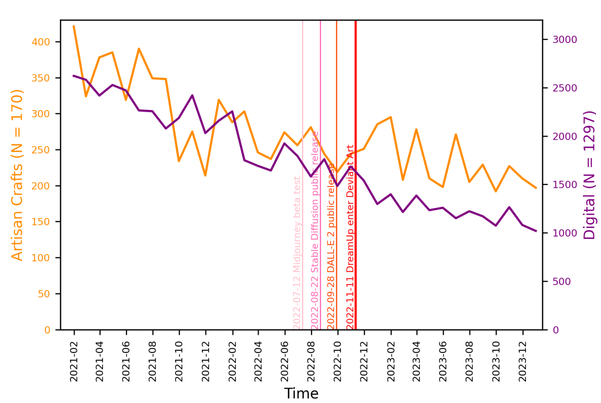
where $Artwork_{it}$ is the number of artworks published by artist i in month t. $Post_t$ is the dummy variable which equals to 1 if the observation is after October 31, 2022. $Treated_i$ is the dummy variable which equals to 1 if the artist's specialty is "Digital Art", equals to 0 if it's "Artisan Crafts". δ_i is the artist fixed effect and δ_t is the month fixed effect. To study the extensive margin of the quantity change, I also run a similar regression where the dependent variable $Stay_{it}$ is a dummy variable indicating whether an artist has at least one publication from month t until December 2023.

Figure 3: Time Trend of Control and Treatment Group

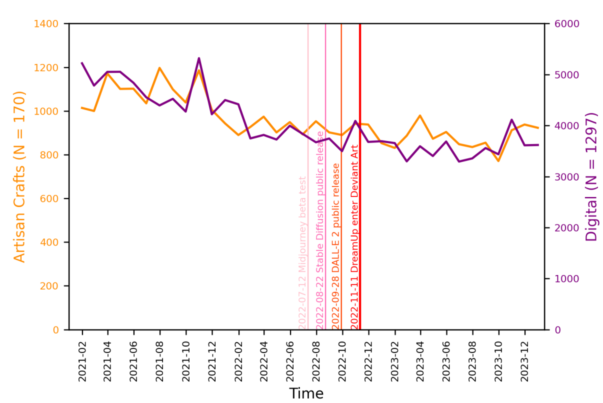
(a) DeviantArt Publication (All artists)



(b) DeviantArt Publication (Multi-homing artists)



(c) Instagram Publication (Multi-homing artists)



Notes: Figure (a) only uses the publication records of non-AI artists specialized in digital art and artisan crafts. They are the subset of the 6,315 non-AI artists displayed in Figure 1. Note that in this graph, only non-AI artists who entered Daily Deviation section before 2021-01-01 are included.

4.2 Impact on Quality of Art

To examine whether digital artists publish fewer high-quality artworks or low-quality artworks after the shock, I use the sub-sample of non-AI digital artists used in the difference-in-differences analysis and focus on Instagram users. I use their publication records on Instagram to see if there is a change in quality difference between the artworks also published on DeviantArt and the artworks only on Instagram. The specification is as follow:

$$y_{ijt}^{Ins} = \beta_1 Post_t \times Matched_j + \beta_2 Matched_j + \mu_i + \mu_t + \epsilon_{ijt} \tag{2}$$

where y_{ijt}^{Ins} represents the performance of artwork j by artist i published in month t on Instagram. Here I use number of likes and comments on Instagram as the measure of performance. $Matched_j$ is a dummy variable that equals 1 if the same artwork is also published on DeviantArt, 0 if it is only published on Instagram. δ_i is the artist fixed effect and δ_t is the month fixed effect.

The coefficient β_2 represents the difference in quality between artworks that are hidden on DeviantArt and those displayed on both DeviantArt and Instagram. If artists selectively present higher-quality artworks on DeviantArt, β_2 will be significantly positive. The coefficient β_1 indicates whether this difference changes after the introduction of DreamUp. If artists begin to withhold higher-quality artworks on DeviantArt after the shock, we would expect β_1 to be significantly negative. This analysis assumes that, given the total amount of attention an artist receives in a given month, higher-quality artworks attract more likes and comments on Instagram. The comparison is made across artworks by the same artist within a given month.

The quality of artworks can change not only at the intensive margin but also at the extensive margin. If better artists withhold a greater number of their works compared to worse ones, the overall quality of artworks may decline, leading to a potential decrease in the quality of publicly available knowledge in the future. To study the extensive margin of quality change, it is essential to measure artist quality. I use artist fixed effects to measure the quality of artists who entered daily featured section on DeviantArt before 2021-01-01. Specifically, I focus on their publications between 2020-01-01 and 2021-01-01, measuring artist quality based on their performance prior to the regression period. The equation below describes the measurement of artist fixed effects:

$$y_{ijt}^{DA} = f(date_{ijt}^{collect\ data} - date_{ijt}^{publish}) + \mu_i + \mu_{month} + \epsilon_{ijt} \tag{3}$$

where y_{ijt}^{DA} represents the performance of artwork j by artists i published in month t on DeviantArt. $f(date_{ijt}^{collect\ data} - date_{ijt}^{publish})$ is the polynomial of time elapsed since the publish date. The idea is that the number of downloads or favorites accumulates over time for a given artwork. Whatever cannot be explained by time elapsed since publication can be decomposed to artist average quality μ_i, calendar month effect μ_{month} and artwork specific component ϵ_{ijt}. If the quality is measured using the number of downloads, for example, the fixed effect μ_i represents the average number of downloads for an artist's artworks between 2020-01-01 and 2021-01-01, after removing the time effect. I then examine whether high-quality artists reduce publication more than low-quality ones using the following equation:

$$Artwork_{it} = \sum_{m \in \{High, Median, Low\}} \beta^m Post_t \times Treated_i^m + \delta_i + \delta_t + \epsilon_{it} \tag{4}$$

where $Treated_i^{High}$, for example, is an indicator of whether the artist belongs to the high-quality group. If better artists are withholding more, we should expect to see $\beta^{High} < \beta^{Median} < \beta^{Low} < 0$.

5 Results

5.1 Decrease in the Number of Artworks Published on DeviantArt

5.1.1 Digital Artists Reduced the Number of Artworks Published

Table 2 presents the result of equation (1). In the main analysis, I employ both Poisson Pseudo Maximum Likelihood (PPML) estimation and OLS estimation. Given that the dependent variable, the number of artworks published by a given artist in a month, is a count number that is non-negative, skewed, and contains approximately 70% zeros, Chen and Roth (2024) points out that average treatment effect can be estimated consistently using Poisson regressions in this context. There are primarily two benefits of using Poisson regressions than taking logarithm of the dependent variables (e.g., log(1+Y)). First, the estimation does not depend on the units of the dependent variable. Second, the coefficient expresses the average treatment effect in levels as a percentage and can be easily interpreted. Poisson regressions can be estimated using the Pseudo Maximum Likelihood method described in Gourieroux, Monfort, and Trognon (1984) and Silva and Tenreyro (2006). The percentage change of the dependent variable is calculated using $e^{\beta} - 1$.

This analysis specifically focuses on non-AI artists and compares the publication behaviors of digital artists with those of artisan crafts artists. Column (1) shows that, compared to artisan crafts artists, digital artists publish $e^{-0.24} - 1 = -21\%$ of artworks on DeviantArt after the shock. Column (2) shows a similar scale of reduction, at 12%. Note that this is likely to be an underestimation of the true effect. By the time I started collecting the data, some artists have already deactivated their accounts. Figure 5 shows the pre-trend of column (1).

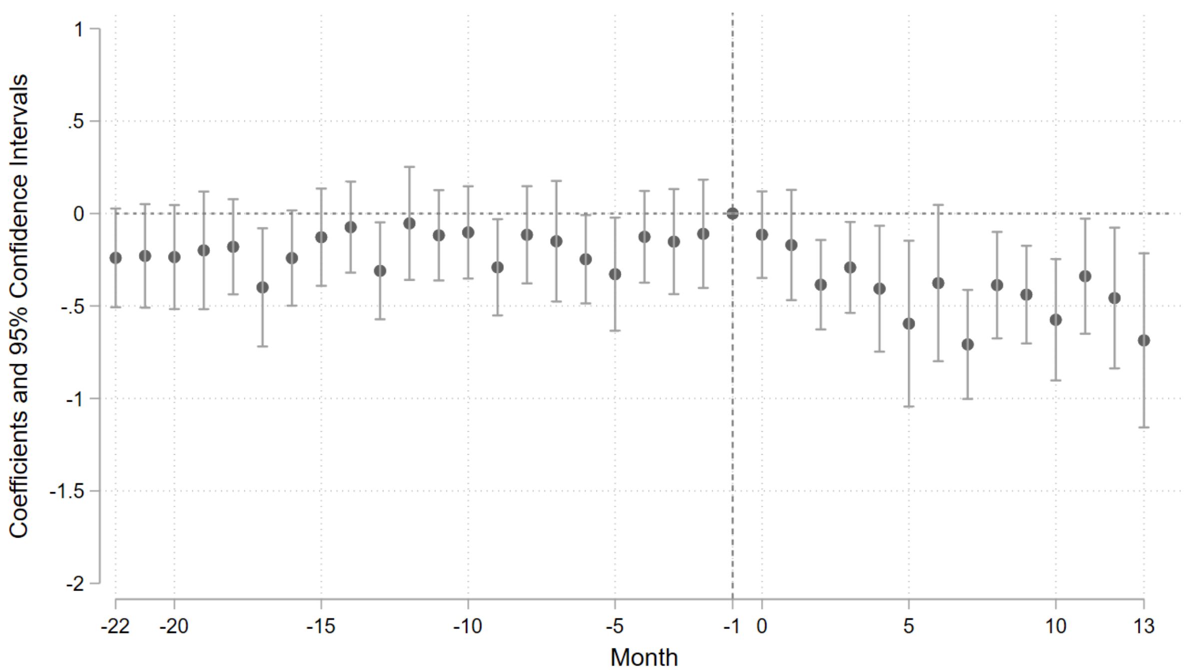
Columns (3) to (6) serve as robustness checks for the baseline estimations. The behavior of outliers can significantly impact the results of linear regression. Columns (3) and (4) apply winsorization to the dependent variable at the 99th percentile, replacing values higher than this threshold with the 99th percentile value itself. Columns (5) and (6) exclude artists

