From Creation to Caution: The Effect of Generative AI on Online Art Market*

Sijie Lin†

Version: April 2024

Abstract

Can generative Artificial Intelligence (AI) disincentivize innovators from disclosing information? In this paper, 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, 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 reduce the publication volume by 22% following the introduction of AI 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 reducing disclosure of 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 Joshua Gans, Siyuan Liu, Irisa Zhou, Tommaso Alba and participants in student workshop and Rotman IO group at University of Toronto for their helpful comments. 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, much of which consists of copyrighted material, leading to 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 for 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 innovation is scarce. This paper finds 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 introduction of DreamUp[1], an AI image generator, to an influential online art platform DeviantArt[2] disincentivizes artists from publishing new artworks on this platform. 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. 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. Although artists were provided the option to be "opted-out" of the training data set at the beginning, and

[1]www.deviantart.com/dreamup

[2]www.deviantart.com

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.

My research questions are the following. Firstly, how do copyright concerns disincentivize artists from disclosing their artwork? To address this, I collect the historical publishing data of 6835 artists from daily featured section on DeviantArt from January 2020 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 22%. 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 888 multi-homing artists in the treatment group (digital artists) and control group (artisan crafts artists) and show that digital artists only reduce disclosure artworks on DeviantArt, not on Instagram, compared to artisan crafts artists.

In addition to influencing the volume of published artwork, copyright concerns related to generative AI could also 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 publication volume, the quality of artworks published does not change for a given digital artist.

This paper provides empirical evidence for the theory of how weak intellectual property rights may lead to more trade secrets and less information disclosure. The findings suggest that when intellectual property rights are perceived as weak or inadequately protected, there is a tendency towards increased secrecy and reduced disclosure of information.

These findings also shed light on knowledge spillover. A decline in art publication volume will likely have a negative effect on 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.

Related Literature

Literature shows, both theoretically and empirically, that firms do not always patent and disclose their innovation. Some early survey papers show that firms use secrecy and other alternatives to protect intellectual property apart from patents (Levin et al., 1987, Cohen, Nelson, and Walsh, 2000). Moser (2012) shows that when reverse engineering became possible in the chemistry industry, firms were more likely to patent than to use secrecy to protect innovation, which implies a significant fraction of unpatented innovation before the shock.

Horstmann, MacDonald, and Slivinski (1985) constructs a theoretical model and shows that firms strategically choose how much to patent and disclose to reveal the profitability of innovation. This paper contributes to this literature by providing empirical evidence of innovators strategically choosing whether to disclose innovations. As intellectual property rights protection decreases, creators have less incentive to disclose their innovation.

Research about when intellectual property rights protection decreases, whether the quality of public innovation drops are diverse. Some find bigger innovations are more likely to be protected by secrecy than patent and disclosure (Anton and Yao, 2003). Some find bigger innovations are more likely to be patented (Moser, 2012). Waldfogel (2012), on the other hand, finds no evidence of a reduction in the product quality as file-sharing became available and copying music became significantly easier. This paper has similar results as Waldfogel (2012). I show that when artists consider AI image generators as potential "theft" of their artworks, the publication volume decreases. However, there is no evidence that the quality of artworks generated by non-AI users has changed since the introduction of AI.

The most related theory paper is Gans, 2024, in which the author argues that copyright protection can improve social welfare when the content providers are capable of negotiation with AI providers. But the welfare implication is ambiguous when the AI provider is large enough such that it's infeasible for the AI providers to track the provenance of any particular piece o f content.

A recent empirical paper, Huang, Fu, and Ghose, 2023, also finds similar effects of this paper. Huang et al., 2023 uses regression discontinuity in time and finds a 0.15% of decline in publication volume after LOFTER, an online art platform, introduced generative AI. This paper uses non-AI artists specialized in artisan crafts as control group and non-AI artists specialized in digital art as treatment group and finds 22% of decline in publication volume. I also use behavior of multi-homing artists and show that they only withhold artworks on DeviantArt, not on Instagram. Artists only reduce information disclosure, not production. Abeillon, Haese, Kaiser, Kiouka, and Peukert, 2024 finds that after an online photograph platform selected images for commercial use, treated photographers left the platform at a higher rate and slowed down the rate of new uploads. This paper shows that while some artists reduce their publication volume on DeviantArt, some artists adopt generative AI and increase their uploads. The welfare implication of the introduction of generative AI is ambiguous.

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 of art publication volume 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 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, 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

[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

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. 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.The volume of discussion about artworks being "stolen" by AI increased dramatically on social media and DeviantArt in November 2022. Moreover, many artists announced the deactivation of DeviantArt accounts or stopped updating new artworks from November 2022 to January 2023.

