Investor Experience Matters: Evidence from Generative Art Collections on the Blockchain*

Sebeom Oh[†], Samuel Rosen[‡], Anthony Lee Zhang[§]

February 2022

Abstract

In the market for non-fungible tokens on the blockchain, experienced investors systematically outperform inexperienced investors. Controlling for holding period, experienced investors make 10 percentage points more per trade. NFT collections purchased by experienced investors sell out more often and more quickly in primary markets, and experience higher price growth in secondary markets. Our results suggest that NFT markets are characterized by high degrees of informational inefficiency, allowing investors with informational advantages to systematically extract profits.

*We are grateful to Dean Karlan for helpful comments, and Wanran Zhao for excellent research assistance. We also thank Michael Bassett and Josh Rosen for sharing their insights into the NFT market. Financial support from the Young Scholars Interdisciplinary Forum at Temple University is gratefully acknowledged.

[†]Temple University Fox School of Business; sebeom.oh@temple.edu.

[‡]Temple University Fox School of Business; sir@temple.edu.

[§]University of Chicago Booth School of Business; anthony.zhang@chicagobooth.edu.

1 Introduction

Non-fungible tokens (NFTs) are "digital collectibles": unique, indivisible, durable digital assets on blockchains, often used to represent works of visual art. The NFT market has experienced explosive growth, increasing from $94.9 million in trading volume in 2020 to $24.9 billion in 2021,[1] and a number of traditional non-crypto firms have started initiatives to sell NFTs.[2] Many well-known NFT collections have generated outsized returns for investors. For example, Bored Ape Yacht Club NFTs sold for 0.08 ETH in primary markets in April 2021 (roughly $160 USD at the time) and the cheapest Bored Ape NFT is currently listed in February 2022 for sale at 93 ETH (roughly $255,000 USD). As a result, many investors have flooded into NFT markets in the hopes of achieving similar returns.

In many ways, NFTs are an extremely opaque asset class. Most NFTs are one-of-a-kind meaning that there is a vast number of different assets available for purchase. NFT creation is also entirely unregulated, leading to a proliferation of low-quality projects. Public information on new NFT collections is often sparse, and information about the potential value of new collections instead tends to percolate through informal social networks. As a result of these features, a commonly held view is that experienced investors have a substantial advantage in investing in NFT markets due to their superior information about the quality and potential of NFT projects. This hypothesized advantage can similarly be described in terms of skill: experienced investors may have a better understanding of what new collections will perform well in secondary markets for NFTs.

Do experienced investors outperform inexperienced investors in NFT markets? How much do they outperform, and what are the mechanisms that drive their outperformance? This paper addresses these questions using a comprehensive dataset of NFT transactions on the Ethereum blockchain. We find that experienced investors systematically outperform: they attain roughly 10 percentage points higher returns on each trade compared to inexperienced investors for trades with similar holding periods. NFT collections purchased by experienced investors are more likely to sell out in primary markets, sell out faster, and experience subsequent higher price growth in secondary markets. Moreover,

[1] See Reuters.

[2] Some examples are the NBA, the Australian Open, the British Museum, and Adidas.

changes in the share of collection items owned by experienced investors also predict changes in collection prices.

Given that the NFT market essentially emerged in 2021, our empirical investigation requires substantial data collection. We begin by compiling a comprehensive list of NFT collections featured on OpenSea, the most popular NFT marketplace. From this list, we restrict attention to 692 "generative" collections (henceforth "GCs") that comprise 2.6 million individual NFTs. Roughly speaking, we define an NFT collection as a GC if the associated digital artwork consists of unique pictures based on a common theme and the NFTs are created through a public primary market sale. In practice, NFTs from GCs appear to provide value to their owners as verifiable and tradable status goods. For example, they are commonly used as profile pictures on social media. We focus on GCs so that the NFT collections in our sample are comparable to each other. Importantly, our sample of GCs are a nontrivial subset of the broader NFT market making up nearly half of primary market sales. Additionally, they include many of the most successful and well-known NFT collections, such as the Bored Ape Yacht Club.

A key benefit of studying digital blockchain-based assets is the availability of comprehensive transaction-level data. We rely on dataset from Moonstream that includes all on-chain transactions for a large set of NFT collections between April 1, 2021, and September 25, 2021. The GC-filtered dataset has over 4.4 million transactions of which approximately 60% are primary market sales and the remainder are secondary market transactions. Importantly, the data include the wallet addresses for both the seller and buyer in each transaction that allows us to perform our investor-based analysis.

Our main finding is that that experienced investors substantially outperform inexperienced investors in their returns from investing in GCs. We define experienced investors as the subset of wallets that are most active within the set of all GCs. Controlling for holding period, we find that experienced investors earn 10 percentage points higher returns on average for each realized trade compared to inexperienced investors. We document even greater outperformance when focusing on differential returns within the same collection. Using estimates of unrealized gains, we show that our outperformance findings are not driven by experienced investors being more likely to realize their positive gains. In sum, our findings support the view that experienced investors harbor skills or advantages that allow them to achieve higher returns in the GC market in general.

To further assess the hypothesis that experienced investors are skilled, we analyze whether experienced investor participation in an NFT collection is associated with subsequent collection success. The first simple measure of success we consider is whether an NFT collection "mints out" – that is, whether the collection is able to sell the entire quantity of NFTs originally allocated to be sold in primary markets. Successful mints are essentially a precondition for post-mint price growth, as successful mints are a signal of strong demand for the NFT collection, and unsold inventory in primary markets remains available for investors to purchase, imposing an effective cap on prices in secondary markets. We find that collections which have a larger fraction of experienced investors participating in the mint stage are more likely to mint out, and mint out more quickly. The magnitudes of these associations are quantitatively large. For example, our estimates imply that a collection with a 1 p.p. higher fraction of experienced investors is also 1.206 p.p. more likely to mint out. We also note that the fraction of experienced investors is quantitatively important in explaining the variation in our outcome variables according to associated R^2 values.

Beyond the success of the primary market itself, we show that collections with a larger fraction of experienced buyers at the mint stage also experience greater price appreciation after the mint stage. This association persists in secondary markets too: when the share of items in a collection owned by experienced investors increases, the price of the collection also tends to increase. This latter finding in particular demonstrates that experienced investors' asymmetric knowledge applies not only to the mint phase of new collections, but also to collections that have minted out and are only trading in secondary markets.

As a final step, we analyze features of collections that predict experienced investor participation. Here we find that experienced investors tend to buy collections with featured artists who also have established online presences. On the other hand, experienced investors avoid collections with roadmaps, that advertise "rare" items, and that are derivatives of well-known collections.

This paper is closely related to a literature which analyzes return differences between experienced and inexperienced investors, in asset classes characterized by high degrees of asset heterogeneity and asymmetric information. A number of papers have analyzed persistent differences in returns across VC and PE funds. Sørensen (2007) shows that companies funded by more experienced VCs are more likely to go public. Relatedly, Nahata (2008) show that firms backed by more reputable VCs are more likely to successfully exit. Kaplan and Schoar (2005) show that there are large and persistent differences in the performance of different partnerships in private equity. In the online fundraising space, Dmitri and Risteski (2021) study the investment behavior of serial and large investors in initial coin offerings, while Kim and Visawanathan (2019) study the role of experienced early investors on a crowdfunding platform.

There is also a literature on differences in performance of different investors in housing markets. Kurlat and Stroebel (2015) show that, when the composition of home sellers in a neighborhood shifts towards more informed agents, neighborhood prices tend to decline, suggesting that a subset of market participants have superior information about common values of the asset. Chinco and Mayer (2016) show that out-of-town home buyers behave like misinformed speculators, driving up prices, but achieving lower than average returns. Bayer et al. (2020) show that experienced house flippers substantially outperformed speculators who entered the housing market during the housing boom. DeFusco, Nathanson and Zwick (2021) show that, over the 2006 housing boom and bust, cities which experienced a larger increase in the share of short-term buyers had larger price booms and busts. Cvijanović and Spaenjers (2021) show that out-of-country buyers in the housing market of Paris buy at higher prices and sell at lower prices than local investors.

This paper is also related to a body of work studying the properties of art as a financial asset. Renneboog and Spaenjers (2013) measures returns on a large dataset of art transactions. Korteweg, Kräussl and Verwijmeren (2016) shows that accounting for selection into sale is important for quantifying the returns on art investments. Lovo and Spaenjers (2018) constructs a model of trading in art markets. Penasse and Renneboog (2021) shows evidence of speculative bubbles in the art market, and Pénasse, Renneboog and Scheinkman (2021) shows evidence that an artist's death is associated with permanent increases in price and volumes of the art. Our paper is also related to the emerging literature on NFTs in particular. Nadini et al. (2021) analyzes statistical properties of the network of NFT transactions, as well as using machine learning algorithms to predict NFT prices.

Relative to the literature, our contribution is to demonstrate evidence for asymmetric information and returns in a new and rapidly growing market. We find that experienced investor behavior is strongly predictive of collection price growth. To our knowledge, we are also the first to show that experienced traders' behavior predicts returns in art markets more generally.

The remainder of our paper proceeds as follows. We describe the institutional background for NFTs and our data sources in Section 2. We describe our data in Section 3 including several stylized facts and descriptions for how we measure our key variables of interest. Section 4 contains our empirical results. We conclude in Section 5.

2 Institutional Background and Data

2.1 Institutional Background

Non-fungible tokens (NFTs) are digital assets that exist on a blockchain. Like other blockchain-based digital assets, an NFT is necessarily associated with a blockchain-based digital wallet (henceforth "wallet") at any given point in time. Each wallet has a public address, as well as a private key that only the wallet owner is supposed to know. For any given wallet, any person who knows its public address can view its contents. However, the wallet's private key is needed in order to spend cryptocurrencies, and buy, sell, or transfer NFTs.

As their name implies, NFTs differ from cryptocurrencies, such as Bitcoin and Ethereum, in that each NFT is indivisible and distinct from other NFTs. The distinct nature of any given NFT can most clearly be seen in its unique identifier on the blockchain, but this is not the only aspect that makes it unique. NFTs generally represent pieces of digital artwork by embedding metadata about the associated file.[3] NFTs are most often meant to uniquely represent their associated digital artwork.[4] In this sense, owning an

[3]In theory, NFTs can represent any digital good but digital artwork is by far the most common in practice.

[4]In some cases, an artist will create multiple NFTs for the same piece of digital artwork. Each of these NFTs will have a unique address on the blockchain although they clearly do not uniquely represent the associated artwork in this case. This situation would be like if an artist painted multiple copies of the

NFT is like having the unique digital certificate of authenticity for the associated artwork.[5]

On their surface, NFTs have a lot in common with both collectibles and art. They are identifiable and scarce goods whose only tangible benefits can be tied to the ownership claim itself. As such, one key way that NFTs appear to provide value to their owners are as verifiable and tradable status goods. For example, NFTs are often used as profile pictures on social media. In fact, Twitter introduced an NFT profile picture integration feature in January 2022 that allows users to demonstrate the blockchain ownership of their profile picture NFTs. This feature works by presenting verified NFT profile pictures with a hexagonal border in contrast to the circular shape of non-NFT profile pictures. Further, one can click on an NFT profile picture to obtain a description of the NFT collection and related links. The Twitter profile picture integration thus allows NFT owners to verifiably signal ownership of high-value NFTs. NFTs from a given collection often grant access to exclusive virtual social groups,[6] and there have also a number of in-person events restricted to verified owners of NFTs from certain collections.[7] Finally, NFTs can be displayed in virtual art galleries in the "metaverse."[8]

NFT Collections. Individual NFTs are usually associated with a broader collection, an aspect that also makes NFTs similar to collectibles and art. NFT collections are often formalized through a smart contract on the blockchain (i.e., a piece of software code) that is connected to each NFT within the collection. The fact that many NFTs are formally

same object that visually appeared identical. In our analysis, we focus on NFTs that are truly unique in representing their associated artwork.

[5]There is an ongoing debate in the legal world about whether and how NFTs can be seen as a legitimate ownership claim on the associated artwork. Our empirical analysis and conclusions do not require taking a stand in this debate as all of our main findings remain after controlling for unobservable aggregate factors (time fixed effects) and unobservable collection-level features (collection fixed effects). By controlling for these factors, we are also accounting for potential time-varying beliefs about whether ownership of a given NFT or collection would entail intellectual property rights as well.

[6]For example, the Bored Ape Yacht Club collection has a private chat group which require verified ownership of a Bored Ape NFT to enter. It is also very common for NFT collections to have private chat groups on Discord gated to verified token holders: examples of collections which have such groups are Doodles, Cool Cats, and Pudgy Penguins. The mechanism through which these chat groups work is that the NFT owner must "sign" a message, proving private-key ownership of a wallet which can be publicly proven to possess a certain NFT, in order to join the private chat groups.

[7]One prominent example is that there have been a number of in-person meetups for members of the Bored Ape Yacht Club. Another example is that VeeFriends token holders get access to a multi-day exclusive event hosted by the creator called VeeCon.

[8]For example, Sotheby's has a virtual gallery in Decentraland.

(a) #3733, 8.88 ETH (b) #6670, 4.10 ETH (c) #1699, 4.09 ETH

Figure 1: SupDucks Example: Select Items Traded on September 25, 2021
Notes. The items displayed in this figure are 3 examples from among the 25 items from the SupDucks GC that traded on September 25, 2021. The 25th percentile price value on this date was 4.10. The captions include the specific item numbers within the collection and corresponding prices in ethereum (ETH) observed in the trades.

assigned to collections provides two benefits for our empirical analysis. The first is that we can easily identify and group together NFTs in our data. The second is that we will be able to control for common collection-level features across sets of NFTs.

Our analysis in this paper focuses on "generative" NFT collections (henceforth "GCs"). These are collections of roughly 5,000-10,000 NFTs around a common theme. We provide a formal definition for these collections in Section 2.2. A specific example of a GC is SupDucks, which consists of 10,000 pictures of cartoon ducks. We provide a few examples of NFTs from the SupDucks GC in Figure 1. Other GCs are often similar in nature except that they are based on a different central object (e.g., apes). We rely on SupDucks as a concrete example throughout our data description in Section 2.2.