Table 2: Effect on Number of Artworks Published on DeviantArt

	Baseline Estimation		Winsorize 99% of Dep Var		Drop 1% Largest SD. Artists	
	(1)	(2)	(3)	(4)	(5)	(6)
	PPML	OLS	PPML	OLS	PPML	OLS
$Post_t \times Treated_i$	-0.24***	-0.17	-0.15**	-0.13*	-0.21**	-0.24**
	(0.09)	(0.12)	(0.06)	(0.07)	(0.09)	(0.10)
Pre-Treatment Mean		1.43		1.27		1.40
Implied %Change	-21%	-12%	-14%	-10%	-19%	-17%
Artist FE	Y	Y	Y	Y	Y	Y
Month FE	Y	Y	Y	Y	Y	Y
N(Artist-Month)	178,092	178,092	178,092	178,092	176,616	176,616
N(Artist)	4,947	4,947	4,947	4,947	4,906	4,906
R^2		0.57		0.51		0.39
Pseudo R^2	0.52		0.45		0.48	

Notes: *** denotes significance at 1 percent, ** at 5 percent, and * at 10 percent. Standard errors are clustered at artist level. To implement Poisson estimations with high-dimensional fixed effects, I employed method described in Correia, Guimarães, and Zylkin (2020).

Figure 4: Pre-Trend of Difference-in-Differences Analysis



Notes: This figure shows the pre-trend of column (1) in Table 2. They are the estimates of β_t with a 95% confidence interval in the OLS regression $Artwork_{it} = \sum_t \beta_t Treated_i \times Month_t + \delta_i + \delta_t + \epsilon_{it}$. Pre-trends of other columns are in the appendix.

Table 3: Effect on Number of Artworks Published at Extensive Margin

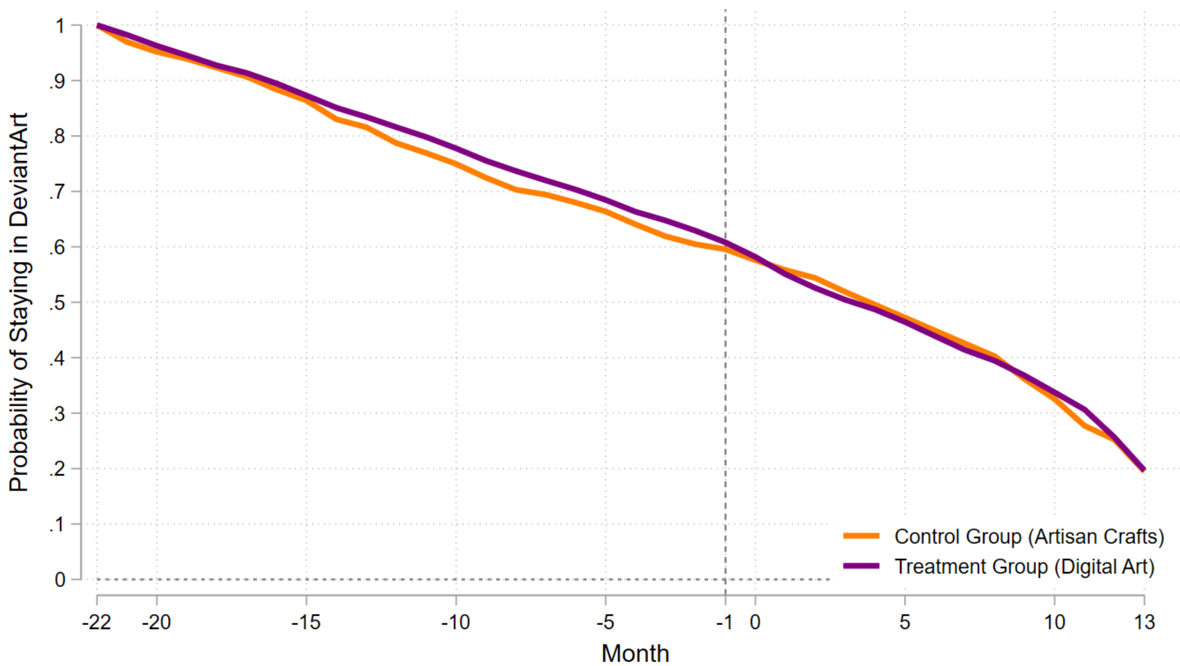
	(1) Linear Probability Model	(2) Logit
$Post_t \times Treated_i$	-0.02	-32.82
	(0.01)	(36453.91)
Artist FE	Y	Y
Month FE	Y	Y
N(Artist-Month)	178,092	143,100
N(Artist)	4,931	3,966
R^2	0.66	
Pseudo R^2		0.46

Notes: *** denotes significance at 1 percent, ** at 5 percent, and * at 10 percent. Standard errors are clustered at artist level. In logit model, 972 artists are omitted because they never exit. Therefore, the dependent variable is always positive for them.

exhibiting the largest 1% standard deviation in the dependent variable, based on panel data from January 2020 to December 2020. It is important to note that the estimations utilize panel data from January 2021 to December 2023, meaning the sample selection is based on behavior prior to the estimation period.

Table 3 shows the extensive margin of the reduction in the number of artworks published.. The specification is still equation (1), except the dependent variable is a dummy variable $stay_{it}$, where one represents a given artist has at least one publication from month t till December 2023. Figure 4 illustrates the time trend of the probability of staying in DeviantArt by group, with both the control group and the treatment group displaying very similar patterns. The results suggest that, despite reducing the number of artworks published on DeviantArt, artists remain active on the platform. One explanation is that artists still want to signal their ability on DeviantArt. Another possibility is that those who chose to exit DeviantArt had already deactivated their accounts before the data collection date, which was ten months after the shock. It's likely that more digital artists left the platform than artisan craft artists, but this cannot be detected in my sample.

Figure 5: Time Trend of Probability of Staying



Notes: This figure shows the probability of staying in DeviantArt over time. An artist is defined as staying if he publishes at least one artworks in the subsequent months.

Table 4: Effect on Number of Artworks Published on Instagram

(a) Artworks on DeviantArt

	Baseline Estimation		Winsorize 99% of Dep Var		Drop 1% Largest SD. Artists	
	(1)	(2)	(3)	(4)	(5)	(6)
	PPML	OLS	PPML	OLS	PPML	OLS
$Post_t \times Treated_j$	-0.27*	-0.26	-0.21*	-0.19	-0.39***	-0.44**
	(0.14)	(0.20)	(0.11)	(0.14)	(0.13)	(0.17)
Pre-Treatment Mean		1.60		1.43		1.55
Implied %Change	-24%	-16%	-19%	-13%	-32%	-28%
Artist FE	Y	Y	Y	Y	Y	Y
Month FE	Y	Y	Y	Y	Y	Y
N(Artist-Month)	52,812	52,812	52,812	52,812	52,308	52,308
N(Artists)	1,467	1,467	1,467	1,467	1,453	1,453
R^2		0.35		0.49		0.35
Pseudo R^2	0.44		0.41		0.43	

(b) Artworks on Instagram

	Baseline Estimation		Winsorize 99% of Dep Var		Drop 1% Largest SD. Artists	
	(1)	(2)	(3)	(4)	(5)	(6)
	PPML	OLS	PPML	OLS	PPML	OLS
$Post_t \times Treated_j$	-0.06	0.14	-0.04	0.21	-0.07	0.06
	(0.07)	(0.36)	(0.06)	(0.32)	(0.07)	(0.33)
Pre-Treatment Mean		3.34		3.14		3.15
Implied %Change	-6%	4%	-4%	7%	-7%	2%
Artist FE	Y	Y	Y	Y	Y	Y
Month FE	Y	Y	Y	Y	Y	Y
N(Artist-Month)	52,812	52,812	52,812	52,812	52,128	52,128
N(Artist)	1,467	1,467	1,467	1,467	1,448	1,448
R^2		0.57		0.55		0.50
Pseudo R^2	0.46		0.41		0.42	

Notes: *** denotes significance at 1 percent, ** at 5 percent, and * at 10 percent. Standard errors are clustered at artist level. Pre-trends of this table are in appendix. There are 5 more artists in column (5) and (6) in panel (b) than in panel (a) because these 5 artists do not have publication from 2020-01-01 to 2021-01-01 on DeviantArt. But they have volatile publication record on Instagram from 2020-01-01 to 2021-01-01.

