Constructing a NFT Price Index and Applications

Hugo Schnoering
Napoleon Group
hugo.schnoering@napoleon-group.com

Hugo Inzirillo
Napoleon Group
CREST - Institut Polytechnique de Paris
hugo.inzirillo@napoleon-group.com

***Abstract*—We are witnessing the emergence of a new digital art market, the art market 3.0. Blockchain technology has taken on a new sector which is still not well known, Non-Fungible tokens (NFT). In this paper we propose a new methodology to build a NFT Price Index that represents this new market on the whole. In addition, this index will allow us to have a look on the dynamics and performances of NFT markets, and to diagnose them.**

I. Introduction

Last year has been the year of the democratization of NFTs. The number of owners (figure 1) and the number of transactions (figure 2) on Ethereum have been exponentially increasing since January 2021. Moreover, the increasing interest of famous brand names for them also suggests that their adoption keeps growing. Online marketplaces like Opensea (https://opensea.io) or LooksRare (https://looksrare.org) have become focal points for trading these new assets.

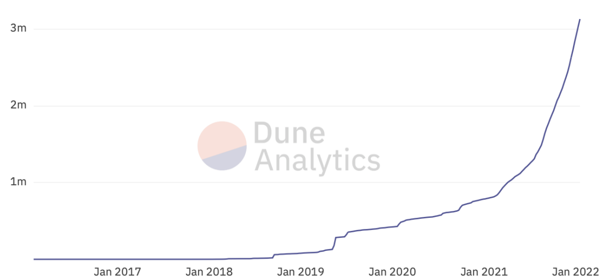

Fig. 1: Cumulative number of wallet that have ever owned an ETH ERC-721 or ERC-1155 NFT [1].

NFTs or *Non-Fungible Tokens* are transferable assets secured by a *blockchain*. A blockchain is an ordered list of blocks that are linked together using cryptography. Blocks contain data about transactions, and are validated and added to the chain through a consensus protocol. Blockchain are used as public transaction ledgers of most cryptocurrencies. For instance bicoin or ethereum respectively with the network Bitcoin [2] or Ethereum [3]. The need for consensus of blockchain solves the double-spending problem using a cryptographic proof instead of a trusted central authority.

Designed as a medium of exchange, cryptocurrencies are transferable assets as well. Because they are defined by their value, cryptocurrencies are perfectly fungible assets, for example a coin can be substituted to another. In opposition to cryptocurrencies, NFTs are uniquely identified by an id and a set of properties, and cannot be interchangeable or divisible. A NFT is thus the perfect way to represent anything unique on blockchains. They can be used for several purposes, but are often considered as digital art due to their properties. As art, most collectors acquire NFTs for aesthetic reasons, acquiring a social status or as a mean of investment [4].

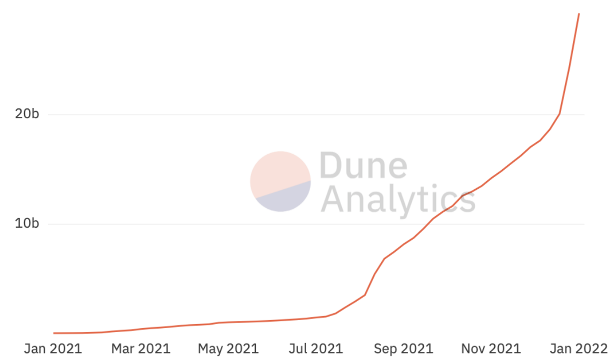

Fig. 2: Cumulative number of NFT traded volumes in USD [1].

NFTs are created or *minted* by the execution of a *smart contract* setting in stone the creation of tokens. Several NFTs can be minted from the same contract, a collection is defined as a set of tokens minted from the same contract. Tokens minted from the same contract collect remain however perfectly identifiable and may differ from each others by their id and their properties or *traits*. To illustrate this, we represent figures 3 and 4 respectively the tokens CryptoPunk #1463 and CryptoPunk #1466 of the Ethereum collection CryptoPunk.

Besides exceptional trade volumes, sales of NFTs have reached record prices, sometimes up to several tens of millions of dollars [5, 6]. Record breaking sales thus suggest a price explosion, but the dynamics of NFT markets are more complex. Indeed, if some top collections are steadily becoming more and more popular and expensive, e.g. the Ethereum CyberPunks and Bored Ape Yacht Club, prices of most collections vary widely. For this reason, it is difficult to have a global vision of the NFT market. In this work, we undertake the hedonic model to build a price index from

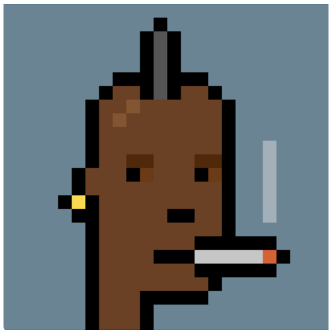

Fig. 3: CryptoPunk #1463. Properties and corresponding frequencies in the collection: male (60%), cigarette (10%), earring (25%), Mohawk thin (4%).