3 Data

3.1 Data Description

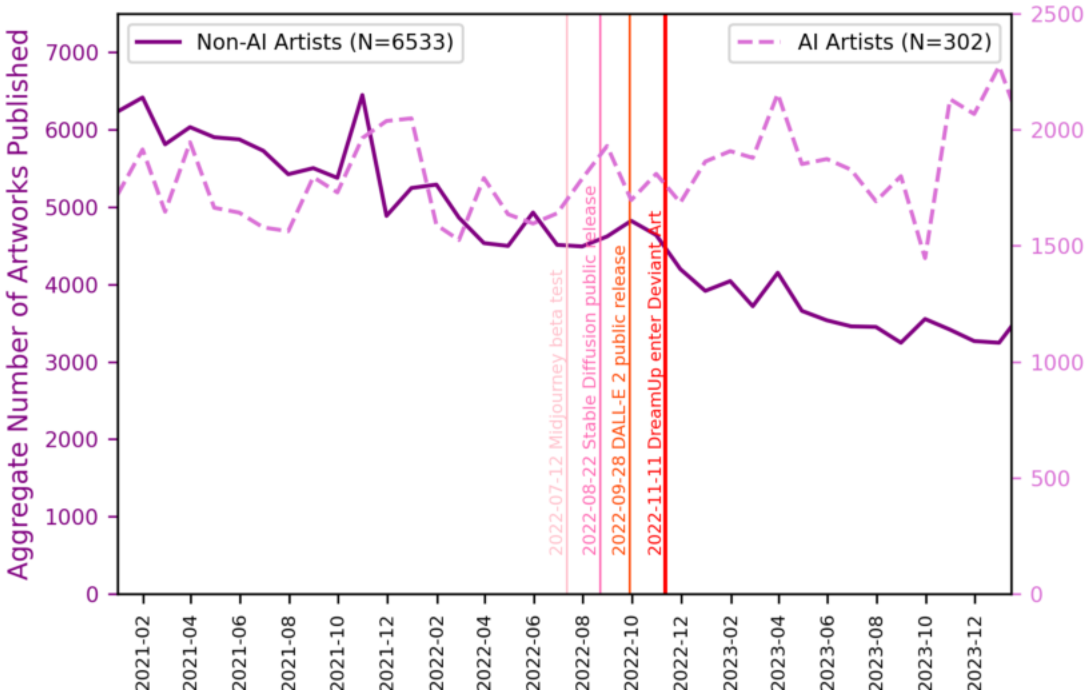
I collect artwork-level data on artists that have artwork selected to be displayed on the daily featured section on DeviantArt using its API. DeviantArt displays around 10 to 20 artworks on this section and they are usually high quality artworks. For the main analysis, I select artists that have artwork on the daily featured section before January 2021, and have been active since then to construct a balanced panel. We should consider them as artists who can produce artwork good enough to enter the daily feature section before AI image generators are available. The full sample contains 6835 artists, including the user profile and their history of publication of artworks on DeviantArt. Table 6 in the appendix presents the summary statistics of artists and artworks in the main analysis.

Figure 1 shows the time trend of historical aggregate publication volume on DeviantArt of 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", "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. There are 4.4% of AI artists and 95.6% of non-AI artists in the sample. Figure 1 shows that AI artists increase their publication volume after the shock. Conversely, the publication frequency of non-AI artists shows a decrease in publication volume after the introduction of Dreamup. One potential concern about these keywords is whether the artists are truth-telling. 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 publication volume

[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.

Figure 1: Overall Time Trend



Notes: This figure shows the number of artworks published by all artists in the data. Since Dreamup was introduced to DeviantArt, there has been an increase in the number of artworks published. This increase is mainly driven by the 4.4% AI artists, while the other 95.6% non-AI artists slightly decrease the number of publications.

of non-AI artists, which does not exist. Therefore, keywords should be a reliable way to identify AI-generated artworks.

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 3050 artists used in the difference-in-differences estimation, 70% 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, which are accessible through the official Meta Developer Graph API. I am able to construct balanced panel data using 888 of the artists.

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

Distribution of Multi-homing Artists	Artists%
Instagram	63.7%
Twitter	48.4%
Facebook	36.4%
YouTube	21.3%
Tumblr	18.1%
Total	100%

Notes: The platforms are not mutually exclusive. Many artists are multi-homing on more than two platforms. Instagram covers over 63% of the artists used in difference-in-differences analysis.

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 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 Publication Volume

I use a difference-in-differences approach to identify a decrease in publication volume after DreamUp was introduced to 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,

[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%.