5.1.2 Reduction in the Number of Artworks Published on DeviantArt, not Reduction in Production

In the section above, I show that artists reduce the number of artworks published on DeviantArt after the introduction of DreamUp. But here I show that artists only reduce on DeviantArt, not on Instagram. Table 4 only use the sub-sample of non-AI artists with a professional or business account on Instagram. The dependent variable for panel (a) is the number of artworks published on DeviantArt per month; panel (b) is the number of artworks published on Instagram per month. There is a significant decrease in the number of artworks published on DeviantArt after the shock, while publication activity appears to remain unchanged on Instagram.

These results imply that the multi-homing artists only publish fewer artworks on DeviantArt, but not on Instagram. They are not gradually shifting towards Instagram either since the coefficients in Panel B are not positive. The underlying explanation is the copyright concern of artists towards DreamUp, which is disincentivizing artists from disclosing their artworks. Such reduction could lead to lower knowledge spillover to both AI and non-AI creators. For AI models, the size and quality of human-generated training data are crucial. In fact, AI models can collapse if they are trained recursively on AI-generated content. Shumailov et al. (2024) published on Nature suggests that it "causes irreversible defects in the resulting models". If there is significantly less human-generated content compared to AI-generated content in the future, there is a risk that the power of AI image generators diminishes. Furthermore, artists rely on viewing and learning from the works of their peers. A decrease in the volume of published artworks could therefore result in reduced opportunities for learning and inspiration, potentially hindering the knowledge spillover among non-AI artists.

5.2 Impact on Quality

5.2.1 Artists Are Not Withholding High-Quality Artworks

To examine whether artists publish fewer high-quality artworks on DeviantArt, I use the Instagram users' data. The focus is on comparing the performance of artworks that are posted on both DeviantArt and Instagram to those exclusively posted on Instagram. The measure of artwork quality is the number of likes and comments each artwork gains on Instagram.

Table 10 in appendix validates the use of likes and comments on Instagram as a quality measure. For a given artist, if an artwork has more favorites, comments, downloads, and views on DeviantArt, it will also have more likes and comments on Instagram. This implies viewers have similar tastes across platforms, and the quality measure is reliable.

Table 5 shows the results of equation (2). The positive coefficients before $Matched_j$ indicate that artists select higher quality artworks and publish them on DeviantArt. But the insignificant coefficients before $Post_t \times Matched_j$ imply that such quality difference does not significantly change after the introduction of DreamUp. The results suggest that for a given artist, he is not withholding higher quality artworks after the shock.

5.2.2 Better Artists Do Not Withhold More

Although there's no evidence suggesting artists are withholding high-quality artworks on DeviantArt, if better artists withhold more artworks, this reduction can still lead to an overall decreased in the quality of disclosed knowledge. To understand whether better artists tend to reduce their number of artworks published more, I obtain artists fixed effects using non-AI digital artists described in equation (3). Artists are then devided evenly into high-quality, median-quality and low-quality groups. Table 6 shows the results of estimation of equation 4.

Artists of different quality reduce the number of artworks they publish at a very similar

Table 5: Performance Comparison of Artworks Only on Instagram and on Both Platforms

Dep Var	$Likes^{Ins}$		$Comments^{Ins}$	
	(1)	(2)	(3)	(4)
	PPML	OLS	PPML	OLS
$Post_t \times Matched_j$	0.07	125.19	-0.03	-0.30
	(0.10)	(215.51)	(0.06)	(0.93)
$Matched_j$	0.09*	184.43*	0.11***	1.63***
	(0.05)	(111.12)	(0.03)	(0.50)
Pre-Treatment Mean		2237		16
Implied %Change of $Post_t \times Matched_j$	7%	6%	-3%	-2%
Artist FE	Y	Y	Y	Y
Month FE	Y	Y	Y	Y
N(Artwork)	145,202	145,280	145,265	145,280
R^2		0.43		0.30
Pseudo R^2	0.77		0.62	

Notes: *** denotes significance at 1 percent, ** at 5 percent, and * at 10 percent. Standard errors are clustered at artist level. Only use the matched artworks across DeviantArt and Instagram. They are from non-AI artists who have a professional or business Instagram account, including only artists specialized in digital art. Pre-Treatment Mean is the average number of likes or comments per artwork on Instagram before the introduction of DreamUp, including both matched artworks and unmatched artworks. In column (1) and (3), some observations are dropped because of separation. For a given artist, if the dependent variable is always 0, this artist is dropped.

scale, and the differences are not statistically significant. This suggests that better artists are not withholding more artworks than others. When using the number of views or comments as a quality measure, the results seem to vary depending on the model employed. However, it is important to note that the number of views and comments are ambiguous measures of quality compared to the number of downloads and favorites. For instance, the number of views can increase simply if a consumer scrolls down the screen without clicking on the artwork. Similarly, the number of comments can be high if the artwork is controversial, regardless of its quality.

Table 5 and Table 6 jointly suggest that the overall quality of artworks published by non-AI artists remains unchanged after the introduction of DreamUp. Despite withholding artworks from DeviantArt, artists are not selectively withholding high-quality artworks, and better artists do not appear to withhold more than others. One explanation for this behavior is that artists still have incentives to signal their ability or obtain recognition from DeviantArt.

Table 6: Similar Effects Between Artists of Different Quality

Dep Var	$Artwork_{it}$							
Quality Measured By	Downloads		Favorites		Views		Comments	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	PPML	OLS	PPML	OLS	PPML	OLS	PPML	OLS
$Post_t \times Treated_i^{High}$	-0.24**	-0.28**	-0.25***	-0.27**	-0.22**	-0.23*	-0.18**	-0.23*
	(0.10)	(0.13)	(0.09)	(0.12)	(0.09)	(0.12)	(0.09)	(0.12)
$Post_t \times Treated_i^{Median}$	-0.32***	-0.40***	-0.23**	-0.26**	-0.27***	-0.21*	-0.31***	-0.41***
	(0.10)	(0.13)	(0.09)	(0.13)	(0.10)	(0.13)	(0.10)	(0.14)
$Post_t \times Treated_i^{Low}$	-0.17*	-0.18	-0.25**	-0.33**	-0.24**	-0.42***	-0.24**	-0.22
	(0.10)	(0.14)	(0.11)	(0.15)	(0.10)	(0.15)	(0.11)	(0.14)
Artist FE	Y	Y	Y	Y	Y	Y	Y	Y
Month FE	Y	Y	Y	Y	Y	Y	Y	Y
N(Artist-Month)	140,364	140,364	140,364	140,364	140,364	140,364	140,364	140,364
R^2		0.62		0.62		0.62		0.62
Pseudo R^2	0.52		0.52		0.52		0.52	

Notes: *** denotes significance at 1 percent, ** at 5 percent, and * at 10 percent. Standard errors are clustered at artist level. Dependent variable is number of artworks published by artist i in month t. Quality is measured using artworks between 2020-01-01 and 2021-01-01. 1,048 digital artists are dropped compared to Table 2. Among the dropped artists, 246 of them have only one artwork during the measure period; 801 of them only have artworks before 2020-01-01 and after 2021-01-01; 1 artist does not have any artwork before 2021-01-01.

5.3 Insignificant Change in Attention

One potential reason for the reduction in the number of artworks published by non-AI artists may be the reduced attention from audiences due to competition from AI-generated artworks. The introduction of DreamUp reduced production cost of visual art, which stimulate an increase in the publication of artworks on DeviantArt. This significant increase of publication can divert consumer attention away from the incumbent non-AI artists, making artworks by non-AI artists less likely to be seen.

However, Table 7 suggests that the reduction in the number of artworks published is unlikely to be caused by AI artists diverting consumer attention away from non-AI artists. In this table, I examine whether an artwork is less likely to be seen or downloaded if it's published by a digital artist compared to an artisan crafts artist after the introduction of DreamUp. The comparison is made across different time periods for artworks by the same artist. Section 5.2.1 demonstrates that artists are not selectively withholding high-quality artworks on DeviantArt following the shock. Holding the quality of artworks published by a given digital artist unchanged, changes in the number of views and downloads reflect shifts in consumer attention to that artist over time.

Columns (1) to (4) in panel (a) show that the number of views and downloads per artwork of digital artists did not decline significantly relative to artisan crafts artists. Artworks published after the shock are equally likely to be seen and downloaded by consumers. Columns (6) and (8) show significantly negative coefficients, primarily driven by a small fraction of superstar artworks. Given the skewed distribution of the number of favorites and comments, and considering they are count variables, more attention should be paid to the results of PPML estimation. Panel (b) shows that after the shock, the likelihood of an artwork being downloaded or favorited remains unchanged, conditional on being seen.