Primary Market ("Minting"). When we refer to the primary market, we are referring to the process in which NFTs are initially created on the blockchain and sold to investors. An NFT is generated on a blockchain in either one of two broad methods. The first method is that a creator can simply generate an NFT, associated with any image, into their own wallet. From this point, the creator can sell or transfer this NFT to another wallet as they would in any secondary market transaction. The second method is that the creator can set up a smart contract, which serves essentially as a vending machine for NFTs. Buyers can then "mint" NFTs from the smart contract, by sending a pre-specified cryptocurrency amount (i.e., the "mint price") to the smart contract. The smart contract then creates the NFT and sends it to the purchaser's wallet.

The GCs that we analyze in this paper use the latter "minting" method to sell NFTs. Collections have websites with key details of the collection, such as the price per NFT and the start date of the public sale. Buyers can initiate the smart contract transaction to mint the NFT simply by clicking a "mint" button on the NFT website, and will then pay the mint price and receive a random NFT from the collection. Thus, at the mint stage, buyers purchase from collections, but cannot target specific elements from the collection.

GC primary market sales can differ along a few dimensions. Perhaps most importantly, the number of NFTs within the collection and the pre-specified mint price can vary substantially. Together, these choices determine the amount of funds that the creator hopes to raise by selling their NFT collection in the primary market. GC creators can also choose to restrict the set of potential purchasers to set of pre-determined wallet addresses (i.e., a "whitelist"), which clearly reduces the "public" nature of the sale. The decision to do so is often motivated by the desire to reward early investors and active community members. In practice, earning a whitelist invitation usually involves some level of participation in the GC's Discord channel prior to the public sale.[9]

Secondary Market. After being minted, NFTs can be traded in a secondary market. As of the time of writing, the largest NFT secondary market platform is OpenSea and therefore we will use it to summarize the secondary market in general.[10] OpenSea serves both as a catalog of the NFT universe and a platform through which buyers and sellers can initiate trades. OpenSea organizes the NFTs by collection and reports key collection-level statistics on the associated collection page. For example, OpenSea reports the "floor" price (i.e., the lowest currently listed price for an NFT from a given collection) as a way

[9] Discord is a popular instant messaging app that effectively serves as a large chat room. It has become the default choice among NFT issuers to facilitate communication with potential investors before and after a primary market sale.

[10] OpenSea is an example of a decentralized NFT exchange in which buyers and sellers can interact without any intermediary. Other examples of decentralized NFT exchanges include Rarible and SuperRare. An example of a centralized NFT exchange is Nifty Gateway, which is owned and operated by the centralized cryptocurrency exchange and custodian Gemini.

to communicate the cheapest price at which an investor can buy into a collection. To initiate a trade, potential NFT purchasers and sellers first connect their wallets to OpenSea by showing their public address, which allows OpenSea to detect all NFTs and funds within their wallets. NFT sellers can then list each NFT they wish to sell at a specified price. Listed offers are binding: buyers can immediately purchase any listed NFT at the posted price. Buyers can also make binding unsolicited bids on NFTs, and owners can immediately sell any of their NFTs at the bid price. In exchange for its services, OpenSea charges a flat 2.5% transaction fee for each realized trade.

In addition to the OpenSea transaction fee, there are two important fees that investors pay in the secondary market. The first is the royalty fee, which determines the share of the transaction price paid back to the creator. Most GCs specify a royalty rate that is 2.5%, 5%, 7.5%, or 10%. If present, the royalty rate is specified directly in the collection-level smart contract so that it is automatically paid in every secondary market transaction captured on the blockchain. The underlying technology for NFTs (i.e., programmable smart contracts), however, makes such royalties both feasible and convenient. Proponents of the NFT space argue that the easy ability to charge royalty rates is an important innovation that "supports new business and profit models" (Kaczynski and Kominers, 2021). The presence of automatically generated royalties also incentivizes NFT creators by giving them a continuing stake in the success of their collections.

The second important additional fee that investors pay in secondary market transactions are "gas" fees. Gas refers to the transaction fees which must be paid on any interaction with the Ethereum blockchain: they must be paid when NFTs are minted, as well as when they are traded on secondary markets. These fees are paid to Ethereum "miners", computer nodes which solve computationally hard problems in order to embed transactions into the blockchain through a "proof-of-work" process. Gas fees tend to be high when there is high demand for transactions on Ethereum. These fees are a potentially important factor in NFT investor decisions. Unfortunately, our current transaction-level data do not include a breakdown gas fees. We account for this omission in our empirical analysis by including time fixed effects and focusing on relative return levels rather than absolute return levels. By doing so, we effectively control for time-varying gas fees that may be influencing our empirical estimates. In future analysis, we plan to gather gas fee data and explicitly account for these costs in our calculations.

2.2 Data

As noted in the previous subsection, we focus on a set of collections we call "generative" NFT collections (henceforth "GCs"). In simple terms, we define an NFT collection as a GC if the associated digital artwork features a common theme and each individual NFT represents a unique variation on that theme. As an example, the associated artwork for SupDucks are 10,000 unique pictures of cartoon ducks (see, e.g., Figure 1) that feature various sets of characteristics combined essentially randomly and combinatorially. Additionally, we require GCs to mint their NFTs through a public sale in which which buyers pay a fixed amount to receive a random NFT within the collection. See Appendix A for our complete formal GC definition as well as justifications for each individual restriction. As a general point, the overarching reason for carefully defining GCs as we do is so that the NFT collections in our sample are comparable to each other. GCs are also a relatively popular form of NFT collection, accounting for approximately half of the amount of funds raised in the primary market in our full sample (Appendix B).

The first step in assembling the data for our analysis is to identify the universe of GCs. We first compile the full list of NFT collections featured on OpenSea, the most popular NFT marketplace as noted in Section 2.1. This step, which we performed on a few dates in October 2021, generated an initial list of 7,987 NFT collections. After applying the filters from our GC definition, we find 692 GCs in total (see Appendix A a more detailed description of this process). Throughout the remainder of this section, we document the ways in which this relatively small set of GCs actually represents a relatively large share of the broader NFT market. For example, many of the largest and well-known NFT collections are GCs in our sample such as the Bored Ape Yacht Club, Cool Cats, World of Women, and Pudgy Penguins (Appendix Table A.3).[11]

Our primary data source is a transaction-level dataset is from Moonstream.[12] The Moonstream data include all on-chain transactions for the GCs in our sample between April 1, 2021, and September 25, 2021. The GC-filtered dataset has over 4.4 million transactions of which approximately 60% are mints and the remainder are secondary

[11] A notable exception is that the Crypto Punks collection is not in our sample because its minting period occurred well before the period covered in our transaction-level data.

[12] See Moonstream (2021) and https://github.com/bugout-dev/moonstream/tree/main/datasets/nfts.

market transactions (Table 1). Importantly, the data include the wallet addresses for both the seller and buyer in each transaction that allows us to perform our investor-based analysis in Section 4. See Appendix A for additional details regarding the Moonstream data.

Table 1: Overview of Transaction-Level Data

Notes. In this table, we describe the sample size of the transaction-level data available for the GCs in our sample. "Mint" is the common term in practice to refer to the primary market sale and on-chain creation of a new item. "Transfer" refers to any observed post-mint transaction.

	N	Mean
Is Mint	4,467,156	0.59
Is Transfer	4,467,156	0.41
Positive Price if Mint	2,648,747	0.91
Positive Price if Transfer	1,818,409	0.77

We supplement our transaction-level dataset with data on collection-level features. Specifically, we manually gather these data from GC-specific websites and Twitter accounts. They include variables such as whether the collection has a dedicated Twitter profile and whether the specific artist(s) of the associated digital artwork are explicitly named. We describe these features in more detail in Appendix A.3. The main purpose of gathering these data is to use them as control variables in our collection-level analyses in Section 4.2. We provide summary statistics for these characteristics in Appendix Table A.2.

3 Stylized Facts and Measurement

In this section, we document several stylized facts about the GC market using our data. We also explain how we construct the key variables used in our empirical analysis.

3.1 Secondary Market Activity

First, we discuss trading patterns of GC NFTs. In Figure 2, we show the distribution of all GC NFTs according to their cumulative secondary market trading activity in our sample. Here, we observe that 63.6% of the total 2.6 million GC NFTs minted never trade in the secondary market within our sample period. In other words, only 37.4% of GC NFTs have ever traded in the secondary market.

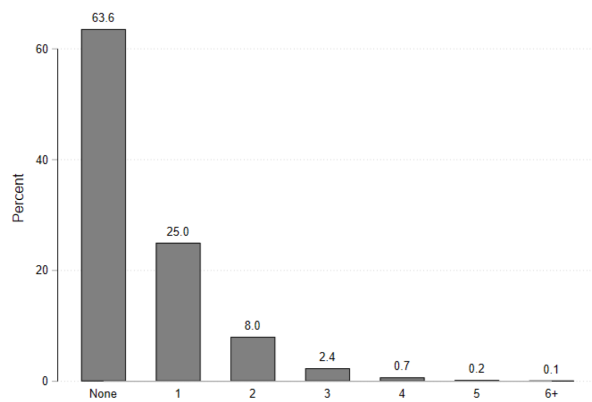


Figure 2: GC Sample: Shares of All NFTs by Number of Times Traded

Notes. This figure reports the shares of all GC NFTs according to the amount of times the item is traded during our sample period. We only consider an observed transaction to be a trade if it the price is nonzero. "None" refers to the case in which an item never trades after its mint.

In Table 2, we summarize measures of secondary market activity at the collection level. Our first takeaway from this table is that the volume of trading varies substantially across GCs. For example, the typical GC experiences at least 5 trades on around half of the days following its minting period. This figure, however, ranges from 0% to 100% when we look across GCs. In fact, roughly 100 GCs never experienced a day with at least 5 trades. We will occasionally refer back to the statistics in Table 2 throughout our empirical analysis in Section 4.

Table 2: GC Sample: Trading Period Variables

Notes. In this table, we summarize variables pertaining the trading period of a GC. With the exception of royalty rate, which was manually gathered, all of these variables are computed from transaction-level data. We only consider an observed transaction to be a trade if it the price is nonzero. Royalties earned are estimated as the royalty rate times the volume traded. Total funds raised is the sum of funds raised through minting and royalties earned.

	N	Mean	SD	Min	10%	50%	90%	Max
N Trades and Transfers	692	2,627.76	4,619.80	0.00	25.00	479.00	8,750.00	34,966.00
N Trades	692	2,032.13	3,579.73	0.00	7.00	291.00	6,984.00	24,229.00
N Trades / N Items	692	0.32	0.39	0.00	0.02	0.14	0.90	2.73
N Trades / N Days	692	69.66	143.49	0.00	0.29	9.21	184.57	1,536.86
Frac. Items Ever Traded	692	0.23	0.23	0.00	0.02	0.13	0.61	0.94
N Days with At Least 5 Trades	692	17.91	22.98	0.00	0.00	9.00	51.00	148.00
Frac. Days with At Least 5 Trades	692	0.51	0.41	0.00	0.00	0.47	1.00	1.00
Frac. Days with At Least 5 Trades (> 0)	592	0.60	0.37	0.01	0.07	0.67	1.00	1.00
Volume Traded (ETH)	692	1,047.91	7,050.97	0.00	0.50	21.16	1,720.76	157,484.57
Royalty Rate	692	0.04	0.03	0.00	0.02	0.05	0.09	0.10
Royalties Earned (ETH)	692	35.79	190.62	0.00	0.00	0.70	61.50	3,937.11
Royalties Earned to Total Funds Raised	692	0.06	0.13	0.00	0.00	0.01	0.16	0.86

3.2 Collection-Level Price Indexes

Next, we analyze NFT prices from secondary market transactions and propose a method to compute collection level price indexes. GCs essentially consist of two types of items: rare and common. The select "rare" items trade at prices much higher than others in the collection while the remaining "common" items tend to trade around the same much lower price. To demonstrate this fact, we regress log NFT prices on collection-date fixed effects:

$$\log p_{j,c,t} = \nu_{tc} + \epsilon_{j,c,t} \tag{1}$$

where $p_{j,c,t}$ is the price in ETH for NFT j from GC c sold on date t. Figure 3 plots the distribution of the exponentiated price residuals, $\exp\left(\epsilon_{j,c,t}\right)$, from specification (1). The distribution is noticeably right-skewed (skewness value of 4.3). This means that a small number of NFTs trade at prices much higher than others in the same collection, but few NFTs trade substantially below the median price. Quantitatively, our estimates from (1) imply that the 90th percentile NFT price for a given collection-day (1.85) is roughly 100% higher than the median (0.93), whereas the 10th percentile price (0.62) is only roughly 33% lower.

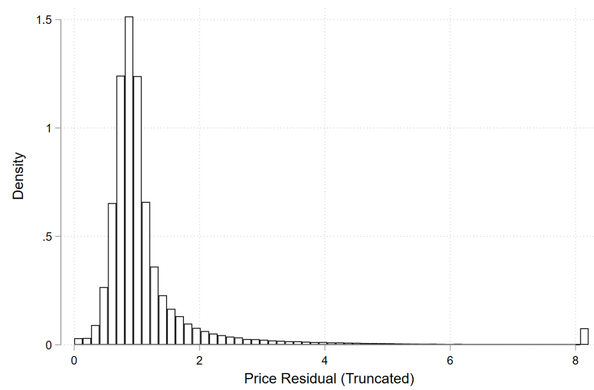


Figure 3: Price Regression Results

Notes. This figure reports the distribution of residuals from the regression specification in (1) of log NFT prices on date-collection fixed effects. The residuals used for plotting are truncated at their 99th percentile value and the distribution statistics discussed in the text are also computed from these data.

