DOI: 10.1111/eufm.12506

ORIGINAL ARTICLE

EUROPEAN FINANCIAL MANAGEMENT WILEY

Pricing dynamics and herding behaviour of NFTs

Gilbert Fridgen[1] | **Roman Kräussl**[2] | **Orestis Papageorgiou**[1] | **Alessandro Tugnetti**[3]

[1]SnT - Interdisciplinary Center for Security, Reliability and Trust, University of Luxembourg, Luxembourg

[2]Bayes Business School (formerly Cass), Hoover Institution at Stanford University and CEPR, London

[3]Department of Finance, University of Luxembourg, Luxembourg

Correspondence
Orestis Papageorgiou, SnT - Interdisciplinary Center for Security, Reliability and Trust, University of Luxembourg, 29JFK, Blvd JF Kennedy, L-1855 Luxembourg.
Email: orestis.papageorgiou@uni.lu

Funding information
Luxembourg National Research Fund (FNR), Grant/Award Numbers: 16326754, NCER22/IS/16570468/NCER-FT; PayPal, PEARL, Grant/Award Number: 13342933/GilbertFridgen; Luxembourg Institute for Advanced Studies (IAS), Young Academics scheme, DATART project

Abstract

This paper analyzes the sales of 875,389 art non-fungible tokens (NFTs) on the Ethereum blockchain to identify the key determinants influencing NFT pricing and market dynamics. We find that market liquidity and trade volume are strong predictors of NFT prices. Contrarily, social media activity negatively correlates with prices. Introducing an artist ranking system, our study reveals a "superstar effect", with a few artists dominating sales, and herding behaviour within the NFT market.

KEYWORDS

art market, cryptocurrency, Ethereum blockchain, herding, liquidity, network, nonfungible tokens, price index, social media, speculation

This research was funded in part by the Luxembourg National Research Fund (FNR), grant reference 16326754 and NCER22/IS/16570468/NCER-FT, by PayPal, PEARL grant reference 13342933/Gilbert Fridgen and by the Luxembourg Institute for Advanced Studies (IAS), Young Academics scheme, DATART project. For the purpose of open access and in fulfillment of the obligations arising from the grant agreement, the author has applied a Creative Commons Attribution 4.0 International (CC BY 4.0) license to any Author Accepted Manuscript version arising from this submission.

JEL CLASSIFICATION
C55, G11, Z11

1 | INTRODUCTION

Nonfungible tokens (NFTs) rose in prominence in 2021, amassing more than $44 billion in traded volume and attracting the attention of both institutional and retail investors (Chainalysis, 2022). In less than a year, NFTs transitioned from being recognized by only a small community of blockchain enthusiasts to becoming the subject of full-length articles in major news outlets, being described as a revolution not only in the art industry but also in various other sectors (Wilson et al., 2021). This prominence even led to Collins Dictionary choosing "NFT" as the 2021 word of the year (Collins, 2021). Such popularity inevitably prompted a considerable amount of scrutiny regarding the utility of NFTs and whether they can truly reform the digital asset sector. One might argue that NFTs are not merely a fleeting trend but rather part of a broader technological evolution in the blockchain space. Some perceive the crash in 2022 as a correction through the underlying value of NFTs, aligning with a more nuanced understanding of market dynamics. Others may contend that the drop in value does not diminish the inherent properties that make NFTs unique; instead, it could potentially pave the way for a more sustainable growth pattern that aligns with long-term financial models.

Despite the hype subsiding and the trading volume falling, NFTs have recently found applications ranging from empowering business models innovation (Hartwich et al., 2023; Li & Chen, 2023) to serving as the primary component in several initial coin offerings used by firms to raise capital (Holden & Malani, 2022). Besides these applications, digital collectibles and artworks represent the most prevalent category within the NFT marketplace. Probably the most influential art NFT sale came in March 2021 with Christie's auction house first-ever sale of a digital artwork: Beeple's "Everydays: The First 5000 Days" was sold for $69.3 million, positioning the artist among the most valuable living creators. This, combined with other notable sales such as Pak's "The Merge", the most expensive NFT sold to date, and certain CryptoPunks selling for more than $10 million, propelled NFTs to the forefront of the art world in 2021, dominating the digital art space (Financial Times, 2021a) and generating almost as much volume as the traditional art space (Financial Times, 2021b). Even in 2022, when every other NFT category experienced a dramatic decrease in the number and value of sales, art NFTs have seen a slight increase in the average sale value (McAndrew, 2022). Despite facing market volatility, NFTs with actual use-cases, such as those in gaming, cultural preservation and fractional ownership, are emerging as particularly promising investment avenues, signalling a market evolution that increasingly prioritizes utility-driven assets over speculative collectibles and thereby suggesting a maturing landscape where authentic value will likely catalyze future growth, thus providing opportunities for both creators and investors to strategically position themselves in sectors of tangible utility and enduring value (Hategan, 2023).

Regardless of growth, relatively little is understood regarding the attributes that lend value to art NFTs and influence their selling probability. Although scarcity is widely acknowledged, both by academics (Alizadeh et al., 2023; Bamakan et al., 2022; Bao & Roubaud, 2022; Mekacher et al., 2022) and by practitioners (Sotheby's, 2021; Valeonti et al., 2021), as an essential trait, other factors, such as the impact of the artist, and the collection characteristics, remain vague, while academic research articles tend to focus on a few collections and attributes (Cong et al., 2022; Huang & Goetzmann, 2023; Sockin & Xiong, 2022).

Dowling (2022a) provides a pioneering exploration of Decentraland LAND NFTs, drawing a mental association between virtual and physical land, and identifying possible inefficiencies in pricing behaviour akin to early cryptocurrency pricing. This notion of inefficiency is further examined from a broader perspective by Cheah and Fry (2015) and Urquhart (2016), underlining the nascent nature of NFT markets. Schaar and Kampakis (2022) take a quantitative approach to investigate the CryptoPunks collection, highlighting an impressive average monthly return of 34.19% over 3 years and the significant role that rarity plays in determining prices. This emphasis on rarity is also central to the work of Mekacher et al. (2022), who delve into 410 collections to demonstrate that rarer NFTs command higher prices and are less risky. Horky et al. (2022) employ a combination of econometric tools and machine learning in their study of the digital art market through SuperRare, finding that hedonic pricing models furnish valuable insights into NFT prices, independent of cryptocurrencies.

Broadening the scope to encompass the entire market, Dowling (2022b) work stands out by interrogating the connection between NFT pricing and cryptocurrencies, uncovering limited volatility transmission effects but a notable comovement between the two. Ante (2022) adds nuance to this relationship by revealing how Bitcoin price shocks trigger an increase in NFT sales, while Ether price shocks have a converse effect. Borri et al. (2023) take an ambitious step by constructing a comprehensive data set for the overall NFT market, unearthing the nuanced interplay between creator fees, rarity, visual characteristics and prices. Kong and Lin (2021) enrich this perspective by emphasizing the role of well-connected investors in the pricing dynamics of NFTs. Mazur (2021) shifts the focus to the risk and return characteristics of NFT-based start-ups, illustrating a staggering investment multiple of 40 over the long term. In their seminal paper, Nadini et al. (2021) map the structure and evolution of the entire NFT market, identifying sale history and visual features as key price predictors. Hemenway Falk et al. (2022) introduce a novel dimension by exploring the value of creator royalties, uncovering their potential to reshape the NFT market through risk-sharing, dynamic pricing and price discrimination.

While the existing literature on NFT pricing and determinants has provided a comprehensive view of the market, the exploration of herding behaviour in the context of art NFTs presents a unique and uncharted territory. Herding behaviour, as observed in traditional financial markets (Banerjee, 1992; Bikhchandani et al., 2021; Shiller, 1995), refers to the inclination of investors to follow the majority's actions rather than making independent decisions based on intrinsic value. Art NFTs offer a perfect primer to study this phenomenon, primarily due to their parallel nature with the physical art market where the phenomenon of herding behaviour has been well-documented. Azarmi and Menny (2013) study of the fine arts market explores the tendency of investors to gravitate towards well-established artists, leaving the quality of the work secondary to the historical auction performance of the artist. They found that a small fraction of artists dominates financial activity and that contemporary and less-established artists are subject to more herding behaviour. This observation reflects a pattern where investors follow trends and the actions of others, often leading to inflation in the value of specific artists or styles. Art NFTs amplify these dynamics by adding layers of complexity. Unlike physical art, where authenticity, provenance, and physical condition play a role in valuation, art NFTs operate on a digital platform where these factors are replaced by digital scarcity, network effects and integration with cryptocurrencies. The rapid rise of art NFTs, by significant price volatility and intense media coverage, has many resemblances to the herding behaviour observed in the traditional art market. Investors and collectors rush towards certain NFTs, often driven more by hype and the actions of others than a deep understanding of the digital art's inherent value (Nadini et al., 2021).

Moreover, the volatility and correlation with the broader cryptocurrency market further amplify the complexity of art NFTs. The fluctuations in their prices, often driven by external factors such as changes in Ethereum or Bitcoin values, create a turbulent environment where herding behaviour could thrive. The allure of quick profits, media attention and the novelty of owning a unique digital asset can lead to a cascade effect where one investment decision triggers another, often without a rational analysis of the asset's underlying value. The work of Bao et al. (2023) provides the first empirical evidence on herding in the NFT market. They identify three waves of herding in the NFT market, with daily market returns becoming more volatile during these periods. Their findings also reveal that herding is more likely to emerge as the proportion of newcomers increases and that media exposure drives investors' attention when herding arises. They further find a connection with the return on Ethereum but a diminishing effect with the return on Bitcoin. Yousaf and Yarovaya (2022) examine herding behaviour across three cryptocurrency classes, including NFTs. Their time-varying analysis identified herding in conventional cryptocurrencies during the most recent bubble of 2021 but failed to demonstrate evidence of herding in NFTs during various market conditions. This contrast presents an intricate picture of the NFT market, suggesting that herding behaviour may manifest differently across various aspects of the broader cryptocurrency space. The speculative nature of NFTs and DeFi markets, often perceived as bubble behaviours, has been explored by Wang et al. (2022). They document that both NFT and DeFi markets exhibit speculative bubbles, with NFT bubbles being more recurrent and having higher explosive magnitudes. The bubbles in these markets are correlated with market hype and broader cryptocurrency market uncertainty, though they also recognize periods where bubbles are not detected, indicating intrinsic value in these markets.

Our study's approach merges financial, descriptive and social network features of art NFTs. We aim to explore how market frictions operate in a digital asset environment, and how they may affect pricing, liquidity and investor behaviour. Specifically, this paper investigates the relationship between art NFT pricing and market herding behaviour. The primary objective is to construct a comprehensive price index for art NFTs, integrating multifaceted attributes, such as financial, artistic and social network influences. This setting aims to challenges conventional asset pricing models, acknowledging the unique, nonfundamental value characteristics of art NFTs (Taleb, 2021). The methodology employs hedonic and repeat-sales regression (RSR) models, analyzing differentiated market trends and isolating intrinsic item qualities to understand price dynamics over time.

Second, we examine investor herding in the art NFT market. We assess whether collective investor behaviour, influenced more by group dynamics than individual rationale, plays a significant role in art NFT valuation cycles. Our research uses Cross-Sectional Absolute Deviation (CSAD) to analyze price comovements, particularly in the context of low liquidity and asymmetric information typical of the art NFT domain. This analysis aims to uncover deviations from fundamental valuation driven by investor herding, providing insights into price formation and market behaviour in the evolving NFT landscape. We focus on the entire art NFT market, analyzing every art collection and transaction performed on Ethereum, the blockchain with the most NFT transactions and home to many successful token offerings from firms. Our sample comprises 875,389 art NFTs that have been deployed on the Ethereum blockchain and have been involved in 385,884 sales.

Our findings confirm the positive correlation between the number of trades and the floor price of NFTs with their average price, elucidating critical elements of the NFT market's pricing structure. We find an unexpected inverse relationship between social media activity and NFT prices, presenting a challenge to popular conceptions of the NFT market. Remarkably, our exploration of herding behaviour revealed a small group of artists who command most of the market's activity, mirroring trends seen in traditional art markets, while also spotlighting the market's susceptibility to potential speculative bubbles. Our results also demonstrate that in instances of high market volatility, investors diverge from herding, adopting more risk-averse strategies. New entrants initially exert a stabilizing effect on market prices, but eventually contribute to increased financial volatility. Notably, a bullish broader cryptocurrency market acts as a moderating variable, diverting capital away from NFTs, thus mitigating herd-driven price inflation in this asset class.

These insights into the art NFT market's pricing dynamics and liquidity constraints contribute to the existing financial literature, providing a rich context for understanding a new asset class within a blockchain-enabled environment. The link between herding behaviour, speculative bubbles and the inherent value of NFTs opens up avenues for understanding the art market's financial dynamics and the broader economic landscape. The integration of insights from traditional herding behaviour theories (e.g., Banerjee, 1992; Shiller, 1995) with contemporary studies on NFTs and art markets provides a nuanced understanding of market dynamics. This interdisciplinary approach is instrumental in drawing parallels between classical assets, like, securities, physical art and other collectibles, and emerging digital assets, like, NFTs.

What sets NFTs apart is the unprecedented access to a wealth of real-time data about transactions and financial characteristics of these digital assets. Unlike traditional markets where information might be fragmented, delayed or obscured by various market frictions, the blockchain technology underpinning NFTs ensures that every transaction is transparent, timestamped and publicly accessible. The ability to track and analyze these transactions in real-time opens up new avenues for research, allowing for a more nuanced understanding of market dynamics, pricing mechanisms and investor behaviour. As an example, researchers can study market frictions in a way that was previously unattainable. Market frictions, such as transaction costs, information asymmetry and liquidity constraints, play a critical role in asset pricing and investment strategies. They have been studied extensively in traditional financial markets by authors, like, Amihud and Mendelson (1986), Stiglitz (1989) and Vayanos and Wang (2012). In the context of NFTs, these frictions might manifest differently, and their impact on market dynamics could be distinct from what is observed in traditional markets.

The remainder of this paper is structured as follows. Section 2 discusses the data extraction techniques. Section 3 describes the econometric models used to construct the price indexes and to study the herd bias. Section 4 discusses our findings and Section 5 concludes.

2 | DATA

Blockchain-related analyses are often classified into two broad categories: on-chain and off-chain. On-chain research involves data retrieved directly from the blockchain's public ledger, whereas off-chain analysis utilizes data sources outside the blockchain, such as price-tracking websites. For our analysis, we rely on both categories to obtain as much data as possible on NFTs and their pricing. Specifically, by leveraging the granular data available through NFT transactions, we strive to provide a more comprehensive understanding of the art NFT market. Please, refer to Appendix A for a detailed explanation of data extraction, cleaning and preparation processes.

Our sample comprises 875,389 art NFTs that have been deployed on the Ethereum blockchain and have been involved in 385,884 sales between 15 July 2018 and 10th 10 February 2022. To facilitate a deeper understanding of this extensive data set, we have organized the information into two distinct tables.