6 AI Adopters Behavior Change

6.1 Incumbent AI Artists Publish More Following the Introduction of DreamUp

Although copyright concerns regarding generative AI's training data disincentivize artists from disclosing more artworks on DeviantArt, AI image generators increase the publication frequency of adopters. Figure 12 panel (a) in the appendix shows the time trend of AI digital artists and non-AI artisan crafts artists. AI adopters are systematically different from non-adopters. Therefore, using non-AI artisan crafts artists as a control group would violate the parallel trend assumption. To address this issue, I employ propensity-score matching difference-in-differences. One way to think about the propensity score matching is by imagining a counterfactual world where new technology, like 3D printing, allows artists to produce artisan crafts at an extremely low cost. In this scenario, the matched artisan crafts artists are more likely to adopt this technology. Table 8 presents the results of baseline estimation and the propensity score matching diff-in-diff proposed by Rosenbaum and Rubin (1983) and described in Caliendo and Kopeinig (2008). To ensure that variables used to calculate

Table 7: Effects on Attention

(a) Views, Downloads, Favorites, and Comments

Dep Var	Views		Downloads		Favorites		Comments	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	PPML	OLS	PPML	OLS	PPML	OLS	PPML	OLS
$Post_t \times Treated_i$	0.02	-720.44	1.10	-2.10	-0.06	-43.91***	-0.04	-0.78***
	(0.15)	(1537.54)	(0.89)	(8.56)	(0.09)	(9.85)	(0.08)	(0.23)
Pre-Treatment Mean		18629		19		196		6
Implied %Change	2%	-4%	200%	-11%	-0.06%	-22%	-0.04%	-13%
Artist FE	Y	Y	Y	Y	Y	Y	Y	Y
Month FE	Y	Y	Y	Y	Y	Y	Y	Y
N(Artwork)	233,482	233,482	157,028	233,482	233,459	233,482	231,061	233,482
R^2		0.33		0.47		0.74		0.43
Pseudo R^2	0.59		0.81		0.82		0.52	

(b) Probability of Download, Favorite, Comment Conditional on Being Seen

Dep Var	(1) $\frac{Downloads}{Views}$	(2) $\frac{Favorites}{Views}$	(3) $\frac{Comments}{Views}$
$Post_t \times Treated_i$	-0.82	-11.60	29.05*
	(3.50)	(13.07)	(17.17)
Pre-Treatment Mean	17.59	243.80	13.37
Artist FE	Y	Y	Y
Month FE	Y	Y	Y
N(Artwork)	233,482	233,482	233,482
R^2	0.23	0.54	0.16

Notes: *** denotes significance at 1 percent, ** at 5 percent, and * at 10 percent. Standard errors are clustered at artist level.

In panel (a) column (3)(5)(7), some observations are dropped because of separation. For a given artist, if the dependent variable is always 0, this artist is dropped. The result of column (3), for example, shows that among the artists whose artworks have been downloaded at least once, the number of downloads per artwork does not change after the shock. And those whose artworks have never been downloaded, the effect on number of downloads is zero. Therefore, dropping these artists does not change our understanding of the effect on number of downloads.

Panel (b) shows the results of linear regressions. Dependent variable is multiplied by 10,000 for table readability.

Table 8: AI Artists Publication Change

	Baseline Estimation		Propensity Score Matching	
	(1)	(2)	(3)	(4)
	PPML	OLS	PPML	OLS
$Post_t \times Treated_i$	0.44**	1.93	0.47*	2.08
	(0.20)	(1.22)	(0.25)	(1.34)
Pre-Treatment Mean		5.26		5.37
Implied %Change	55%	37%	60%	39%
Artist FE	Y	Y	Y	Y
Month FE	Y	Y	Y	Y
N(Artist-Month)	39,528	39,528	21,528	21,528
N(Artist)	1,098	1,098	598	598
N(Artist in Control)	559	559	99	99
N(Artist in Treatment)	539	539	499	499
R^2		0.33		0.27
Pseudo R^2	0.65		0.63	

Notes: *** denotes significance at 1 percent, ** at 5 percent, and * at 10 percent. Standard errors are clustered at artist level. Single nearest neighbor matching is used in column (3)(4). I use artist demographic characteristics, including country, gender and age, to calculate the propensity score.

propensity score are unaffected by the treatment, I use artist demographic characteristics, including country, gender, and age. The results show that AI adopters increase the number of artworks they publish by 55% to 60% after the shock. Pr-trends are in Figure 12 in the appendix.

6.2 Few but Very Productive Entrants

Panel (c) from Figure 6 shows that there is no significant increase in number of entrants each month after the introduction of DreamUp. However, panel (d) indicates that entrants who adopted AI and joined after the shock published significantly more frequently than those who joined before the shock. For example, artists who joined DeviantArt in August 2023 published around 1,500 artworks per month on average within their first three months. I limit the analysis to the first three months to ensure that I do not include more months for earlier entrants, avoiding capturing potential life cycle patterns of artists' publication behavior.

Figure 7 shows a similar pattern from another sample on DeviantArt, specifically the

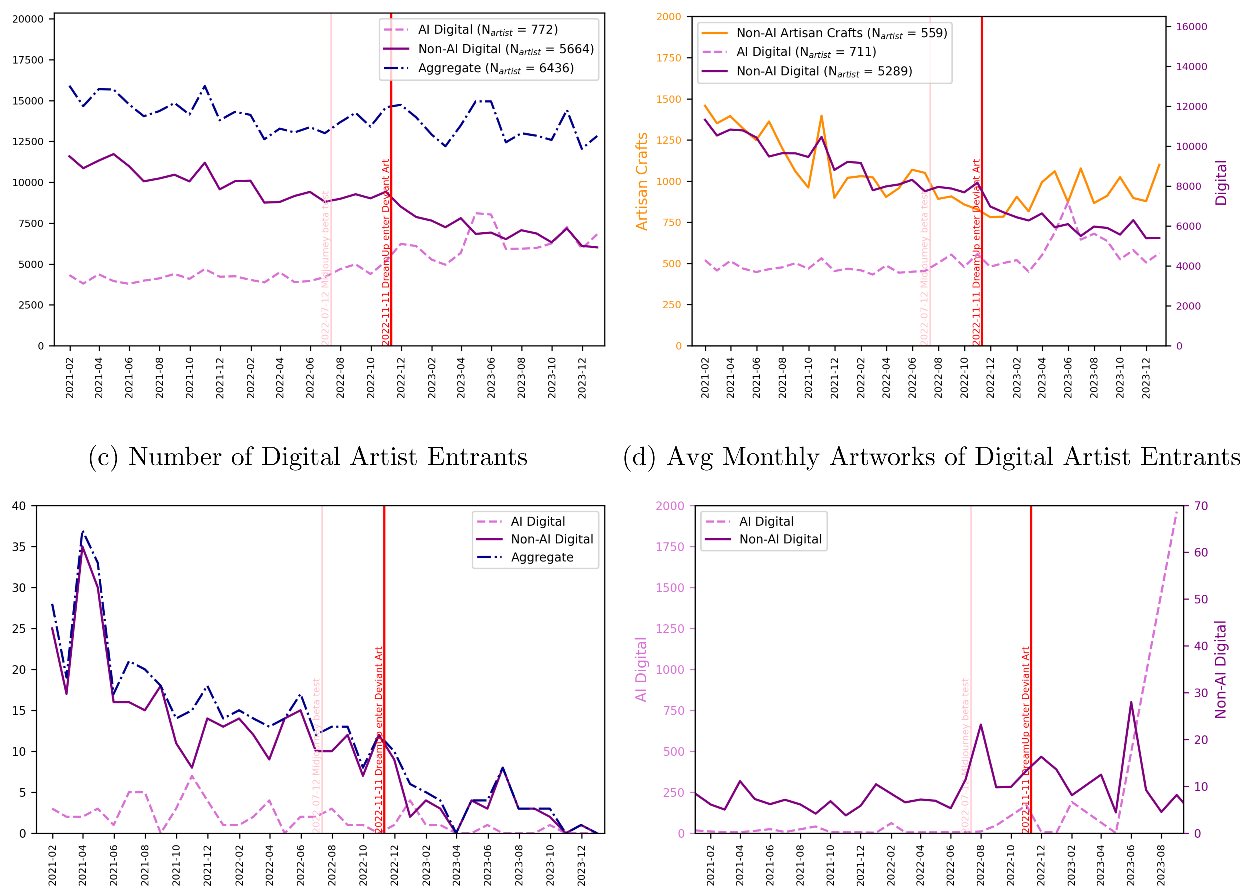
Figure 6: AI Adopters in Daily Deviation Section Time Trends

(a) Aggregate Artworks of All Digital Artists

(b) Aggregate Artworks of Incumbents

(c) Number of Digital Artist Entrants

(d) Avg Monthly Artworks of Digital Artist Entrants



Notes: Panel (b) uses artists entered DeviantArt before 2021-01-01. Artisan Crafts artists are from daily deviation section in table 2. They are artisan crafts artists enter the daily deviation section before 2021-01-01.
Panel (d) shows the average monthly publications by entrants within the first three months.

"Most Popular Artists in This Century" section. Panel (a) suggests that the number of artworks published increased significantly after the shock, primarily driven by AI artists. This increase is mainly driven by AI artists. Panel (b) shows that for artists who joined the platform before January 1, 2021, non-AI digital artists reduced the number of artworks they published compared to artisan craft artists, while AI adopters increased the number of artworks they published. Panel (c) indicates that the number of entrants does not increase dramatically after the shock. Instead, AI adopters who entered after the shock published hundreds of artworks per month within the first three months of their entry, as shown in panel (d). The significant increase in the number of artworks published in panel (a) is mainly driven by these productive entrants.

