Modeling NFT Investor Behavior Using Belief Dissensus

Percy Venegas

King's College London/Economy Monitor

Fernand Gobet

London School of Economics and Political Science

Acknowledgements: We thank SimilarWeb Ltd for access to web panel data at the US state level.

Abstract

Investment into non-fungible tokens (NFTs) has skyrocketed in 2021. Since NFTs are issued on blockchains, the underlying operation is that of a social computer—therefore, modeling social cognition in NFT markets becomes relevant. Market participants (collectors, speculators, and investors) may display different levels of expertise that serve as "social labels." However, do users of NFT marketplaces care about price or community? Besides the operational consensus mechanism of the blockchain (which also provides the hard judiciary and settlement layers of the system), we must consider the soft consensus of the internet communities that drive their attention towards NFT marketplaces where the monetary assets are listed. In this research note, we propose an approach that offers a window on human cognition and collective intelligence, but that can inform the development of artificial systems that help develop policies to protect the public interest of investors.

Keywords: non-fungible tokens, NFTs, cryptomarkets, evolutionary algorithms, cognitive science

Modelar el comportamiento de los inversores NFT utilizando el desacuerdo de creencias

Resumen

La inversión en tokens no fungibles (NFT) se disparó en 2021. Dado que los NFT se emiten en cadenas de bloques, la operación subyacente es la

doi: 10.18278/jpcs.7.2.2

de una computadora social; por lo tanto, el modelado de la cognición social en los mercados de NFT se vuelve relevante. Los participantes del mercado (coleccionistas, especuladores e inversores) pueden mostrar diferentes niveles de experiencia que sirven como “etiquetas sociales”. Sin embargo, ¿a los usuarios de los mercados NFT les importa el precio o la comunidad? Además del mecanismo de consenso operativo de la cadena de bloques (que también proporciona las capas judiciales y de liquidación duras del sistema), debemos considerar el consenso blando de las comunidades de Internet que dirigen su atención hacia los mercados NFT donde se enumeran los activos monetarios. En esta nota de investigación, proponemos un enfoque que ofrece una ventana a la cognición humana y la inteligencia colectiva, pero que puede informar el desarrollo de sistemas artificiales que ayuden a desarrollar políticas para proteger el interés público de los inversores.

Palabras clave: tokens no fungibles, NFT, criptomercados, algoritmos evolutivos, ciencia cognitiva

使用信念岐见对非同质化代币投资者行为进行建模

摘要

对非同质化代币（NFTs）的投资在2021年激增。鉴于NFT在区块链上发行，潜在的操作则是社会计算机操作，因此，对NFT市场中的社会认知进行建模一事便具有相关性。市场参与者（收集者、投机者和投资者）可能展现不同程度的专业性，这种专业性充当不同的“社会标签”。不过，NFT市场用户真的关心价格或社区吗？除了区块链的操作共识机制（为该系统提供稳固的司法层面和解决层面），我们必须衡量互联网社区的软共识，该社区将关注转向将货币资产包括在内的NFT市场。在该研究纪要中，我们提出一项措施，该措施为人类认知和集体智慧提供窗口，并能影响人工系统的开发，帮助发展一系列保护投资者公共利益的政策。

关键词：非同质化代币，NFTs，加密货币市场，演化算法，认知科学

Introduction

The technology behind non-fungible tokens (NFTs) has been around since 2017, but interest in NFTs as collectives and investable assets only gained traction during 2021, with sales over USD 10 Billion in 2021 (see Table 1).

Table 1. *NFT market size. Source: Nonfungible.com*

Key figures of NFT market as of November 12, 2021	All-time as of November 12 2021	Last 30 days as of November 12 2021
Sales value	10.2bn USD	1.69bn USD
Primary sales value	2.35bn USD	0.43bn USD
Secondary sales value	7.84bn USD	1.26bn USD
Average sales value	927.71 USD	1831.41 USD
Number of sales	10.99 million	0.92 million
Active market wallets	799623	241337
Unique buyers	756135	215227
Unique sellers	304798	88969

Effectively, NFTs are tokens that provide access to communities of like-minded investors. Those subcultures not only become investable but also offer an intrinsic rewards system. In their paper on social identification and investment decisions, Bauer and Smeets (2015) argue that investors get non-financial utility if investments fit their (desired) social identity. However, to what extent are NFT investors interested in community versus only prices?

Expertise theory

Nadini et al. (2021) have studied traders and NFTs networks and found that most traders are specialized: measuring how individuals distribute their trades across collections, they found that at least 73% of traders' transactions are performed in their top collection, and at least 82% in their top two collections. Such studies that use exchange and marketplace data indeed shed light on the microeconomics of the market. However, to understand the market operation at the macroeconomic level, it is necessary to use off-chain data.

A prevalent albeit imperfect definition of expertise considers *expertise as experience*—that is, the achievement of expert status is related to the amount of time an individual has spent in a domain (Gobet, 2016). Table 2 shows the estimated visit duration in minutes at the site Opensea.io, one of the largest NFT marketplaces in the world. There are two interesting observations to make: first, we can confirm an increase in interest during the last half of 2021;

second, time spent on the site increased prominently among desktop users. Professional and semi-professional investors are desktop users, so having the time on site doubling from May to August is a strong indication of increased expertise.