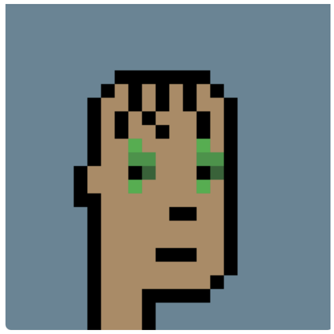

Fig. 4: CryptoPunk #1466. Properties and corresponding frequencies in the collection: male (60%), stringy hair (5%), clown eyes green (4%).

thousands of NFT transaction records. Subsequent applications presented in this study include:

- the detection of price explosion periods through statistical tests
- the computation of the correlation matrix between the NFT returns with the return of the cryptocurrencies market
- the detection of undervalued and overvalued assets

The hedonic regression framework is a well known model in Economics. Hedonic models decompose an item into its core characteristics and study the contribution of each of them to its value. These models are commonly used in Real Estate Economics [7, 8]. A hedonic index is an index computed from a fitted hedonic regression model, Triplett reviews in [9] the different methodologies that can be used. Hedonic indices are often used as proxies of price indices, such as consumer price indices (CPI). A price index represents the aggregate price of a basket of items, and tracks how the prices of these items, taken as a whole, change over time. The hedonic methodology is particularly appraised because it can be applied to illiquid markets. Moreover, it can even handle the removal, replacement and addition of items over time. For all these reasons, several previous studies have applied it to diagnose art markets [4, 10, 11]. Some works also used the hedonic model to price NFTs, but focused on a single collection: CryptoKitties [12], CryptoPunks [13] or Decentraland [14]. To the best of our knowledge, our article is the first proposal of a global NFT price index.

II. Methods

We will use the following notations:

- $\mathcal{C}$ is the set of collections
- $\mathcal{A}$ is the set of assets
- $\{0, 1, ..., T\}$ is the set of sale dates
- $I(a,b) \triangleq \{a, a+1, ..., b-1, b\}$
- $I(b) \triangleq I(0,b)$

A. Definitions

An *asset* a is defined by a collection C^a, a token id i^a and a dictionary of traits $\mathcal{P}^a$. $\mathcal{P}^a$ is a set of (trait, value) characterizing a within its collection. $\mathcal{P}^a$ is usually used to quantify the scarcity of a w.r.t. other assets of C^a, and, then, to price a. Intuitively the scarcer an asset is, the more expensive it should be. It is therefore straightforward to deal with frequencies of traits within a collection when it comes to build a pricing model. Given an asset a, a couple $(p,v) \in \mathcal{P}^a$, the frequency of the value v for the trait p is defined as follows:

$$f^a_{(p,v)} \triangleq \frac{1}{|C^a|} \sum_{a' \in C^a} \sum_{(p',v') \in \mathcal{P}^{a'}} 1_{p'=p} 1_{v'=v} \tag{1}$$

Since the number of traits may differ from an asset to another, even within the same collection, we construct three aggregate quantities: the minimum frequency $f^a_{\min}$ (equation 2), the average frequency f^a_{avg} (equation 3), and the maximum frequency $f^a_{\max}$ (equation 4). All these quantities are set equal to 1 if $\mathcal{P}^a$ is empty.

$$f^a_{\min} \triangleq \min_{(p,v) \in \mathcal{T}^a} f^a_{(p,v)} \tag{2}$$

$$f^a_{\text{avg}} \triangleq \frac{1}{|\mathcal{P}^a|} \sum_{(p,v) \in \mathcal{P}^a} f^a_{(p,v)} \tag{3}$$

$$f^a_{\max} \triangleq \max_{(p,v) \in \mathcal{P}^a} f^a_{(p,v)} \tag{4}$$

A *sale* is defined by an asset a, a date t and a price P^a_t in USD equivalent.

B. Pricing model

The pricing model that we develop in this section is the result of three empiric observations:

- as mentioned in the introduction, the price of an asset is impacted by the popularity and hype around its collection,
- assets within the same collection may differ from each others by their traits, the scarcer the traits of an asset are, the more appreciated it is in its collection,
- the NFT market seems to experience periods during which prices of NFT are globally impacted positively or negatively.

We finally come up with the multiplicative pricing model of equation (5).

$$P^a_t = P \times f(C^a) \times g(a) \times h(t) \times \epsilon(a,t) \tag{5}$$

where:

- $P \in \mathbb{R}^{+*}$ defines a scale price,
- $f : \mathcal{C} \rightarrow \mathbb{R}^{+*}$ impacts the price of an asset according to its collection,
- $g : \mathcal{A} \rightarrow \mathbb{R}^{+*}$ impacts the price of an asset according to the scarcity of its traits within its collection,
- $h : \{0, 1, ..., T\} \rightarrow \mathbb{R}^{+*}$ impacts the prices according to the global state of the NFT market,

- $\epsilon : \{0, 1, ..., T\} \times \mathcal{A} \to \mathbb{R}^{+*}$ is a noise term explaining price fluctuations.

We model f, g, h and ϵ as follows:

$$f : C \to \exp \sum_{C'} \alpha_{C'} 1_{C'=C} \tag{6}$$

$$g : a \to \exp\left(\beta_{\min} f^a_{\min} + \beta_{\text{avg}} f^a_{\text{avg}} + \beta_{\max} f^a_{\max}\right) \tag{7}$$

$$h : t \to \exp \sum_{t'=0}^{T} \gamma_{t'} 1_{t'=t} \tag{8}$$

$$\epsilon : (a, t) \to \exp \chi(a, t) \tag{9}$$

where $\chi(a, t) \overset{i.i.d.}{\sim} \text{Normal}(0, \sigma^2)$.