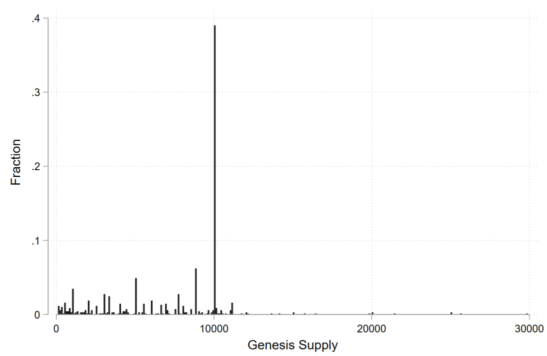
In order to calculate total returns to investors accounting for unsold NFTs, we need to value unsold NFT inventory. Doing so can be tricky, however, given that that most NFTs never trade (Figure 2). This problem exists at the collection level too where we observe similar fraction-ever-traded numbers (Table 2). To account for this issue, we rely on the above takeaway from Figure 3 that most NFTs in a collection tend to trade around the same price. This fact can be quantified using the regression result that GC-level fixed effects captures over 60% in the variation in log prices, according to the R^2 statistic from the estimates of Specification (1). Further, we observe that date-GC-level fixed effects capture roughly 80% of the variation. These results imply that a collection-level price index would be a reasonable approximation for the value of any NFT in that collection.

Based on the above facts, we propose to measure daily collection-level price indexes as the 25th percentile price provided that there are at least 5 trades. The choice of 25th percentile is to ensure that we grab a price from the middle to lower end of the distribution, which is where the majority of NFTs in a given collection appear to valued

(Figure 3). The choice of 5 trades as a minimum is to ensure that the 25th percentile value is reasonably well estimated. This approach to pricing a collection is similar to the concept of a "floor" price, which is reported on OpenSea and often discussed in NFT market commentary. The key difference is that the "floor" price is based on the lowest available offers to purchase an item at any given moment, while our measure is based on realized trades that are observable on the Ethereum blockchain.

3.3 Primary Market Activity and Outcomes

Next, we analyze the primary market for GCs. Recall that the simultaneous purchase and creation of an NFT is commonly referred to as "minting," and therefore we will use this term synonymously with primary market activity. We report summary statistics for variables that characterize the minting periods of our GCs in Table 3 and Figure 4.



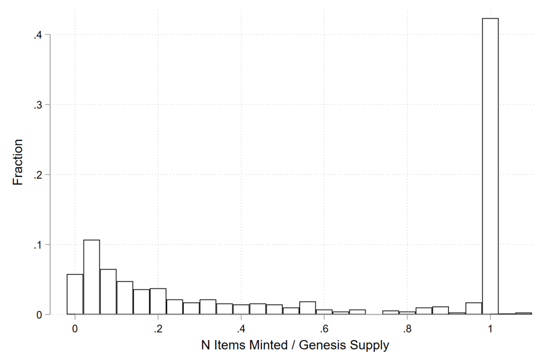


Figure 4: Distribution of genesis supply and fraction of supply sold
Notes. This figure shows the distribution of the genesis supply, and the fraction of the genesis supply sold during the minting period (N Items Minted / Genesis Supply). Each data point is one GC.

Our first takeaway from Table 3 is that GCs experience different degrees of success in selling their collection of NFTs in the primary market. Here we focus on the fraction of NFTs sold by the GC relative to the initial set that it planned to sell, which we refer to as the "genesis supply." Only 42% of GCs are successful in selling their entire genesis supply. Many GCs are very unsuccessful: conditional on not selling their entire genesis

Table 3: GC Sample: Minting Period Variables

Notes. In this table, we summarize variables pertaining the minting period of a GC. With the exception of genesis supply, which was manually gathered from GC-specific webpages, all of these variables are computed from transaction-level data. Weighted average mint price is the total amount of ETH raised in mint transactions divided by the total number of items minted. Average items minted per wallet is the total number of items minted divided by the number of minting wallets. Days to mint full collection is only computed for GCs that raised over 99% of their collection. It is measured in fraction of days and the ending time is the time of the mint that pushes the GC over the 99% minted threshold.

	N	Mean	SD	Min	10%	50%	90%	Max
N Items Minted	692	3,827.32	3,992.92	6.00	236.00	1,957.50	10,000.00	25,000.00
Genesis Supply	692	7,388.49	3,933.11	99.00	1,111.00	8,888.00	10,000.00	29,886.00
N Items Minted / Genesis Supply	692	0.57	0.42	0.00	0.04	0.58	1.00	1.09
N Items Minted / Genesis Supply (< 99%)	403	0.27	0.28	0.00	0.02	0.16	0.76	0.99
Frac. Minted at Price > 0	692	0.88	0.19	0.05	0.61	0.96	1.00	1.00
Dummy Minted All Genesis	692	0.42	0.49	0.00	0.00	0.00	1.00	1.00
Weighted Average Mint Price (ETH)	692	0.07	0.20	0.00	0.02	0.05	0.09	3.33
Funds Raised through Minting (ETH)	692	244.70	897.47	0.09	7.84	84.89	635.92	22,070.66
Implied Funds Raised Goal (ETH)	692	495.70	1,539.11	2.46	50.00	293.58	788.48	22,070.66
Number of Minting Wallets	692	805.23	899.54	1.00	72.00	527.00	1,884.00	8,207.00
Average Items Minted per Wallet	692	10.08	126.97	1.00	2.16	3.91	8.29	3,333.00
Max Frac. Items Minted by Wallet	692	0.10	0.14	0.00	0.02	0.05	0.23	1.00
Days to Mint Full Collection	289	6.71	12.76	0.00	0.11	2.16	16.46	124.40

supply, a GC only sells 27% of their supply on average. This dichotomy in outcomes can be further validated by the bimodal cross-sectional distribution for the fraction of genesis supply sold variable presented below Table 3. Of course this variation in success can be similarly seen in other outcome variables of interest. For example, conditional on selling their entire genesis supply, GCs differ in how quickly they sell out. The 90th percentile successful collection takes over 16 days to sell out whereas the 10th percentile GC takes less than one day.

The second takeaway from Table 3 is that GCs aim to earn different amounts of funds in their primary market sales. The average and median collection have an implied fundraising goal of roughly 500 ETH and 300 ETH, respectively. Using an exchange rate of $3,000 per ETH, these numbers imply that the typical GC aims to raise between $0.9 and $1.5 million through the sale of their NFTs. Given low rates of success, however, the median GC in our sample actually raised 84 ETH (approximately $250,000) through minting.[13] The implied fundraising goal of a GC is calculated as the product of two GC-specific choices: the average mint price and the genesis supply. In our sample, we find that the typical GC has a mint price around 0.05 ETH, or $150 based on an exchange rate of $3000 per ETH.[14] We also find that approximately 40% of GCs plan to sell 10,000 NFTs, although this number ranges from around 100 to almost 30,000.

The success of primary market sales is important because GCs which mint out successfully and quickly tend to experience higher price growth. This occurs for two reasons. First, minting out quickly is a signal to the market that the NFT is in high demand. Second, if a GC does not mint out, the primary market serves as competition for the secondary market. If the secondary market floor price rises above the mint price, an investor can simply mint a new NFT from the GC rather than buying one from the secondary market. Figure 5 shows how minting period success is associated with price growth using the ratio of the post-minting GC price index to average mint price. In the left panel, we compare the distributions for this ratio between the set of collections that successfully and the set of collections that did not. In line with our expectation, we find that secondary market prices tend to be higher than mint prices for collections that mint out, and lower for collections that do not. Focusing only on collections that successfully minted out, the right panel shows a binned scatter plot of the price index ratio against the time it takes for a collection to mint out. Collections that mint out more quickly also experience higher price growth relative to mint prices.

[13] At the largest end, we observed that a GC raised the equivalent of roughly $66 million. This GC is Meebits, which was launched in May 2021 by the same company that launched the first wildly successful NFT collection in 2017 named CryptoPunks. In general, the largest GCs are among the largest NFT collections that can be observed in the full Moonstream data as well. The three largest NFT collections in the Moonstream data that do not meet our full GC definition (and hence are not in our sample) are Art Blocks, Mutant Ape Yacht Club, and VeeFriends.

[14] In practice, mint prices denominated in ETH are established in advance of the minting period. However, there are two complications in determing a representative mint price for any given GC. The first is that GC creators have the ability to mint items for free. These are typically done as part of giveaways and related promotions to generate interest in the GC. We find that the typical GC mints 90%–95% of its collection at a positive price. The second complication is that there can be a schedule of mint prices that are based on factors such as number of items minted. Given these two complications, we compute the weighted average mint price as the total amount of ETH raised in mint transactions divided by the total number of items minted. This average price has the helpful property that multiplying it by the number of NFTs actually minted yields the total amount of funds raised.

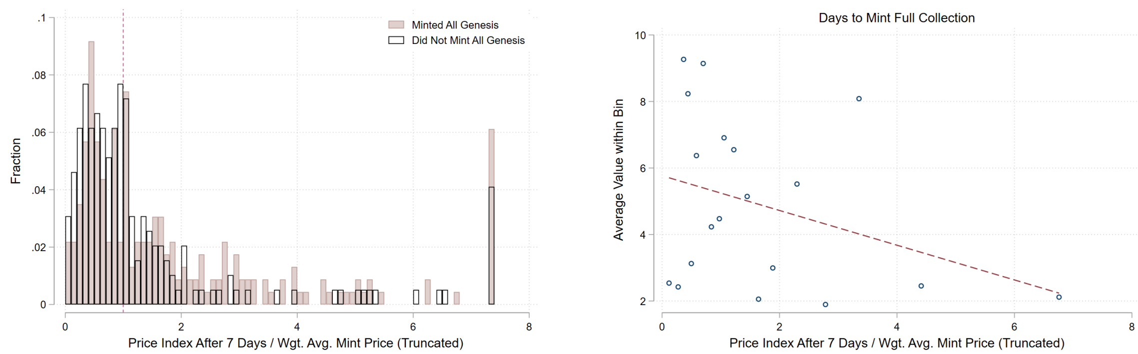


Figure 5: GC Sample: Post-Mint Price Index Growth by Minting Period Success

Notes. The left panel reports the cross-sectional distributions of the ratio of the GC's price index 7 days after minting began compared to the weighted average mint price. The underlying values are truncated at the 95th percentile value for the ratio in the entire sample for visual purposes. These distributions are reported separately for the set of GCs that sold over 99% of their genesis supply and the set of GCs that did not. The right panel reports a binned scatter plot for the same two variables but only using data for GCs that sold over 99% of their genesis supply.

In sum, the above observations motivate using various measures of minting success in our analysis: a dummy for whether a collection mints out, the time it takes for a collection to mint out, and the total amount of ETH raised in the mint process.

On the investor side, we note that the minting period for the typical collection includes between 500-800 wallet addresses. For the purposes of our analysis, we consider each wallet to effectively be a unique investor. This assumption is mainly so we can use the term "investors" rather than "wallets" throughout our exposition, which makes the intuition for our findings more clear. Combined with the number of items minted, these wallet address counts imply that the typical investor in any GC mints between 4 to 10 items. To provide a sense of concentration, we also measure the maximum share of items minted across minting wallets within each GC. Here, we find that the typical GC minting period features a largest investor that purchases between 5% and 10% of the entire collection.

3.4 Defining Experienced Investors

The main object of our analysis is to study how experienced investors differ from inexperienced investors in the NFT market. In short, we will define experienced investor wallets as those that conducted a relatively large number of transactions. For the purposes of our analysis, we consider each wallet to effectively be a unique investor. This assumption is mainly so we can use the term "investors" rather than "wallets" throughout our exposition, which makes the intuition for our findings more clear. The are over 300,000 unique wallet addresses that appear in our transaction-level data.

Figure 6 demonstrates the high degree of concentration for both NFT minting and trading activity among investors. We obseve that a relatively large share of NFT market activity is attributable to a relatively small fraction of wallets. In the top left panel, the x-axis displays the number of distinct GCs minted from by a given wallet, and the y-axis shows the fraction of all GC mints executed by wallets that minted from at most the given number of GCs. For example, the value of the line at $x = 6$ is around 50%, implying that half of all NFT mints are performed by wallets that minted from at least 6 GCs in our data. Analogously, the right plot shows the number of distinct GCS that a wallet traded within on the x-axis, and the cumulative fraction of trades on the y-axis. Half of all NFT trades are performed by wallets that traded within at least 11 trades in our data.

We propose to identify a subset of "experienced" investors as those that minted from at least 6 GCs and traded within at least 11 GCs. These thresholds are precisely those described above and shown in Figure 6. According to these values, around 12,000 (or 3.6%) of investors are "experienced." In Figure 6, we also report the corresponding figures based on either threshold. Here, we see that the requirement to meet both thresholds instead of at least one cuts the implied share of experienced investors by more than half. We consider this dual requirement important, however, because the notion of experience should imply an understanding of both the primary and secondary markets.