Figure 2: Illustration of Artisan Crafts and Digital Art

(a) Artisan Crafts

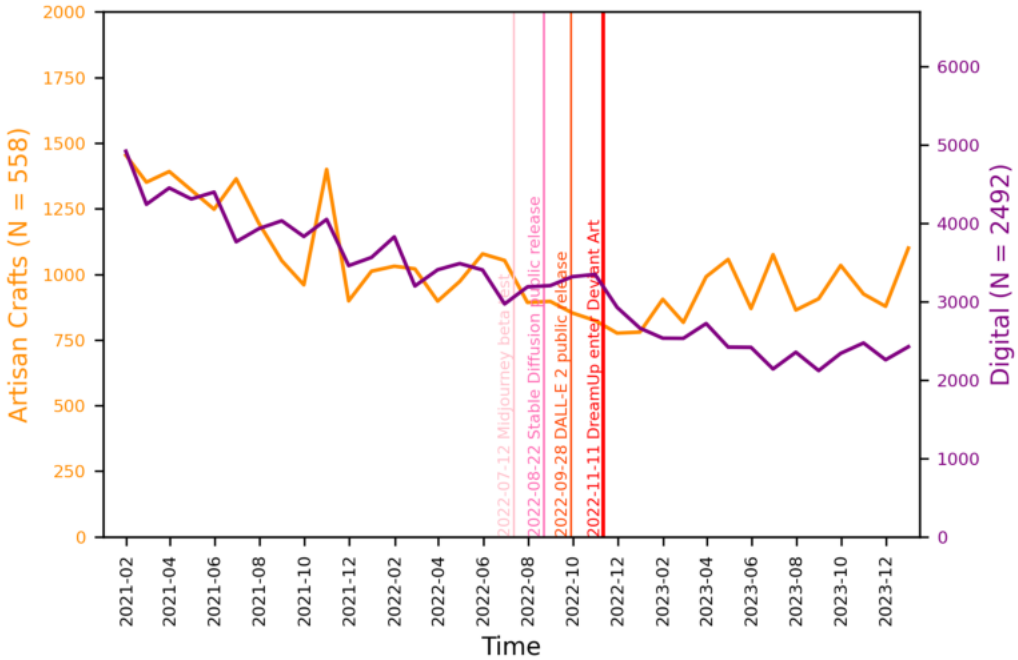
(b) Digital Arts

Notes: I use artists specialized in artisan crafts as the control group and artists specialized in digital art as the treatment group. There are 558 non-AI artisan artists and 2492 non-AI digital artists in my sample.

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. Observations are at the artist-month level.

Figure 3 presents the time trend of aggregate publication volume of both the control group (non-AI artisan crafts artists) and the treatment group (non-AI digital artists). These artists are the subset of the 6533 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. 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

Figure 3: Time Trend of Control and Treatment Group



Notes: This figure only uses the publication records of non-AI artists specialized in digital art and artisan crafts. They are the subset of the 6533 non-AI artists displayed in Figure 1. Note that in the full sample, there are many other specializations of artists.

artists because generative AI is mostly used to generate digital artworks.

The main specification uses two-way fixed effects:

$$Artwork_{it} = \beta_0 + \beta_1 Post_t \times Treated_i + \delta_i + \delta_t + \epsilon_{it} \tag{1}$$

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.

4.2 Impact on Quality of Art

To examine whether digital artists reduce disclosure of 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:

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

where y_{ijt}^{Ins} is the performance of an artwork j of 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.

5 Results

5.1 Decrease of Publication Volume on DeviantArt

Table 2 shows the result of equation (1). I employ both Poisson Pseudo Maximum Likelihood (PPML) estimation and OLS estimation in the main analysis. 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, we should focus more on the results of the PPML estimation. 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 reduce 22% of artworks published on DeviantArt after the shock. Column (2) shows a similar scale of reduction, at 14%. Note that this is likely to be an underestimation of the true effect. By

Table 2: Effect on Artist Publication Volume

	Baseline Estimation		Winsorize 99% of Dep Var		Drop 1% Largest SD. Artists	
	(1)	(2)	(3)	(4)	(5)	(6)
	PPML	OLS	PPML	OLS	PPML	OLS
$Post_t \times Treated_i$	-0.25***	-0.21*	-0.17**	-0.17**	-0.27***	-0.30***
	(0.09)	(0.12)	(0.07)	(0.07)	(0.09)	(0.11)
Pre-Treatment Mean		1.50		1.34		1.48
Artist FE	Y	Y	Y	Y	Y	Y
Month FE	Y	Y	Y	Y	Y	Y
N(Artist-Month)	109,800	109,800	109,800	109,800	108,936	108,936
N(Artists)	3,050	3,050	3,050	3,050	3,027	3,027
Adjusted R^2	0.53	0.58	0.46	0.50	0.49	0.37