Equivalently,

$$\begin{aligned} \log P^a_t &= \log P + \sum_{C'} \alpha_{C'} 1_{C'=C} \\ &+ \beta_{\min} f^a_{\min} + \beta_{\text{avg}} f^a_{\text{avg}} + \beta_{\max} f^a_{\max} \\ &+ \sum_{t'=0}^{T} \gamma_{t'} 1_{t'=t} + \chi(t, a) \end{aligned} \tag{10}$$

The model of equation (10) is also known as the *time dummy variable* version of the hedonic regression model [9]. Hedonic coefficients P, $\boldsymbol{\alpha}$ and $\boldsymbol{\beta}$ are inferred only once using all dates $\{0, 1, ..., T\}$, this approach is called the *multi-period pooled regression* [9].

C. Price Index Construction

Since we used the time dummy variable method to train our pricing model, the hedonic index I is defined as follows [9]:

$$\frac{I_{t+1}}{I_t} = \frac{\exp \gamma_{t+1}}{\exp \gamma_t} \tag{11}$$

As a result,

$$\begin{aligned} I_t &= A \prod_{t'=0}^{t-1} \left(1 + \left(\frac{I_{t'+1}}{I_{t'}} - 1\right)\right) \\ &= A \prod_{t'=0}^{t-1} \left(1 + \left(\frac{\exp \gamma_{t'+1}}{\exp \gamma_{t'}} - 1\right)\right) \\ &= A \prod_{t'=0}^{t-1} \frac{\exp \gamma_{t'+1}}{\exp \gamma_{t'}} \\ &= A \exp(\gamma_t - \gamma_0) \end{aligned} \tag{12}$$

Thus, the return of I between t and $t+1$ is $\frac{\exp \gamma_{t'+1}}{\exp \gamma_{t'}} - 1$.

The construction of I can be justified as follows, if I is proportional to the geometric mean of prices over all assets of $\mathcal{A}$, then:

$$\begin{aligned} \frac{I_{t+1}}{I_t} &= \frac{\prod_{a \in \mathcal{A}} (P^a_{t+1})^{\frac{1}{|\mathcal{A}|}}}{\prod_{a \in \mathcal{A}} (P^a_t)^{\frac{1}{|\mathcal{A}|}}} \\ &= \prod_{a \in \mathcal{A}} \left(\frac{P^a_{t+1}}{P^a_t}\right)^{\frac{1}{|\mathcal{A}|}} \\ &= \frac{\exp \gamma_{t+1}}{\exp \gamma_t} \prod_{a \in \mathcal{A}} \exp\left(\chi(a, t+1) - \chi(a, t)\right)^{\frac{1}{|\mathcal{A}|}} \\ &= \frac{\exp \gamma_{t+1}}{\exp \gamma_t} \exp\left(\frac{1}{|\mathcal{A}|} \sum_{a \in \mathcal{A}} \chi(a, t+1) - \chi(a, t)\right) \end{aligned} \tag{13}$$

Since $\chi(a, t+1) - \chi(a, t) \overset{i.i.d.}{\sim} \text{Normal}(0, 2\sigma^2)$, according to the strong law of large numbers:

$$\frac{1}{|\mathcal{A}|} \sum_{a \in \mathcal{A}} \chi(a, t+1) - \chi(a, t) \underset{|\mathcal{A}| \to +\infty}{\to} 0, \text{ a.s.} \tag{14}$$

Thus,

$$\exp\left(\frac{1}{|\mathcal{A}|} \sum_{a \in \mathcal{A}} \chi(a, t+1) - \chi(a, t)\right) \underset{|\mathcal{A}| \to +\infty}{\to} 1, \text{ a.s.} \tag{15}$$

because exp is continuous. Finally, if $\mathcal{A}$ is large enough:

$$\frac{I_{t+1}}{I_t} \approx \frac{\exp \gamma_{t+1}}{\exp \gamma_t} \tag{16}$$

D. Bubble Detection

A bubble is often defined as a period of explosive or mildly explosive behavior in the price dynamic [15]. In particular, autoregressive dynamics can be observed during speculative bubbles. For this reason, unit root tests are useful for making a market diagnosis, e.g. for detecting market excesses or mispricing. Augmented Dickey-Fuller (ADF) tests are usually used to determine whether a series y is stationary or not, but they can also be used to detect explosive behaviors. Under the null hypothesis H_0, the process is autoregressive and has an unit root. Under the alternative hypothesis H_1, the process has an explosive root. Hypothesis H_0 is wide and must be specified in order to derive an asymptotic distribution and the critical values useful for the ADF test. To this end, Phillips et al. assume that the process y is a random walk with an unit root and an asymptotically negligible drift [16], i.e.:

$$y(t+1) = d(T+1)^{-\eta} + \theta y(t) + \epsilon(t) \tag{17}$$

where $\theta = 1$, $\eta > 1/2$ and $\epsilon(t) \overset{i.i.d.}{\sim} \text{Normal}(0, \sigma^2)$.