We characterize experienced GC investors further by comparing their entry dates into the sample and relative trading activity in Figure 7. There are two main takeaways from this figure. First, experienced GC investors entered the sample much earlier than inexperienced in general (see the left panel). This finding, however, is simply a consequence of the way we define experienced investors. Those entering the sample

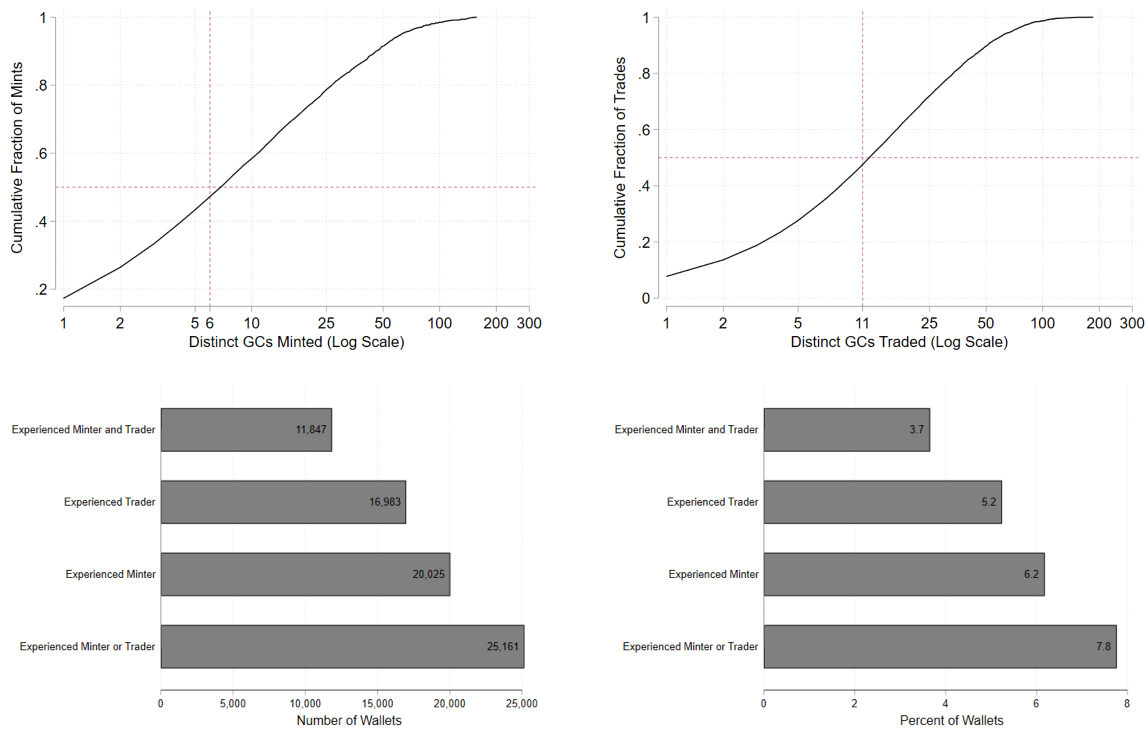


Figure 6: Investor Concentration in GC Activity

Notes. This figure reports the cumulative shares of mints (trades) when wallets are grouped and ordered by the number of distinct GCs with which they minted (traded) in the top left (top right) panel. We only consider an observed transaction to be a trade if it the price is nonzero. The vertical dashed line in the top left (top right) panel denotes the maximum number $\overline{\text{N}}$ such that at least 50% of GC items minted (traded) were done so by wallets that had minted from (traded within) $\overline{\text{N}}$ or more GCs. The bottom left (bottom right) panel reports the number of (percent of) wallets that satisfy both, one, or either distinct GC threshold.

earlier have had more time to interact with different GCs, and therefore are more likely to meet our "experienced" definition. The second takeaway is that, controlling for entry date, experienced investors simply trade more (see the right panel). While this feature is also related to how we define them, it is interesting to note that the bulk of experienced investors only trade on 20%-80% of the days since they entered the sample meaning that they are not necessarily trading every day. Thus it does not appear that experienced investors we identify are simply high-frequency traders or "bots."

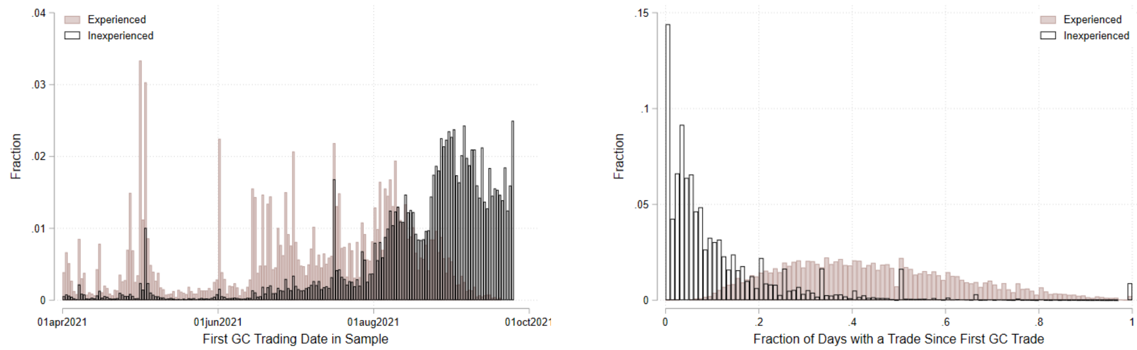


Figure 7: Characterizing Investor Groups: Experienced versus Inexperienced

Notes. This figure reports the cross-sectional distributions by investor type for GC sample transaction data entry date and fraction of days with a trade since entering the sample. Experienced investors are defined as those with sufficient minting and trading activity (see Figure 6). We only consider an observed transaction to be a trade if it the price is nonzero.

4 Results

In this section, we present the results from our empirical analysis. We begin by investigating differences in returns achieved by experienced investors as defined in Section 3.4 compared to all other GC investors in our data (i.e., "inexperienced" investors). These differences are the focus of this paper, and we find that returns for experienced investors are significantly higher on average using several different approaches. In the remainder of the section, we attempt to understand and explain these findings by investigating the relationships between experienced investor activity, GC outcomes, and GC characteristics.

4.1 Realized Returns

Using our transaction-level data, we compute the realized return attributable to any given transaction as follows:

$$r^{\text{realized}}_{i,j,c,t,\tau} \equiv \frac{\text{Price}^{\text{Sold}}_{i,j,c,t} - \text{Price}^{\text{Purch}}_{i,j,c,\tau}}{\text{Price}^{\text{Purch}}_{i,j,c,\tau}} \tag{2}$$

where $\text{Price}^{\text{Sold}}_{i,j,c,t}$ is the price received by investor i when they sell NFT j from collection c on date t that they had previously purchased on date τ at $\text{Price}^{\text{Purch}}_{i,j,c,\tau}$. Note that the previous purchase price is observed in a different transaction that we need to connect to the ultimate sale. Specifically, we identify the investor i and the previous purchase price as the wallet address and price observed in the most recent transaction with a positive price for the given NFT j. In other words, before computing realized returns we drop all transactions with a zero price, which we interpret as transfers between wallets of the same investor.

Before delving into a formal regression approach at the trade level, it is instructive to first look at aggregate figures. Specifically, we can follow the same method described above to compute realized returns at the investor group level. The only difference here is that we sum up all sold prices and purchase prices across all realized trades by members of a specified investor group. By doing so, we are computing the weighted average realized return for all investors in that group. In the left panel of Figure 8, we compare these aggregated realized returns between experienced and inexperienced investors. There are two clear takeaways. The first is that aggregate returns to investors were very high during our sample, with values near 150% for investors as whole. It is important to note when interpreting them, however, that these return values do not account for gas fees, transaction fees, and royalties as discussed in Section 2.1. The second is that experienced investors appear to have substantially outperformed inexperienced investors. In the right panel of Figure 8, we report aggregate realized profits, which are the sum of the numerators used in all realized return observations. Here we find that total realized profits in the GC market were over 400,000 ETH (or \$1.2 billion using an exchange rate of \$3,000 per ETH). This amount represents the net additional inflow of investor capital into GCs beyond the funds raised in primary market sales.

To formally assess the apparent outperformance of experienced investors shown in Figure 8, we estimate regression specifications of the following form:

$$r^{\text{realized}}_{i,j,c,t,\tau} = \beta \times \text{Experienced Seller Dummy}_i + \gamma \log(t-\tau) + \mu_c + \eta_t + \xi_{ct} + \zeta_j + \epsilon_{i,j,c,t,\tau} \quad (3)$$

where the dependent variable is the realized return to investor i for NFT j in collection

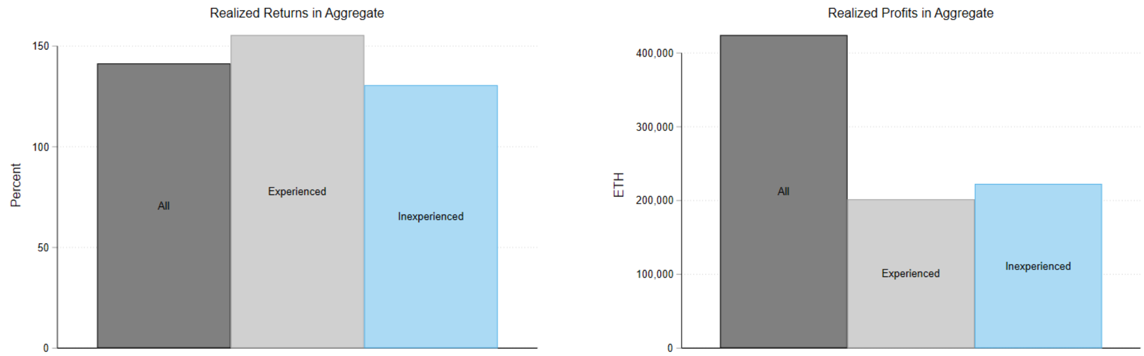


Figure 8: Aggregate Realized Returns and Profits by Investor Type

Notes. This figure reports the weighted average returns for all realized trades within each investor group. Realized profits are computed using the sum of prices received in selling transactions minus the sum of prices paid in the original purchase transactions $\left(\sum \mathrm{Price}^{\mathrm{Sold}} - \sum \mathrm{Price}^{\mathrm{Purch}}\right)$. Realized returns are computed as the realized profits divided by the sum of prices paid in the orignal purchase transactions $\left(\left(\sum \mathrm{Price}^{\mathrm{Sold}} - \sum \mathrm{Price}^{\mathrm{Purch}}\right) / \sum \mathrm{Price}^{\mathrm{Purch}}\right)$. When determining the investor type of the seller in a realized trade, we use the wallet address of the purchaser from the most recent transaction with a non-zero price.

c as defined in (2). The right-hand side variables include a dummy for whether the associated investor is in our experienced group and the log of the fractional number of days the position was held. We also include sets of fixed effects that vary across specifications, for collections, dates, collection-dates, and individual items.

The results from our trade-level realized return regressions are shown in Table 4. The positive coefficient value for the experienced seller dummy in the first column shows that the unconditional average realized return for experienced investors is indeed higher than that of inexperienced investors. The result in column (2) shows that, for trades with similar holding periods, experienced investors earn roughly 10 percentage points higher returns than inexperienced investors. Experienced investors tend to have much shorter holding periods than inexperienced investors. This fact tends to lower experienced investors' average returns per trade, because returns are higher for longer holding periods when prices are rising as the GC market experienced during most of our sample. As a result, experienced investors' outperformance is larger when controlling for holding period.

Table 4: Regressions at Trade Level: Realized Returns

Notes. In this table, we report the results from estimates of specification (3) in which we regress realized returns for each NFT on an experienced seller dummy, the log of the holding period, and various fixed effects. We only include realized return values where the purchase price was 0.01 ETH or more, and these values are further winsorized at the 1st and 99th percentile level. Holding period is the fractional number of days the position was held. Standard errors are heteroskedasticity-consistent. t-statistics are in parentheses. $^{*}p < 0.10$; $^{**}p < 0.05$; $^{***}p < 0.01$.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Experienced Seller Dummy	0.037***	0.100***	0.115***	0.430***	0.414***	0.435***	0.395***
	(3.70)	(9.88)	(11.32)	(45.94)	(44.10)	(47.70)	(22.62)
ln(Days to Realize)		0.122***	0.122***	0.203***	0.314***	1.008***	0.277***
		(54.86)	(53.78)	(98.90)	(126.59)	(230.90)	(59.68)
Date FE	No	No	Yes	No	Yes	No	Yes
GC FE	No	No	No	Yes	Yes	No	No
Date-GC FE	No	No	No	No	No	Yes	No
Item FE	No	No	No	No	No	No	Yes
R^2	0.000	0.002	0.026	0.205	0.229	0.345	0.411
N	1,347,099	1,344,092	1,344,092	1,344,081	1,344,081	1,341,811	707,902

In columns (3) to (6), we add date, collection, and date-collection fixed effects. The coefficient on the experienced seller dummy actually increases substantially with the inclusion of collection fixed effects. For example, the result in column (4) implies that experienced investors achieve 43.5 percentage point higher returns than inexperienced investors conditional on trading items from the same collection. Finally, the result in column (7) shows that experienced seller outperformance remains even after including NFT-level fixed effects. When an experienced investor and an inexperienced investor trade the exact same NFT at different points in time, the experienced investor obtains 39.5 percentage point higher returns on average than the inexperienced investor.

One concern about our realized return results is that they may be biased by selection into sale: perhaps experienced investors tend to sell NFTs that are doing well, whereas inexperienced investors tend to hold on to them. In Appendix C.1, we explore this possibility by calculating unrealized returns on NFTs held until the end of the sample using our collection-level price indexes. Using analogous approaches as in Figure 8 and

Table 4, we find similar outperformance by experienced investors in terms of unrealized returns compared to realized returns. These findings suggest that selection into NFT sales is not driving our experienced investor outperformance results. Rather, they support the view that experienced investors harbor skills or advantages that allow them to achieve higher returns in the GC market in general. We explore the potential sources of these skills or advantages in the remainder of this section.

4.2 Experienced Investors and NFT Collection Success

One potential advantage that experienced investors may have is that they understand better which new GCs are going to be successful after their primary market sale. Under this hypothesis, we would expect to see more experienced investors participating in the primary market sales of those GCs. In this section, we test this hypothesis using our measures of GC success described in Subsection 3.3 and a measure of primary market participation by experienced investors.

The first step in this process is to define an explicit measure of experienced investor involvement for each GC. We propose the share of experienced investor wallets that participate in the GC's primary market sale:

$$\text{Frac. Minted by Experienced} = \frac{\text{NFTs Minted by Experienced}}{\text{All NFTs Minted}}. \tag{4}$$

We need to be careful, however, about how we identify experienced wallets for this type of measure. A potential concern if we use the exact definition in Section 3.4 is that our measure would incorporate ex post information from the perspective of any given GC because both the GC activity thresholds and investor levels of trading activity would be based on the full sample. To mitigate this concern, we rely on ex ante indicators from the perspective of a given GC for identifying experienced wallets when computing its measure in (4). See Appendix A.4.1 for more details on the construction of this measure and comparisons to a version based on the full sample.

Before proceeding, it is instructive to first visually assess the relationship between our measure of experienced investor involvement in (4) and collection-level measures of success. We do so using binned scatter plots in Figure 9. The top left panel shows that collections in which a larger fractions of investors are experienced investors are much more likely to "mint out" (i.e., sell their entire genesis supply). The relationship is very strong: the highest quantile of experienced investor participation is around 80% likely to mint out, whereas the lowest quantile is associated with only a 10% probability of minting out. The top right panel shows that this result also holds if we use a continuous measure for the dependant variable, which is the fraction of genesis supply that is sold. The bottom panel shows that collections with more experienced participation mint out faster. Specifically, collections with around 60% experienced participation mint out in a few days on average whereas collections with less than 10% usually take closer to a month.