Table 1 presents the descriptive statistics, summarizing key features such as median, standard deviation and other relevant statistical measures that provide an overall picture of the art NFT market. These statistics offer insights into the general trends, distributions and characteristics that define the market landscape. Furthermore, they are instrumental in constructing a comprehensive price index for the art NFT market. The table contains financial variables such as the returns of the Bloomberg Galaxy Crypto Index (*BGCI_return*) since past research has indicated the valuation of NFTs can move in tandem with broader cryptocurrency markets, suggesting that investor sentiment in cryptocurrencies could spill over into NFT valuations. To understand the market's valuation floor, the table includes the lowest selling price of an NFT in a collection (*floor_price*), which sets the baseline for the collection's value. Complementing this, the most recent selling price (*last_price*) offers a snapshot of the current market demand. The number of times an NFT has been sold (*num_trades*) is tracked to assess liquidity and investor interest, while the time span between the first and last sales (*timediff*) provides insight into the NFT's market presence and activity over time. Additionally, it contains visual analytics, quantifying the aesthetic attributes of NFTs, such as the dominance of specific colours within the image (*white, black and so forth*), the total number of colours in each image (*num_colours*) and the complexity of the image expressed by the normalized Shannon's entropy (*norm_shannon_entropy*). These visual analytics reflect the NFT's aesthetic uniqueness, which can influence its market value (Nadini et al., 2021). Lastly, the table captures the presence of NFT collections on Twitter, with metrics like the number of accounts each collection follows (*following_count*) and the monthly average of quotes received (*quote_count_month*). These metrics serve as indicators of the collection's community engagement and outreach, which are essential for understanding the social standing and potential influence of NFT collections. Primarily, the rationale behind leveraging Twitter data, as opposed to other social media, such as Reddit, Google and Facebook, hinges on its relative accessibility and the ease of an expedited data aggregation and analysis process, consistent with the current literature on digital asset pricing (Kapoor et al., 2022). Furthermore, Twitter stands out as a primary source of insights about the NFT market. As noted by Nadini et al. (2021), a predominant proportion of NFT transactions, happens on platforms like *OpenSea*, where artists predominantly link their Twitter accounts. This unique aspect of the NFT market makes Twitter a more relevant and comprehensive source for gauging market sentiment and trends, compared with other social media platforms. This extensive inclusion is aimed at enhancing the index's scope and accuracy, drawing on methodologies from the hedonic and repeat-sales literature (Bailey et al., 1963; Borri et al., 2023).

Table 1 shows that the distribution of most variables is negatively skewed; for many of them, this remains around 0 up to the second quartile and then increases exponentially towards the third and fourth quartiles. The application of the Jarque–Bera (JB) and Kwiatkowski–Phillips–Schmidt–Shin tests to our data uncovers intricate statistical properties. Notable deviations from normality and the existence of unit roots in continuous variables such as *BGCI_return* emphasize the nonlinear characteristics of the art NFT market. The JB statistics reveal pronounced skewness and kurtosis, confirming the fragmented nature of the NFT market (Caporale et al., 2021).

The variables *floor_price* and *last_price* highlight the extensive range of pricing that characterizes the art NFT market. This dynamic pricing behaviour correlates with observations in

TABLE 1 Descriptive statistics art nonfungible tokens (NFTs).

This table reports a detailed statistical overview of art NFTs. Panel A focuses on the continuous variables, providing a comprehensive statistical summary of the data extracted. This includes the median, standard deviation, minimum, maximum, Jarque–Bera (JB) test statistic for normality and the Kwiatkowski–Phillips–Schmidt–Shin (KPSS) test for stationarity across both the aggregate and transaction data sets. The aggregate data set encompasses augmented data per NFT, including data from Discord and Twitter, as well as color and image complexity data. The transaction data set contains detailed information pertaining to every sale of the extracted art NFTs. Panel B details the discrete variables in both data sets, listing the unique values and the top frequent value with its corresponding frequency. The information compiled in this table stems from a rigorous data extraction process, leveraging on-chain and off-chain sources as described in Appendix A, providing a rich and diversified analysis of the NFTs, encompassing both their visual characteristics and market behaviour. As a final note, following data extraction, the variables measuring total and monthly messages and unique users in both the announcement and general channels of Discord servers have a number of missing values greater than 70%. For this reason, they will not be reported or used in the analysis. $^{*}p \le 0.1$; $^{**}p \le 0.05$; $^{***}p \le 0.01$.

Panel A: Continuous variables

Variable	Median	σ	Minimum	Maximum	JB Statistic	KPSS test
Aggregate data set						
BGCI return	0	0.03	−0.26	0.22	7.91e + 04***	0.94**
black	0.04	0.27	0	1	1.15e + 05***	4.91***
blue	0.04	0.23	0	1	2.23e + 05***	14.38***
floor price	616.01	2.38e + 04	1	6.25e + 06	1.13e + 13***	2.49***
gray	0.01	0.2	0	1	6.19e + 05***	10.62***
green	0	0.17	0	1	1.44e + 06***	7.56***
last price	728.68	2.82e + 04	1	6.25e + 06	3.59e + 12***	4.29***
norm shannon entropy	0.29	0.12	0	0.53	1.69e + 04***	5.38***
orange	0.03	0.2	0	1	4.22e + 05***	17.01***
purple	0	0.13	0	1	6.34e + 06***	10.13***
red	0.05	0.21	0	1	3.96e + 05***	4.16***
timediff	48.94	242.59	0.01	1307.19	5.50e + 05***	16.89***
white	0.01	0.25	0	1	2.10e + 05***	2.66***
yellow	0	0.1	0	1	1.23e + 07***	1.64***
Transaction data set						
average price	773.42	2.23e + 04	1	6.25e + 06	1.35e + 13***	0.73***
floor price	479.99	2.00e + 04	1	6.25e + 06	3.08e + 13***	0.6**
last price	775.45	2.85e + 04	1	6.25e + 06	3.96e + 12***	0.77***

(Continues)

TABLE 1 (Continued)

Panel A: Continuous variables						
Variable	**Median**	**σ**	**Minimum**	**Maximum**	**JB Statistic**	**KPSS test**
max price	972.04	3.17e + 04	1	6.25e + 06	1.25e + 12***	1.06***
usd amount	701.84	2.41e + 04	1	6.25e + 06	8.60e + 12***	0.7***

Panel B: Discrete variables			
Variable	**Unique values**	**Top frequent value**	**Frequency of top value**
Aggregate data set			
collection name	506	Foundation (FND)	21.96
deployer creator generalities	8	Company	40.03
deployer creator name	287	Pak	3.49
discord account	2	1	61.85
discord server	237	SuperRare	7.67
followers count	408	334,093	21.96
following count	349	854	21.96
like count month	404	10,452	21.96
listed count	209	1	23.35
marketplace collection	2	0	70.09
nft type	2	erc721	97.76
num of colors	9	9	33.01
num tweets	377	2397	21.96
num of owners	156	2	77.91
num of trades	149	1	77.87
platform of last sale	6	OpenSea	69.49
quote count month	380	300	21.96
reply count month	392	1274	21.96
retweet count month	401	1284	21.96
twitter account	2	1	97.58
twitter handle	417	FND	21.96
verified	2	0	64.11
Transaction data set			
buyer	93,184	0 × 8888888888e9997e...	0.63
collection name	506	FND	15.64
nft	238,420	0x73da73ef3a6982109c4.../8	0.69
num of owners	156	2	53.82

TABLE 1 (Continued)

Panel B: Discrete variables			
Variable	**Unique values**	**Top frequent value**	**Frequency of top value**
num of trades	149	1	53.73
platform of last sale	6	OpenSea	75.11
seller	72,618	0x8c9f364bf7a56ed058...	1.11

the financial literature, and it frequently results from complex factors, such as rarity and artist recognition. Rarity acts as a recognized determinant of value in both conventional and digital art markets, creating a perception of exclusivity and uniqueness (Renneboog & Spaenjers, 2013; Schaar & Kampakis, 2022). In the realm of NFTs, rarity's definition further refines and authenticates via the blockchain technology, thereby augmenting its attraction. Conversely, artist recognition plays an instrumental role in the valuation of art pieces, and eminent artists demand elevated prices that reflect their consolidated reputation and brand (Mandel, 2009).

Examining the discrete variables panel provides insights into the structural composition of the NFT ecosystem. From the six NFT marketplaces included in our analysis, namely, *OpenSea*, *Rarible*, *NFTX*, *LooksRare*, *Foundation* and *SuperRare*, the dominance of *OpenSea* is prevalent, capturing 69.49% of all transactions. The platform's dominance is congruent with academic research on network effects, where platforms that captivate more users gain advantages in terms of augmented liquidity and information dissemination, thereby fortifying their market dominance (Parker & Van Alstyne, 2005; Wilson et al., 2021). We observe that most NFTs are created through companies, accounting for 40.03% of total NFT creations. Notably, *Foundation* stands out by capturing 21.96% of the total number of created NFTs. This could be attributed to the user-friendly interface provided by *Foundation* and similar marketplaces, which allow NFT creation without the need for programming skills. *Foundation*'s influence extends to the art NFT sales, where it accounts for 15.29% of the market. This is probably because it curates pieces from distinguished artists, which aligns with scholarly research concerning consumer behaviour, where branded collections often evoke greater trust and thus attract a larger number of buyers (Rojas-Lamorena et al., 2022).

Analogously, Table 2 illustrates the dynamics of herding behaviour in the art NFT market, providing a detailed analysis of the patterns and tendencies among investors. This table provides descriptive statistics for the continuous and discrete variables used in the study of herd bias. It examines how investor decisions are influenced by the actions of others, leading to collective market trends. Finally, Table A1 in Appendix A provides a brief description of the quantitative variables presented in Tables 1 and 2.

Following the data extraction, we rely on exploratory data analysis to identify key characteristics of our data set. By analyzing the number of sales and sales volume in USD per year, we observe that art NFTs followed the overall market trend and skyrocketed in 2021. Specifically, 2018 and 2019 account for just 0.72% of the number of sales and 0.02% of the USD volume, while 2021 accounts for 93.88% of sales and 99.21% of the volume. Focusing on 2020 and 2021, we see a similar trend when examining the average monthly volume and sales price. The volume exhibits a strong uptrend for both years, however, the average price, despite initially following a similar pattern, starts plateauing after February 2021. This indicates that the volume increase in the art market in 2021 was mostly due to an increase in the number of sales and not in price.

TABLE 2 Herding behaviour descriptive statistics.

This table reports descriptive statistics for the continuous and discrete variables used in the study of herd bias. The continuous variables are split into two data sets: the Artists data set and the Transactions data set. Each variable's median value is displayed, along with its standard deviation, minimum and maximum values. The Jarque–Bera statistic (JB Stat.), which tests for normality in the data distribution, is also provided. A triple asterisk denotes a significance level of $p \le 0.01$. The Kwiatkowski–Phillips–Schmidt–Shin (KPSS) test results for stationarity are also shown, with a single asterisk denoting $p \le 0.1$. For instance, in the Artist Data Set, *sold ratio*, which shows that the average artist sells only 2% of their work, reveals a great disparity as some artists manage to sell all their work. The wide-ranging *average price* highlights the diversity in the pricing strategy across artists, with the median being 556, while some nonfungible tokens (NFTs) are priced, on average, as high as 1.35e + 06. Trade activity, represented by the *num of trades*, is generally limited, given the fact that 25% of NFTs are sold twice. The *buyer seller pair* variable confirms this trend, as despite the top frequent value being just 1 unique pair, some artists attract thousands of unique pairs. NFT selling times, captured by *timediff*, typically do not exceed three months, but some artworks may require over three years to find a buyer. The *Volume* variable, a measure of an artist's monthly market activity, further supports this imbalance: while the median market activity is 4492.12, the most popular artists can stir activity levels up to 1.87e + 07. *Share* suggests that the NFT market can be dominated by a few artists, with some accounting for up to 82% of the total monthly market capitalization, despite the median share being zero. Finally, the *Demand* variable summarizes the shifting nature of consumer preferences in the NFT market, with a wide range from 0.97 to 3570.16. In the Transactions data set, discrete variables such as *discord account* and *deployer artist* provide additional insights. Most artists appear to have a Discord account, while independent artists constitute the majority. $^{*}p < 0.1$; $^{**}p < 0.05$; $^{***}p < 0.01$.

Panel A: Continuous variables

Variable	Median	σ	Minimum	Maximum	JB Stat.	KPSS test
Artists data set						
average price	556	4.05e + 04	2.06	1.35e + 06	3.44e + 07***	0.15*
cum score	1.77e + 07	1.95e + 07	788.19	7.13e + 07	259.43***	7.9***
cum score year rank	443.5	467.63	1	1576	228.6***	7.94***
market cap	4492.12	1.01e + 06	2.41	1.87e + 07	2.08e + 06***	0.28***
demand	0.97	114.36	0	3570.16	1.85e + 07***	0.42*
score	1170.45	1.71e + 05	62.28	3.12e + 06	1.98e + 06***	0.28*
share	0	0.05	0	0.82	1.16e + 06***	0.11*
sold ratio	0.02	0.16	0	1	2.45e + 04*	0.14*
timediff	88.6	186.77	1.04	1231.58	7891.99***	0.32*
volume	4492.12	1.01e + 06	2.41	1.87e + 07	2.08e + 06***	0.28*
Transactions data set						
CSAD	0.06	0	0.03	0.1	16.29***	3.03***
market return	0.02	0.01	0	0.04	19.72***	2.92***
newcomer proportion	0.24	0.1	0.04	0.67	26.51***	1.25***

Panel B: Discrete variables

Variable	Unique values	Top frequent value	Frequency of top value
Artists data set			
buyer seller pair	344	1	13.09

TABLE 2 (Continued)

Panel B: Discrete variables			
Variable	**Unique values**	**Top frequent value**	**Frequency of top value**
deployer creator generalities	2	Individual	87.08
discord account	2	Yes	54.07
num of trades	842	2	25.38

We note that 71.9% of the examined NFTs have never been sold. Of those sold, only 22.1% were sold more than once. However, NFTs that are sold more than once have a higher chance of being sold a third time, with the probability increasing to 27.1%, which further increases to 30.71% for the fourth sale.

The observed illiquidity and fragmentation in the art NFT market resonate with characteristics commonly associated with emerging markets (Amihud & Mendelson, 1986; Kirilenko et al., 2017; So & Wang, 2014) where high volatility and uncertainty prevail.

These trends suggest a dynamic market shaped by a blend of intrinsic and extrinsic factors. In particular, extrinsic factors, such as market hype and overall crypto market trends, add further complexity to pricing dynamics. Similar to phenomena observed during technology bubbles (Ofek & Richardson, 2003), the influence of social media and broader market sentiment can create rapid price swings and speculative behaviours.

3 | APPROACH

The methodology section of our analysis is split into two distinct parts, each addressing a unique aspect of the art NFT market. In the first part, we undertake an evaluation of the art NFT pricing structures. Indeed, the evolving landscape of financial markets, particularly with the emergence of new asset classes like art NFTs, has brought to the forefront a critical reassessment of traditional valuation models. In financial literature, there's a burgeoning consensus that these novel assets defy conventional categorization and valuation metrics (Chanel et al., 1996; Nadini et al., 2021; Rosen, 1974), primarily due to their distinct blend of financial, artistic and social attributes. Art NFTs, unlike traditional assets, encapsulate a unique intersection of digital scarcity, artistic expression and community-driven value, challenging the traditional paradigms of intrinsic value assessment. This divergence from established valuation norms necessitates a more adaptive and multifaceted approach to understand their market behaviour:

> **Hypothesis 1.** A holistic and robust price index for art NFTs can be developed by integrating financial, artistic and social network features, providing a more accurate reflection of their value than traditional financial metrics alone.

This hypothesis arises from recognizing that such assets lack intrinsic fundamental value (Taleb, 2021). Our approach seeks to establish a pricing model for art NFTs that goes beyond traditional financial valuation, encapsulating the complex interplay of artistic merit and social influence. To this end, we adopt a two-sided approach that employs both hedonic and RSR models. The hedonic pricing model is a revealed preference method used to estimate the influence of various features on the price of a good, and it is particularly suited for studying markets where the goods being transacted are highly differentiated (Rosen, 1974), like, the art NFT market. The RSR model, on the other hand, focuses on items that have been sold more than once, thus allowing us to control for the intrinsic qualities of each item and isolate the pure price effect over time. This method is designed to comprehensively capture the intricate dynamics among financial, artistic and social aspects that jointly determine the valuation of art NFTs.