The ADF testing procedure is applied to the regression model of equation (18).

$$\Delta y(t) = \mu + \nu y(t-1) + \sum_{i=1}^{k} \psi^i \Delta y(t-i) + \epsilon(t) \tag{18}$$

where $\Delta y(t) = y(t) - y(t-1)$, $k \in \mathbb{N}$ is a lag order and $\epsilon \overset{i.i.d.}{\sim} \text{Normal}(0, \sigma^2)$. The ADF statistic is defined as the t-statistic of the coefficient ν of the regression model.

Under the null hypothesis (equation 17), the process y has an unit root, thus $\nu = 0$. Under the alternative hypothesis of an explosive root, $\nu > 0$. For this reason, we use the right-sided version of the ADF test, i.e. if ADF ¿ critical : H_0 is rejected and H_1 accepted, else : H_0 is accepted and H_1 rejected

In order to detect local bubbles and not to be fooled by pseudo-stationary behaviors, the ADF test can be performed on continuous sub-sample of y. We denote by $ADF_{t_1 \to t_2}$ the ADF statistic computed on $\{y_{t_1}, y_{t_1+1}, ..., y_{t_2}\}$. Phillips et al. develop in [16] a methodology to detect multiple bubbles. For this purpose, they derive the BSADF statistic from the ADF statistic (equation 19).

$$BSDAF(t) = \sup_{t' \in I(0, t-w)} ADF_{t' \to t} \tag{19}$$

where w is a hyper-parameter defining the minimum window size on which ADF tests can be performed. The beginning and the end of a bubble, denoted respectively by t_b and t_e, are detected as follows:

$$t_b \triangleq \inf\{t \in I(w, T), BSDAF(t) > v_w^c(t)\} \tag{20}$$

$$t_e \triangleq \inf\{t \in I(t_b + \delta, T), BSDAF(t) < v_w^c(t)\} \tag{21}$$

where δ is the minimum duration of a bubble and $v_w^c(t)$ is a critical value defined as the $(1 - c)$ quantile of the distribution of the statistic $\sup_{t' \in \{w,...,t\}} ADF_{0 \to t'}$. Critical values $v_w^c(t)$ can be estimated using Monte-Carlo simulations.

III. Data

Several blockchains support NFTs including Ethereum, Solana, Flow, Tezos or Polygon. However, most of the NFT trade volume is concentrated on Ethereum (figure 5). For this reason, we decide to focus our work on Ethereum NFT collections, more precisely on NFTs satisfying the standard ERC-721.

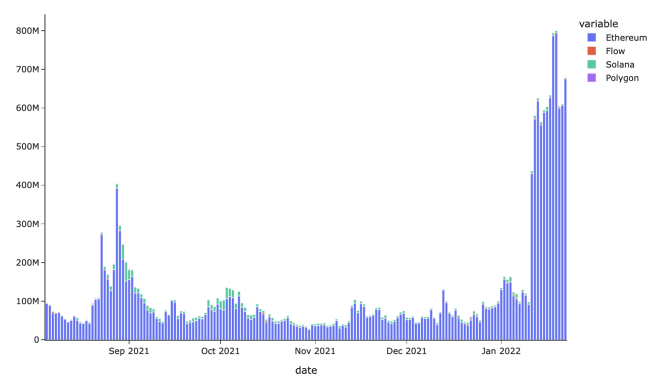

Fig. 5: Daily sale volumes in dollar on four blockchains (data from [17]).

We have selected 59 Ethereum collections among the most traded collections on OpenSea during the second semester

Coefficient	Value
$\beta_{\min}$	-0.176
β_{avg}	-0.009
$\beta_{\max}$	-0.025

TABLE I: Asset-related coefficients of the model.

of 2021. The list of collection names and corresponding Ethereum smart contracts can be found in the appendix (section VII-B).

We use the official Opensea API (https://docs.opensea.io) to download the properties of all tokens in the selected collections, as well as all sale transactions concerning these collections from the 1st June 2021 to the 15th January 2022. Our dataset totals up 597198 transactions.

IV. Materials and Applications

In order to train the model of equation (18) we use the Huber Regressor, a linear regression model that is robust to outliers. We train two models: a model with all collected collections in order to build the NFT price index, and another model with only metaverse-related collections to build the metaverse price index. Collections used for both indices can be found in the appendix (section VII-B). The first value of our indices, i.e. the hyper-parameter A, is set equal to 100. We report on table I the inferred parameters $\boldsymbol{\beta}$ of g, the asset rarity impact function, for the NFT price index.

We plot in figures 6 and 7 the 7-day moving averages of the NFT and Metaverse Index respectively. Raw indices can be found in the appendix (section VII-A). We also plot in figures 6 and 7 the Google Trends (GT) signals for the queries ”nft” and ”metaverse” between the 1st June 2021 and the 16th January 2022 [18, 19]. Since the GT signals for q is proportional to the Google search volume for q, both GT signals will be used to chart the fad for NFTs and the metaverse.

A. Correlation with cryptocurrency markets

In table II, we report the correlation coefficients between the daily returns of our indices, ETH, BTC and SOL.

	NFT	Metaverse	ETH	SOL	BTC
NFT		0.09	0.16	0.18	0.10
Metaverse			0.30	0.17	0.20
ETH				0.59	0.82
SOL					0.48

TABLE II: Correlation matrix of daily returns. Correlation coefficients are computed using the data from the 1st June 2021 to the 15th January 2022.