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

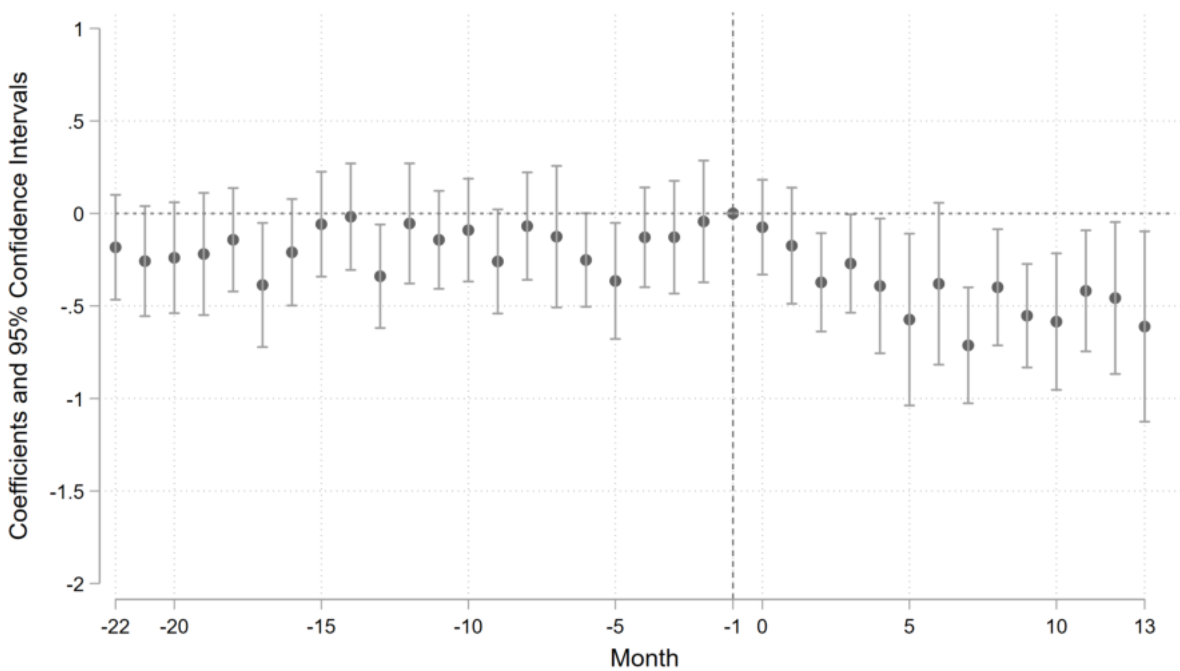
the time I started collecting the data, some artists have already deactivated their accounts. Figure 4 shows the pre-trend of column (1).

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 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.

However, artists only reduce publication volume on DeviantArt, not on Instagram. Table 3 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 publication volume on DeviantArt after the shock, while the publication volume appears to remain unchanged on Instagram.

These results imply that the multi-homing artists only reduce disclosure of 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

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 Artist Publication Volume on Instagram

	Panel 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
$Post_t \times Treated_j$	-0.29**	-0.31	-0.22*	-0.23	-0.33**	-0.38*
	(0.15)	(0.20)	(0.12)	(0.14)	(0.14)	(0.19)
Pre-Treatment Mean		1.70		1.54		1.66
Artist FE	Y	Y	Y	Y	Y	Y
Month FE	Y	Y	Y	Y	Y	Y
N(Artist-Month)	31,968	31,968	31,968	31,968	31,716	31,716
N(Artists)	888	888	888	888	881	881
Adjusted R^2	0.44	0.33	0.42	0.48	0.44	0.36
	Panel B: Artworks on Instagram					
	Baseline Estimation		Winsorize 99% of Dep Var		Drop 1% Largest SD. Artists	
	(7)	(8)	(9)	(10)	(11)	(12)
	PPML	OLS	PPML	OLS	PPML	OLS
$Post_t \times Treated_j$	-0.09	-0.04	-0.04	0.18	-0.06	-0.03
	(0.08)	(0.41)	(0.07)	(0.34)	(0.07)	(0.37)
Pre-Treatment Mean		3.65		3.35		3.39
Artist FE	Y	Y	Y	Y	Y	Y
Month FE	Y	Y	Y	Y	Y	Y
N(Artist-Month)	31,968	31,968	31,968	31,968	31,644	31,644
N(Artists)	888	888	888	888	879	879
Adjusted R^2	0.49	0.59	0.43	0.56	0.44	0.55

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.