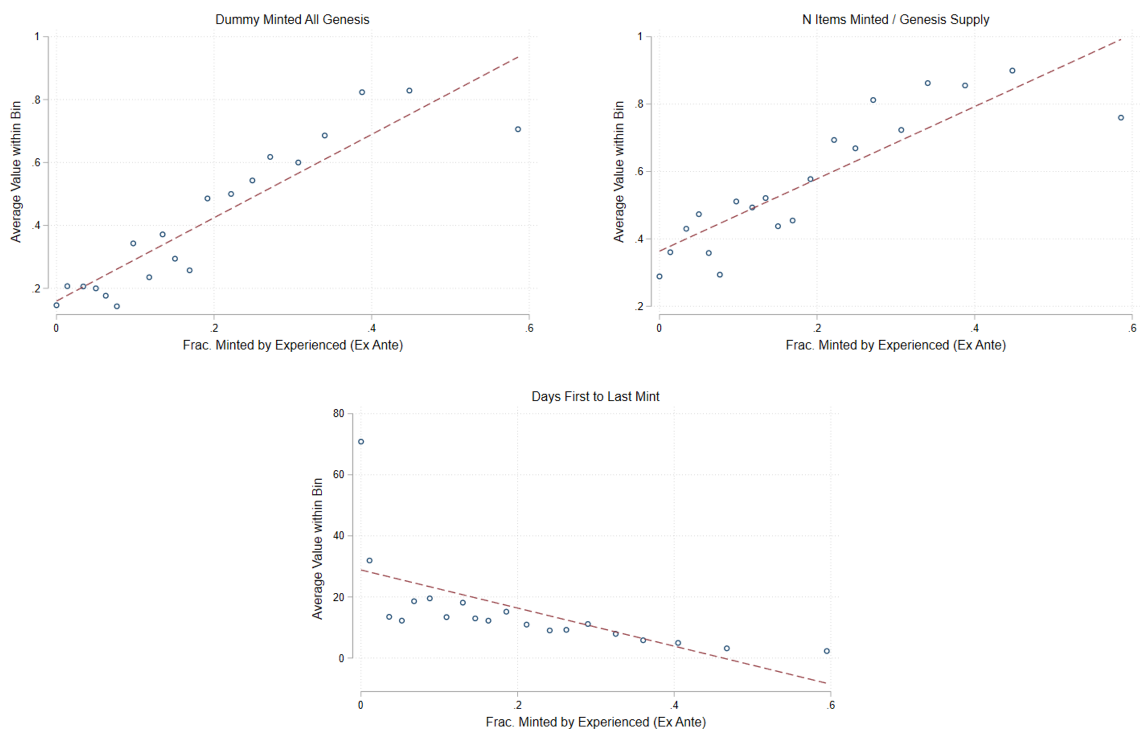


Figure 9: Fraction of Experienced Investors and Minting Period Success

Notes. The figure reports binned scatter plots to visualize the relationship between our measure of experienced investor involvement from (4) and collection-level measures of success.

We next estimate cross-sectional regression specifications of the following form:

$$y_{c,t} = \beta \times \mathrm{Frac.\,Minted\,by\,Experienced}_c + \Gamma' X_c + \nu_t + \epsilon_{c,t} \tag{5}$$

where the dependent variable is a collection-level outcome from the minting period of GC c that started during week t. Specifically, we consider the measures of GC success described in Subsection 3.3 as well as post-minting price index returns. The key explanatory variable is our collection-level measure of experienced investor involvement as defined in (4). We also control for other observable features of the collection and its minting period in addition to including fixed effects for the week in which the GC's primary market sale began.

The regression results reported in Table 5 confirm the suggestive findings from Figure 9. Specifically, we find that higher experienced investor participation is robustly associated with greater minting period success across all of our key measures controlling for many collection-level features. For example, our estimate in column (2) implies that a collection with a 1 p.p. higher fraction of experienced investors is also 1.206 p.p. more likely to sell its entire genesis supply in its primary market sale (i.e., "mint out"). Additionally, we note that the fraction of experienced investors explains the majority of the variance in the minting period outcome variables according to the R^2 values without and with the other control variables.

In Table 6, we report our cross-sectional regression results using GC post-mint price index returns as the dependent variables. These measures, which use mint price as the reference level, are meant to capture the initial success of a GC in the weeks following its minting period. Recall that our daily collection-level price indexes are computed as the 25th percentile prices observed on days with at least 5 trades (see Section 3.2). Therefore we are measuring the hypothetical return to an investor who minted from a collection and then sold it at the "common" collection price after N days. Across horizons up to 28 days, we find that higher experienced investor participation is associated with collections that experience higher post-mint price growth.

We interpret the findings in Tables 5 and 6 as joint support for the view that experienced investors are more skilled at picking successful GCs. Given this view, we are also interested in understanding what collection-level characteristics are associated with expe-

Table 5: Predicting Minting Period Success

Notes. In this table, we report the results from the cross-sectional regression specified in (5) where the dependent variable is a minting period outcome for a GC. The key explanatory variable is our collection-level measure of experienced investor involvement as defined in (4). See Section 2 for variable descriptions. Standard errors are heteroskedasticity-consistent. t-statistics are in parentheses. $^{*}p < 0.10$; $^{**}p < 0.05$; $^{***}p < 0.01$.

	Dummy Minted All Genesis		N Items Minted / Genesis Supply		ln(Days to Mint Full)	
	(1)	(2)	(3)	(4)	(5)	(6)
Frac. Minted by Experienced (Ex Ante)	1.389***	1.206***	1.104***	0.870***	-5.349***	-6.175***
	(10.91)	(8.99)	(9.71)	(7.36)	(-5.22)	(-7.88)
Frac. Minted at Price > 0		0.498***		0.229**		-0.410
		(4.56)		(2.39)		(-0.52)
Max Frac. Items Minted by Wallet		-0.438***		-0.716***		-0.002
		(-2.67)		(-4.25)		(-0.00)
Average Items Minted per Wallet		0.000*		0.000***		0.002***
		(1.71)		(2.72)		(4.23)
Has Twitter		0.100		-0.002		1.606*
		(0.94)		(-0.02)		(1.66)
Has Website		0.049		0.117		-0.063
		(0.57)		(1.34)		(-0.11)
Has Discord		0.040		0.052		-0.109
		(0.83)		(1.23)		(-0.33)
Has Roadmap		-0.085**		-0.113***		0.309
		(-2.10)		(-3.32)		(1.22)
Has Charity Description		-0.076*		-0.072*		-0.034
		(-1.77)		(-1.93)		(-0.12)
Advertises Rare Items		0.020		0.043		-0.166
		(0.56)		(1.44)		(-0.69)
ln(Weighted Average Mint Price)		-0.031		-0.007		-0.255*
		(-1.47)		(-0.35)		(-1.88)
Royalty Rate		-0.712		-0.222		-5.423
		(-1.05)		(-0.38)		(-1.23)
Has Named Artist		0.016		0.021		0.261
		(0.38)		(0.57)		(0.90)
Named Artist Has Twitter/Website		0.061		0.046		-0.193
		(1.24)		(1.12)		(-0.60)
Art is 3-D		0.022		0.045		0.107
		(0.58)		(1.46)		(0.48)
Art is Animated		-0.053		-0.030		0.765**
		(-0.77)		(-0.53)		(2.27)
Art Has Music		0.077		-0.021		-0.574
		(0.63)		(-0.20)		(-0.81)
Art Is Cute		0.001		-0.059		0.358
		(0.01)		(-1.06)		(1.03)
Art Is Punk Derivative		-0.037		-0.061		0.210
		(-0.44)		(-0.86)		(0.33)
Art Is BAYC Derivative		0.053		0.063		-0.162
		(0.48)		(0.72)		(-0.28)
Art Is Loot Derivative		-0.125		-0.070		-1.685**
		(-1.05)		(-0.51)		(-2.53)
Week FE	Yes	Yes	Yes	Yes	Yes	Yes
R^2	0.235	0.313	0.214	0.320	0.232	0.327
N	688	688	688	688	283	283

rienced investor involvement. To assess this question, we run cross-sectional specifications in which the dependent variable is our measure of experienced investor involvement in (4)

Table 6: Predicting Post-Minting-Period Price Index Returns

Notes. In this table, we report the results from the cross-sectional regression specified in (5) where the dependent variable is the post-minting-period price index return for a GC relative to its weighted average mint price. The key explanatory variable is our collection-level measure of experienced investor involvement as defined in (4). See Section 2 for variable descriptions. Standard errors are heteroskedasticity-consistent. t-statistics are in parentheses. $^{*}p < 0.10$; $^{**}p < 0.05$; $^{***}p < 0.01$.

	1 Day		7 Days		14 Days		21 Days		28 Days	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Frac. Minted by Experienced (Ex Ante)	1.123***	1.107***	1.079***	1.029**	1.670***	1.754***	1.980***	1.719***	2.274***	1.349*
	(3.47)	(3.20)	(2.70)	(2.31)	(3.56)	(3.64)	(3.69)	(2.70)	(3.02)	(1.68)
Frac. Minted at Price > 0		-1.849***		-1.100**		-0.586		-1.604**		-0.520
		(-5.39)		(-2.11)		(-1.02)		(-2.52)		(-0.82)
Max Frac. Items Minted by Wallet		-0.657		-0.128		1.165		-1.592*		-0.873
		(-1.44)		(-0.11)		(1.35)		(-1.93)		(-0.99)
Average Items Minted per Wallet		0.000		0.000		-0.000		-0.000		0.067**
		(0.38)		(0.63)		(-1.39)		(-0.17)		(2.34)
Has Twitter		0.460		0.535		1.192*		3.389***		2.850***
		(1.13)		(1.20)		(1.89)		(5.16)		(3.52)
Has Website		-0.271		-0.061		0.278		0.413		0.564**
		(-0.82)		(-0.20)		(0.67)		(1.08)		(2.01)
Has Discord		0.038		0.516**		0.588**		0.261		0.446
		(0.25)		(2.21)		(2.15)		(0.72)		(1.19)
Has Roadmap		-0.031		0.098		0.069		0.004		-0.050
		(-0.29)		(0.63)		(0.42)		(0.02)		(-0.23)
Has Charity Description		0.019		-0.013		-0.344*		-0.590**		-0.752**
		(0.16)		(-0.08)		(-1.91)		(-2.48)		(-2.48)
Advertises Rare Items		-0.006		-0.076		-0.336**		-0.258		-0.229
		(-0.06)		(-0.56)		(-2.33)		(-1.46)		(-1.08)
Royalty Rate		1.965		3.092		2.103		4.238		9.339***
		(0.92)		(1.16)		(0.67)		(1.24)		(2.74)
Has Named Artist		0.001		-0.126		-0.236		0.270		-0.314
		(0.01)		(-0.83)		(-1.11)		(1.27)		(-1.15)
Named Artist Has Twitter/Website		0.097		0.177		0.424**		0.174		0.722***
		(0.70)		(1.00)		(2.03)		(0.82)		(2.67)
Art is 3-D		0.037		0.016		-0.096		-0.268		-0.230
		(0.38)		(0.12)		(-0.64)		(-1.45)		(-1.03)
Art is Animated		-0.210		-0.042		0.061		0.298		0.380
		(-1.21)		(-0.19)		(0.27)		(1.13)		(1.44)
Art Has Music		0.170		0.339		0.767		0.714		0.715
		(0.90)		(0.89)		(1.03)		(1.41)		(1.21)
Art Is Cute		0.138		-0.072		0.462*		0.132		0.779**
		(0.94)		(-0.23)		(1.93)		(0.30)		(2.17)
Art Is Punk Derivative		-0.112		-0.287		-0.459		-0.410		-0.658
		(-0.65)		(-0.95)		(-0.97)		(-0.91)		(-0.73)
Art Is BAYC Derivative		0.348*		0.419		0.590		1.172***		1.256***
		(1.74)		(1.50)		(1.53)		(4.39)		(3.82)
Art Is Loot Derivative		-0.422		1.172**		-3.672***		0.000		0.000
		(-0.74)		(2.38)		(-15.58)		(.)		(.)
Week FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
R^2	0.075	0.205	0.053	0.116	0.090	0.227	0.125	0.239	0.114	0.302
N	429	429	417	417	368	368	302	302	245	245

and the explanatory variables include only the collection-level characteristics known prior to the minting period. We present these results in Table 7. Using both our ex ante and full sample versions, we find that experienced investors are less likely to participate in the primary market sales of GCs that have a roadmap, advertise rare items, or are a derivative of the Crypto Punks collection. On the other hand, they more likely to participate in the sales for projects that use an artist with a web presence.

For our final set of empirical tests, we explore the impact of experienced investor trading in secondary markets on collection-level price indexes. If experienced investors are better skilled at assessing the value of GCs, we would expect their movement into or out of a collection to predict future price increases or decreases, respectively. To formally test this hypothesis, we estimate panel regressions of the following form:

$$\begin{aligned}
&\Delta \ln \left(\mathrm{PriceIndex}_{c,t}\right) = \\
&\qquad \sum_{\tau=0}^{M} \left[\beta_{\tau,1} \ln \left(\mathrm{Trades}_{c,t-\tau}\right) + \beta_{\tau,2} \Delta \left(\mathrm{Frac.\,Owned\,by} \times \mathrm{Experinced}_{c,t-\tau}\right)\right] + \\
&\qquad \sum_{\tau=0}^{M} \left[\beta_{\tau,3} \times \ln \left(\mathrm{Trades}_{c,t-\tau}\right) \times \Delta \left(\mathrm{Frac.\,Owned\,by} \times \mathrm{Experinced}_{c,t-\tau}\right)\right] + \\
&\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad \eta_t + \eta_c + \epsilon_{k,t} \qquad (6)
\end{aligned}$$

where the dependent variable is the daily change in the log of the collection-level price index as described in Section 3.2. The explanatory variables include the log of the number of trades on a given date demeaned by collection, the change in the fraction of NFTs within a collection owned by experienced investors during a given date, the interaction between the two, date fixed effects, and GC fixed effects. Note that, relative to the measure used in the analysis thus far, we are using a time-varying measure of experienced investor ownership that is updated each day at midnight. We focus this analysis on the subset of GCs that were successful in selling their entire genesis supply (i.e., minted out) and those that had at least 7 days with sufficient trading to define price indexes.