The second part of this methodology section analyzes the herd bias within the art NFT market. In this setting, it becomes imperative to consider the possibility that investor rationality, a cornerstone assumption in traditional financial markets (Amihud & Mendelson, 1986; Campbell & Shiller, 1988), may not hold true in this novel and evolving landscape. The art NFT market, characterized by its nascent stage, digital nature and unique asset properties, presents a scenario where traditional notions of market behaviour and investor rationality may be challenged or redefined.

The concept of herding behaviour, which implies investors making decisions based on the actions of others rather than based on independent analysis or changes in intrinsic asset values, becomes particularly important in this context. This is compounded by the unique attributes of the NFT market, such as its relatively low liquidity and high information asymmetry (Li & Chen, 2023). These factors can create an environment conducive to herding, where investors are more likely to follow market trends and collective sentiment rather than basing their investment decisions on fundamental analysis:

Hypothesis 2. Herding behaviour significantly influences the valuation cycles of art NFTs, with investor actions conforming to group dynamics, especially in scenarios of high market volatility and information asymmetry.

We utilize the CSAD as a key analytical tool to analyze herding behaviour and artist-driven market trends within the art NFT market. CSAD is instrumental in revealing how market participants react to volatility and conform to prevalent market trends, often prioritizing these over fundamental analysis. This trend is especially noticeable in markets fraught with uncertainty and information asymmetry, characteristics intrinsic to the art NFT sector (Bao et al., 2023). The low liquidity of the NFT market can actually amplify herding effects because a few investors' actions can disproportionately influence prices, prompting others to follow (Azarmi & Menny, 2013; Banerjee, 1992; Demirer & Kutan, 2006). As such, a small number of transactions can excessively influence the market, leading to a scenario where prices comove more due to investor behaviour rather than underlying fundamentals. The strength of the CSAD model lies in its ability to identify herding by assessing the comovement of NFT prices, observing when prices move more uniformly than fundamentals would suggest. The model's efficacy lies in its ability to detect this uniformity in price movements. In a normal market scenario, where decisions are made independently, price movements of different assets would be more varied and less correlated. However, in situations where herding is present, there is a notable convergence in the way prices of different assets move, indicating that investors are likely following similar cues or strategies, rather than basing their decisions on changes in intrinsic asset values. The CSAD model's applicability in low liquidity markets like NFTs is underscored by studies that have examined herding behaviour in similar contexts, such as small-cap stocks or emerging market equities (Chang et al., 2000; Park & Sabourian, 2011; Yousaf & Yarovaya, 2022).

3.1 | Pricing dynamics

The hedonic regression approach applied to the study of art NFTs price dynamics is rooted in the consumer theory of utility maximization, where NFTs are considered bundles of characteristics, and the price reflects the consumer's willingness to pay for these characteristics (Lancaster, 1966; Rosen, 1974). This approach has been widely used to study price determinants in various markets, such as real estate (Muehlenbachs et al., 2015), traditional art (Adams et al., 2021; Aubry et al., 2023; Chanel et al., 1996; Renneboog & Spaenjers, 2013) and more recently, in the analysis of digital assets like NFTs (Borri et al., 2023; Horky et al., 2022).

For the analysis of art NFTs in this study, we commenced by employing a hedonic regression model, described by the following equation:

$$\ln(y_i + 1) = \beta_0 + \boldsymbol{\beta}\mathbf{X_i} + \gamma D + \delta T + \epsilon_i, \tag{1}$$

where y_i is the average price of the NFT, $\mathbf{X}$ represent a vector of continuous variables, D denotes control variables, T stands for time fixed effects and ϵ is the error term. Here, the continuous variables include features such as colour proportions, Shannon's entropy and floor price, reflecting both aesthetic and complexity aspects of the NFTs. The control variables represent a set of dummy variables accounting for intrinsic characteristics (e.g., the number of colours in each image) of the digital assets and presence on social media. Time coefficients are incorporated to construct a price index, accounting for temporal variations and market trends.

To correct the potential selection bias in observed price data, we performed a Heckman two-stage regression. Selection bias arises when the sample selection for observation is not random, and there is a correlation between the observable selection and the unobservable factors affecting the dependent variable (Korteweg et al., 2016). In the context of art NFTs, this bias can occur when only the prices of traded NFTs are observed, while unsold NFTs remain unobserved. This nonrandom selection can create a bias in estimating the relationship between price and characteristics of the NFTs, as the unobserved factors affecting the decision to trade might also influence the price. To correct for this, the inverse Mills ratio (IMR) is derived from the Probit estimation and encapsulates the likelihood of the NFT being traded. This IMR is incorporated into the hedonic regression as a new variable, λ, ensuring unbiased estimates:

$$\lambda = \frac{\phi(\gamma\mathbf{X})}{\Phi(\gamma\mathbf{X})}. \tag{2}$$

The model corrects for the selection bias by capturing the unobserved factors affecting both the selection process and the dependent variable:

$$\ln(y_i + 1) = \beta_0 + \boldsymbol{\beta}\mathbf{X_i} + \gamma D + \delta T + \theta\lambda_i + \epsilon_i. \tag{3}$$

The two-stage approach ensures that the relationship between the price and characteristics of NFTs is estimated without the bias induced by nonrandom selection.

We addressed multicollinearity using the Variance Inflation Factor method, ensuring that the explanatory variables were not highly correlated (O'Brien, 2007). Furthermore, we assume that our data respect all the assumptions of the hedonic and Heckman models, such as linearity, homoscedasticity and the normality of error terms and residuals.

In the context of hedonic modelling for art NFT prices, the time-varying and invariant variables are crucial in determining the underlying price index. However, it is vital to ensure that the hedonic regression parameters are stable over time. If not, this instability can lead to biased estimates of time dummy coefficients and consequent price indexes (Kuminoff et al., 2010).

To mitigate this issue, we employ the chained Fisher index, derived from sectional hedonic regressions, following the methodology used in classical financial literature (Diewert, 1976; Triplett, 2004). Specifically, we construct the Fisher index by first calculating the Laspeyres and Paasche indexes and then obtaining the geometric mean of these two indexes. Given the estimated β coefficients from the hedonic regression and the normalized weights q for each characteristic, the Laspeyres index for a given period t is calculated as

$$Laspeyres_{\,t} = \frac{\sum_{j=1}^{J} \beta_{j,t+1} q_{j,t}}{\sum_{j=1}^{J} \beta_{j,t} q_{j,t}}. \tag{4}$$

Similarly, the Paasche index for the same period is determined as

$$Paasche_{\,t} = \frac{\sum_{j=1}^{J} \beta_{j,t+1} q_{j,t+1}}{\sum_{j=1}^{J} \beta_{j,t} q_{j,t+1}}. \tag{5}$$

The Fisher index is then the geometric mean of these two indexes:

$$F_{t+1} = \sqrt{Laspeyres_t \times Paasche_{\,t}}. \tag{6}$$

The Fisher index can be chained across time to generate a consistent price index for art NFTs.

In the final step, we employ RSRs. This technique focuses specifically on the items that have been sold multiple times, thereby controlling for the unobservable heterogeneity of those assets (Bailey et al., 1963; Korteweg et al., 2016; Mei & Moses, 2002). This approach offers a complementary perspective to hedonic regression by capturing price changes for identical NFTs, thus eliminating the need to control for detailed characteristics. The model is estimated with the regression:

$$\ln(y_{i,t}) = \beta_0 + \sum_{t=1}^{T} \beta_t D_{i,t} + \epsilon_{i,t}, \tag{7}$$

where $y_{i,t}$ represents the ratio between the prices of two consecutive sales, β_t the coefficients for the dummy variable $D_{i,t}$ with a value of 1 in the period when the resale occurs, −1 in the period of the previous sale, and 0 otherwise and $\varepsilon_{i,t}$ is the error term.

However, a significant concern in RSRs is, again, the potential selection bias and the heteroskedasticity of the error term. Indeed, as specified before, in RSR, the sample only consists of NFTs that have been sold multiple times. This specific selection can introduce bias since NFTs that are sold more than once may not be a random subset of all NFTs. They might have unique characteristics that make them more or less likely to be resold, leading to a systematic deviation from the overall population. To correct for this bias and account for the varying holding periods between repeat sales, we utilize the Case Shiller three-stage regression method (Case & Shiller, 1989; Korteweg et al., 2016).

We begin by conducting the RSR as previously described, extracting the residuals. This initial step lays the foundation for modelling the heteroskedasticity in the error term. Such nonconstant variability could be related to factors like the holding period, which might have varying effects on the error term across different observations. We regress the squared residuals against a constant and the specific holding periods of art NFTs:

$$u_t^2 = \gamma_0 + \gamma_1 \times holding\_period + \zeta_t. \tag{8}$$

By modelling the squared residuals in this manner, we capture the pattern of heteroskedasticity and its relationship with the holding period. Finally, we fit the original RSR using Generalized Linear Models with weights derived from the reciprocal of the square root of the fitted values from the second stage.

3.2 | Herd bias

The notion of herding behaviour refers to the propensity of market participants to conform to the investment decisions of a larger group, often at the expense of their own available information. In the context of the art NFT market, this can manifest in the form of preferences aligning with recognized or leading artists, thereby creating self-reinforcing patterns of demand. Such mechanisms can lead to a superstar phenomenon where a small number of artists dominate the market. This phenomenon has been empirically investigated in traditional art markets (Rosen, 1981), but the exploration in digital markets is new.

Our analytical framework encompasses three primary dimensions: (1) artistic quality as reflected by market metrics, (2) herding behaviour rooted in classical financial theories and (3) the intricate relationship between historical performance, popularity and influence factors like social media engagement. This last dimension, particularly, provides insights into how external elements, such as an artist's Discord account presence or the artist status, interact with traditional market dynamics to shape an artist's ranking and demand. By leveraging the OLS estimation approach of the multinomial logit model of Azarmi and Menny (2013), we collectively analyze consumer choices based on the individual market shares of artists rather than focusing on separate purchasing decisions.

Our analytical representation takes the following form:

$$\log(\Pi_{it}) = \alpha_i + \alpha_t + \sum_{j=1}^{J} \beta_j (X_{ijt} - \bar{X}_{jt}) + \epsilon_{it}, \tag{9}$$

where the intrinsic value of an artwork created by artist i is symbolized by a_i, the j-th explanatory variable is denoted by X_{ijt} and ε_{it} refers to the error term of the model. The corresponding arithmetic means of X_{ijt} are represented by $\overline{X}_{jt}$. Π_{it} symbolizes the selection made by art enthusiasts and is rooted in the relative demand attributed to artist i at time instance t. We define this relative demand as

$$\Pi_{it} = \frac{S_{it}}{\widetilde{S}_t} = \frac{\frac{V_{it}}{V_t}}{\sqrt[n]{\prod_{j=1}^{n} \frac{V_{jt}}{V_t}}}. \tag{10}$$

In this model, the relative demand Π_{it} is derived by dividing an artist's market share S_{it} by the geometric mean of the individual market shares across all artists, represented by $\widetilde{S}_t$. The term S_t is ascertained by dividing the monthly market capitalization V_{it} specific to an artist by the cumulative market capitalization V_t for all n artists at time t. We define the market capitalization for each artist similarly to how it's calculated in the stock market, by multiplying the number of artworks sold by the average price of these artworks for each month and summing the total for each artist.

We assume that consumers do not have predetermined preferences for a specific artist or art style within a given sample category. For instance, an individual evaluating works from a particular NFT collection would exhibit equal interest in diverse artistic expressions, irrespective of the creator.

As an explanatory variable, the cumulative score of the annual ranking lagged by one period captures the inertia in an artist's popularity, reflecting how past performance continues to affect current standings. The yearly rank of the artist and the previous monthly rank provide insights into the temporal dynamics of an artist's market position. The interaction terms in the regression model serve specific purposes in understanding the dynamics of the art NFT market. The complexity of the art NFT market is further dissected through specific interaction terms in the model. The term involving yearly rank and Discord account ($year * discord\_account$) captures the influence of social media engagement on an artist's rank, reflecting how an active presence on platforms like Discord can amplify the artist's visibility and market appeal. The interaction between yearly rank and the artist status ($year * deployer\_artist$) distinguishes between artists who are independent creators and those associated with companies. In the art NFT market, the interaction effects help explain how information cascades might form. For instance, a high yearly rank, coupled with active engagement on Discord, might create a positive feedback loop where popularity leads to more visibility, further enhancing demand. Conversely, the artist's status might affect how information about the artist disseminates, with individual artists potentially benefiting from a more personalized connection with collectors.

We complement and enrich the study of herding behaviour and artist stardom by examining the CSAD. This method serves as an analytical tool for discerning how market participants adapt to fluctuations within the market. By evaluating the cumulative sum of deviations from a certain trend or mean, we identify instances of herding behaviour, where investors are inclined to follow prevailing trends rather than base their decisions on fundamental analysis. Such herding behaviour forms a common thread between the stardom effect and cascade theory, as well as the study of CSAD. These interconnected areas pertain to the manner in which investors or market participants might succumb to trends, overlook private information or make determinations grounded in the conduct of others, rather than a rational evaluation of value. In the context of NFTs, being relatively nascent and heavily influenced by social dynamics, celebrity endorsements and market hype, often leads to behaviours where individuals align with the crowd instead of pursuing independent analysis. We recognize this alignment as a phenomenon where a few artists dominate the market.

Initially, we calculate daily returns of individual art NFTs as in Bao et al. (2023) and we adapt the CSAD measure, as formulated by Chang et al. (2000):

$$CSAD_t = \frac{1}{N}\sum_{i=1}^{N}|R_{i,t} - R_{m,t}|. \tag{11}$$

In the NFT market, which is characterized by low liquidity, the price movements are often more pronounced due to the significant impact of trades in the market. When herding occurs, investors tend to move collectively, leading to a decrease in the dispersion of returns across different assets, that is, they start moving more in sync with each other. In such scenarios, the CSAD value tends to decrease as the absolute deviations of individual asset returns from the mean return diminish. This uniformity in price movements, despite the low liquidity, is a key indicator of herding.

The effectiveness of CSAD in capturing herding in low liquidity environments lies in its sensitivity to the convergence of asset returns. When investors herd, the correlation between individual asset returns and the market average increases, reducing the CSAD value. This contrasts with a market dominated by independent decision-making, where one would expect a higher level of dispersion in returns, and hence, a higher CSAD value. Further, to enhance the robustness of the findings and capture the dynamic interactions in the market, we extend the methodology by introducing a lagged dependent variable, as advised by Fu and Wu (2021):

$$CSAD_t = \alpha + \gamma_1|R_{m,t}| + \gamma_2 R_{m,t}^2 + \gamma_3 CSAD_{t-1} + \varepsilon_t. \tag{12}$$

The convergence between individual and market returns in this model reflects the manifestation of artist stardom, where market preferences tend to converge towards specific artists.

To delve deeper into the dynamics of herding, we employ the fixed transition probabilities Markov-switching model (MSM) (Diebold & Rudebusch, 1999), integrating the methodology developed by Bao et al. (2023). This approach enables the distinction between different market regimes, revealing how herding behaviour fluctuates over time:

$$CSAD_t = \begin{cases} \alpha + \gamma_{1,1}\,|R_{m,t}| + \gamma_{2,1}R_{m,t}^2 + \gamma_3 CSAD_{t-1} + \varepsilon_{1,t} & \text{if } s_t = 1, \\ \alpha + \gamma_{1,2}\,|R_{m,t}| + \gamma_{2,2}R_{m,t}^2 + \gamma_3 CSAD_{t-1} + \varepsilon_{2,t} & \text{if } s_t = 2, \end{cases} \tag{13}$$

with transition probabilities expressed as

$$P(s_{t+1} = i \mid s_t = j) = p_{ij}, \quad i, j \in \{1, 2\}.$$

The transition matrix for this first-order Markov-switching process takes the form

$$P = \begin{pmatrix} P_{1,1} & 1 - P_{2,2} \\ 1 - P_{1,1} & P_{2,2} \end{pmatrix},$$

where $0 < P_{i,i} < 1$ and $0 < P_{j,j} < 1$. In this context, Regime 1 represents the state with herding, characterized by a negative $\gamma_{2,1}$, while Regime 2 denotes the state without herding, indicated by positive or insignificant $\gamma_{2,2}$.