Table 4: High Performance Correlation Across Platforms

Dep Var	$Likes^{Ins}$				$Comments^{Ins}$			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
$Favorites^{DA}$	2.73***				0.01***			
	(0.56)				(0.00)			
$Comments^{DA}$		74.76***				0.50***		
		(13.32)				(0.06)		
$Downloads^{DA}$			1.75				0.01	
			(1.08)				(0.01)	
$Views^{DA}$				0.01***				0.00***
				(0.00)				(0.00)
Artists FE	Y	Y	Y	Y	Y	Y	Y	Y
Month FE	Y	Y	Y	Y	Y	Y	Y	Y
N(Artworks)	22,690	22,690	22,690	22,690	22,690	22,690	22,690	22,690
Adjusted R^2	0.46	0.46	0.44	0.45	0.50	0.51	0.48	0.49

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.

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 on AI-generated content. 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: Fewer High-Quality Artworks?

To examine whether artists reduce disclosure of 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 4 validates the use of likes and comments on Instagram as a quality measure.

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
$Post_t \times Matched_j$	0.08	164.11	-0.04	-0.55
	(0.06)	(139.15)	(0.05)	(1.10)
$Matched_j$	-0.03	-48.88	0.08***	1.33*
	(0.03)	(86.03)	(0.03)	(0.68)
Pre-Treatment Mean		2058.08		15.06
Artist FE	Y	Y	Y	Y
Month FE	Y	Y	Y	Y
N(Artworks)	87,769	87,840	87,836	87,840
Adjusted R^2	0.77	0.34	0.61	0.17

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.

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 $Matched_j$ in column (3) and (4) indicate that artists select higher quality artworks and post 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.

6 Conclusion

This paper studies the effect of the introduction of AI into online art platforms on art publication volume and quality of artworks. I use a difference-in-differences approach to show that compared to non-AI artisan artists, non-AI digital artists reduce over 22% of art publication volume each month after the introduction of DreamUp to DeviantArt. By looking into the multi-homing artists with a professional or business account on Instagram, I show that digital artists only reduce disclosure of artworks on DeviantArt, not on Instagram. The reason is that the introduction of DreamUp implies weak copyright protection on DeviantArt, and artists show concern about their artwork entering the AI company's training data without consent.

By comparing artworks only displayed on Instagram to those also on DeviantArt of multi-homing digital artists, I show that this reduction of disclosure is not biased towards high-quality information or low-quality information 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 innovators to disclose information in an AI-dominated era due to weak copyright protection. But the quality of published knowledge does not appear to decline. This raises important follow-up questions: what can we do to encourage innovators to disclose their knowledge, and 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 publication volume affects future productivity and, if so, by how much.

7 Appendix

7.1 Summary Statistics of Artists

Table 6 shows the summary statistics of artists.

7.2 Pre-trend of Table 2

Figure 5 shows pretrends of Table 2.

7.3 Pre-trend of Table 3

Figure 6 shows pretrends of Table 3.

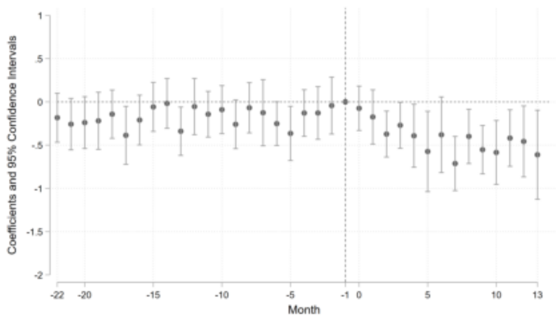
Table 6: Summary Statistics

	Artisan Crafts Artists				Digital Art Artists			
Panel A: Artists Used in Difference-in-Differences								
	Mean	Min	Max	sd	Mean	Min	Max	sd
Monthly Pre-Period Artworks	1.97	0	369	11	1.50	0	474	5.26
Profile Pageviews	1.04e+05	639	3.25e+06	2.30e+05	3.23e+05	799	5.38e+07	1.69e+06
Followers	1964	11	6.31e+04	4724	7053	11	6.76e+05	2.32e+04
$Views^{DA}$ per Artwork	5598	15	1.09e+06	4.06e+04	1.81e+04	14	5.49e+06	7.41e+04
$Downloads^{DA}$ per Artwork	5.85	0	1253	34	18	0	2.25e+04	118
$Favourites^{DA}$ per Artwork	39	0	2486	89	189	0	1.09e+04	432
$Comments^{DA}$ per Artwork	2.40	0	151	6.49	6.21	0	2887	15
$N(Artist)$	558				2492			
Panel B: Artists Used in Instagram Difference-in-Differences								
	Mean	Min	Max	sd	Mean	Min	Max	sd
Monthly Pre-Period Artworks	1.78	0	67	4.88	1.70	0	271	5.04
Profile Pageviews	1.37e+05	1668	3.25e+06	3.35e+05	3.64e+05	799	1.04e+07	9.98e+05
Followers	2853	24	6.31e+04	7200	1.00e+04	33	3.43e+05	2.66e+04
$Views^{DA}$ per Artwork	1.18e+04	35	1.09e+06	5.40e+04	2.72e+04	34	1.64e+06	9.30e+04
$Downloads^{DA}$ per Artwork	19	0	955	60	25	0	4089	128
$Favourites^{DA}$ per Artwork	90	0	2486	144	263	0	1.09e+04	523
$Comments^{DA}$ per Artwork	2.49	0	90	4.57	7.88	0	348	16
$Likes^{Ins}$ per Artwork	787	0	2.23e+05	4315	2015	0	1.29e+06	8985
$Comments^{Ins}$ per Artwork	13	0	1.17e+04	73	14	0	7906	63
$N(Artist)$	170				718			