Are these artworks published by AI-adopting entrants after the shock as good as those published by incumbents? Figure 8 plots the coefficient of $\mathbf{1}(entry\ month_i = m)$ of the following equation:

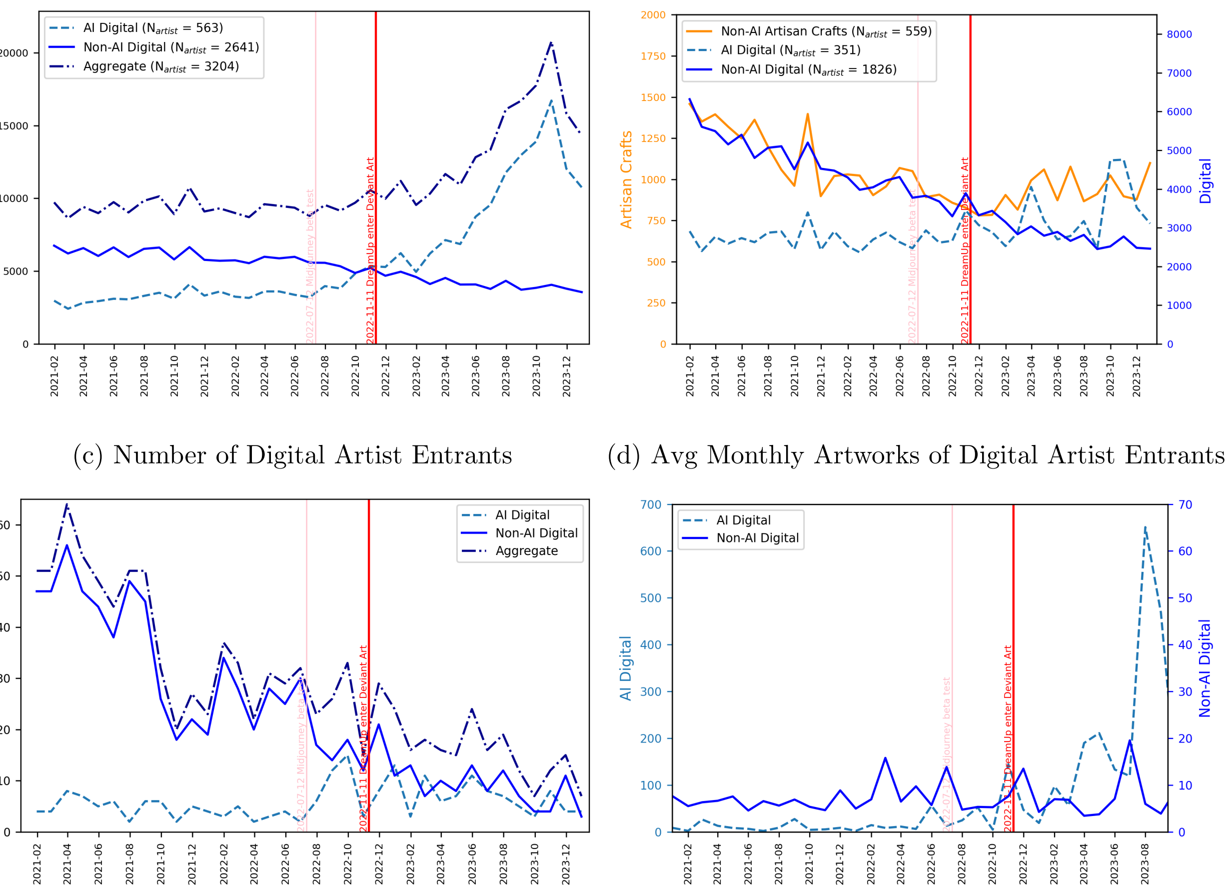
$$Downloads_{ijt} = \sum_m \beta_m \mathbf{1}(entry\ month_i = m) + \delta_t + \epsilon_{ijt} \tag{5}$$

where coefficient β_m represents the average number of downloads of artworks published by artists entered in month m. The figure indicates that artists who entered after the shock and publish hundreds of artworks per month seem to receive fewer downloads than previous entrants.

7 Conclusion

This paper studies the impact of AI introduction on online art platforms, focusing on both the number of artworks published and the quality of the artworks. I use a difference-in-differences approach to show that compared to non-AI artisan artists, non-AI digital artists reduced the number of artworks they publish by over 21% each month after the introduction of DreamUp on DeviantArt. By looking into the multi-homing artists with a professional

Figure 7: AI Adopters in Most Popular Artists Section Time Trends



(a) Aggregate Artworks of All Digital Artists

(b) Aggregate Artworks of Incumbents

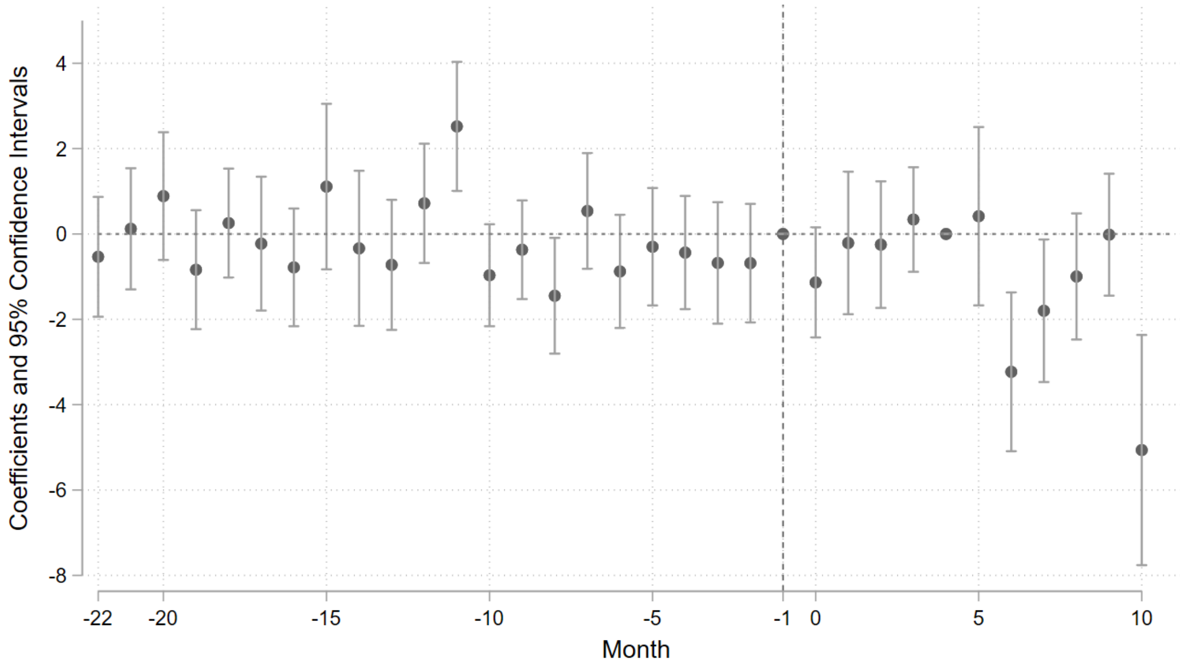
(c) Number of Digital Artist Entrants

(d) Avg Monthly Artworks of Digital Artist Entrants

Notes: Panel (b) uses digital artists entered DeviantArt before 2021-01-01. Artisan Crafts artists are from daily deviation section in table 2. They are artisan crafts artists enter the daily deviation section before 2021-01-01. I did not use artisan crafts artists from "Most Popular Artists" section because there are only 54 such artists.

Panel (d) shows the average monthly publications by entrants within the first three months.

Figure 8: Downloads of Entrants



or business account on Instagram, I show that digital artists only publish fewer artworks on DeviantArt, not on Instagram. This is consistent with the fact that the introduction of DreamUp implies weak copyright protection on DeviantArt, and artists are concerned about their artwork being included in AI training data without their consent.

By comparing artworks only displayed on Instagram to those also on DeviantArt of multi-homing digital artists, I show that this reduction of publication is not biased towards high-quality artworks or low-quality artworks on DeviantArt. There is no evidence suggests that artists are withholding high quality artworks from DeviantArt after the introduction of DreamUp. These findings highlight the potential disincentive for content creators to publish in an AI-dominated era due to weak copyright protection. However, the quality of published knowledge does not appear to decline.

This paper is not suggesting to ban generative AI. Instead, the question is how to benefit from this technology without inducing too much creative work withholding. A natural idea would be to create a market between artists and AI companies where artists can decide whether to sell the artworks to AI companies to train their models. This raises important follow-up questions: If there exists a market between innovators and AI companies, what should the price scheme look like? Another open question is whether this decrease in the number of artworks published affects future productivity and, if so, by how much. However, such a market will not be able to exist if intellectual property rights are not protected in the AI training data.

8 Appendix

8.1 Summary Statistics of Artists

Table 9 shows the summary statistics of artists.

8.2 Pre-trend of Table 2

Figure 9 shows pretrends of Table 2.

8.3 Pre-trend of Table 4

Figure 10 shows pretrends of Table 4.