The constant term α serves as a nonswitching parameter, symbolizing the long-term average of $CSAD_t$, while the transition probabilities, $P(s_{t+1} = i|s_t = j)$, govern the progression between these two states. The probability $1 - P_{2,2}$ specifically elucidates the likelihood of the emergence of herding behaviour.

TABLE 3 Hedonic and Heckman two-stage model estimations.

This table reports the estimation results of hedonic models (Models 1, 2 and 3) and Heckman two-stage models (Models 1', 2' and 3') applied to the average price of art nonfungible tokens (NFTs) collected from February 1, 2020, to February 10, 2022. The dependent variable is the logarithm of the average price plus one. The models progressively include additional controls: Models 1 and 1' do not include social network controls (*i.e., verified, twitter account, discord account, deployer creator generalities*) or market dummy variables (*i.e., marketplace collection, platform of last sale, nft type*), Models 2 and 2' introduce social network controls and Models 3 and 3' further include market dummy variables. Month fixed effects (FE) are included in all models. We acknowledge that the variables *Mar-21*, *Apr-21*, *Jul-21*, *Aug-21*, *Sep-21*, *Oct-21*, *Nov-21*, *Dec-21*, *Jan-22*, *Feb-22*, *marketplace collection*, *black* and *white* exhibited VIF values greater than 10. Upon reevaluation with these time FEs and other variables removed, the magnitudes of the regression coefficients for the vast majority of variables remained practically unchanged. This suggests that the high VIFs mainly influenced the redistribution of effects among the time FEs. Furthermore, the coefficients on the colors *gray*, *green*, *blue*, *yellow*, *purple*, *orange* and *red* are not reported in the table. These have coefficients of magnitude and sign similar to those obtained for *black* and *white*. Each entry in the table is the coefficient estimate from the respective model, with the standard error reported in parentheses. Asterisks denote significance levels: *denotes significance at the 10% level, **at the 5% level and ***at the 1% level. The number of observations for all models is 238,420. The adjusted R^2, residual standard error and F statistics are reported at the bottom of the table. The *mills ratio* row represents the Inverse Mills Ratio, which is included only in the Heckman models (1', 2' and 3'). $^{*}p \leq 0.1$; $^{**}p \leq 0.05$; $^{***}p \leq 0.01$.

Dependent variable: log(*average price* + 1)						
	(1)	**(1′)**	**(2)**	**(2′)**	**(3)**	**(3′)**
const	−0.495***	−0.204***	−0.494***	−0.181***	−0.738**	−0.400***
	(0.021)	(0.013)	(0.021)	(0.014)	(0.022)	(0.016)
black	−0.266***	−0.264***	−0.259***	−0.257**	−0.264***	−0.261***
	(0.011)	(0.011)	(0.011)	(0.011)	(0.011)	(0.011)
white	−0.257***	−0.257***	−0.252***	−0.250***	−0.248**	−0.247***
	(0.011)	(0.011)	(0.011)	(0.011)	(0.011)	(0.011)
Norm_Shannon_entropy	−0.006	−0.007	0.025***	0.020**	0.047**	0.042***
	(0.009)	(0.009)	(0.009)	(0.009)	(0.009)	(0.009)
Num_of_colors	0.001	0.001	0.001	0.000	−0.001*	−0.001
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
BGCI_return	0.086***	0.252***	0.077***	−0.027	0.070**	−0.001
	(0.023)	(0.052)	(0.023)	(0.053)	(0.023)	(0.053)

TABLE 3 (Continued)

Dependent variable: log(*average price* + 1)						
	(1)	**(1′)**	**(2)**	**(2′)**	**(3)**	**(3′)**
log(*number_of_trades* + 1)	1.065***	1.067***	1.068**	1.071***	1.094***	1.097***
	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)
log(*floor_price* + 1)	0.935***	0.936***	0.931***	0.932***	0.925**	0.927***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
log(*following_count* + 1)	−0.003***	−0.003***	−0.007**	−0.007***	−0.003***	−0.004***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
log(*listed_count* + 1)	0.024***	0.023***	0.025**	0.025***	0.018**	0.017***
	(0.000)	(0.000)	(0.001)	(0.000)	(0.000)	(0.001)
log (*quote_count_month* + 1)	0.001	−0.000	−0.015***	−0.015***	−0.008***	−0.008***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
log(*timediff* + 1)	0.030**	0.030**	0.023***	0.023***	0.024***	0.024***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Mills_ ratio		−0.044***		0.023*		0.015
		(0.012)		(0.012)		(0.012)
Social network controls		×		✓		✓
Market FE		×		×		✓
Month FE		✓		✓		✓
Other colors controls		✓		✓		✓
Observations				238,420		
Adjusted R^2	0.933	0.933	0.933	0.933	0.934	0.934

4 | DISCUSSION OF RESULTS

4.1 | Price indices design

For the construction of hedonic indices, we select sold non-wash traded NFTs starting from Janurary 1, 2020, due to the rarity of the transactions before that date, and we are left with 238,420 unique NFTs. We identify several significant pricing determinants for art NFTs. Table 3 illustrates our hedonic regression results. A pivotal outcome, observed consistently across all our models, is the positive and significant coefficient on the log transformed variables *number_of_trades* and *floor_price*, with values around 1.07 and 0.93, respectively. This suggests that NFTs with a higher number of trades and a higher floor price are likely to command a higher average price. This finding aligns with existing literature on asset pricing, such as Campbell and Shiller (1988) and Fama and French (1993), highlighting the relevance of these factors in financial markets. In particular, the variable *floor_price* may be capturing the price floor effect, suggesting that the minimum acceptable price for the NFT could be a significant determinant of its average price. As Kahneman and Tversky (1979) illustrate, this minimum price often sets a psychological benchmark or reference point for transactions, influencing final prices. In a similar vein, the role of a base price or minimum price as a determinant of the final price of an asset has been emphasized by Campbell and Shiller (1988). In the context of our study, a higher floor price could convey a higher perceived value of the NFT, resulting in a higher average price. Further, the floor price might encapsulate the price floor effect. In auction theory, Milgrom and Weber (1982) underscore the importance of the reserve price (the lowest price at which the seller is willing to sell the item) in determining the final price. This concept could apply to NFTs as well, where the floor price acts as a reserve price and impacts the final price of the NFT.

On the other hand, the negative coefficients on the log transformed variables *following_count* and *quote_count_month* are indeed unexpected, given that one would typically expect higher levels of social media activity to positively influence the price of NFTs. However, our regression results suggest that this is not the case for the art NFT market, at least within the context of the observed data. One possible explanation for these counterintuitive results is the over-saturation hypothesis: When an NFT or a collection of NFTs is excessively quoted or posted about on social media, it could lead to an over-saturation of information. This phenomenon could reduce the perceived rarity associated to the NFT, thereby driving down its price. This idea aligns with the findings of Dellarocas (2003) and BabićRosario et al. (2020), which discuss the potential downsides of excessive online exposure.

The application of Heckman's two-step model, as exhibited in models (1′), (2′) and (3′), reveals insights about the potential selection bias in the study of art NFT pricing dynamics. The IMR (λ) for the model (1′), derived from the Probit model in the first step of the Heckman process, is significant and negative in magnitude, confirming the presence of selection bias in the sample of traded NFTs. The negative sign of the IMR suggests that the unobserved factors influencing the decision to trade NFTs are inversely correlated with the error term in the average price equation. This implies that those NFTs less likely to be traded (due to unobserved factors) tend to have higher prices than what could be predicted based solely on their observable characteristics. Higher-priced NFTs, which are less likely to be traded, might contain superior artistic or unique features not captured by the observable characteristics.

In terms of the robustness of the results, the consistency of most variable coefficients across both the hedonic and Heckman models indicates that our findings are resilient to adjustments for selection bias. However, some variables, such as the return on the Bloomberg Galaxy Crypto Index (*BGCI_return*), demonstrate significant changes in their coefficients when transitioning from the hedonic models to the Heckman models. The coefficients' substantial increase in models (1) and (1′) suggests that the relationship between these variables and NFT prices is subject to the influence of selection bias. As such, the return on the Galaxy Crypto Index could be capturing some of the unobserved factors affecting the decision to trade NFTs. Previous studies have found a significant impact of selection bias on the relationship between asset prices and market-wide indicators (Ang et al., 2020; Bianchi & Babiak, 2022). The pronounced positive coefficient in our model suggests a correlation between the performance of cryptocurrencies, as measured by the Bloomberg Galaxy Crypto Index (BGCI), and the pricing of NFTs.

The application of the RSR model drastically reduced the number of observations from 238,420 to 107,101. This is expected as the RSR model only considers assets that have been traded more than once, which is a crucial characteristic of the RSR model that leverages repeated sales data to control for the unobserved time-invariant characteristics of the assets. Figure 1 draws the estimated art NFT price indices.[1] We observe that all four indices—the hedonic, hedonic-Heckman, repeat sales and repeat sales–Case Shiller—display a similar trend of price growth over the period from August 2020 to February 2022. There is a particularly noticeable spike in price growth around February 2021. This surge in price indices coincides with several significant events in the NFT market. For instance, in March 2021, a digital artwork by the artist Beeple was sold as an NFT at Christie's auction for a staggering $69 million. This sale not only marked a historic moment for digital art but also significantly boosted the perceived value of NFTs. Furthermore, in the same month, the first-ever tweet by Twitter CEO Jack Dorsey was sold as an NFT for $2.9 million. These high-profile sales events likely contributed to the increased attention and perceived value of NFTs, which is reflected in the sharp increase in the price indices around this time.

In the subsequent period, from April to June 2021, the indices had a sudden drop. This could be linked to several notable developments in the NFT market. Between February 2021 and June 2021, the NFT market, experienced a significant decline due to an unsustainable surge in prices driven by speculative investments, which led to a market correction as initial excitement waned. This period also saw increasing regulatory scrutiny and uncertainty, with the rapid advancements in NFT technology outpacing regulatory frameworks, creating instability and eroding investor confidence. Concurrently, the broader cryptocurrency market experienced a downturn, with major cryptocurrencies like Bitcoin and Ethereum seeing significant price drops, which further impacted NFT transactions. Additionally, the market was flooded with low-quality projects and a growing realization that many NFTs had little intrinsic value, leading to oversupply and decreased demand, causing many NFTs to become virtually worthless. The sharp spike in the indices around July 2021 coincides with significant investments in the NFT space. At that time, OpenSea, one of the largest NFT marketplaces, announced that it had raised $100 million in a Series B funding round led by AH Capital Management. This

[1]Since the RSR model does not effectively handle sales occurring within the same time period—relying instead on time variation between sales to identify price changes and control for time-invariant characteristics—we acknowledge this limitation in our data. To mitigate this, we constructed repeat sales indices on monthly, weekly, and daily bases. We opted for the daily index because it minimized the number of discarded transactions (35.50%). In Figure 1, we aggregated the repeat sales indices to a monthly level by averaging the parameters within each timeframe.

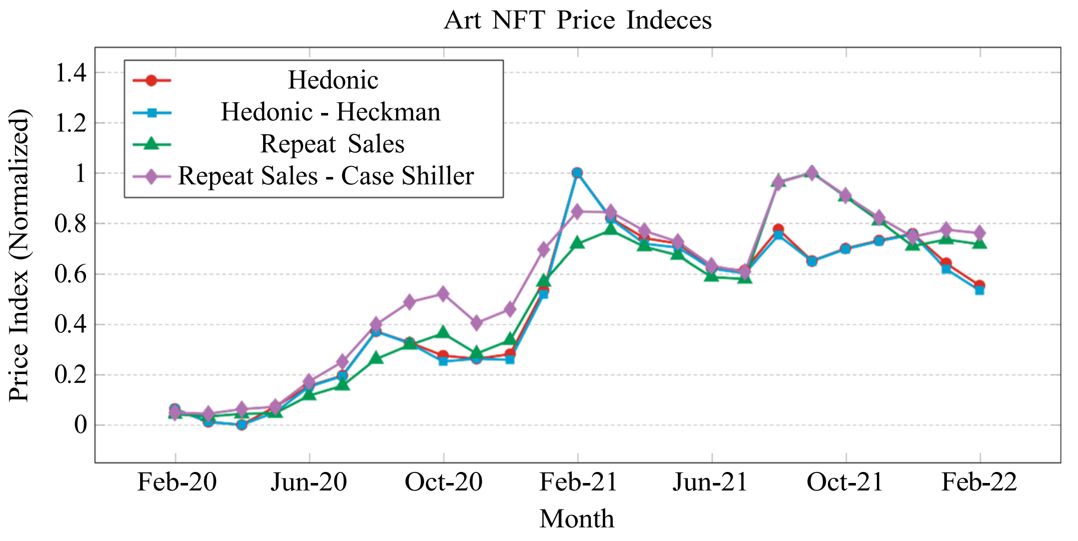

FIGURE 1 Art nonfungible token (NFT) price indices. This figure presents the four different art NFT price indices: the hedonic, hedonic-Heckman, repeat sales and repeat sales-Case Shiller. These indices trace the evolution of art NFT prices from February 2020 to February 2022. The hedonic model, for instance, evaluates the prices of art NFTs based on their observable characteristics, while the repeat sales model focuses on price changes for identical art NFTs traded over time. The hedonic-Heckman model augments the hedonic model by correcting for potential selection bias. The repeat sales-Case Shiller model extends the repeat sales model by accounting for both heteroskedasticity and selection bias. From August 2020 to February 2022, the hedonic, hedonic-Heckman, repeat sales and repeat sales-Case Shiller indices all displayed similar price growth trends, with a noticeable spike around February 2021. This spike appears to coincide with significant events in the NFT market, such as Beeple's $69 million NFT sale at Christie's and the $2.9 million sale of Jack Dorsey's first tweet, potentially boosting NFT values. From April to June 2021, the indices dropped, likely resulting from a market correction driven by speculative investments, regulatory scrutiny and a broader cryptocurrency downturn. In July 2021, a spike in indices may be attributed to OpenSea's $100 million funding round. [Color figure can be viewed at wileyonlinelibrary.com]

investement valued the company at $1.5 billion. This massive investment likely boosted market confidence and contributed to the surge in NFT prices reflected in the price indices.

After correcting for selection bias, Figure 1 shows that the repeat-sales–Case Shiller model, which corrects for both heteroskedasticity and selection bias, is slightly more volatile, particularly after February 2021. This could be attributed to the increased sensitivity of this model to changes in repeat-sales data. This unique sensitivity stems from the model's design, which leverages repeat-sales data to capture pure price changes, while simultaneously accounting for potential heteroskedasticity and selection bias. As a result, any substantial variations in repeat-sales data are likely to be more pronounced in this model, leading to increased volatility in the derived price index. Overall, from June 2021 to October 2021, the Repeat Sales index was generally above the Hedonic index, possibly due to speculative trading where investors resold NFTs at inflated prices driven by enthusiasm and potential speculative bubbles. In contrast, the Hedonic index remained relatively stable, suggesting that the intrinsic characteristics of NFTs didn't change much in value.

The hedonic and RSR offer robust analytical frameworks for understanding the pricing dynamics of the digital assets in our sample. However, they are not exempt from potential issues. A significant concern in these models is the time instability of the parameters, which can lead to biased estimates of the price indices (Kuminoff et al., 2010). This instability can stem from changing market dynamics, evolving consumer preferences or shifts in the mix of characteristics present in the traded NFTs over time.