We report in table III the realized return of our indices, ETH, BTC, and SOL between the 1st June 2021 and the 15st January 2022.

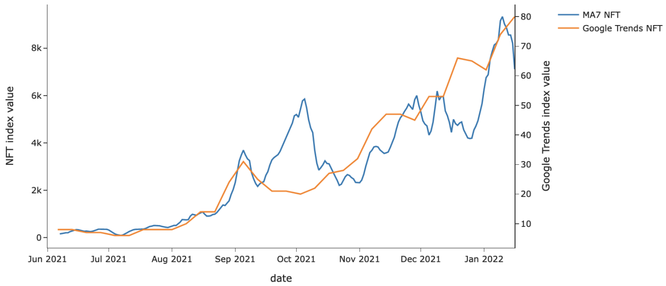

Fig. 6: 7-day-MA NFT Index.

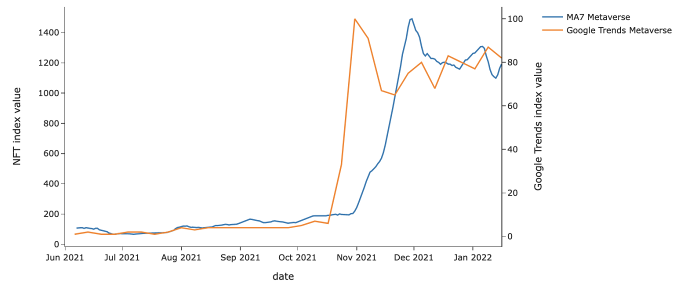

Fig. 7: 7-day-MA Metaverse Index.

	Return
NFT	7011%
Metaverse	1064%
ETH	26%
BTC	6%
SOl	372%

TABLE III: Realized return computed from prices in USDT between the 1st June 2021 and the 15st January 2022.

c	$v^c_{w,t}$
90%	1.08
95%	1.36
99%	1.97

TABLE IV: Critical values $v_w^c(t)$ of SADF with $t = T = 230$ and $w = 40$.

B. Bubble Detection

The minimum length w to compute an ADF test is fixed equal to 40. We use the methodology of Phillips et al. presented in section II-D to detect the start and the end of bubbles. c-critical values $v_w^c(t)$ of $\sup_{t' \in I(w,t)} ADF_{0\to t'}$ are estimated using Monte-Carlo simulations of $N_{MC} = 5000$ random walks with an asymptotically negligible drift (equation 17) where we set $\eta = \sigma = d = 1$. $v_w^c(t)$ are estimated for $t \in I(w,t)$. We report in particular the 90%, 95% and 99% critical values when $t = T = 230$ in table IV. We plot on figures 8 and 9 the $t \to BSADF(t)$ signals (equation (19)) computed for both indices. We have also plotted the 95% and 99% critical value signal for the statistics $t \to \sup_{t' \in I(w,t)} ADF_{0\to t'}$. We recall that an explosive behavior is detected in the price dynamic if the BSADF statistic exceeds the 99% critical value signal.

C. Undervalued and Overvalued Assets

Once the parameters P, $\boldsymbol{\alpha}$, $\boldsymbol{\beta}$, $\boldsymbol{\gamma}$ and σ of equation (10) are inferred after a training phase, any asset a can be priced at time $t \in \{0, ..., T\}$ if at least one asset of the same collection has been used during the training phase. According to the model of equation (10), the log-price of a at time t follows the distribution $\mathrm{Normal}(\log \bar{P}_t^a, \sigma^2)$ where $\log \bar{P}_t^a$ is computed as follows:

$$\log \bar{P}_t^a = \log P + \alpha_{C^a} + \beta_{\min} f_{\min}^a + \beta_{\text{avg}} f_{\text{avg}}^a + \beta_{\max} f_{\max}^a + \gamma_t \tag{22}$$

We define the probabilities of being undersold and oversold in equations (23) and (24), respectively.

$$p_{\text{under}}(P_t^a) \triangleq \frac{1}{2}\left(1 - \operatorname{erf}\left(\frac{\log(P_t^a) - \log(\bar{P}_t^a)}{\sigma\sqrt{2}}\right)\right) \tag{23}$$

$$p_{\text{over}}(P_t^a) \triangleq \frac{1}{2}\left(1 + \operatorname{erf}\left(\frac{\log(P_t^a) - \log(\bar{P}_t^a)}{\sigma\sqrt{2}}\right)\right) \tag{24}$$

where erf is the Gauss error function. According to the definitions (23) and (24), it is straightforward to derive the following properties:

- $p_{\text{over}} + p_{\text{under}} = 1$,
- at $P_t^a = \bar{P}_t^a$, $p_{\text{over}}(P_t^a) = p_{\text{under}}(P_t^a) = \frac{1}{2}$,
- p_{under} is strictly increasing, $\lim_{0^+} p_{\text{under}} = 1$ and $\lim_{\infty} p_{\text{under}} = 0$.

The simplest investment strategy could consist in investing at time t in an asset a for which the listing price P_t^a gives a high undersold probability $p_{\text{under}}(P_t^a)$ below $1/2$, and listing it at $\bar{P}_t^a$. One problem, however, is that equation (22) assumes to have estimated γ_t, i.e. the market impact coefficient of today. To deal with it, γ_t can be estimated using earlier in the day transactions, or it can be approximated by a moving average of the market impact coefficients over recent previous days.