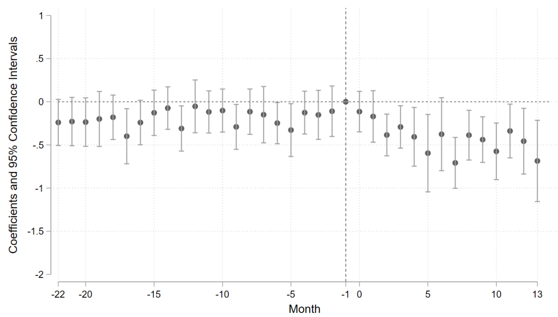
8.4 High Performance Correlation Across Platforms

8.5 Price and Revenue

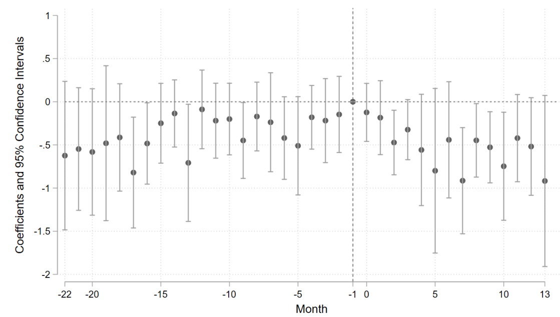
8.6 AI Adopters Time Trends

Figure 9: Pre-Trends of Table 2

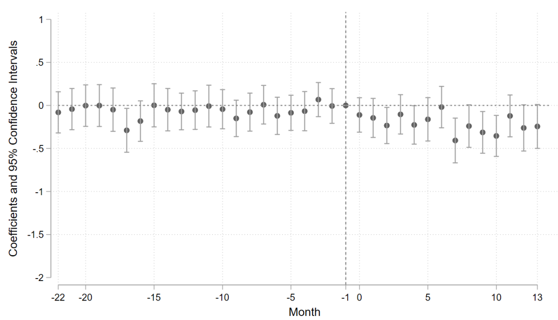
(a) Baseline Estimation PPML



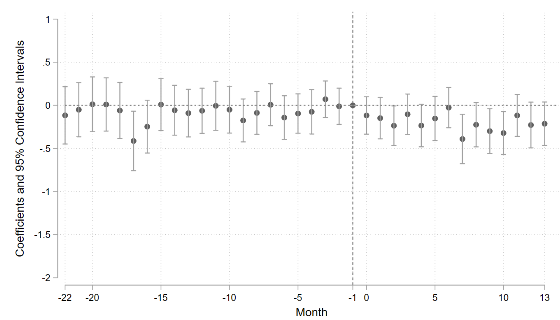
(b) Baseline Estimation OLS



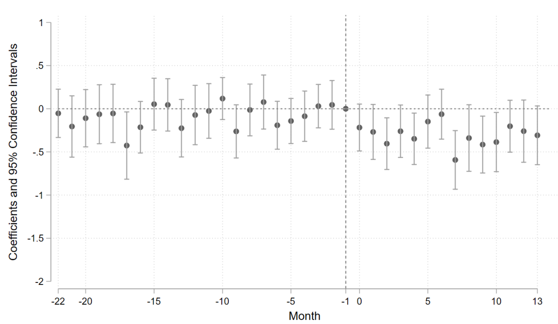
(c) Winsorize PPML



(d) Winsorize OLS



(e) Drop 1% largest SD PPML



(f) Drop 1% largest SD OLS

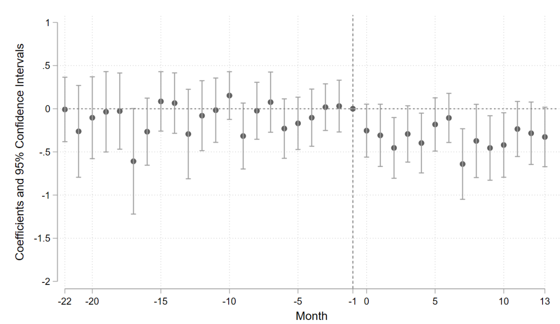
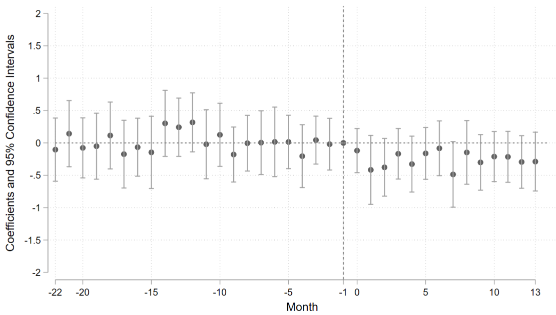
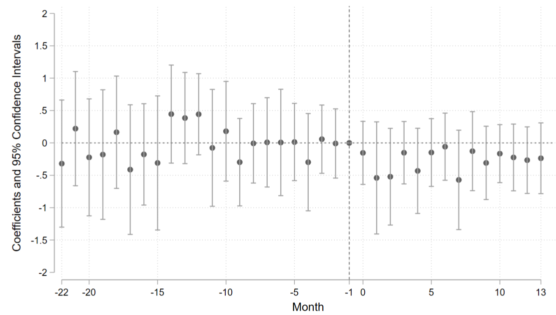


Figure 10: Pre-Trends of Table 4 Panel A: Artworks on DeviantArt

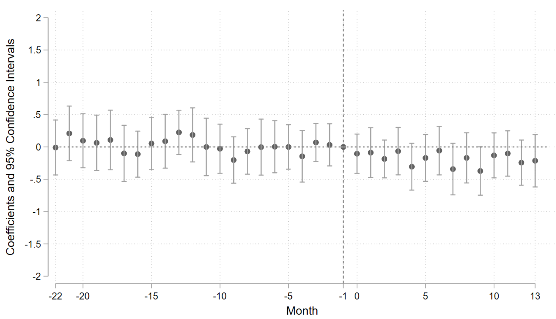
(a) Baseline Estimation PPML



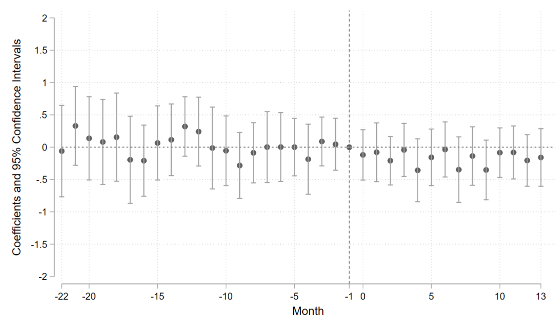
(b) Baseline Estimation OLS



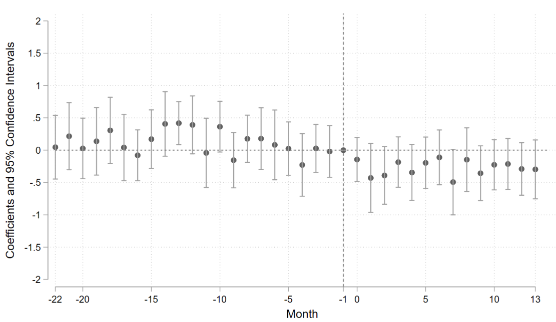
(c) Winsorize PPML



(d) Winsorize OLS



(e) Drop 1% largest SD PPML



(f) Drop 1% largest SD OLS

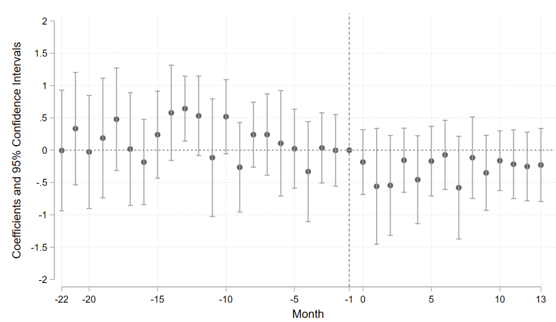
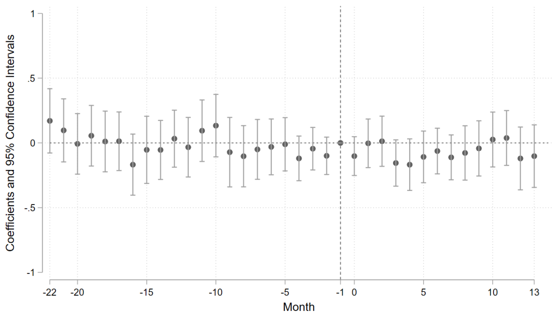
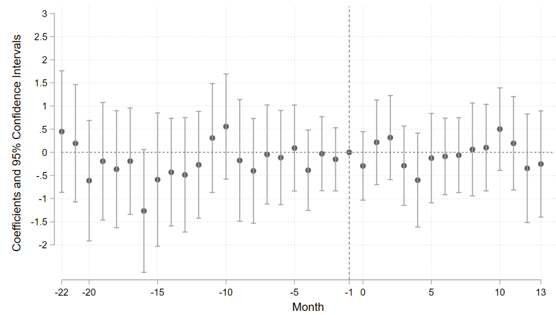