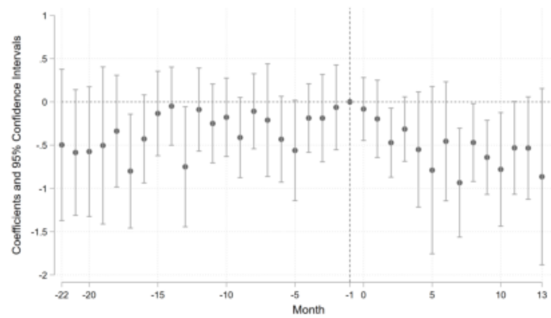
Notes: Use panel from January 2021 to December 2023.

Figure 5: Pre-Trends of Table 2

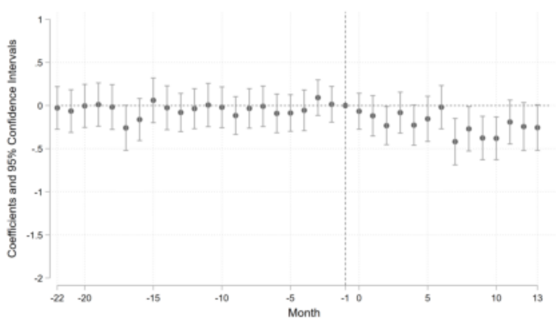
(a) Baseline Estimation PPML



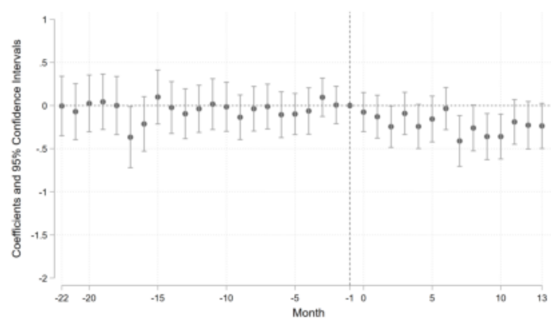
(b) Baseline Estimation OLS



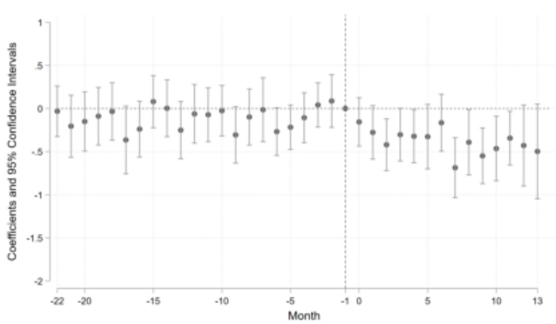
(c) Winsorize PPML



(d) Winsorize OLS



(e) Drop 1% largest SD PPML



(f) Drop 1% largest SD OLS

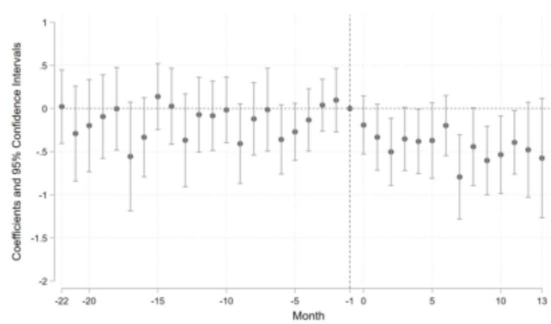
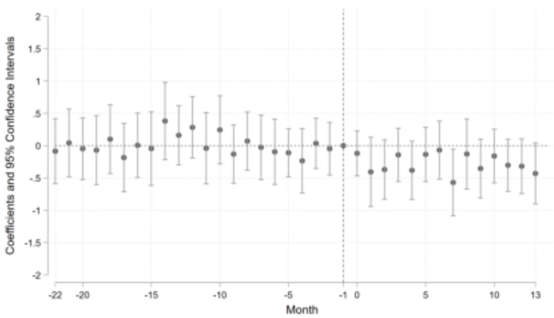
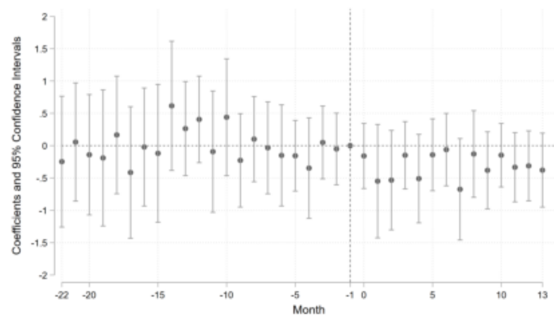