We present our panel regression results in Table 8, and there are two key takeaways. The first is that we confirm that trading volume is a significant predictor of future price index returns. These relationships are positive for volume on the contemporaneous or prior day. The second is that increases in experienced investor ownership are robustly associated with increases in the collection price index. If we focus on the full specification

Table 7: Predicting Experienced Investor Involvement

Notes. In this table, we report the results from the cross-sectional regressions where the dependent variable is the collection-level measure of experienced investor involvement as defined in (4). See Section 2 for variable descriptions. Standard errors are heteroskedasticity-consistent. t-statistics are in parentheses. $^{*}p < 0.10$; $^{**}p < 0.05$; $^{***}p < 0.01$.

	Frac. Minted by Experienced (Ex Ante)		Frac. Minted by Experienced (Full Sample)	
	(1)	(2)	(3)	(4)
Has Twitter	0.011		0.010	
	(0.29)		(0.32)	
Has Website	0.047		0.036	
	(1.23)		(1.11)	
Has Discord	-0.017		-0.019	
	(-0.78)		(-1.02)	
Has Roadmap	-0.059***	-0.062***	-0.056***	-0.059***
	(-3.57)	(-4.21)	(-4.10)	(-4.93)
Has Charity Description	-0.005		-0.000	
	(-0.28)		(-0.02)	
Advertises Rare Items	-0.032**	-0.029**	-0.028**	-0.026**
	(-2.30)	(-2.05)	(-2.47)	(-2.32)
ln(Weighted Average Mint Price)	-0.001		0.002	
	(-0.09)		(0.22)	
Royalty Rate	0.452		0.354	
	(1.62)		(1.55)	
Has Named Artist	-0.002		-0.008	
	(-0.10)		(-0.63)	
Named Artist Has Twitter/Website	0.059***	0.060***	0.059***	0.055***
	(3.00)	(3.33)	(3.57)	(3.54)
Art is 3-D	0.026*		0.018	
	(1.71)		(1.46)	
Art is Animated	0.059**		0.023	
	(2.03)		(0.98)	
Art Has Music	-0.022		0.003	
	(-0.42)		(0.07)	
Art Is Cute	-0.032		-0.026	
	(-1.09)		(-1.17)	
Art Is Punk Derivative	-0.110***	-0.104***	-0.082***	-0.075***
	(-3.28)	(-3.10)	(-2.93)	(-2.72)
Art Is BAYC Derivative	0.038		0.016	
	(1.09)		(0.59)	
Art Is Loot Derivative	-0.025		-0.027	
	(-0.53)		(-0.89)	
Week FE	Yes	Yes	Yes	Yes
R^2	0.280	0.259	0.228	0.213
N	688	688	688	688

in column (6), this result can be seen both for experienced ownership on its own and its joint effect when combined with trading volume. The former effect can be seen in the positive coefficient on the change in experienced ownership the day before. The latter effect can be seen in the positive and significant coefficient value on the interaction term between trading volume and the change in experienced ownership observed during the same day. In other words, the association between price index changes and volume is stronger when that volume is comprised of experienced investors.

Table 8: Predicting Collection-Level Price Index Returns

Notes. In this table, we report the results from the panel regression specified in (6) where the dependent variable is the daily change in the log of the collection-level price index as described in Section 3.2. The explanatory variables include the log of the number of trades on a given date demeaned by collection, the change in the fraction of NFTs within a collection owned by experienced investors during a given date, the interaction between the two, date fixed effects, and GC fixed effects. Both the number of trades and change in experienced ownership are measured during the 24-hour period each day ending at midnight UST. The sample only includes the subset of GCs that were successful in selling their entire genesis supply (i.e., minted out) and those that had at least 7 days with sufficient trading to define price indexes. See Appendix C.2 for additional description of the underlying sample. Standard errors are heteroskedasticity- and autocorrelation-consistent and clustered at the date level. t-statistics are in parentheses. $^{*}p < 0.10$; $^{**}p < 0.05$; $^{***}p < 0.01$.

	(1)	(2)	(3)	(4)	(5)	(6)
ln(Trades) [t]	0.056***	0.074***		0.064***	0.055***	0.062***
	(6.68)	(6.11)		(5.51)	(4.29)	(5.16)
ln(Trades) [t-1]		0.054***		0.057***	0.008	0.061***
		(4.20)		(4.29)	(0.68)	(4.45)
ln(Trades) [t-2]		-0.093***		-0.084***		-0.084***
		(-11.00)		(-9.39)		(-9.14)
Change in Frac. Owned by Experienced [t]			0.134**	0.079	0.128*	0.094
			(2.04)	(1.21)	(1.93)	(1.50)
Change in Frac. Owned by Experienced [t-1]			0.348***	0.200**	0.259***	0.202**
			(4.33)	(2.48)	(4.35)	(2.44)
Change in Frac. Owned by Experienced [t-2]			0.122	0.064		0.063
			(1.49)	(0.80)		(0.79)
ln(Trades) [t] x Change in Frac. Owned by Experienced [t]					0.137**	0.117**
					(2.33)	(2.22)
ln(Trades) [t-1] x Change in Frac. Owned by Experienced [t-1]					0.109***	0.039
					(3.02)	(0.83)
ln(Trades) [t-2] x Change in Frac. Owned by Experienced [t-2]						-0.032
						(-0.63)
Date FE	Yes	Yes	Yes	Yes	Yes	Yes
GC FE	Yes	Yes	Yes	Yes	Yes	Yes
R^2	0.006	0.019	0.004	0.018	0.011	0.019
N	7497	7211	7016	7002	7260	7002

In sum, the results in Table 8 suggest that experienced investors' asymmetric knowledge applies not only to the mint phase of new collections, but also to collections that have minted out and are only trading in secondary markets.

4.3 Robustness and Alternative Explanations

One factor which may contribute to explaining experts' outperformance is that, as we discussed in Section 2.1, many projects – in a small sample we checked, almost 90% – restrict mint at early stages (or reserve a set of NFTs to mint) to "whitelists" of wallet addresses. Experienced investors may outperform partially because they have better access to investments, rather than better information: that is, because they are more likely to be whitelists of projects which are likely to be successful. We do not have data on whitelists, so we cannot fully distinguish between the access and information hypotheses. However, since the vast majority of projects appear have whitelists, experienced investors are unlikely to have worse access to projects which are unlikely to succeed; thus, experts must have some information about projects likely to succeed in order to outperform at the mint stage. Moreover, whitelists and access alone would not explain our findings in Table 8, that changes in experienced investor ownership are also associated with price changes in secondary markets, where whitelists are not relevant.

5 Conclusion

In this paper, we analyzed the outperformance of experienced investors in the NFT market. After controlling for holding periods, experienced investors make 10% more per trade. Collections with higher experienced investor participation are more likely to mint out, mint out faster, and experience higher post-mint price growth. Our analysis points to the presence of substantial information inefficiencies in the NFT market, and suggests that the returns from inexperienced new entrants in the NFT market are unlikely to reflect average returns in the NFT market.

References

Bayer, Patrick, Christopher Geissler, Kyle Mangum, and James W Roberts. 2020. "Speculators and middlemen: The strategy and performance of investors in the housing market." *The Review of Financial Studies*, 33(11): 5212–5247.

Chinco, Alex, and Christopher Mayer. 2016. "Misinformed speculators and mispricing in the housing market." *The Review of Financial Studies*, 29(2): 486–522.

Cvijanović, Dragana, and Christophe Spaenjers. 2021. ""We'll always have Paris": Out-of-country buyers in the housing market." *Management Science*, 67(7): 4120–4138.

DeFusco, Anthony A, Charles G Nathanson, and Eric Zwick. 2021. "Speculative dynamics of prices and volume." National Bureau of Economic Research.

Dmitri, Boreiko, and Dimche Risteski. 2021. "Serial and Large Investors in Initial Coin Offerings." *Small Business Economics*, 57: 1053–1071.

Kaczynski, Steve, and Scott Duke Kominers. 2021. "How NFTs Create Value." Harvard Business Review Working paper.

Kaplan, Steven N, and Antoinette Schoar. 2005. "Private equity performance: Returns, persistence, and capital flows." *The journal of finance*, 60(4): 1791–1823.

Kim, Keongtae, and Siva Visawanathan. 2019. "The 'Experts' in the Crowd: The Role of Experienced Investors in a Crowdfunding Market." *MIS quarterly*, 347–372.

Korteweg, Arthur, Roman Kräussl, and Patrick Verwijmeren. 2016. "Does it pay to invest in art? A selection-corrected returns perspective." *The Review of Financial Studies*, 29(4): 1007–1038.

Kurlat, Pablo, and Johannes Stroebel. 2015. "Testing for Information Asymmetries in Real Estate Markets." *The Review of Financial Studies*, 28(8): 2429–2461.

Lovo, Stefano, and Christophe Spaenjers. 2018. "A model of trading in the art market." *American Economic Review*, 108(3): 744–74.

Moonstream. 2021. "An analysis of 7,020,950 NFT transactions on the Ethereum blockchain." Working paper.

Nadini, Matthieu, Laura Alessandretti, Flavio Di Giacinto, Mauro Martino, Luca Maria Aiello, and Andrea Baronchelli. 2021. "Mapping the NFT revolution: market trends, trade networks and visual features." *arXiv preprint arXiv:2106.00647*.

Nahata, Rajarishi. 2008. "Venture capital reputation and investment performance." *Journal of financial economics*, 90(2): 127–151.

Penasse, Julien, and Luc Renneboog. 2021. "Speculative trading and bubbles: Evidence from the art market." *Management Science*.

Pénasse, Julien, Luc Renneboog, and José A Scheinkman. 2021. "When a master dies: Speculation and asset float." *The Review of Financial Studies*, 34(8): 3840–3879.

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

Sørensen, Morten. 2007. "How smart is smart money? A two-sided matching model of venture capital." *The Journal of Finance*, 62(6): 2725–2762.

Appendix

A Supplementary Material for Section 2

This section describes additional details of our data and cleaning steps.

A.1 Data Sources and GC Sample Overview

Our primary data source is a transaction-level dataset is from Moonstream.[15] The Moonstream data include all on-chain transactions for the GCs in our sample between April 1, 2021, and September 25, 2021. The Moonstream transaction-level data includes the following variables: transaction hash, which is a unique identifier on the Ethereum blockchain; transaction date and time; collection-level contract address; item ID, which is a number that identifies an item within a collection; the wallet addresses of the seller and buyer; and transaction value. There are over 4.4 million transactions. Roughly 60% of these transactions are mints, and the remaining 40% are secondary market transactions.

We compute a transaction price variable based on the transaction value as follows. First, we convert the transaction value from Wei to ETH by dividing it by 10^{18}. Wei is simply the smallest denomination of ETH, the native digital asset on the Ethereum blockchain. Next, we divide the ETH values by the number of items reported with the same transaction hash. We do this because the value provided is for the whole group when there are multiple items in the same transaction. We would therefore be necessarily overstating the true (but unobserved) prices for each item if we do not adjust for the number of items. By dividing the value equally, we are assuming that each item in a transaction has the same implied price.

A.2 Defining Generative Collections

As we discuss in Subsection 2.2, we restrict attention to generative collections. The specific set of filters we use to select collections is as follows.

[15] See Moonstream (2021) and https://github.com/bugout-dev/moonstream/tree/main/datasets/nfts.

1. **Each item corresponds to a unique piece of digital art.** Technically, all NFTs are unique in the sense that they have unique identifiers on the blockchain, hence their "non-fungible" nature. However, some NFT collections will include multiple items that refer to the same digital art file, which would be like an artist creating multiple copies of the same painting.

2. **Items are variations on the same object/theme.** This condition ensures a degree of consistency across the items in a collection. It is admittedly a subjective feature that we determine during our data collection process.

3. **There exists a collection-level ERC-721 smart contract.** This collection-level smart contract not only formally ties together the items on the blockchain, but plays a crucial role in the initial crowdsale and governance of a GC as we describe later in this section. This condition also effectively restricts our sample to GCs on the Ethereum blockchain. Note that "ERC-721" refers to a "free, open standard that describes how to build non-fungible or unique tokens on the Ethereum blockchain."[16]

4. **Predetermined and fixed initial supply of items.** In these cases, this initial supply is common referred to as the "genesis supply." In addition to characterizing the contents of a collection, this condition provides a predetermined tangible goal that the creator is trying to attain in the initial crowdsale.

5. **Items in the genesis supply are sold on the primary market through a public sale.** This condition excludes collections in which the creator generates all the items on the blockchain and then sells them through the secondary market.

6. **Investors in the initial public sale receive a random item.** This condition further restricts the nature of the public sale, although it is quite common within the set of collections that meet the above conditions. It ensures that primary market investors are buying into the collection more broadly, not an individual item of interest.

We construct our sample of GCs, and implement these filters, through the following process. First, we scraped the rankings tables on the website OpenSea.io ("OpenSea"),

[16]See http://erc721.org/.

the most popular NFT marketplace. This step, which we performed on a few dates in October 2021, generated an initial list of 7,987 NFT collections. We consider this set to represent the universe of NFT collections created until that time given the popularity of OpenSea. Moreover, we are not concerned about survivorship bias because we observe so many NFT collections in this sample that effectively failed (i.e., no secondary market activity and prices close to zero).[17] Second, we visually assessed the items in each collection to determine whether they met our GC criteria in terms of being (i) unique and (ii) variations on the same object/theme.[18] We are left with 2,545 potential GCs after this step. Third, we check each collection for whether or not there exists a central ERC-721 smart contract, which further restricts our list to 1,376 potential GCs. This step drops both Ethereum-based NFT collections without a central contract and also those on non-Ethereum blockchains. The latter group must ultimately be dropped regardless of our GC definition because the transaction-level data described in the next section only includes NFT collections on Ethereum. In future versions of our analysis, we will consider gathering the transaction data for and including non-Ethereum GCs.