Figure 11: Pre-Trends of Table 4 Panel B: Artworks on Instagram

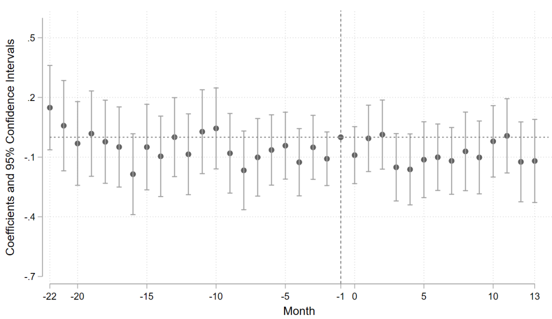
(a) Baseline Estimation PPML



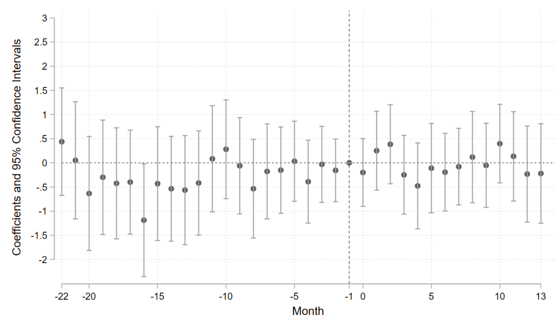
(b) Baseline Estimation OLS



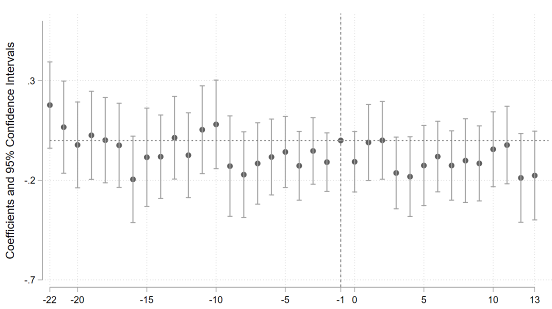
(c) Winsorize PPML



(d) Winsorize OLS



(e) Drop 1% largest SD PPML



(f) Drop 1% largest SD OLS

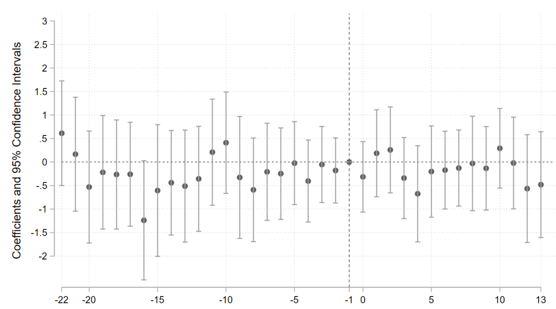
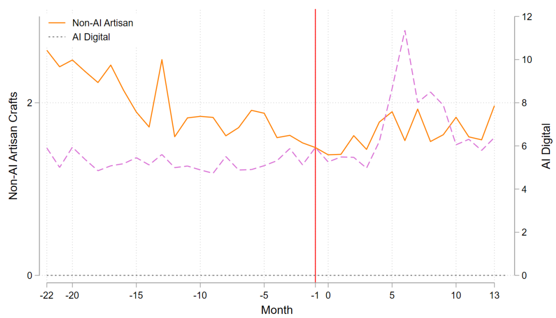
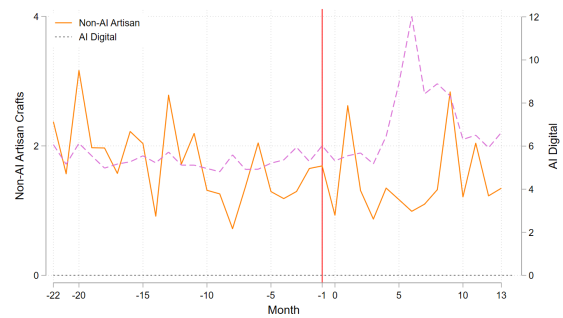


Figure 12: AI Adopters Time Trends

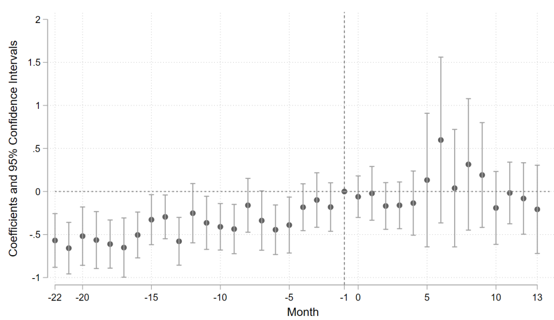
(a) Before Propensity Score Matching



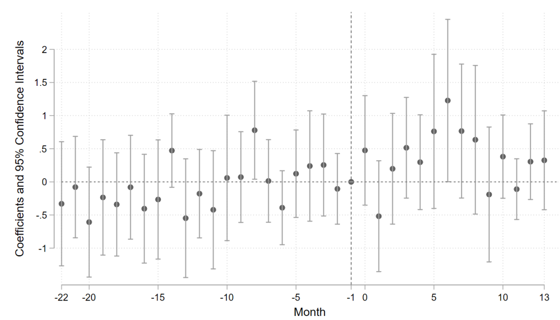
(b) After Propensity Score Matching



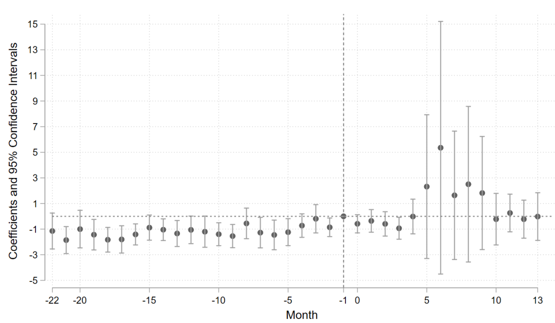
(c) Baseline Estimation PPML



(d) Propensity Score Matching PPML



(e) Baseline Estimation OLS



(f) Propensity Score Matching PPML OLS

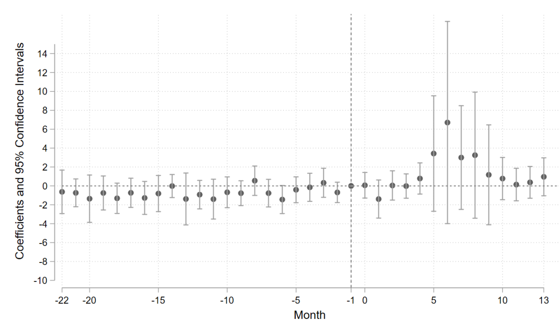


Table 9: Summary Statistics

(a) Panel A: Artists Used in Difference-in-Differences

	Artisan Crafts Artists				Digital Art Artists			
	Mean	Min	Max	sd	Mean	Min	Max	sd
Monthly Pre-Period Artworks	1.97	0	369	11	1.43	0	474	5.00
Profile Pageviews	1.05e+05	639	3.25e+06	2.30e+05	2.81e+05	799	5.38e+07	1.14e+06
Followers	1977	11	6.31e+04	4727	6613	5.00	6.76e+05	1.96e+04
$Views^{DA}$ per Artwork	5586	15	1.09e+06	4.04e+04	1.80e+04	14	5.49e+06	7.42e+04
$Downloads^{DA}$ per Artwork	5.85	0	1253	34	16	0	2.25e+04	102
$Favourites^{DA}$ per Artwork	39	0	2486	89	182	0	1.09e+04	400
$Comments^{DA}$ per Artwork	2.42	0	151	6.50	6.07	0	2887	14
N(Artist)	559				4388			

(b) Artists Used in Instagram Difference-in-Differences

	Artisan Crafts Artists				Digital Art Artists			
	Mean	Min	Max	sd	Mean	Min	Max	sd
Monthly Pre-Period Artworks	1.75	0	67	4.87	1.60	0	271	4.86
Profile Pageviews	1.39e+05	1668	3.25e+06	3.35e+05	3.35e+05	799	1.04e+07	9.22e+05
Followers	2889	24	6.31e+04	7203	9252	33	3.43e+05	2.35e+04
$Views^{DA}$ per Artwork	1.19e+04	35	1.09e+06	5.42e+04	2.79e+04	19	1.64e+06	9.63e+04
$Downloads^{DA}$ per Artwork	19	0	955	60	20	0	7042	116
$Favourites^{DA}$ per Artwork	91	0	2486	144	256	0	1.09e+04	503
$Comments^{DA}$ per Artwork	2.51	0	90	4.59	7.82	0	373	16
$Likes^{Ins}$ per Artwork	800	0	2.23e+05	4362	2024	0	1.29e+06	9114
$Comments^{Ins}$ per Artwork	13	0	1.17e+04	73	15	0	7906	67
N(Artist)	170				1297			

Notes: Use panel from January 2021 to December 2023.