The Laspeyres and Paasche indices, as shown in Figure 2a, provide a clear picture of the monthly price dynamics of the NFT market. Despite observable fluctuations in the market, these indices remain relatively stable and mirror each other closely throughout the sample period. This stability suggests that the cost of purchasing a specific bundle of NFT characteristics has remained consistent over time, a finding in line with the assumptions of the hedonic model (Rosen, 1974). However, a significant boom in the NFT market was observed starting from September 2021. Despite this surge, the Laspeyres and Paasche indices remained stable, prompting us to conduct a more detailed investigation into the underlying causes.

We plot the changes in the β coefficients and quantities (intended as the median of characteristics in our data set) over time in Figure 2b. The β coefficients, indicative of the implicit prices of NFT characteristics, show substantial variability throughout the period under consideration. Conversely, the quantities of the NFT characteristics, denoted by q, follow a

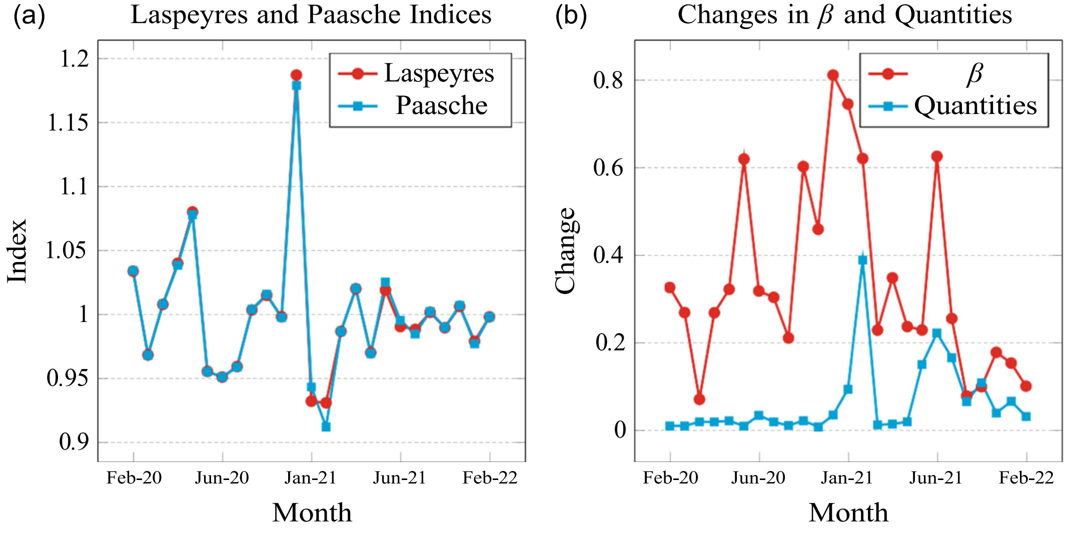

FIGURE 2 Fisher index. (a) The Laspeyres and Paasche indices (components of the overall Fisher price index) over the examined period from February 2020 to February 2022. These indices reflect the monthly price dynamics of the nonfungible token (NFT) market. Despite observable fluctuations in the market, particularly a significant surge in September 2021, these indices remain relatively stable. This stability suggests a consistent cost for a specific bundle of NFT characteristics, underpinning the assumptions of the hedonic model. (b) The changes in the β coefficients and quantities over the same time period. The β coefficients, which represent the implicit prices of NFT characteristics, show substantial variability, capturing the evolving nature of the NFT market and changes in consumer preferences. Conversely, the quantities, denoted by q, remain relatively constant, indicating a stable mix of NFT characteristics sold over time. This contrast between the variability of β coefficients and the stability of q suggests that the changes in prices are generally balanced by changes in quantities. This balance further highlights the stability of the Laspeyres and Paasche indices, and consequently the Fisher index, even during the market boom in September 2021. [Color figure can be viewed at wileyonlinelibrary.com]

smoother trajectory, suggesting that the mix of characteristics in the sold NFTs has remained relatively constant over time. This stability in quantities, coupled with the variability in β coefficients, causes the steady nature of the Laspeyres and Paasche indices during the observed market boom. Essentially, the stability of the Fisher index, which is a geometric mean of the Laspeyres and Paasche indices, implies that changes in prices (β) are generally balanced by changes in quantities (q).

The graphical representation of the number of trades and floor price trends in Figure 3 further underscores the dynamic nature of the NFT market. We observe a significant increase in the number of trades, or market activity, towards the end of the period, which coincides with a rise in the floor price. This pattern suggests a potential correlation between market activity, prices and the quantities of characteristics traded. In other words, as the volume of trade and average prices increased (evidenced by the rise in the floor price), the quantities of NFT characteristics sold also increased, effectively stabilizing the price indices.

4.2 | Herding behaviour

Following classical financial theories, we believe that consumers tend to follow the crowd, aligning their preferences and demands with popular artists, thereby creating a cascade effect. This herding behaviour is manifested in the NFT market, leading to a superstar phenomenon where a handful of artists dominate the market, as proposed by Rosen (1981).

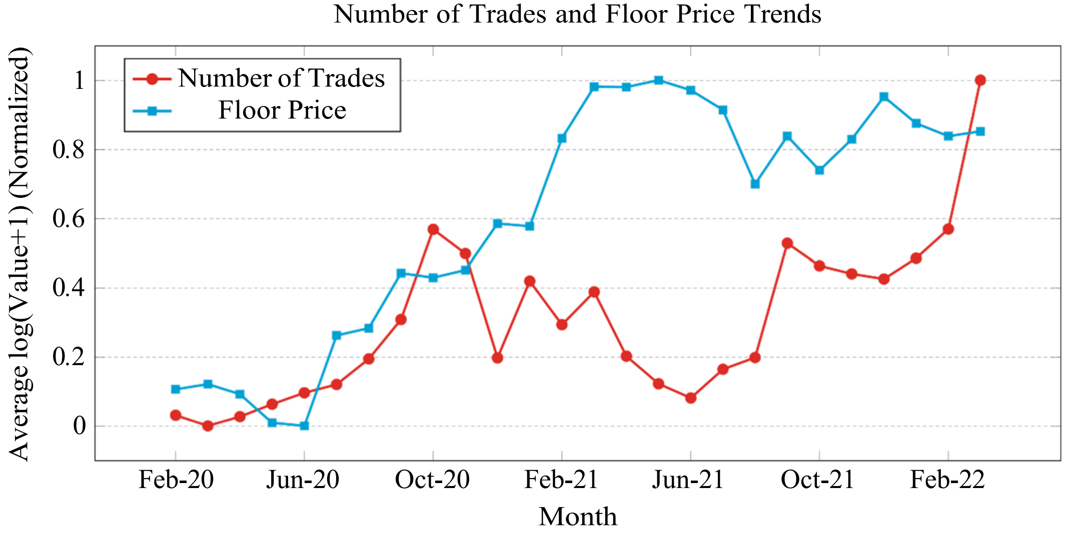

FIGURE 3 Number of trades and floor price trends. This graph presents the evolution of the number of trades and floor price trends in the art nonfungible token (NFT) market from February 2020 to February 2022. The log(*num of trades* + 1) (depicted in red) and log(*floor price* + 1) (depicted in blue) are represented on a normalized, monthly averaged scale. An observable upswing in both metrics is evident from around October 2021, suggesting a significant market boom during this period. This upward trend likely reflects evolving consumer preferences, technological advancements, increased public awareness and broader market acceptance of NFTs as digital assets. Importantly, as trading volume and average prices increased, so did the quantities of NFT characteristics sold, effectively balancing the price indices. [Color figure can be viewed at wileyonlinelibrary.com]

We employ a comprehensive ranking system for artists. This system works by considering a diverse range of market metrics and social factors which encompass various aspects of an artist's presence in the NFT market, allowing us to evaluate the artists' market positions and popularity. The ranking process is performed monthly to reflect the dynamic nature of the NFT market. Specifically, we leverage the previous analysis database and select all deployers based on their Twitter handles who are not companies or marketplaces. We group the NFTs by artist, leading to a total of 413 unique NFT artists based on Twitter handles.

We factor in the ratio of an artist's sold to her unsold artworks, which gives us an understanding of the market demand for her works. The higher the ratio, the greater the demand for the artist's work. Similarly, we account for the average price of sold NFTs. A higher average price typically indicates a higher perceived value or desirability of an artist's work. The market capitalization for each artist, calculated on a monthly basis, is also considered. This measure, defined as the total value of her sold artworks within a given month, provides an insight into the artist's overall market presence and financial impact. We also take into account the average number of trades for each sold NFT. This measure provides an indication of the liquidity and secondary market activity of an artist's work. More trades generally suggest greater interest and engagement from collectors and investors. Additionally, we look at the number of unique buyer–seller pairs, which gives us an idea of the diversity of interactions associated with an artist's works. A greater number of unique pairs usually indicates a broader collector base and more varied market activity. Lastly, we consider the selling time of each sold NFT. This measure indicates the average time it takes for an artist's work to sell. Shorter selling times often suggest higher demand and faster market turnover. By considering all these equally weighted factors, we compute a composite score for each artist. This score, where each factor has an equal weight, serves as the basis for our monthly ranking of artists.

To contextualize the variables employed in our study, we start with an examination of the descriptive statistics illustrated in Table 2. The substantial variation in *volume*—with a median value of 4492.12—exposes the disparity between artists, highlighting an uneven playing field where a few artists garner significant market activity, while the majority struggle to gain traction. As this variable reflects the market activity for individual artists in a given month, this could potentially signal illiquidity issues, where the market is concentrated around a small number of artists, akin to the "superstar" phenomenon. Moreover, *share*, which represents the proportion of the total monthly market capitalization attributable to an individual artist, further accentuates this fragmentation. The median value of 0 indicates that the market is disproportionately dominated by a few artists, suggesting a significant skew in market power. Lastly, the *demand* variable, representing an artist's popularity relative to other artists, underscores the fluctuating and potentially illiquid nature of the art NFT market. With a median value of 0.97 and a wide range of values, it demonstrates how rapidly consumer preferences can shift in the art NFT market. This suggests a market prone to sudden swings and potentially fragile liquidity, as demand can quickly pivot from one artist to another.

Table 4 presents the findings from our regression models, dividing the artists into three distinct categories based on their monthly rankings. Artists are grouped by the month they are active, and their rankings are evaluated within these monthly groups. They are then divided into three quantile-based categories based on their rank, ensuring that each category has roughly the same number of artists where possible. Since rankings can vary from month to month due to changes in the metrics that determine an artist's rank, it is common for an artist to appear in different categories across different months. This method allows for each group to contain about 780 artists, even though there are only 413 unique artists, because the

TABLE 4 Least square estimation herd bias.

This table reports the findings from a comprehensive regression analysis based on a least square estimation of the model of Azarmi and Menny (2013), investigating the determinants of demand for artworks in the nonfungible token market. We divide artists into three categories according to their monthly rankings: the highest-ranked artists (Model 1), midtier artists (Model 2) and lower-tier artists (Model 3). The dependent variable in each model is the logarithm of demand for artworks by each artist category. A diverse array of explanatory variables is included, reflecting historical performance metrics, artist rankings and social media engagement. Yearly rankings are consistently negative and significant, indicating that higher rankings correlate with lower demand, suggesting a preference for niche or emerging artists. Furthermore, lagged monthly rankings show that previous high ranks reduce current demand for top and bottom artists but increase it for midtier artists, likely due to a "momentum effect". Nevertheless, engagement on Discord does not boost demand for low-tier artists, possibly due to a smaller follower base, but enhances visibility and demand for midtier artists. Finally, independently deployed works by top-ranked artists see higher demand, reflecting a preference for authenticity. However, this strategy is less beneficial for midtier artists, likely due to the lack of support systems. Each entry in the table delineates the coefficient estimates, with standard errors provided in parentheses. Artist and month fixed effects (FE) are included to control for unobserved time-invariant artist-specific characteristics and common temporal shocks. The total number of observations, the adjusted R^2 and the F statistic are reported for each artist category. Significance levels are denoted as follows: $^*p \le 0.1$; $^{**}p \le 0.05$; $^{***}p \le 0.01$

	Dependent variable: log($demand_{it}$)		
	(1)	**(2)**	**(3)**
const	2.9905***	−0.0320***	−2.9255***
	(0.0189)	(0.0124)	(0.0233)
$cum\ score\ year\ rank_{t-1}$	−0.0002*	−0.0001	0.0001
	(0.0001)	(0.0001)	(0.0001)
$year\ rank_t$	−0.0087***	−0.0058***	−0.0073***
	(0.0006)	(0.0002)	(0.0003)
$month\ rank_{t-1}$	−0.0017*	0.0003	−0.0008
	(0.0012)	(0.0006)	(0.0009)
$year * discord\ account_t$	−0.0001	0.0002*	−0.0001
	(0.0003)	(0.0001)	(0.0001)
$year * deployer\ artist_t$	0.0013**	−0.0005***	0.0000
	(0.0005)	(0.0002)	(0.0002)
Artist FE	✓	✓	✓
Month FE	✓	✓	✓
Observations	787	777	796
R^2	0.7346	0.7855	0.6280
	297.30***	329.64***	172.53***
F statistic (robust)	(df = 538)	(df = 450)	(df = 511)

same artists can appear in multiple categories across various months. The categories are as follows: the highest-ranked artists (Model 1), the midtier artists (Model 2) and the lower-tier artists (Model 3).

We find that $year\_rank_t$ is negative and statistically significant across all models, suggesting that artists with higher yearly rankings have lower demand, potentially indicating art enthusiasts' preference for niche artists or their inclination to discover new talent rather than following the crowd.

The variable $month\_rank_{t-1}$ further underscores the complexity of the art choice dynamics. It is negative for the highest- and lowest-ranked artists (Models 1 and 3), suggesting that a higher rank in the previous month decreases the current demand. However, the relationship is positive for midtier artists (Model 2), which could be attributed to a "momentum effect" where midranking artists who perform well in 1 month continue to attract attention in the subsequent month.

The interaction term $year_t * discord\_account$ although not significant for the lowest-ranked artists (Model 3), suggests that active engagement on Discord does not necessarily boost demand for these artists' works. This observation might seem counterintuitive, especially considering the common belief that social media engagement generally increase artist visibility and market appeal, as highlighted by Bao et al., 2023. However, it could be that for low-tier artists, the benefit of additional exposure on platforms like Discord are limited by factors such as a smaller existing follower base or less recognition in the broader art market. In contrast, for mid-tier artists, where the interaction term is positive and significant, engagement on Discord likely plays a crucial role in enhancing visibility and demand. This suggests that mid-tier artists, who may already have some market presence but have not yet reached saturation, can significantly leverage social media to broaden their audience and increase market appeal.

Turning our attention to the interaction term $year_t * deployer\_artist$, we find that it is positive for the highest-ranked artists (Model 1). Artists who independently deploy their work, as opposed to those associated with or acting on behalf of a company, are met with higher demand. This finding could be indicative of art enthusiasts' preference for the authenticity, creativity and individuality that independent artists often embody. However, the dynamics shift when we look at midtier artists (Model 2). For these artists, the term is negative, implying that independent deployment might not be as beneficial. This could be due to the additional challenges and responsibilities that come with operating independently. Lesser-known artists might lack the resources, network or reputation that companies provide, thereby facing hurdles in establishing themselves in the competitive art NFT market. Furthermore, art enthusiasts may perceive company-associated artists as more reliable or credible.

Our findings support the hypothesis of herding behaviour. The yearly score shows a clear and consistent picture: The auction performance of artists (as represented by their rank) has a profound influence on consumer choice. This is particularly pronounced for the high and mid-tier artist samples (Model 1 and Model 2), where herding has a stronger impact on consumer behaviour, and the effect of current performance is even more apparent.