V. Discussion

From figures 6 and 7, we detect a strong upward trend for both price indices indicating a global price rise. This increase in prices appears to follow the rising interest for both themes. The NFT price index seems to have experienced several brief explosive periods, while the metaverse price index experienced an unique explosive period beginning in November 2021. This comes a short time after the renaming of Facebook's parent company to Meta reveiling their interest for the metaverse.

As shown in table I, the parameters of g are all non positive. As a consequence, g, and then the price, is strictly

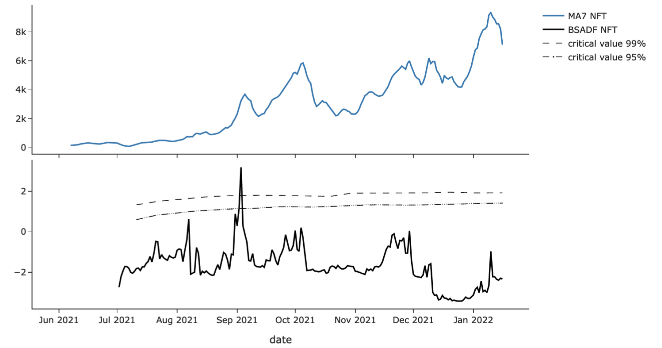

Fig. 8: 7-day-MA NFT Index (above). BSADF statistic signal for the NFT Index (below). 95% and 99% critical values (below).

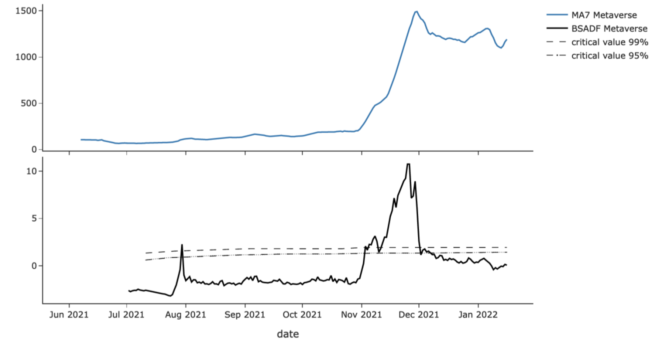

Fig. 9: 7-day-MA Metaverse Index (above). BSADF statistic signal for the Metaverse Index (below). 95% and 99% critical values (below).

decreasing with the quantities $f_{\min}$, f_{avg} and $f_{\max}$. It statistically confirms our intuition that scarcer assets are more expensive.

According to table II, returns of both indices are positively correlated with the returns of the top-tier cryptocurrencies BTC, ETH and SOL, but correlation coefficients remain relatively small compared to the ones between BTC, ETH and SOL. From table III, we observe that the selected collections have, on the whole, dramatically outperformed these cryptocurrencies. Putting these results together suggests that blue ship NFTs have, taken as a whole, offered both-crypto-and-NFT investors high returns while diversifying their risk.

According to figure 8, only a brief bubble has been detected by the presented methodology. Thus, the NFT price index has globally never experienced any explosive price dynamic, but it doesn't mean that some collections have not. Indeed, as shown in figure 8, a bubble has been detected in the metaverse index between the 2. November 2021 and the 2. December 2021 following the renaming of Facebook.

Concerning our methodology, two weaknesses have been identified. First, the selection of collections is highly subjective and includes a look-ahead bias. Secondly, hedonic coefficients $\boldsymbol{\alpha}$ and $\boldsymbol{\beta}$ are hold fixed over time, thus, our model can not take into account the time varying popularity of collections to explain price dynamics. Both of these problems could be solved by implementing a quantitative rule to update the collection set on a regular basis, and, by using the *adjacent dummy variable* method [9], an alternative training procedure for the time dummy variable hedonic model. It could also be interesting in a future work to consider the temporal dependence of the trends according to the collections, or to detect NFT market regimes. Finally, it would also be appropriate to test and compare other bubble detection models.

VI. Acknowledgement

We would like to thank Opensea for providing us an API key to collect all the necessary data to conduct this study.

VII. Conclusion

In this work, we have proposed a methodology to construct price indices for the NFT markets. In contrary to previous works, our indices aggregate transactions from varied collections, this allows us to better represent this new art market 3.0. These indices have allowed us to analyse the dynamics and performances of NFT markets, and to perform diagnostic tests. In particular, we have used statistical tests to detect speculative bubbles. Finally, we have demonstrated that simple intuitive investment strategies could be derived from the pricing model used for the contruction of our indices.

References

[1] D. Analytics and @thomas_m. Nft market overview. [Online]. Available: https://dune.xyz/thomas_m/NFT-stats

[2] S. Nakamoto, "Bitcoin: A peer-to-peer electronic cash system," *Decentralized Business Review*, p. 21260, 2008.

[3] V. Buterin *et al.*, "Ethereum white paper," *GitHub repository*, vol. 1, pp. 22–23, 2013.

[4] R. Kräussl, T. Lehnert, and N. Martelin, "Is there a bubble in the art market?" *Journal of Empirical Finance*, vol. 35, pp. 99–109, 2016.