Table 10: High Performance Correlation Across Platforms

Dep Var	$Likes^{Ins}$					$Comments^{Ins}$				
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
$Favorites^{DA}$	2.765***					0.014***				
	(0.420)					(0.002)				
$Comments^{DA}$		61.213***					0.535***			
		(13.099)					(0.074)			
$Downloads^{DA}$			1.573**					0.010*		
			(0.789)					(0.005)		
$Views^{DA}$				0.005***					0.000***	
				(0.001)					(0.000)	
$Price^{DA}$					0.052					0.000
					(0.036)					(0.000)
Artists FE	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
Month FE	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
N(Artwork)	35,655	35,655	35,655	35,655	727	35,655	35,655	35,655	35,655	727
R^2	0.48	0.47	0.47	0.47	0.73	0.47	0.48	0.46	0.46	0.62

Notes: *** denotes significance at 1 percent, ** at 5 percent, and * at 10 percent. Standard errors are clustered at artist level. Only use the matched artworks across DeviantArt and Instagram. They are from non-AI artists who have a professional or business Instagram account, including artists specialized in digital art and artisan crafts. Column (5) and (10) use only premium download artworks. Premium downloads are high-resolution versions of artworks that consumers can access by paying a fixed price. Without payment, only low-resolution versions can be viewed. Sales can be measured by the number of downloads for each artwork.

Table 11: Price Change

Dep Var	Price	
	(1)	(2)
	$\ln(Y)$	Y
$Post_t \times Treated_i$	0.09	112.96
	(0.07)	(111.53)
Pre-Treatment Mean		16
Artist FE	Y	Y
Month FE	Y	Y
N(Artwork)	4,591	4,591
N(Artist)	252	252
R^2	0.83	0.33

Notes: *** denotes significance at 1 percent, ** at 5 percent, and * at 10 percent. Standard errors are clustered at the artist level. I only use premium downloads. Premium downloads are high-resolution versions of artworks that consumers can access by paying a fixed price. Without payment, only low-resolution versions can be viewed.

References

Aguiar, L., & Martens, B. (2016). Digital music consumption on the Internet: Evidence from clickstream data. *Information Economics and Policy*, *34*, 27–43. doi:10.1016/j.infoecopol.2016.01.003

Biasi, B., & Moser, P. (2021). Effects of Copyrights on Science: Evidence from the WWII Book Republication Program. *American Economic Journal: Microeconomics*, *13*(4), 218–260. doi:10.1257/mic.20190113

Blackburn, D. (2004). On-line Piracy and Recorded Music Sales. *Unpublished manuscript*, 1–60.

Caliendo, M., & Kopeinig, S. (2008). Some Practical Guidance for the Implementation of Propensity Score Matching. *Journal of Economic Surveys*, *22*(1), 31–72. _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1111/j.1467-6419.2007.00527.x. doi:10.1111/j.1467-6419.2007.00527.x

Chen, J., & Roth, J. (2024). Logs with Zeros? Some Problems and Solutions*. *The Quarterly Journal of Economics*, *139*(2), 891–936. doi:10.1093/qje/qjad054

Correia, S., Guimarães, P., & Zylkin, T. (2020). Ppmlhdfe: Fast Poisson Estimation with High-Dimensional Fixed Effects. *The Stata Journal: Promoting communications on statistics and Stata*, *20*(1), 95–115. arXiv:1903.01690 [econ]. doi:10.1177/1536867X20909691

Galasso, A., & Schankerman, M. (2015). Patents and Cumulative Innovation: Causal Evidence from the Courts *. *The Quarterly Journal of Economics*, *130*(1), 317–369. doi:10.1093/qje/qju029

Gans, J. S. (2024, April). Copyright Policy Options for Generative Artificial Intelligence. SSRN Scholarly Paper. doi:10.2139/ssrn.4707911

Giorcelli, M., & Moser, P. (2020). Copyrights and Creativity: Evidence from Italian Opera in the Napoleonic Age. *Journal of Political Economy*, *128*(11), 4163–4210. Publisher: The University of Chicago Press. doi:10.1086/710534

Givon, M., Mahajan, V., & Muller, E. (1995). Software Piracy: Estimation of Lost Sales and the Impact on Software Diffusion. *Journal of Marketing*, *59*(1), 29–37. Publisher: American Marketing Association. doi:10.2307/1252012

Gourieroux, C., Monfort, A., & Trognon, A. (1984). Pseudo Maximum Likelihood Methods: Applications to Poisson Models. *Econometrica*, *52*(3), 701–720. Publisher: [Wiley, Econometric Society]. doi:10.2307/1913472

Guo, S., Jiang, X., Thornton, T., & Saunders, D. (2019, February). Approximate String Matching of Power System Substation Names. In *2019 IEEE Power and Energy Conference at Illinois (PECI)* (pp. 1–6). doi:10.1109/PECI.2019.8698905

Huang, H., Fu, R., & Ghose, A. (2023, December). Generative AI and Content Creators: Evidence from Digital Art Platforms. SSRN Scholarly Paper. doi:10.2139/ssrn.4670714

Hui, K.-L., & Png, I. (2003). Piracy and the Legitimate Demand for Recorded Music. *The B.E. Journal of Economic Analysis & Policy*, *2*(1), 1–24. Publisher: De Gruyter. Retrieved July 29, 2024, from https://econpapers.repec.org/article/bpjbejeap/v_3acontributions.2_3ay_3a2003_3ai_3a1_3an_3a11.htm

Mahmood, K., Konagurthu, A. S., Song, J., Buckle, A. M., Webb, G. I., & Whisstock, J. C. (2010). EGM: Encapsulated gene-by-gene matching to identify gene orthologs and homologous segments in genomes. *Bioinformatics*, *26*(17), 2076–2084. doi:10.1093/bioinformatics/btq339

Munkres, J. (1957). Algorithms for the Assignment and Transportation Problems. *Journal of the Society for Industrial and Applied Mathematics*, *5*(1), 32–38.

Murray, F., & Stern, S. (2007). Do formal intellectual property rights hinder the free flow of scientific knowledge? An empirical test of the anti-commons hypothesis. *Journal of Economic Behavior & Organization*, *63*(4), 648–687. Place: Netherlands Publisher: Elsevier Science. doi:10.1016/j.jebo.2006.05.017

Nagaraj, A. (2018). Does Copyright Affect Reuse? Evidence from Google Books and Wikipedia. *Management Science*, *64*(7), 3091–3107. Publisher: INFORMS. doi:10.1287/mnsc.2017.2767

Oberholzer-Gee, F., & Strumpf, K. (2007). The Effect of File Sharing on Record Sales: An Empirical Analysis. *Journal of Political Economy*, *115*(1), 1–42. Publisher: The University of Chicago Press. doi:10.1086/511995

Peukert, C., Abeillon, F., Haese, J., Kaiser, F., & Staub, A. (2024, April). Strategic Behavior and AI Training Data. arXiv:2404.18445 [cs, econ, q-fin]. doi:10.48550/arXiv.2404.18445

Peukert, C., Claussen, J., & Kretschmer, T. (2017). Piracy and box office movie revenues: Evidence from Megaupload. *International Journal of Industrial Organization*, *52*, 188–215. doi:10.1016/j.ijindorg.2016.12.006

Reimers, I. (2019). Copyright and Generic Entry in Book Publishing. *American Economic Journal: Microeconomics*, *11*(3), 257–284. Publisher: American Economic Association. Retrieved January 19, 2024, from https://www.jstor.org/stable/26754090

Rob, R., & Waldfogel, J. (2006). Piracy on the High C's: Music Downloading, Sales Displacement, and Social Welfare in a Sample of College Students. *The Journal of Law and Economics*, *49*(1), 29–62. Publisher: The University of Chicago Press. doi:10.1086/430809

Rob, R., & Waldfogel, J. (2007). Piracy on the Silver Screen. *The Journal of Industrial Economics*, *55*(3), 379–395. Publisher: Wiley. Retrieved July 29, 2024, from https://www.jstor.org/stable/4622391

Rosenbaum, P. R., & Rubin, D. B. (1983). The central role of the propensity score in observational studies for causal effects. *Biometrika*, *70*(1), 41–55. doi:10.1093/biomet/70.1.41

Shumailov, I., Shumaylov, Z., Zhao, Y., Papernot, N., Anderson, R., & Gal, Y. (2024). AI models collapse when trained on recursively generated data. *Nature*, *631*(8022), 755–759. Publisher: Nature Publishing Group. doi:10.1038/s41586-024-07566-y

Silva, J. M. C. S., & Tenreyro, S. (2006). The Log of Gravity. *The Review of Economics and Statistics*, *88*(4), 641–658. doi:10.1162/rest.88.4.641

Waldfogel, J. (2012a). Copyright Protection, Technological Change, and the Quality of New Products: Evidence from Recorded Music since Napster. *The Journal of Law & Economics*, *55*(4), 715–740. doi:10.1086/665824

Waldfogel, J. (2012b). Music Piracy and Its Effects on Demand, Supply, and Welfare. *Innovation Policy and the Economy*, *12*, 91–110. Publisher: The University of Chicago Press. doi:10.1086/663157

Williams, H. L. (2013). Intellectual Property Rights and Innovation: Evidence from the Human Genome. *Journal of Political Economy*, *121*(1), 1–27. Publisher: The University of Chicago Press. doi:10.1086/669706

Yang, S. A., & Zhang, A. H. (2024, February). Generative AI and Copyright: A Dynamic Perspective. SSRN Scholarly Paper. doi:10.2139/ssrn.4716233

Zentner, A. (2006). Measuring the Effect of File Sharing on Music Purchases. *The Journal of Law and Economics*, *49*(1), 63–90. Publisher: The University of Chicago Press. doi:10.1086/501082