The clear result obtained from our regression analysis is the conspicuous inequalities in the art NFT market: the choices of artworks are highly skewed towards the highest and mid-ranked artists. This result implies a heavy concentration of market influence amongst a small number of artists. Indeed, the presence of stardom in art markets is characterized by the fact that a select group of artists disproportionately commands both attention and revenue. These "star" artists dominate the marketplace, often overshadowing their less well-known peers. Their works are sought after by collectors and command high prices, contributing to a skewed distribution of wealth within the artist community. To visually illustrate this concentration, we turn to the 2021 market capitalization Lorenz curve for the top 200 artists in Figure 4a. The Lorenz curve's substantial deviation from the equality line—indicative of a scenario where every artist shares an equal proportion of the market cap—provides a clear visualization of the pronounced inequality in the distribution of market capitalization among artists. The curve's steep incline towards the end of the distribution highlights that a small fraction of artists command a disproportionately large share of the market cap, reinforcing our regression findings.

Figure 4b shows the Gini Index, a single measure derived from the Lorenz curve that quantifies the degree of inequality over time. The relatively high and fluctuating Gini Index values across all three samples underscore a persistent and varying level of inequality in the

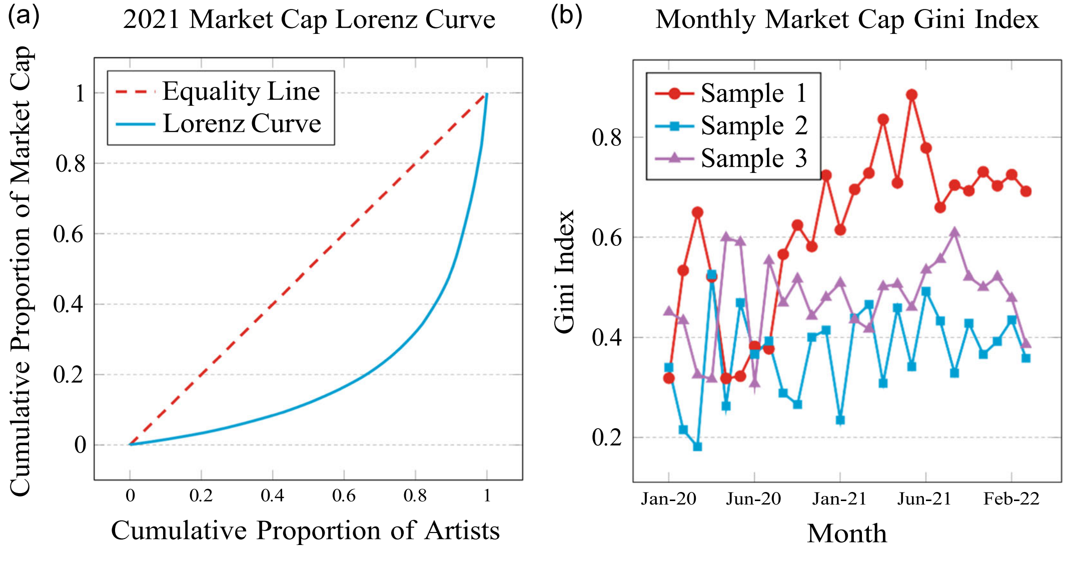

FIGURE 4 Lorenz curve and Gini index for art nonfungible tokens (NFTs). In the analysis of the 2021 Market Cap Lorenz Curve for the top 200 artists, a striking deviation from the line of equality is evident. This line of equality serves as a baseline, symbolizing an idealistic scenario where each artist enjoys an equal proportion of market capitalization—an equitable distribution of wealth and influence. The Lorenz curve, however, tells a different story, one marked by stark disparities and profound inequality. The Lorenz curve's *x*-axis measures the cumulative proportion of artists, a metric that ranges from the least to the most successful. The *y*-axis quantifies the cumulative proportion of the market capitalization, thereby reflecting the accumulated wealth and influence within the artist community. The substantial bowing of the Lorenz curve away from the line of equality indicates the pronounced inequality in the distribution of market capitalization among artists. The curve takes a steep ascent towards the end of the distribution, suggesting that a minor fraction of artists wield a disproportionately large share of the market cap. The monthly market capitalization Gini Index investigates the degree and temporal dynamics of this inequality. The Gini Index, a derivative measure from the Lorenz curve, encapsulates the inequality within a single, powerful metric. Plotted over time, it offers a chronological snapshot of inequality, tracing its ebb and flow through different periods. Three different samples were analyzed across a series of months (top, mid and low tier artists), each revealing its own tale of inequality. However, the common thread across all three is the high Gini Index values that fluctuate over time, underscoring a persistent and variable level of inequality in the market capitalization distribution among artists. Among these, Sample 1 exhibits the most pronounced inequality, reflecting the dominance of the highest-ranked artists. (a) Lorenz curve and (b) Gini index. [Color figure can be viewed at wileyonlinelibrary.com]

market capitalization distribution among artists. The most pronounced inequality is observed in the top-tier artists sample, which aligns with our regression analysis indicating the dominance of the highest-ranked artists.

Building on our prior analysis, we aim to identify if and when herding behaviour is likely to occur in the art NFT market of high and mid-tier artists, especially during periods of significant market fluctuations. For a reliable study of CSAD, it is crucial to have a sufficient number of active artists to enhance liquidity and ensure robust results. Therefore, we restrict our sample to artists who have performed at least 10 sales. Additionally, we focus on data starting from January 1, 2021, as this was the year with the majority of NFT transactions. This approach leads to a total of 405 artists, ensuring that our findings are based on a well-defined and active segment of the market.

The ordinary least squares (OLS) regression results in Table 5 suggest that market volatility, represented by the absolute market return, has a positive influence on the CSAD. The significant positive coefficient indicates that increased market volatility is associated with increased herding behaviour: when the market experiences significant ups and downs, investors are more likely to follow the crowd, potentially leading to inflated or deflated asset prices. The lagged CSAD also has a significant positive coefficient, suggesting a persistence of herding behaviour over time. This implies that if herding behaviour was present in the past, it is likely to continue into the future, potentially creating self-reinforcing trends in the NFT market. Most interestingly, however, is the non-significant coefficient of the squared market return, which is unexpected since we typically expect a negative coefficient here. This anomaly suggests that our initial model might not fully capture the underlying dynamics, warranting further analysis. To investigate this further, we use the Markov Switching Model. Essentially, during times of intense market fluctuations, there is a tendency for the dispersion among individual NFT returns to decrease. This suggests that in high-volatility periods, investors are likely following a common narrative or sentiment, causing their trading behaviours to align. Thus, the individual returns of these art NFTs begin to converge around a central tendency, reducing the CSAD (Drehmann et al., 2005; Economou et al., 2011; Park & Sabourian, 2011).

In the MSM, the coefficients under Regime 1 (where herding is present) and Regime 2 (where herding is absent) provide more specific insights into the dynamic nature of herding behaviour in the NFT market. We notice under Regime 1 that the squared market return exhibits a negative coefficient compared with its OLS counterpart. This finding is in alignment with the hypothesis that herding is amplified under extreme market conditions. More specifically, in periods of elevated market returns, the prevalence and intensity of herding seem to be exacerbated. This reinforces the notion that market participants are especially prone to follow the crowd when the stakes are high, subsequently reducing the CSAD of individual asset returns. In contrast, Regime 2 features a positive coefficient for the squared market return, although it is poorly statistically significant. This suggests that extreme market conditions could serve as a deterrent to the formation of herding behaviour when such a behaviour is not already manifest. The model's transition probabilities indicate that once herding is initiated (Regime 1), it tends to persist over time. In this case, herding can be self-reinforcing due to factors, such as information cascades or social influence.

The absolute market return parameter in the MSM serve as complementary layers of interpretation. The significant positive coefficient on absolute market return in Regime 1 suggests that herding is more prominent in periods of higher returns. This can be interpreted as a form of “return-chasing” behaviour, where traders and investors flock to trending assets, thereby enhancing herding.

TABLE 5 Ordinary least squares (OLS) and Markov switching models (MSMs) for herd bias.

This table reports the coefficients of the OLS regression and the MSM to investigate the presence and dynamics of herding behaviour in the art nonfungible token market. The dependent variable is the Cross-Sectional Absolute Deviation (CSAD), a measure of herding. The first column shows the OLS regression results, illustrating the overall dynamics of herding behaviour in the market. Both the absolute market return ($|R_{m,t}|$) and lagged CSAD ($CSAD_{t-1}$) have positive coefficients, indicating that market volatility and past herding behaviour increase the tendency for herding. The next two columns display the coefficients under two different regimes in the MSM: Regime 1, which signifies the presence of herding, and Regime 2, which signifies the absence of herding. In Regime 1, the negative coefficient of the squared market return confirms and reinforces the relationship between market returns and herding. In Regime 2, the squared market return displays a positive, though poorly significant, coefficient, suggesting the difficult formation of herding behaviour when it is not already present. The last column shows the transition probabilities ($p[1 \rightarrow 1]$ and $p[2 \rightarrow 1]$) in the MSM, providing insights into the persistence and transition dynamics of herding behaviour. The table contains 405 observations. The model's goodness of fit is represented by the log-likelihood, Bayesian Information Criterion (BIC) and Akaike Information Criterion (AIC). Significance levels are denoted as follows: ***$p \leq 0.01$; **$p \leq 0.05$; *$p \leq 0.1$.

	Dependent Variable: $CSAD_t$			
		Markov switching model		**Nonswitching Parameters**
	OLS	**Regime 1**	**Regime 2**	
const	0.0006*** (0.000)	0.0012*** (0.000)	0.0003* (0.000)	
$|Rm,t|$	0.1704*** (0.045)	1.6929*** (0.080)	0.0702* (0.037)	
$R^2{}_{m,t}$	0.0830 (0.898)	−7.1029*** (0.017)	0.2868 (0.759)	
$CSAD_{t-1}$	0.9739*** (0.007)	0.4864*** (0.029)	0.9507*** (0.011)	
$p[1\rightarrow1]$				0.8115*** (0.081)
$p[2\rightarrow1]$				0.0185*** (0.008)
Observations		405		
Log likelihood	2064.3		2148.75	
BIC	−4105		−4231.45	
AIC	−4121		−4257.49	

Figure 5 effectively demonstrates the temporal oscillation between regimes of herding behaviour and independent decision-making within the NFT market, as postulated by our model. Specifically, the shaded areas highlight periods when investors, driven by market momentum, decide to follow the crowd, leading to an increase in herding behaviour. These instances are closely aligned with our MSM findings, which reveal a negative relationship between squared market returns and CSAD, hence indicating increased herding behaviour during market volatility.

The shaded regions in Figure 5 align with notable events in the NFT space that likely incited increased market activity and investor interest. For example, Beeple's $69 million

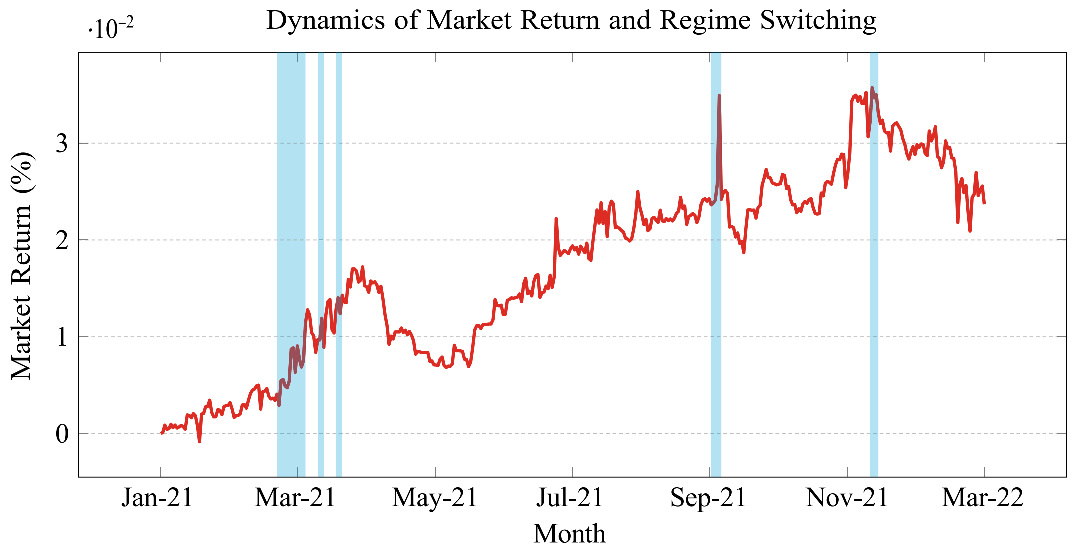

FIGURE 5 Dynamics of market return and regime switching. This graph provides a temporal illustration of the nonfungible token (NFT) market returns and the corresponding prevalence of herding behaviour. The shaded areas represent periods of herding behaviour and correlate with significant market returns. This alignment which indicate a positive relationship between absolute market returns and Cross-Sectional Absolute Deviation, offers a nuanced visual representation of herding behaviour during periods of market volatility. The graph showcases oscillation between herding and independent decision-making, reflecting the complex, nonlinear relationship between market returns and herding behaviour. For instance, high but not extreme market returns still induce herding behaviour, as indicated by the shaded areas. These shaded periods, driven by market momentum, correspond to key events in the NFT space that amplified market activity. In March 2021, the $69 million auction of Beeple's artwork led to increased trading activity. Similarly, the "Loot (for Adventurers)" project in September 2021 ignited a surge in market returns and herding behaviour. [Color figure can be viewed at wileyonlinelibrary.com]

auction watershed event in March 2021 attracted significant media attention and public interest, leading to increased trading activity and a surge in market returns. Similarly, the shaded region in September 2021 coincides with the period when the "Loot (for Adventurers)" project gained significant traction. This project led to a surge in trading activity, manifesting as increased market returns and herding behaviour.

To further enhance our understanding of the mechanisms that drive herding behaviour in the art NFT market, we will investigate in the following how the coefficient on the squared market return is modulated by the presence of investor newcomers and fluctuations in the BGCI. Specifically, we seek to assess whether, and to what extent, the percentage of newcomers in the market and variations in BGCI returns impact the strength and direction of the coefficient on squared market returns, which serves as our proxy for herding behaviour. To do so, we use a weighted least squares (WLS) regression model. We first calculate the proportion of newcomers in the market for each date, defining a newcomer as an investor who has traded only once. This proportion is then merged with the existing dataset containing market returns and CSAD values. We perform weighted regressions across a grid of newcomer proportions, using Silverman's Rule of Thumb to determine the bandwidth for our kernel weights. This allows us to assess how the coefficient on squared market returns varies with different levels of newcomer presence, providing a nuanced view of herding behavior under varying market conditions. The WLS model is well-suited for this analysis as it accounts for heteroskedasticity and provides more reliable estimates by giving different weights to observations based on the newcomer proportion. By plotting the coefficients and their confidence intervals across the newcomer proportion grid, we can visualize and interpret the impact of newcomers and BGCI fluctuations on herding behavior in the NFT art market.

Drawing on insights from Bao et al. (2023), we know that experienced investors tend to secure higher returns per unit of cryptocurrency invested compared with their inexperienced counterparts. This disparity may be attributed to inexperienced investors, often newcomers, purchasing NFTs at higher average prices, possibly due to their limited familiarity with market valuations and dynamics. Given that the majority of these newcomers engage in trading activities only once, their lack of market understanding makes them susceptible to following market consensus. When there is an influx of such newcomers into the art NFT market, the propensity for herding behaviour to manifest is likely heightened.

Simultaneously, the rationale for incorporating BGCI returns stems from the observed high correlation between cryptocurrency markets and NFT markets. Fluctuations in the BGCI can serve as a proxy for general market sentiment within the broader cryptoasset ecosystem, which invariably influences the art NFT space. Thus, BGCI returns can either amplify or attenuate herding tendencies, depending on the prevailing market conditions.