[5] Opensea. Cryptopunk #9998. [Online]. Available: https://opensea.io/assets/0xb47e3cd837ddf8e4c57f05d70ab865de6e193bbb/9998

[6] ——. Cryptopunk #3100. [Online]. Available: https://opensea.io/assets/0xb47e3cd837ddf8e4c57f05d70ab865de6e193bbb/3100

[7] S. Herath and G. Maier, "The hedonic price method in real estate and housing market research: a review of the literature," 2010.

[8] S. Selim, "Determinants of house prices in turkey: A hedonic regression model," *Doğuş Üniversitesi Dergisi*, vol. 9, no. 1, pp. 65–76, 2011.

[9] J. Triplett, “Handbook on hedonic indexes and quality adjustments in price indexes: Special application to information technology products,” 2004.

[10] R. Kräussl and N. v. Elsland, “Constructing the true art market index: A novel 2-step hedonic approach and its application to the german art market,” CFS working paper, Tech. Rep., 2008.

[11] D. Witkowska, “An application of hedonic regression to evaluate prices of polish paintings,” *International Advances in Economic Research*, vol. 20, no. 3, pp. 281–293, 2014.

[12] P. Kireyev and R. Lin, “Infinite but rare: Valuation and pricing in marketplaces for blockchain-based nonfungible tokens,” 2021.

[13] D.-R. Kong and T.-C. Lin, “Alternative investments in the fintech era: The risk and return of non-fungible token (nft),” *Available at SSRN 3914085*, 2021.

[14] M. Goldberg, P. Kugler, and F. Schär, “The economics of blockchain-based virtual worlds: A hedonic regression model for virtual land,” *Available at SSRN 3932189*, 2021.

[15] P. C. Phillips and T. Magdalinos, “Limit theory for moderate deviations from a unit root,” *Journal of Econometrics*, vol. 136, no. 1, pp. 115–130, 2007.

[16] P. C. Phillips, S. Shi, and J. Yu, “Testing for multiple bubbles,” 2012.

[17] Cryptoslam. [Online]. Available: https://cryptoslam.io

[18] Google trends. [Online]. Available: https://trends.google.com/trends/explore?date=2021-01-23%202022-01-23&q=metaverse

[19] Google trends. [Online]. Available: https://trends.google.com/trends/explore?date=2021-01-23%202022-01-23&q=nft

APPENDIX

A. Raw signals

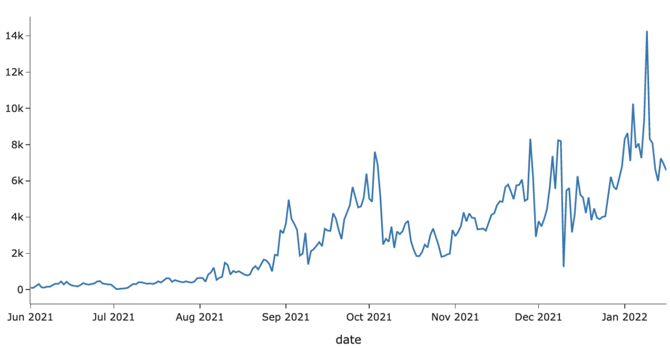

Fig. 10: NFT Index.

B. Smart contracts

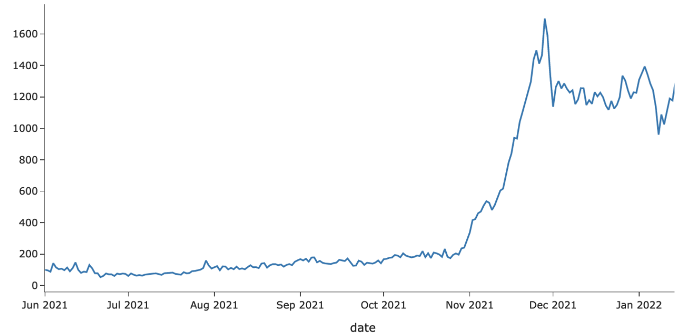

Fig. 11: Metaverse Index.