In the final step of creating our list of GCs, we apply a few data filters that are both consistent with our GC criteria and necessary for our empirical analysis. The main filter is that the NFT collection must have a predetermined genesis supply. We manually gather this piece of information from a collection's OpenSea page, website, Twitter account, and Discord channel, as available. This variable is important to define a key measure of initial GC success: the number of items minted divided by genesis supply. In addition, we keep only GCs for which we have their primary market transaction data, which are required for computing many of our GC-level variables. The most common reasons that these data are unavailable include that the GC was created before the beginning of the

[17]We cannot conclusively say that all failed NFT collections remain on the blockchain and continue to maintain an OpenSea collection page. However, we assume the extent to which any collections were removed from OpenSea was very small at most for two reasons. The first is the aforementioned high observed rate of failures in our sample. The second is that we are not aware of any driving mechanisms in practice to remove stale ERC-721 smart contracts from the Ethereum blockchain or unsuccessful NFT collections from the OpenSea website. As an example, we note that the OpenSea collection page for Evolved Apes Inc remains active on OpenSea as of this manuscript date despite it being a well-documented scam in which the creator disappeared in October 2021 with $2.7 million in funds raised from investors (see, e.g., https://finance.yahoo.com/news/another-nft-rug-pull-evolved-084902519.html).

[18]In many cases, the collection description includes the term "generative" but we do not consider this a sufficient condition to be a GC.

transaction-level dataset (April 1, 2021) or that it had not yet begun its primary sale by the end of the dataset (September 25, 2021). Finally, we only keep collections for which at least 5% of the items sold in their primary sale were done so at a nonzero price. This filter captures our notion that a GC must have a public sale.

In aggregate, GCs raised the equivalent of $0.51 billion through primary market sales over the sample period, which represents nearly half of the total for all Ethereum-based NFT collections (Figure A.4). We must qualify the denominator in this comparion because the Moonstream transaction-level data only includes Ethereum-based NFT collections. Primary market sales represent inflows of capital into the NFT asset class and thus we document that GCs are a particularly attractive form of NFT collection to investors.

A.3 Gathering Collection-Level Variables

We collect other information on GCs from project-specific websites. Summary statistics for these characteristics are listed in Table A.2. We gather data on whether a GC has a Twitter account, an independent website, and a Discord channel.[19] We gather data on whether each GC provides a "roadmap," which is a document that outlines planned future steps for the GC; and whether a GC highlights that certain items in their collection are rare, which is true for roughly one third of GCs. We collect data on whether the artist which created the art is explicitly named, and whether the artist has a linked Twitter and website.

We manually evaluate whether the art in the NFT pictures is 3D, animated, and has music. We evaluate subjectively whether the art is "cute". A number of NFT collections are "derivatives" which clearly build off three popular projects: Crypto Punks, Bored Ape Yacht Club, and Loot. We label collections if they are clearly derivatives of these three projects.

[19]Discord is a chat application, where community members can chat in different groups or "channels" with each other, and there are often private channels which are restricted to verified owners of NFTs within a collection.

A.4 Defining Collection-Level Variables

A.4.1 Fraction Minted by Experienced Investors

For an explicit measure of experienced investor involvement in each GC, we propose the share of experienced investor wallets that participate in the GC's primary market sale:

$$\text{Frac. Minted by Experienced} = \frac{\text{NFTs Minted by Experienced}}{\text{All NFTs Minted}}. \tag{7}$$

We need to be careful, however, about how we identify experienced wallets for this type of measure. A potential concern if we use the exact definition in Section 3.4 is that our measure would incorporate ex post information from the perspective of any given GC because both the thresholds and levels of trading activity would be based on the full sample. To mitigate this concern, we rely on ex ante indicators for identifying experienced wallets when computing the measure in (7).

To serve as the benchmark for comparison, we first compute the measure proposed in (7) using the GC activity thresholds based on the full sample as described in Section 3.4. Specifically, we define a wallet as corresponding to an experienced investor if it both minted from at least 6 GCs and traded within at least 11 GCs during our full sample. The corresponding indicators are used to compute the numerator in (7).

The ex ante version of (7) that we use in our analysis in Section 4.2 follows the same approach to identify experienced investors for the full sample as described above but with two key differences. The first is that the designation of an investor's experienced status is updated each date of the sample to make it an ex ante indicator. The second is that we impose the thresholds must be at least 2 distinct GCs for minting and trading. These minimum requirements make sure that we do not designate a large fraction of wallets as experienced early in the sample when the corresponding thresholds can be 0 or 1. We report the number and share of experienced investors over time based on the ex ante defintion in Figure A.1. Under the ex ante definition, the share of experienced investors peaks around 10% in July 2021 and then ultimately converges to the full sample share of 3.6%.

To get a sense for how the ex ante version of (7) differs from the full sample version, we plot the cross-sectional distributions of both variables in Figure A.2. Visually, the ex ante definition appears to shift the cross-sectional distribution to the left, but otherwise the shape of the distributions are similar.

B Supplementary Material for Section 3

In aggregate, GCs raised the equivalent of $0.51 billion through primary market sales over the sample period, which represents nearly half of the total for all Ethereum-based NFT collections (Figure A.4). We must qualify the denominator in this comparion because the Moonstream transaction-level data only includes Ethereum-based NFT collections. Primary market sales represent inflows of capital into the NFT asset class and thus we document that GCs are a particularly attractive form of NFT collection to investors.

In Figure A.3, we plot the transaction volume and price index by day for our SupDucks example, which launched in July 2021 so we have a few months of trading data available. We observe an initial spike in volume dirctly after the minting period followed by waves of transaction volume and price increases. Visually, it appears that higher volume may be leading higher prices, and we explore this idea formally in Section 4.

Many GCs experienced massive collection-level returns during our sample period as gauged by their price index. The summary statistics reported in Table 2 do not showcase this right tail. In Table A.3, we report the largest GCs in our sample according to their price index multiplied by number of items, which is a measure of market capitalization. The most valuable GC is the Bored Ape Yacht Club, which minted items for 0.08 ETH but had a price index of 19.50 ETH on September 25, 2021. These numbers implied a total return of 24,351% for any wallet that minted an item and an approximate collection-level value of over $500 million. On the opposite end of the spectrum, we oberve plenty of GCs whose price indexes quickly fall below their weighted average mint price within a week of the minting period. The figures in Table 2 suggest that the price index return for the median collection is about the same as the weighted average mint price after 7 days. Note, however, that this statistic can only be reported for the roughly two thirds of collections with sufficient trading volume to construct a price index.

One clear stylized fact that comes from an analysis of GC investors is that both ownership and trading are quite concentrated. In their initial analysis of their data, Moonstream (2021) find that the top 17% of NFT owners control 81% of the NFTs. Similar to this finding, we note that most GC investors only interact with 1–2 distinct GCs while the top 1% interact with dozens. We report the formal summary statistics across types of interactions in Table A.1.

Table A.1: Overview of Wallets Involved in GC Transactions

Notes. In this table, we describe the sample of wallets that ever interact with a GC in the transaction-level data. Specifically, we summarize the cross-sectional distributions across wallets in terms of the distinct number of GCs interacted with by transaction type.

	N	Mean	50%	90%	95%	99%	Max
Distinct GCs Minted	323,958	1.7	1	4	6	18	157
Distinct GCs Traded	323,958	2.8	1	6	11	27	185

To provide a further description of our two sets of GC investors, experienced and inexperienced, we provide additional summary statistics in Table A.4. In line with figures from Section 3.4, we confirm that experienced investors mint and trade much more in terms of volume. For example, the median experienced investor minted 2.48 ETH (approximately $7,500) worth of items, which is over 30 times the value for the median inexperienced investor. Similar, we find that most experienced investors have realized returns on at least several investments that they made (i.e., they have sold at least several items that they had previously purchased). The median inexperienced investor, on the other hand, has not even realized a return by the end of our sample.

We also observe that experienced investors appear to earn higher weighted average returns on their investments. For example, the median experienced investor earned 134.00% on their realized investments compared to 97.73% for the median inexperienced investor.[20] We formally investigate differences in realized returns for experienced investors in Section 4.1. We find similar outcomes in unrealized returns using our estimates

[20]Note that we are only able to compute realized returns for the subset of inexperienced investors that have realized a return on at least one investment. We further restrict the set of investors for which we compute investor-level realized returns to those that purchased at least 0.01 ETH worth of items to avoid capturing large but economically insignificant return values.

of unrealized gains, which are based on purchase prices and our collection-level price indexes at the end of our sample period. Further, the return on unrealized investments for the median inexperienced investor is negative in contrast to the small and positive value for the median experienced investor.

C Supplementary Material for Section 4

C.1 Unrealized Returns

Our analysis in the main text focuses on realized returns. However, many NFTs are held to the end of our sample; our results may simply reflect some selection bias, from the fact that experienced and inexperienced investors may have different propensities to hold NFTs that are performing well. To account for this possibility, we approximate unrealized returns using our collection-level price indexes (Section 3.2). For NFTs which are held until the end of our sample, we assign a value equal to the price index. Specifically, the end-of-sample price index is the most recent value between September 19 and September 25, 2021. If we are not able to compute a price index for any day during this period due to a lack of trades, we consider the price index to be zero for the purposes of the unrealized gain calculation. We can then calculate unrealized returns as our estimated value divided by the purchase price of the NFT, for every NFT held until the end of our sample period:

$$r^{unrealized}_{i,j,c,T,\tau} \equiv \frac{\text{PriceIndex}_{c,T} - \text{Price}^{Purch}_{i,j,c,\tau}}{\text{Price}^{Purch}_{i,j,c,\tau}} \tag{8}$$

where T is end of sample.

Figure A.5 shows aggregate unrealized returns, overall, and separately for experienced and inexperienced investors. Aggregate unrealized returns are slightly lower than aggregate realized returns, at approximately 30%, suggesting that investors tend to hold on to NFTs that are doing poorly.[21] Experienced investors' unrealized returns were

[21] An alternative explanation is that our price index methodology tends to produce downwards-biased estimates of the expected prices that NFTs would fetch, since we essentially assume all sales track the 25th percentile sale price within the collection.

roughly 60%, which is almost triple inexperienced investors' unrealized returns at 20%.

Next, in Table A.5, we re-estimate the specifications in Table 4, using unrealized returns as the dependent variable. Results are broadly similar to those in the main text: experienced investors attain higher returns than inexperienced investors in all specifications.

C.2 Collection-Level Panel Data

In Table A.6, we present summary statistics for the the panel data used in the regressions whose results are presented in Table 8. The sample only includes the subset of GCs that were successful in selling their entire genesis supply (i.e., minted out) and those that had at least 7 days with sufficient trading to define price indexes.

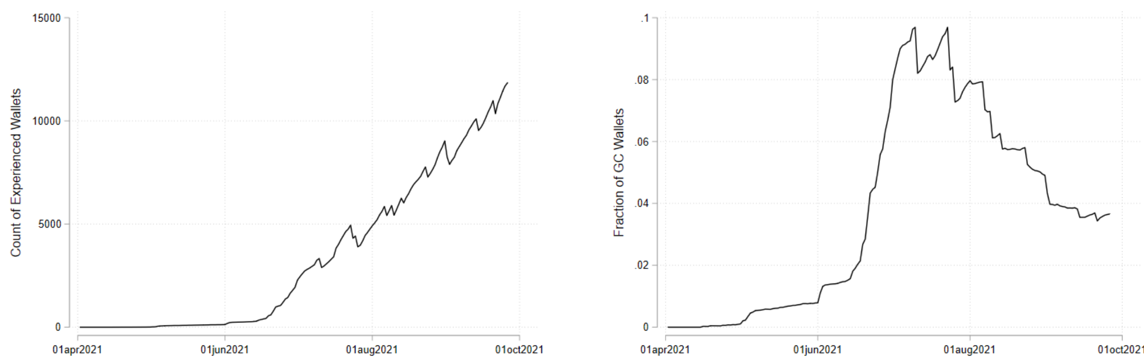
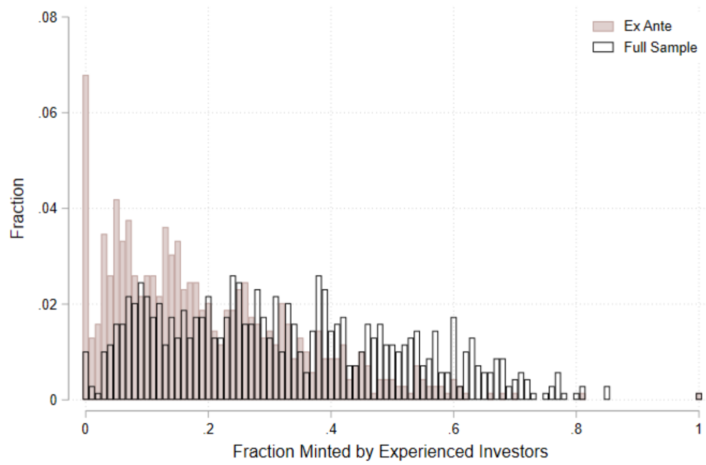


Figure A.1: Experienced Investors Over Time Under Ex Ante Definition

Notes. This figure reports the number and share of experienced investors over time based on our ex ante defintion, which follows the same approach from Section 4.2 with two key differences. The first is that the designation of an investor's experienced status is updated each date of the sample to make it an ex ante indicator. The second is that we impose the thresholds must be at least 2 distinct GCs for minting and trading.



	N	Mean	SD	Min	10%	50%	90%	Max
Frac. Minted by Experienced (Ex Ante)	692	0.19	0.15	0.00	0.03	0.16	0.41	1.00
Frac. Minted by Experienced (Full Sample)	692	0.33	0.20	0.00	0.08	0.30	0.61	1.00

Figure A.2: Fraction of GC Minted by Experienced Investors: Ex Ante versus Full Sample Definition

Notes. The figure reports and summarizes the cross-sectional distributions for fraction of the GC minted by experienced investors for both the ex ante and full sample definitions.