Figure 6: Pre-Trends of Table 3 Panel A: Artworks on DeviantArt

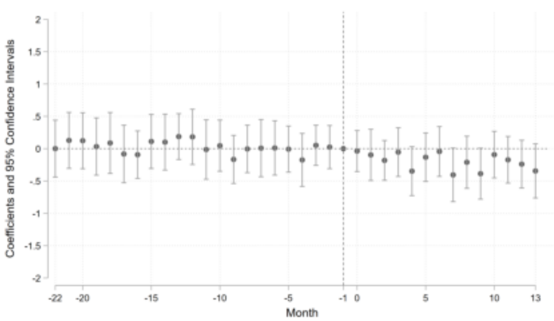
(a) Baseline Estimation PPML



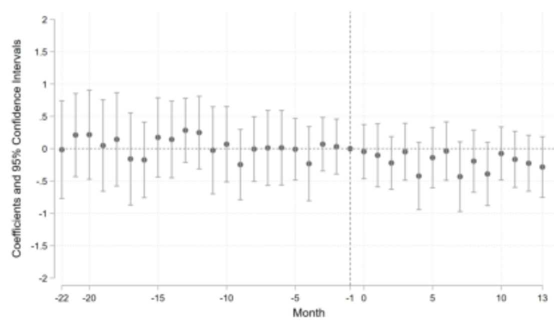
(b) Baseline Estimation OLS



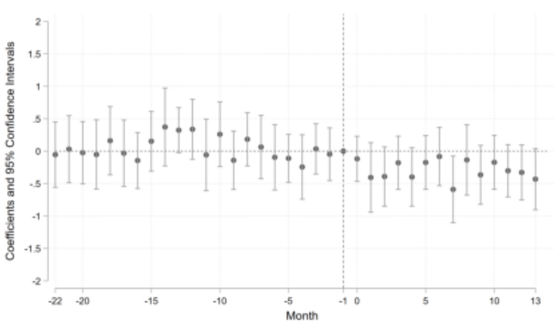
(c) Winsorize PPML



(d) Winsorize OLS



(e) Drop 1% largest SD PPML



(f) Drop 1% largest SD OLS

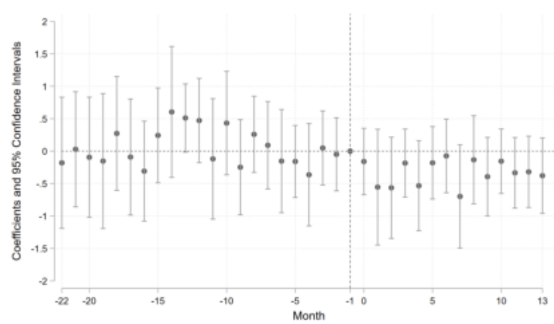
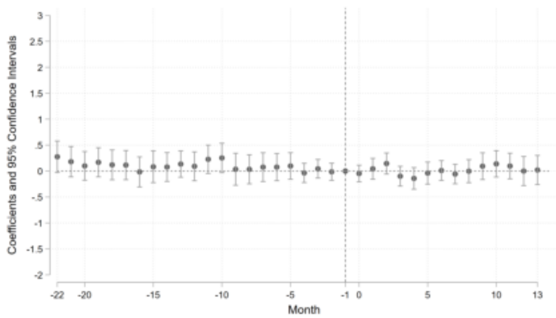
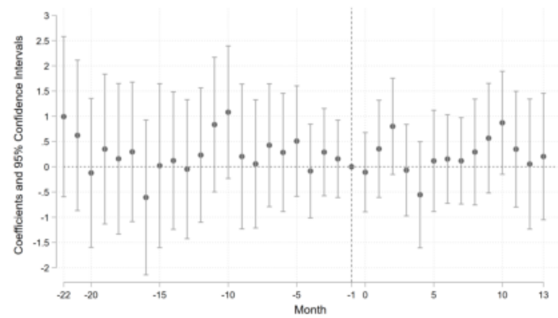