Figure 6a illustrates the relationship between newcomer proportion and the coefficient γ_2 of Equation (13). Initially, when the share of newcomer is lower than 0.30, γ_2 touches 0, suggesting minimal herding behavior. This phase indicates that a smaller influx of newcomers does not significantly drive collective behavior, and market participants tend to act more independently. As the percentage of newcomers increases to the range of 0.35, γ_2 becomes more substantially negative, reflecting a pronounced increase in herding behavior. In this intermediate phase, the actions of newcomers start to align more closely, leading to a stronger collective movement in market behavior. This alignment is indicative of newcomers following similar trends and signals, amplifying the herding effect.

However, as the share of newcomers continues to rise beyond 0.55, γ_2 begins to increase, although it remains negative. This trend might mark the onset of a saturation effect (Bikhchandani et al., 2021; Cong et al., 2021). At this point, the market might experience collective uncertainty, as the proportion of inexperienced participants could become too large. This might lead to a shift towards more individualistic behaviour. Such a shift would be characterized by a less pronounced negative γ_2, indicating reduced herding.

The relationship between γ_2 and BGCI return in Figure 6b presents a non-linear and fluctuating path. Our findings reveal a counterintuitive result. Contrary to expectations that higher returns on the BGCI would amplify herding behaviour in the NFT market, we observe that stronger performance in the cryptocurrency market actually attenuates the prevalence of herding tendencies within the NFT landscape.

According to Demirer and Kutan (2006) and Huo et al. (2023), assets that are the focus of institutional herding initially experience positive abnormal returns. In other words, when institutional investors move en masse to buy a particular asset, its price surges above what would be its intrinsic value under normal market conditions. However, this is a transitory phenomenon; prices tend to revert to their mean or intrinsic value in the long run, nullifying the temporary gains accrued due to herding.

Corbet et al. (2022) elaborate on this by offering a theoretical framework. Their model suggests that the inflation of asset prices is not a random or inexplicable event but a direct

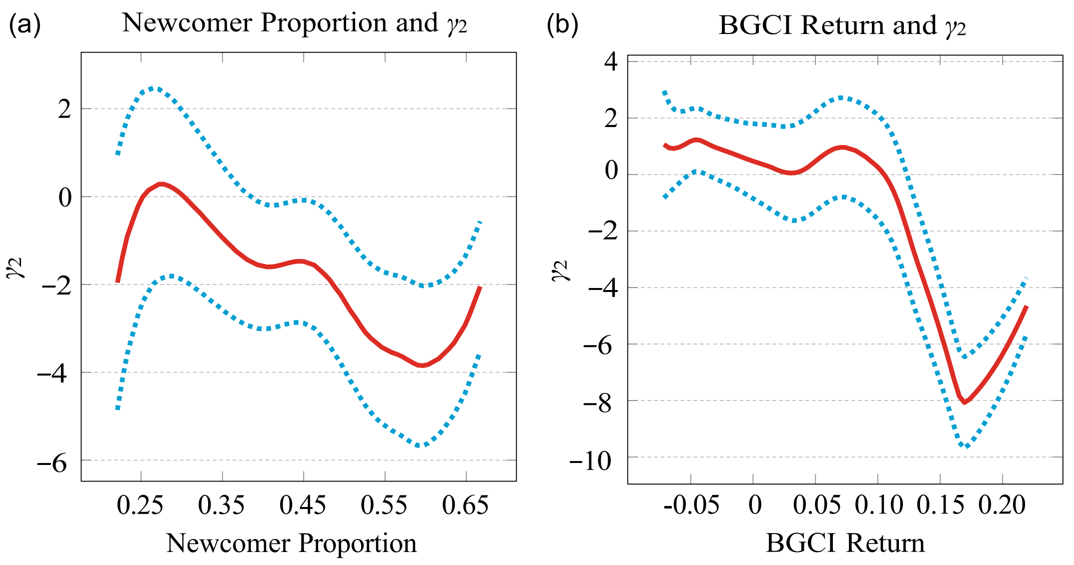

FIGURE 6 γ_2 Dynamics, newcomers and Bloomberg Galaxy Crypto Index (BGCI) return. This graph presents the relationships between the γ_2 coefficient, the percentage of newcomer investors in the nonfungible token (NFT) market and the BGCI returns. The γ_2 coefficient represents the sensitivity of the cross-sectional asset deviation to the squared market returns, a proxy for herding behaviour. The functions in the graphs have been smoothed with a moving average with a window of 30 to reduce noise and highlight underlying trends. Initially, when the share of newcomers is lower than 0.30, γ_2 touches 0, suggesting minimal herding behaviour. This phase indicates that a smaller influx of newcomers does not significantly drive collective behaviour, and market participants tend to act more independently. As the percentage of newcomers increases to the range of 0.35, γ_2 becomes more substantially negative, reflecting a pronounced increase in herding behaviour. In this intermediate phase, the actions of newcomers start to align more closely, amplifying the herding effect. However, as the share of newcomers rises beyond 0.55, γ_2 begins to increase, although it remains negative. This trend marks the onset of a saturation effect. At this point, the market might experience collective uncertainty, leading to a shift towards more individualistic behaviour, characterized by a less pronounced negative γ_2, indicating reduced herding. The relationship between γ_2 and BGCI return is nonlinear and fluctuating. Interestingly, instead of increasing herding behaviour, a strong performance in the cryptocurrency market seems to reduce it in the art NFT market. Typically, assets subject to institutional herding experience initial return spikes, but these gains are short lived as prices correct to intrinsic values. Herding inflates prices temporarily, which then revert over time. A strong BGCI, indicating a healthy cryptocurrency market, leads investors to shift capital from NFTs to cryptocurrencies, thereby reducing herding in the art NFT market. Thus, strong BGCI performance stabilizes investor behaviour and reduces herding-induced distortions in art NFTs. (a) Newcomer proportion and γ_2 and (b) BGCI return and γ_2. [Color figure can be viewed at wileyonlinelibrary.com]

outcome of herding behaviour among institutional investors. Once the herd dissipates or shifts focus, the inflated prices naturally correct themselves, aligning more closely with their true, intrinsic value. This implies a sort of pull effect that intrinsic value exerts on asset prices, drawing them back to equilibrium once the distorting influence of herding has waned.

We consider in the following how these theories interface with our observations regarding the BGCI and the art NFT market. Figure 6b illustrates that the relationship between BGCI return and γ_2 is characterized by significant fluctuations rather than a straightforward trend. This non-linear relationship suggests that a strong performance in the BGCI may moderate herding tendencies in the NFT market but this effect is complex and variable. In practice, when the BGCI is performing well, it signifies strength in the broader cryptocurrency market. As herding in the NFT market ceases and leads to negative returns, investors in the NFT market may seek alternative avenues for investment, often transitioning their capital into more stable or promising assets, like, cryptocurrencies. This inflow of capital positively impacts the BGCI. A strong BGCI performance serves as a signal or even a magnet for capital reallocation away from the NFT market, thereby moderating herding behaviour in NFTs. This mechanism acts as a balancing force across the two correlated but distinct asset classes, art NFTs and cryptocurrencies represented by the BGCI. This balancing act ensures a more stable market environment, reducing the distortions typically caused by herding and allowing for a more rational asset pricing across both markets. Hence, the performance of the BGCI serves not only as a reflection of market sentiment but also as a moderating variable that indirectly influences investor behaviour in the NFT market.

5 | CONCLUSIONS

We have analyzed the value determinants and herding behaviour in the art NFT market, establishing a ranking system for artists and examining the factors that drive NFT prices. Our analysis reveals a nuanced picture of the NFT ecosystem. We find a positive relationship between the number of trades and floor prices with the average price of NFTs, underscoring the influence of market activity and perceived value on pricing. In contrast to prevailing assumptions, our research indicates that social media activity may inversely affect NFT prices, potentially due to market oversaturation and reduced perceived rarity. This challenges the conventional wisdom that increased visibility invariably boosts demand.

We observe a pronounced herding behaviour, particularly among top and mid-tier artists, manifesting in a market heavily skewed towards a select few, creating pronounced inequalities. This finding is critical as it highlights the concentration of market influence and wealth within a small segment of artists, reflecting a broader trend of inequality that transcends the art world into the digital space. Additionally, herding behaviour is influenced by the entrance of new investors and fluctuations in the cryptocurrency market. While newcomers tend to reinforce herding, an increase in their numbers could lead to a saturation effect, ultimately introducing more volatility. Conversely, a strong cryptocurrency market appears to moderate herding within the NFT market, likely due to a reallocation of capital towards more "stable" crypto investments. Furthermore, our study provides insights into the complex dynamics of consumer choices, showing that higher artist rankings do not guarantee continued future demand. This suggests a market that values novelty and the discovery of new talents over established reputations.

Our findings validate a prevalent belief in the crypto space as a whole, that due to the unregulated nature of the industry, cryptocurrency-related endeavours attract swindlers that do more harm than good both in terms of reputation and finances. However, this does not imply that art NFTs have no artistic worth. As speculative activities diminish and the market stabilizes, there is an opportunity for the art NFT sector to foster a more creative and sustainable environment that benefits both artists and collectors. While the art NFT market is subject to the whims of speculation and herding, it also appears to possess the potential for innovation and value creation. The key lies in understanding and navigating the market's complex dynamics, from social media influences to broader economic factors, to harness the potential of NFTs as a novel artistic medium.

CONFLICT OF INTEREST STATEMENT

The authors declare no conflict of interest.

DATA AVAILABILITY STATEMENT

The authors elect not to share the data supporting the findings of this study.

ORCID

Gilbert Fridgen http://orcid.org/0000-0001-7037-4807
Roman Kräussl http://orcid.org/0000-0001-8933-9278
Orestis Papageorgiou http://orcid.org/0000-0002-2245-5079
Alessandro Tugnetti http://orcid.org/0000-0003-3811-2828

REFERENCES

Adams, R. B., Kräussl, R., Navone, M., & Verwijmeren, P. (2021). Gendered prices. *The Review of Financial Studies*, *34*(8), 3789–3839.

Alizadeh, S., Setayesh, A., Mohamadpour, A., & Bahrak, B. (2023). A network analysis of the non-fungible token (NFT) market: Structural characteristics, evolution, and interactions. *Applied Network Science*, *8*(1), 1–18.

Amihud, Y., & Mendelson, H. (1986). Asset pricing and the bid–ask spread. *Journal of Financial Economics*, *17*(2), 223–249.

Ang, A., Liu, J., & Schwarz, K. (2020). Using stocks or portfolios in tests of factor models. *Journal of Financial and Quantitative Analysis*, *55*(3), 709–750.

Ante, L. (2022). The non-fungible token (NFT) market and its relationship with Bitcoin and Ethereum. *FinTech*, *1*(3), 216–224.

Aubry, M., Kräussl, R., Manso, G., & Spaenjers, C. (2023). Biased auctioneers. *The Journal of Finance*, *78*(2), 795–833.

Azarmi, T. F., & Menny, P. R. P. (2013). Star artists and herding in fine arts' market: Theory and empirical evidence. Available at SSRN: https://ssrn.com/abstract=2232307

Babić Rosario, A., de Valck, K., & Sotgiu, F. (2020). Conceptualizing the electronic word-of-mouth process: What we know and need to know about eWOM creation, exposure, and evaluation. *Journal of the Academy of Marketing Science*, *48*(3), 422–448.

Bailey, M. J., Muth, R. F., & Nourse, H. O. (1963). A regression method for real estate price index construction. *Journal of the American Statistical Association*, *58*(304), 933–942.

Bamakan, S. M. H., Nezhadsistani, N., Bodaghi, O., & Qu, Q. (2022). Patents and intellectual property assets as non-fungible tokens; Key technologies and challenges. *Scientific Reports*, *12*(1), 2178.

Banerjee, A. V. (1992). A simple model of herd behavior. *The Quarterly Journal of Economics*, *107*(3), 797–817.

Bao, T, Ma, M., & Wen, Y. (2023). Herding in the non-fungible token (NFT) market. *Journal of Behavioral and Experimental Finance*, *39*, 100837.

Bao, H., & Roubaud, D. (2022). Non-fungible token: A systematic review and research agenda. *Journal of Risk and Financial Management*, *15*(5), 215.

Bianchi, D., & Babiak, M. (2022). On the performance of cryptocurrency funds. *Journal of Banking & Finance*, *138*, 106467.

Bikhchandani, S., Hirshleifer, D., Tamuz, O., & Welch, I. (2021). *Information cascades and social learning* [NBER Working Paper, ID 28887].

Borri, N., Liu, Y., & Tsyvinski, A. (2023). The economics of non-fungible tokens. Available at SSRN: https://ssrn.com/abstract=4052045

Campbell, J. Y., & Shiller, R. J. (1988). The dividend–price ratio and expectations of future dividends and discount factors. *The Review of Financial Studies*, *1*(3), 195–228.

Caporale, G. M., Kang, W.-Y., Spagnolo, F., & Spagnolo, N. (2021). Cyber-attacks, spillovers and contagion in the cryptocurrency markets. *Journal of International Financial Markets, Institutions and Money*, *74*, 101298.

Case, K. E., & Shiller, R. J. (1989). The efficiency of the market for single-family homes. *The American Economic Review*, *79*(1), 125–137.

Chainalysis. (2022). *The Chainalysis 2021 NFT market report*.

Chanel, O., Chanel, O., Gérard-Varet, L.-A., & Ginsburgh, V. (1996). The relevance of hedonic price indices: The case of paintings. *Journal of Cultural Economics*, *20*(1), 1–24.

Chang, E. C., Cheng, J. W., & Khorana, A. (2000). An examination of herd behavior in equity markets: An international perspective. *Journal of Banking & Finance*, *24*(10), 1651–1679.

Cheah, E.-T., & Fry, J. (2015). Speculative bubbles in bitcoin markets? An empirical investigation in the fundamental value of Bitcoin. *Economics Letters*, *130*, 32–36.

Collins. (2021). *Get your crypto at the ready: NFTs are big in 2021*. Collins Dictionary Language Blog.

Cong, L. W., Landsman, W., Maydew, E., & Rabetti, D. (2022). *Tax-loss harvesting with cryptocurrencies* [NBER Working Paper, ID 30716].

Cong, L. W., Li, Y., & Wang, N. (2021). Tokenomics: Dynamic adoption and valuation. *The Review of Financial Studies*, *34*(3), 1105–1155.

Corbet, S., Hou, Y. G., Hu, Y., Larkin, C., Lucey, B., & Oxley, L. (2022). Cryptocurrency liquidity and volatility interrelationships during the COVID-19 pandemic. *Finance Research Letters*, *45*, 102137.

Dellarocas, C. (2003). The digitization of word of mouth: Promise and challenges of online feedback mechanisms. *Management Science*, *49*(10), 1407–1424.

Demirer, R., & Kutan, A. (2006). Does herding behavior exist in Chinese stock markets? *Journal of International Financial Markets, Institutions and Money*, *16*(2), 123–142.

Diebold, F. X., & Rudebusch, G. D. (1999). *Business cycles: Durations, dynamics, and forecasting*. Princeton University Press.

Diewert, W. E. (1976). Exact and superlative index numbers. *Journal of Econometrics*, *4*(2), 115–145.

Dowling, M. (2022a). Fertile LAND: Pricing non-fungible tokens. *Finance Research Letters*, *44*, 102096.

Dowling, M. (2022b). Is non-fungible token pricing driven by cryptocurrencies? *Finance Research Letters*, *44*, 102097.

Drehmann, M., Oechssler, J., & Roider, A. (2005). Herding and contrarian behavior in financial markets: An internet experiment. *The American Economic Review*, *95*(5), 1403–1426.