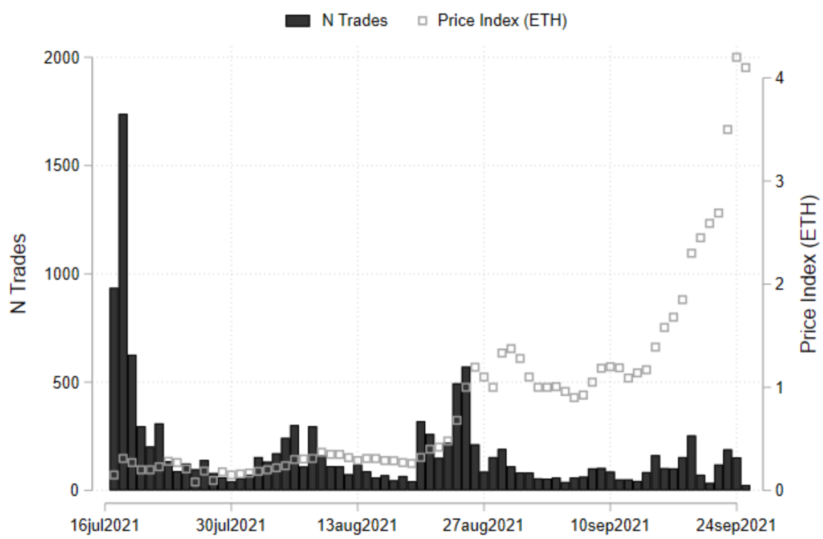


Figure A.3: SupDucks Example: Post-Mint Trading Period

Notes. This figure shows the daily trade volume and price index for the SupDucks GC since its initial primary market sale beginning on July 16, 2021. See Section 3.2 for a description of our price index construction.

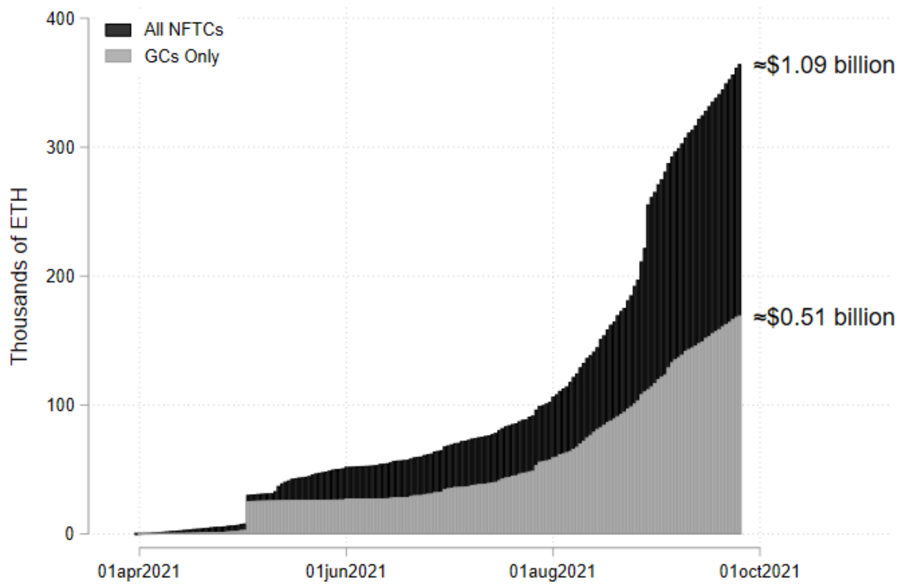


Figure A.4: Cumulative Funds Raised Through Primary Market Sales

Notes. This figure shows the accumulated amount of funds raised through the primary sales by our sample of GCs and the full sample of NFT collections in the Moonstream data. Dollar estimates are based on an exchange rate of $3,000 per ETH, which was the approximate ETH-USD exchange rate at the end of September 2021. In determining the full sample of NFT collections, we exclude a few collections that appear to be related to decentralized finance protocols. We do so because they are both large and do not represent NFT art collections. The specific collection-level contract addresses we exclude are the following: 0xC36442b4a4522E871399CD717aBDD847Ab11FE88 (Uniswap V3 Positions), 0x58A3c68e2D3aAf316239c003779F71aCb870Ee47 (Curve Synth-Swap), 0xb9ed94c6d594b2517c4296e24A8c517FF133fb6d (Hegic ETH ATM Calls Pool), and 0x3AFF7B16489Fcc59483DE44e96Bd9Ec533915924 (BiFi Position).

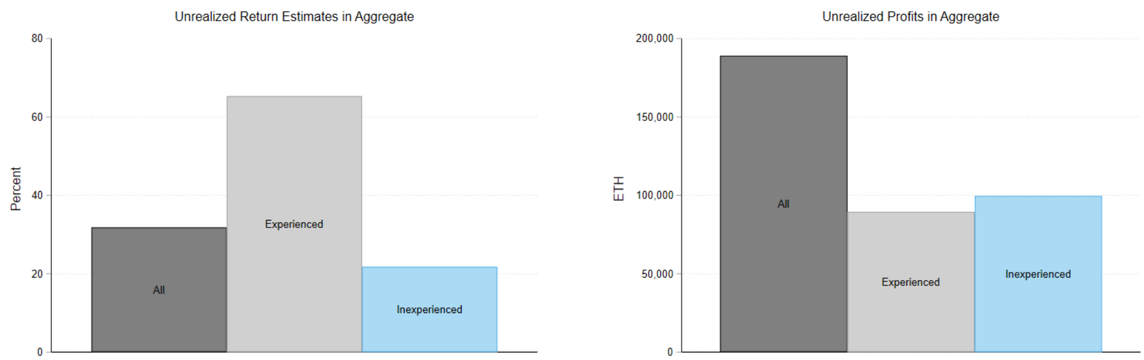


Figure A.5: Aggregate Unrealized Returns and Gains by Investor Type

Notes. This figure reports the weighted average returns for all unrealized positions within each investor group. The unrealized gain for any given position is computed as the difference between the price index at the end of our sample minus the price paid at purchase. Specifically, the end-of-sample price index is the most recent value between September 19 and September 25, 2021. If we are not able to compute a price index for any day during this period due to a lack of trades, we consider the price index to be zero for the purposes of the unrealized gain calculation. Unrealized gain returns are computed using the sum of end-of-sample price indexes minus the sum of prices paid in the original purchase transactions $\left(\left(\sum \mathsf{PriceIndex}^{\mathsf{EndSample}} - \sum \mathsf{Price}^{\mathsf{Purch}}\right) / \sum \mathsf{Price}^{\mathsf{Purch}}\right)$.

Table A.2: Overview of GC Characteristics

Notes. In this table, we summarize the dummy variables that we created for each GC in our sample based on manually gathered data. The sources for our manual data gathering efforts are GC-specific webpages including but not limited to their OpenSea webpage. "Has Website" refers only to independent websites (e.g., an OpenSea webpage does not count). A "roadmap" is a document provided by a GC creator that outlines their planned future steps for the GC. "Has Charity Description" is true as long as the GC claims that at least part of its proceeds will go to a specified charity. The determination of the art characteristics are subjective based on our review.

	N	Count	Mean
Has Twitter	692	670	0.97
Has Website	692	663	0.96
Has Discord	692	593	0.86
Has Roadmap	692	404	0.58
Advertises Rare Items	692	236	0.34
Has Charity Description	692	117	0.17
Has Named Artist	692	296	0.43
Named Artist Has Twitter	692	136	0.20
Named Artist Has Website	692	35	0.05
Art is 3-D	692	221	0.32
Art is Animated	692	65	0.09
Art Has Music	692	14	0.02
Art Is Cute	692	40	0.06
Art Is Punk Derivative	692	22	0.03
Art Is BAYC Derivative	692	16	0.02
Art Is Loot Derivative	692	11	0.02

Table A.3: Top 10 GCs by Implied Collection Value as of Sep. 25, 2021

Notes. In this table, we report the top 10 GCs according to their price-index-implied valuation as of September 25, 2021. Mint price is the weighted average value (i.e., total amount of ETH raised in mint transactions divided by the total number of items minted). See Section 3.2 for a description of our price index construction. Return is the price index divided by the mint price minus 1. Implied value in USD computed as the price index times the number of items times $3,000 per ETH, which was the approximate ETH-USD exchange rate at the end of September 2021.

Rank	Name	First Mint Date	Mint Price ETH	Price Index ETH	Return %	Genesis Supply	Implied Value USD Mlns
1	Bored Ape Yacht Club	04/22/21	0.08	19.50	24,351	10,000	585.1
2	Meebits	05/03/21	1.10	4.82	337	20,000	289.4
3	Cool Cats NFT	06/27/21	0.02	7.50	35,064	9,933	223.5
4	SupDucks	07/16/21	0.07	4.10	5,627	10,001	123.0
5	World of Women	07/27/21	0.07	2.50	3,543	10,000	75.0
6	Sneaky Vampire Syndicate	09/09/21	0.08	2.19	2,666	8,888	58.4
7	Pudgy Penguins	07/22/21	0.03	2.16	7,096	8,888	57.6
8	0N1 Force	08/15/21	0.08	2.00	2,494	7,777	46.7
9	The Doge Pound	07/12/21	0.07	1.50	2,168	10,000	45.0
10	Rumble Kong League	07/27/21	0.08	1.50	1,781	10,000	45.0

Table A.4: Summary Statistics by GC Investor Type

Notes. In this table, we summarize investor-level variables separately for experienced and inexperienced GC investors. Inexperienced investors that never engaged in a transaction with a non-zero price are excluded. See Section 4.1 for a description of how we measure realized returns. We only compute investor-level realized returns for those that purchased at least 0.01 ETH worth of items to avoid capturing large but economically insignificant return values. See Section C.1 for a description of how we measure unrealized gains and returns. As with realized returns, we only compute unrealized returns for those that purchased at least 0.01 ETH worth of items to avoid capturing large but economically insignificant return values.

Panel A. Experienced

	N	Mean	SD	Min	10%	50%	90%	Max
ETH Minted	11,847	6.04	16.96	0	0.69	2.48	13.17	1,130
ETH Traded, Sold	11,847	27.31	67.09	0	1.79	10.84	65.10	4,085
ETH Traded, Purchased	11,847	16.48	40.43	0	1.09	6.53	36.39	1,320
N Positions Realized	11,847	70.69	133.26	0	7.00	33.00	159.00	4,864
Realized Gross Profit (ETH)	11,847	17.03	54.88	-93	0.64	5.79	40.72	3,626
Realized Gross Return (%)	11,653	244.23	592.25	-88	35.57	134.00	466.01	26,775
N Positions Still Unrealized	11,847	71.75	190.05	0	8.00	36.00	148.00	7,492
ETH Spent Positions Still Unrealized	11,847	11.57	33.20	0	0.67	4.51	24.67	1,327
Unrealized Gain Return (%)	11,764	78.52	310.80	-100	-53.67	3.20	255.67	13,150

Panel B. Inexperienced

	N	Mean	SD	Min	10%	50%	90%	Max
ETH Minted	280,255	0.35	3.18	0	0.00	0.07	0.63	639
ETH Traded, Sold	280,255	1.38	13.40	0	0.00	0.00	1.80	2,473
ETH Traded, Purchased	280,255	1.89	14.16	0	0.00	0.10	3.03	2,448
N Positions Realized	280,255	2.01	9.34	0	0.00	0.00	5.00	1,136
Realized Gross Profit (ETH)	280,255	0.79	9.96	-481	0.00	0.00	0.97	2,808
Realized Gross Return (%)	84,154	344.14	1623.03	-100	-15.20	97.73	733.33	245,855
N Positions Still Unrealized	280,255	6.43	21.02	0	1.00	2.00	14.00	3,517
ETH Spent Positions Still Unrealized	280,255	1.63	12.17	0	0.00	0.20	2.70	1,819
Unrealized Gain Return (%)	245,280	65.35	788.75	-100	-84.63	-22.23	156.41	138,615

Table A.5: Regressions at Trade Level: Unrealized Returns

Notes. In this table, we report the results from estimates of specification (3), using unrealized returns, from (8) of Appendix C, as the dependent variable. We regress unrealized returns for each NFT held until the end of the sample period on an experienced seller dummy, the log of the holding period, and various fixed effects. We only include unrealized return values where the purchase price was 0.01 ETH or more. The right-hand side variables include a dummy for whether the associated investor is in our experienced group and the log of the fractional number of days the position was held. We also include sets of fixed effects that vary across specifications. Standard errors are heteroskedasticity-consistent. t-statistics are in parentheses. $^{*}p < 0.10$; $^{**}p < 0.05$; $^{***}p < 0.01$.

	(1)	(2)	(3)	(4)	(5)	(6)
Experienced Buyer Dummy	0.714***	0.413***	0.060***	0.457***	0.161***	0.030***
	(30.53)	(18.20)	(3.28)	(22.81)	(9.83)	(4.23)
ln(Days Held)		1.322***	-1.535***	3.035***	-0.694***	1.089***
		(94.76)	(-47.38)	(123.50)	(-9.65)	(61.04)
Date FE	No	No	Yes	No	Yes	No
GC FE	No	No	No	Yes	Yes	No
Date-GC FE	No	No	No	No	No	Yes
R^2	0.001	0.008	0.383	0.341	0.566	0.929
N	2,370,023	2,370,023	2,370,023	2,370,023	2,370,023	2,367,986

Table A.6: Summary Statistics for Data Used in Table 8 Regressions

Notes. In this table, we summarize the panel data used in the regressions whose results are presented in Table 8. The sample only includes the subset of GCs that were successful in selling their entire genesis supply (i.e., minted out) and those that had at least 7 days with sufficient trading to define price indexes. See Section 3.2 for a description of our price index construction. See Section 4 and Appendix A.4.1 for descriptions of our experienced investor definition and construction of the ex ante version. Both the number of trades and change in experienced ownership are measured during the 24-hour period each day ending at midnight UST.

	N	Mean	SD	Min	10%	50%	90%	Max
Price Index	7,511	0.45	2.41	0.00	0.01	0.07	0.63	48.00
Change in ln(Price Index)	7,511	0.00	0.59	-10.99	-0.26	0.00	0.27	10.95
N Trades	7,511	118.67	269.73	5.00	11.00	46.00	251.00	6,598.00
Frac. Owned by Experienced (Ex Ante)	7,511	0.14	0.13	0.00	0.00	0.10	0.31	0.91
Change in Frac. Owned by Experienced (Ex Ante)	7,511	-0.01	0.14	-0.87	-0.15	-0.00	0.14	0.78