Figure 7: Pre-Trends of Table 3 Panel B: Artworks on Instagram

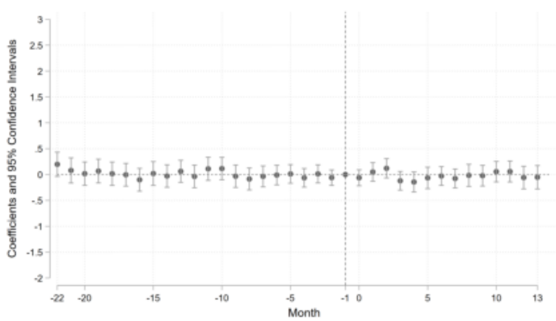
(a) Baseline Estimation PPML



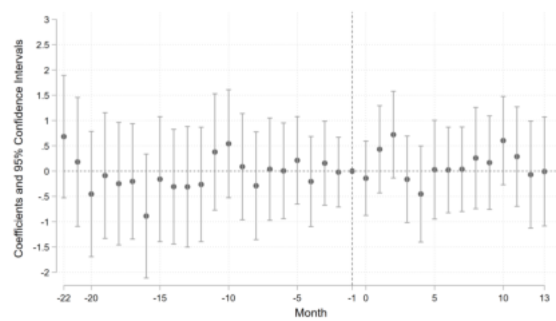
(b) Baseline Estimation OLS



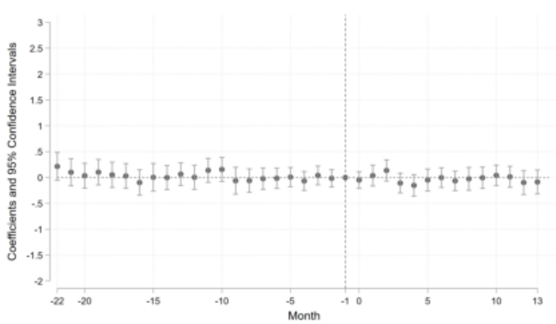
(c) Winsorize PPML



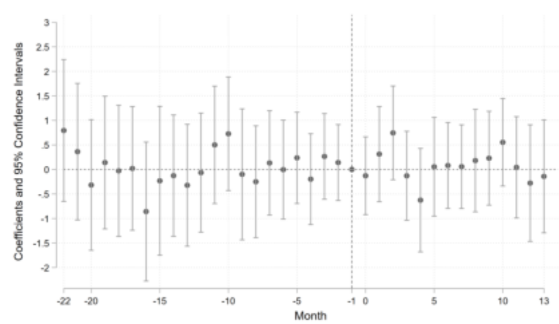
(d) Winsorize OLS



(e) Drop 1% largest SD PPML



(f) Drop 1% largest SD OLS



References

Abeillon, F., Haese, J., Kaiser, F., Kiouka, A., & Peukert, C. (2024). Stratigic behavior and AI training data. *Working paper.*

Anton, J. J., & Yao, D. A. (2003). Patents, invalidity, and the strategic transmission of enabling information. *Journal of Economics & Management Strategy*, *12*(2), 151–178. _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1111/j.1430-9134.2003.00151.x. doi:10.1111/j.1430-9134.2003.00151.x

Cohen, W. M., Nelson, R. R., & Walsh, J. P. (2000). *Protecting their intellectual assets: Appropriability conditions and why u.s. manufacturing firms patent (or not).* National Bureau of Economic Research. doi:10.3386/w7552

Gans, J. S. (2024, January 26). Copyright policy options for generative artificial intelligence. Rochester, NY. Retrieved January 28, 2024, from https://papers.ssrn.com/abstract=4707911

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). 2019 IEEE power and energy conference at illinois (PECI). doi:10.1109/PECI.2019.8698905

Horstmann, I., MacDonald, G. M., & Slivinski, A. (1985). Patents as information transfer mechanisms: To patent or (maybe) not to patent. *Journal of Political Economy*, *93*(5), 837–858.

Huang, H., Fu, R., & Ghose, A. (2023, December 20). Generative AI and content-creator economy: Evidence from online content creation platforms. doi:10.2139/ssrn.4670714

Levin, R. C., Klevorick, A. K., Nelson, R. R., Winter, S. G., Gilbert, R., & Griliches, Z. (1987). Appropriating the returns from industrial research and development. *Brookings Papers on Economic Activity*, *1987*(3), 783–831. doi:10.2307/2534454

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

Moser, P. (2012). Innovation without patents: Evidence from world's fairs. *The Journal of Law and Economics*, *55*(1), 43–74. Publisher: The University of Chicago Press. doi:10.1086/663631

Munkres, J. (1957). Algorithms for the assignment and transportation problems. *Journal of the Society for Industrial and Applied Mathematics*, *5*(1), 32–38.

Waldfogel, J. (2012). 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