Economou, F., Kostakis, A., & Philippas, N. (2011). Cross-country effects in herding behaviour: Evidence from four South European markets. *Journal of International Financial Markets, Institutions and Money*, *21*(3), 443–460.

Fama, E. F., & French, K. R. (1993). Common risk factors in the returns on stocks and bonds. *Journal of Financial Economics*, *33*(1), 3–56.

Financial Times. (2021a). *How digital art dominated 2021*.

Financial Times. (2021b). *How NFTs became a $40bn market in 2021*.

Fu, J., & Wu, L. (2021). Regime-switching herd behavior: Novel evidence from the Chinese A-share market. *Finance Research Letters*, *39*, 101652.

Hartwich, E., Ollig, P., Fridgen, G., & Rieger, A. (2023). Probably something: A multi-layer taxonomy of non-fungible tokens. *Internet Research*, *34*(1), 216–238.

Hategan, V. (2023). *Dead NFTs: The evolving landscape of the NFT market*. dappGambl.

Hemenway Falk, B., Tsoukalas, G., & Zhang, N. (2022). Economics of NFTs: The value of creator royalties. Available at SSRN: https://ssrn.com/abstract=4284776

Holden, R., & Malani, A. (2022). An examination of velocity and initial coin offerings. *Management Science*, *68*(12), 8515–9218.

Horky, F., Rachel, C., & Fidrmuc, J. (2022). Price determinants of non-fungible tokens in the digital art market. *Finance Research Letters*, *48*, 103007.

Huang, D., & Goetzmann, W. N. (2023). *Selection-neglect in the NFT bubble* [NBER Working Paper, ID 31498].

Huo, X., Liu, X., & Zheng, W. (2023). The term structure of mutual fund herding. *European Financial Management*, *29*(3), 901–929.

Kahneman, D., & Tversky, A. (1979). Prospect theory: An analysis of decision under risk. *Econometrica*, *47*(2), 263–291.

Kapoor, A., Guhathakurta, D., Mathur, M., Yadav, R., Gupta, M., & Kumaraguru, P. (2022). TweetBoost: Influence of social media on NFT valuation. *Companion Proceedings of the Web Conference 2022*, *22*, 621–629.

Kirilenko, A., Kyle, A. S., Samadi, M., & Tuzun, T. (2017). The flash crash: High-frequency trading in an electronic market. *The Journal of Finance*, *72*(3), 967–998.

Kong, D.-R., & Lin, T.-C. (2021). Alternative investments in the fintech era: The risk and return of non-fungible token (NFT). Available at SSRN: https://ssrn.com/abstract=3914085

Korteweg, A., Kräussl, R., & Verwijmeren, P. (2016). Does it pay to invest in art? A selection-corrected returns perspective. *The Review of Financial Studies*, *29*(4), 1007–1038.

Kuminoff, N., Parmeter, C., & Pope, J. (2010). Which hedonic models can we trust to recover the marginal willingness to pay for environmental amenities? *Journal of Environmental Economics and Management*, *60*(3), 145–160.

Lancaster, K. J. (1966). A new approach to consumer theory. *Journal of Political Economy*, *74*(2), 132–157.

Li, S., & Chen, Y. (2023). How non-fungible tokens empower business model innovation. *Business Horizons*, *66*(4), 543–544.

Mandel, B. R. (2009). Art as an investment and conspicuous consumption good. *American Economic Review*, *99*(4), 1653–1663.

Mazur, M. (2021). Non-fungible tokens (NFT). The analysis of risk and return. Available at SSRN: https://ssrn.com/abstract=3953535

McAndrew, C. (2022). *Collectors are spending more, not less, on NFT art, according to the art market report*. Art Basel and UBS.

Mei, J., & Moses, M. (2002). Art as an investment and the underperformance of masterpieces. *The American Economic Review*, *92*(5), 1656–1668.

Mekacher, A., Bracci, A., Nadini, M., Martino, M., Alessandretti, L., Aiello, L. M., & Baronchelli, A. (2022). Heterogeneous rarity patterns drive price dynamics in NFT collections. *Scientific Reports*, *12*(1), 13890.

Milgrom, P. R., & Weber, R. J. (1982). A theory of auctions and competitive bidding. *Econometrica*, *50*(5), 1089–1122.

Muehlenbachs, L., Spiller, E., & Timmins, C. (2015). The housing market impacts of shale gas development. *American Economic Review*, *105*(12), 3633–3659.

Nadini, M., Alessandretti, L., Di Giacinto, F., Martino, M., Aiello, L. M., & Baronchelli, A. (2021). Mapping the NFT revolution: Market trends, trade networks, and visual features. *Scientific Reports*, *11*(1), 20902.

O'Brien, R. M. (2007). A caution regarding rules of thumb for variance inflation factors. *Quality & Quantity*, *41*(5), 673–690.

Ofek, E., & Richardson, M. (2003). DotCom Mania: The rise and fall of internet stock prices. *The Journal of Finance*, *58*(3), 1113–1137.

Park, A., & Sabourian, H. (2011). Herding and contrarian behavior in financial markets. *Econometrica*, *79*(4), 973–1026.

Parker, G. G., & Van Alstyne, M. W. (2005). Two-sided network effects: A theory of information product design. *Management Science*, *51*(10), 1494–1504.

Renneboog, L., & Spaenjers, C. (2013). Buying beauty: On prices and returns in the art market. *Management Science*, *59*(1), 36–53.

Rojas-Lamorena, Á.J., Del Barrio-García, S., & Alcántara-Pilar, J. M. (2022). A review of three decades of academic research on brand equity: A bibliometric approach using co-word analysis and bibliographic coupling. *Journal of Business Research*, *139*, 1067–1083.

Rosen, S. (1974). Hedonic prices and implicit markets: Product differentiation in pure competition. *Journal of Political Economy*, *82*(1), 34–55.

Rosen, S. (1981). The economics of superstars. *The American Economic Review*, *71*(5), 845–858.

Schaar, L., & Kampakis, S. (2022). Non-fungible tokens as an alternative investment: Evidence from CryptoPunks. *The Journal of The British Blockchain Association*, *5*(1), 1–12.

Shiller, R. (1995). Conversation, information, and herd behavior. *American Economic Review*, *85*(2), 181–185.

So, E. C., & Wang, S. (2014). News-driven return reversals: Liquidity provision ahead of earnings announcements. *Journal of Financial Economics*, *114*(1), 20–35.

Sockin, M., & Xiong, W. (2022). *Decentralization through tokenization* [NBER Working Paper, ID 29720].

Sotheby's. (2021). *NFTs: Redefining digital ownership and scarcity*.

Stiglitz, J. E. (1989). Markets, market failures, and development. *The American Economic Review*, *79*(2), 197–203.

Taleb, N. N. (2021). Bitcoin, currencies, and fragility. *Quantitative Finance*, *21*(8), 1249–1255.
Triplett, J. (2004). *Handbook on hedonic indexes and quality adjustments in price indexes: Special application to information technology products* [OECD Working Papers Series, 9].
Urquhart, A. (2016). The inefficiency of Bitcoin. *Economics Letters*, *148*, 80–82.
Valeonti, F., Bikakis, A., Terras, M., Speed, C., Hudson-Smith, A., & Chalkias, K. (2021). Crypto collectibles, museum funding and OpenGLAM: Challenges, opportunities and the potential of non-fungible tokens (NFTs). *Applied Sciences*, *11*(21), 9931.
Vayanos, D., & Wang, J. (2012). Liquidity and asset returns under asymmetric information and imperfect competition. *The Review of Financial Studies*, *25*(5), 1339–1365.
Wang, Y., Horky, F., Baals, L. J., Lucey, B. M., & Vigne, S. A. (2022). Bubbles all the way down? Detecting and date-stamping bubble behaviours in NFT and DeFi markets. *Journal of Chinese Economic and Business Studies*, *20*(4), 415–436.
Wilson, K. B., Karg, A., & Ghaderi, H. (2021). Prospecting non-fungible tokens in the digital economy: Stakeholders and ecosystem, risk and opportunity. *Business Horizons*, *65*(5), 657–670.
Yousaf, I., & Yarovaya, L. (2022). Herding behavior in conventional cryptocurrency market, non-fungible tokens, and DeFi assets. *Finance Research Letters*, *50*, 103299.

How to cite this article: Fridgen, G., Kräussl, R., Papageorgiou, O., & Tugnetti, A. (2024). Pricing dynamics and herding behaviour of NFTs. *European Financial Management*, 1–41. https://doi.org/10.1111/eufm.12506

APPENDIX A: DATA EXTRACTION, CLEANING AND PREPARATION

For the on-chain part of our analysis, we use Dune, Etherscan and the Application Programming Interfaces (APIs) of Alchemy and OpenSea, while for the off-chain portion, we employ the price oracle of Chainlink and the APIs of Discord and Twitter.

Using Dune, we extract every on-chain transaction involving ERC-721 and ERC-1155 tokens on six of Ethereum's largest NFT marketplaces, specifically OpenSea, Rarible, NFTX, LooksRare, Foundation and SuperRare. The extraction period spans from the establishment of the ERC-721 standard in January 2018–February 2022.[2] From the retrieved transactions, we identify 14,580 collections that have sold at least one NFT. Using Etherscan and OpenSea, we identify the deployers and the artists behind each collection and gather all the available metrics on a collection level, such as floor price, the total number of owners, collection name, number of transfers, the total volume traded in ETH and the artists' characteristics, like, gender and the total number of artists.

In a second step, we exclude the collections whose NFTs were involved in fewer than 30 sales to establish statistical significance within collections. Following the categorization criteria of https://Nonfungible.com, we divide the collections into six categories: art, collectibles, game, metaverse, utility and other. Since our research focuses on art NFTs, we omit collectibles as they are comprised of a single artwork with minor variations per NFT. We also exclude NFTs associated with games and metaverses since their price is also influenced by the in-game or metaverse mechanics. Further, we remove utility NFTs such as liquidity provider positions in Uniswap V3 and other NFTs that do not fit in any of the preceding categories, like, music NFTs or those with unique features.

[2]The data extraction stopped in February 2022 due to the Russo-Ukrainian war, which introduced substantial levels of systematic volatility, making it impractical to examine the volatility of an individual asset class.

TABLE A1 Variable descriptions.

This table reports all the variables extracted from on and off blockchain sources alongside their short descriptions.

Variable	Description
average price	Average nonfungible token (NFT) price
BGCI return	Daily Bloomberg Galaxy Crypto Index last price in USD
black	Percentage of black color in the NFT image
blue	Percentage of blue color in the NFT image
buyer	ETH address of the NFT buyer
buyer seller pair	Unique buyer–seller combined ETH addresses
collection name	Name of the NFT collection
CSAD	Cross-sectional absolute deviation
cum score	Artist cumulative score
cum score year rank	Artist yearly cumulative score
demand	Demand of NFTs by each artist category, as ratio between *share* and its geometric mean
deployer creator generalities	Deployer type: Single artist (male, female), artists collaboration, company
deployer creator name	Name of the deployer
discord account	Presence of a Discord account representing the artist or the NFT collection
discord server	Name of the artist Discord account
floor price	NFT floor price
followers count	Number of followers per Twitter account
following count	Number of following per Twitter account
gray	Percentage of grey color in the NFT image
green	Percentage of green color in the NFT image
last price	NFT last selling price
listed count	Average number of public lists memberships per Twitter account
market return	Overall mean return across NFT collections
marketplace collection	NFT belonging to one between: Foundation (FND), Editional, KnownOrigin, SuperRare
max price	Maximum NFT price
newcomer proportion	Proportion of newcomers in the NFT market
nft	Single NFT's ETH address
nft type	Belonging to the ERC721 or ERC1155 protocol per single NFT
norm shannon entropy	Normalized Shannon Entropy per NFT image
num of colors	Number of colors in the NFT image
num of owners	Number of NFT's owners overtime

(Continues)

TABLE A1 (Continued)

Variable	Description
num of trades	Number of trades per NFT
orange	Percentage of orange color in the NFT image
platform of last sale	Blockchain platform of last sale per NFT
purple	Percentage of purple color in the NFT image
red	Percentage of red color in the NFT image
quote count month	Average number of "Retweets" with comments per Twitter account per month
score	Artist score
seller	ETH address of the NFT seller
share	Artist market share, computed as ratio between artist specific *volume* and the overall market *volume*
sold ratio	Percentage of NFT sold per artist's collection
timediff	Time difference (in days) between the collection creation and the selling dates per NFT
twitter account	Presence of a Twitter account representing the artist or the NFT collection
twitter handle	Name of the artist Twitter account
usd amount	NFT transaction specific USD amount
volume	Artist's market capitalization
verified	Presence of "verified" badge per Twitter account
white	Percentage of white color in the NFT image
year collection creation	Year of creation per NFT collection
yellow	Percentage of yellow color in the NFT image

On the basis of this data set, we continue with the off-chain data extraction on Twitter and Discord. After manually identifying the Discord Servers and Twitter handles of the NFT collections (or in the case where these were missing their artists), we retrieve all publicly available metrics such as the number of followers, retweets, likes, replies, number of messages as well as discord messages and users that where involved in the general and announcement channels. At the end of the data extraction and cleaning on a collection level, our data set comprises 531 art NFT collections containing a total of 1,460,718 NFTs.

Using the Alchemy and OpenSea APIs, we retrieve all on-chain available information on these NFTs.[3] Consequently, we further exclude the NFTs whose data are missing, those that are in video or animated Scalable Vector Graphics (SVG) formats and those with wrong data formats. In a final step, we download every file associated with each NFT and converted SVG

[3]We opted for the use of both APIs because we identified that each has missing data. By utilizing both, we were able to cover most of the gaps.

images to Portable Network Graphics (PNG) to allow for easier image analysis. Our final sample contains 2.15 terabytes (TB) of PNG, Joint Photographic Experts Group and Graphics Interchange Format (GIF) files of 875,389 NFTs.

From the extracted files, we determine their colour proportions and the total number of colours present by utilizing the Hue, Saturation, Value (HSV) colour model.[4] We group every colour around nine commonly used colours: black, white, grey, red, green, blue, yellow, purple and orange. In addition, we assess the complexity of the files using Shannon's entropy and the block decomposition method, with the use of the coding theorem method, because it is regarded as one of the best approaches both from a statistical and an algorithmic point of view. In the case of GIF files, we average the values of each frame.

So-called wash trading relates to a form of market manipulation where a small number of investors repeatedly buy and sell the same asset, generating an inorganic market activity. It creates artificial trading volume and gives the appearance that the asset is more in demand than it actually is. Wash trading can greatly affect the price, the traded volume and the selling frequency. We exclude in our analysis all individual NFTs that have been involved in wash trading and all collections whose trading volume is greater than 90% due to wash trading, such as Terraforms by Mathcastles. Additionally, we eliminate NFTs whose first sale took place after 01-01-2022, since more than 75% of the total volume in January 2022 and 55% of the total volume in 2022 was a result of wash trading and identifying the genuine transactions would be complicated and risky.

Using Dune, we determine that 875,389 of the aformentioned art NFTs were involved in 385,884 sales. On the basis of these sales, we generate two data sets, one containing data pertaining to every sale of the extracted art NFTs (*Transaction data set* in Table 1), and the other comprising aggregated data per NFT, which was augmented with data from Discord and Twitter, and colour and image complexity data (*Aggregate data set* in Table 1). In both data sets, the sale price is measured in USD, for which we utilize the exchange rates provided by Chainlink at the time of the sale for the individual NFT sales and at the end of the day on 10 February 2022 for the collection floor price.

[4]HSV is a mathematical abstraction that describes how colours can be represented by electronic systems. In this model, colours are represented by a cylindrical geometry whose angle represents the hue, the x-axis determines saturation and the value depends on the y-axis.