Index	Collection	Ethereum Contract
NFT	CryptoPunks	0xb47e3cd837dDF8e4c57F05d70Ab865de6e193BBB
	CloneX	0x49cF6f5d44E70224e2E23fDcdd2C053F30aDA28B
	MutantApeYachtClub	0x60E4d786628Fea6478F785A6d7e704777c86a7c6
	BoredApeYachtClub	0xBC4CA0EdA7647A8aB7C2061c2E118A18a936f13D
	Sandbox	0x50f5474724e0Ee42D9a4e711ccFB275809Fd6d4a
	VOX Series 2	0xf76179bb0924BA7da8E7b7Fc2779495D7A7939d8
	Nanopass	0xf54cC94f1F2f5De012B6Aa51F1E7eBdc43Ef5afC
	Shiba Social Club	0xD692cEd124A474f051f9744a301C26D1017B3D54
	Psychedelics Anonymous Genesis	0x75E95ba5997Eb235F40eCF8347cDb11F18ff640B
	Art Blocks	0xa7d8d9ef8D8Ce8992Df33D8b8CF4Aebabd5bD270
	CoolmansUniverse	0xa5C0Bd78D1667c13BFB403E2a3336871396713c5
	My Pet Hooligan	0x09233d553058c2F42ba751C87816a8E9FaE7Ef10
	Decentraland	0xF87E31492Faf9A91B02Ee0dEAAd50d51d56D5d4d
	Terraforms	0x4E1f41613c9084FdB9E34E11fAE9412427480e56
	Doodles	0x8a90CAb2b38dba80c64b7734e58Ee1dB38B8992e
	Neo Tokyo	0x698FbAACA64944376e2CDC4CAD86eaa91362cF54
	Creature World	0xc92cedDfb8dd984A89fb494c376f9A48b999aAFc
	More Than Gamers	0x49907029e80dE1cBB3A46fD44247BF8BA8B5f12F
	VOX Series 1	0xad9Fd7cB4fC7A0fBCE08d64068f60CbDE22Ed34C
	DinoBabies	0xf9d53E156fE880889E777392585FEb46D8D840f6
	Party Bears	0x35471f47c3C0BC5FC75025b97A19ECDDe00F78f8
	DEGENERATE/REGENERATE	0x7828c811636CCf051993C1EC3157b0B732e55B23
	ALIENFRENS	0x123b30E25973FeCd8354dd5f41Cc45A3065eF88C
	NFT Worlds	0xBD4455dA5929D5639EE098ABFaa3241e9ae111Af
	CyberKongz	0x57a204AA1042f6E66DD7730813f4024114d74f37
	Slotie	0x5fdB2B0C56Afa73B8ca2228e6aB92Be90325961d
	RumbleKongLeague	0xEf0182dc0574cd5874494a120750FD222FdB909a
	Cool Cats	0x1A92f7381B9F03921564a437210bB9396471050C
	Punks Comic 2	0x128675d4FddbC4a0D3f8aA777D8EE0fb8B427C2F
	Wolf Game	0x7F36182DeE28c45dE6072a34D29855BaE76DBe2f
	Little Lemon Friends	0x0B22fE0a2995C5389AC093400e52471DCa8BB48a
	VeeFriends	0xa3AEe8BcE55BEeA1951EF834b99f3Ac60d1ABeeB
	Crypto Bull Society	0x469823c7B84264D1BAfBcD6010e9cdf1cac305a3
	House of Legends	0x8C714199d2eA08CC1f1F39A60f5cD02aD260A1e3
	Meta-Legends	0xF9c362CDD6EeBa080dd87845E88512AA0A18c615
	Gambling Apes	0x90cA8a3eb2574F937F514749ce619fDCCa187d45
	BoredApeKennelClub	0xba30E5F9Bb24caa003E9f2f0497Ad287FDF95623
	Cryptoadz	0x1CB1A5e65610AEFF2551A50f76a87a7d3fB649C6
	apekidsclub	0x9Bf252f97891b907F002F2887EfF9246e3054080
	ExpansionPunks	0x0D0167A823C6619D430B1a96aD85B888bcF97C37
	Elderly Ape Retirement Club	0x9ee36cD3E78bAdcAF0cBED71c824bD8C5Cb65a8C
	Divine Wolves	0xb4e9123bd3Ef4Df17f8cc6EF7C2Be66428CF4931
	Superlative Apes	0x1e87eE9249Cc647Af9EDEecB73D6b76AF14d8C27
	Desperate ApeWives	0xF1268733C6FB05EF6bE9cF23d24436Dcd6E0B35E
	Solarbots	0x8009250878eD378050eF5D2a48c70E24EB2edE7E
	merge.	0xc3f8a0F5841aBFf777d3eefA5047e8D413a1C9AB
	Neo Tokyo Part 3	0x0938E3F7AC6D7f674FeD551c93f363109bda3AF9
	Art blocks	0x059EDD72Cd353dF5106D2B9cC5ab83a52287aC3a
	Champions	0x97a923ed35351a1382E6bcbB5239fc8d93360085
	Pepsi Mic Drop	0xa67D63E68715DCF9b65e45e5118b5fcD1e554b5f
	CROAKZ	0x7caE7B9b9a235D1D94102598E1f23310A0618914
	Meebits	0x7Bd29408f11D2bFC23c34f18275bBf23bB716Bc7
	Cyber Gorillas	0x1E1b4E127A510cafa6d0eAec024a4319a5E18821
	Lost Poets	0x4b3406a41399c7FD2BA65cbC93697Ad9E7eA61e5
	SuperRare	0xb932a70A57673d89f4acfFBE830E8ed7f75Fb9e0
	Lazy Lions	0x8943C7bAC1914C9A7ABa750Bf2B6B09Fd21037E0
	Neo Tokyo Part 2	0xab0b0dD7e4EaB0F9e31a539074a03f1C1Be80879
	Loot	0xFF9C1b15B16263C61d017ee9F65C50e4AE0113D7
	Cryptovoxels	0x79986aF15539de2db9A5086382daEdA917A9CF0C
Metaverse	Decentraland	0xF87E31492Faf9A91B02Ee0dEAAd50d51d56D5d4d
	Sandbox	0x50f5474724e0Ee42D9a4e711ccFB275809Fd6d4a
	Cryptovoxels	0x79986aF15539de2db9A5086382daEdA917A9CF0C
	NFT Worlds	0xBD4455dA5929D5639EE098ABFaa3241e9ae111Af

TABLE V: Composition of the indices.