Table 2. *Visit duration to Opensea.io (hours: mins: secs), source: Similarweb*

Date	Avg. Visit Duration (Mobile Web)	Avg. Visit Duration (Desktop)
01/12/2019	00:02:36	00:09:00
01/01/2020	00:02:06	00:11:39
01/02/2020	00:01:46	00:09:47
01/03/2020	00:01:12	00:11:59
01/04/2020	00:00:56	00:10:19
01/05/2020	00:01:31	00:11:37
01/06/2020	00:01:20	00:08:12
01/07/2020	00:01:20	00:09:16
01/08/2020	00:01:13	00:09:50
01/09/2020	00:01:02	00:07:15
01/10/2020	00:01:25	00:11:13
01/11/2020	00:01:04	00:07:14
01/12/2020	00:01:00	00:07:11
01/01/2021	00:01:09	00:09:15
01/02/2021	00:01:14	00:09:50
01/03/2021	00:01:24	00:10:26
01/04/2021	00:01:13	00:08:59
01/05/2021	00:01:26	00:11:09
01/06/2021	00:01:40	00:14:19

01/07/2021	00:01:40	00:16:54
01/08/2021	00:02:01	00:20:22
01/09/2021	00:01:46	00:19:08
01/10/2021	00:01:41	00:17:29
01/11/2021	00:01:48	00:14:07
01/12/2021	00:01:40	00:14:43
01/01/2022	00:02:02	00:17:39

However, deliberate practice (Ericsson et al., 1993) offers a better measure to understand the acquisition of expertise. Deliberate practice is systematic, focused, and seeks to improve performance. Specialization from this perspective means that the new NFT traders, many technically savvy, had to peruse financial sites (many of which were specialized in crypto) to improve their financial literacy and execute their trading activities. The concentration of topics into finance and technology in Figure 1 confirms this assertion.

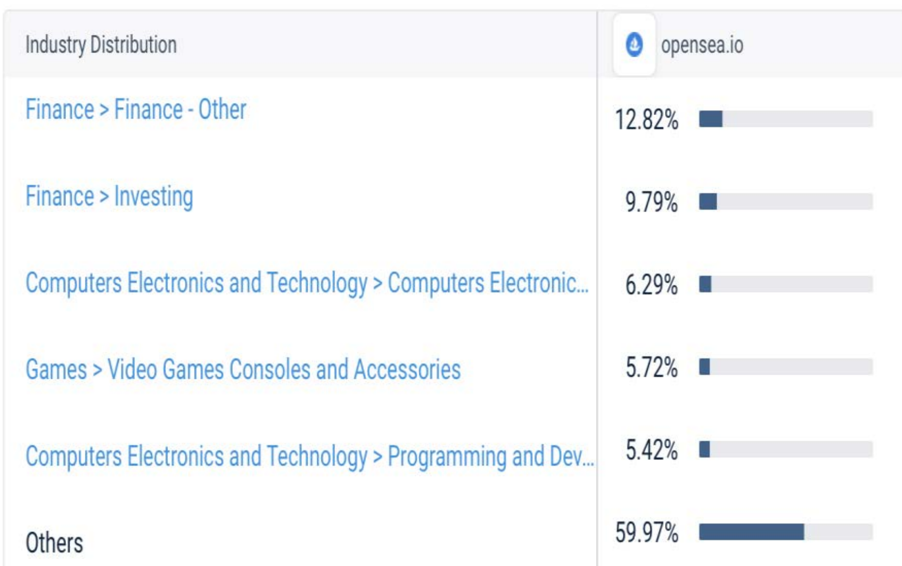

Figure 1. Distribution of site categories visited by the audience of Opensea.io in the U.S. from Dec 2020 to Nov 2021. Source: Similarweb.

Nevertheless, "expertise" can often only be used within a specific context (Stein, 1997). Therefore, we also explore geographical social groups across the United States.

The study

To investigate the response of market participants to movements in price and popularity of NFTs, we use alternative data. Prices are denominated in USD and come from the NFT Index by https://nftindex.tech/, an index that tracks the performance of tokens within the NFT industry. The index is capitalization-weighted and tracks the market performance of decentralized financial assets, if they are significantly used and committed to ongoing maintenance and development. The index tracks assets available in the Ethereum blockchain and is independent of any marketplace, making it suitable for macro-level price monitoring, i.e., keeping the pulse of the NFT markets.

We use the estimated average number of daily pages visited in Opensea.io across a group of U.S. States (Arizona, California, Florida, Georgia, Illinois, Massachusetts, Michigan, New Jersey, New York, North Carolina, Ohio, Pennsylvania, Texas, Virginia, and Washington), since an increment on pages visited signals increased interest and adoption. Assets listed in Opensea are issued as ERC-721 standard compliant, which means that they are predominantly issued on the Ethereum blockchain and priced in Ether.

We use total pages visited on the site https://www.airnfts.com/ as a proxy for the popularity of NFT collectives. To prevent the evolutionary algorithm (see below) from feeding on a self-reinforcing bias, we choose a site operating on a platform different from Ethereum, which should remove at least part of the possible audience overlap. AirNFTs lists assets issued in the Binance Blockchain; Binance is the largest exchange globally in terms of the daily trading volume of cryptocurrencies (Peters, 2021).

The dataset contains daily observations between April 19th and November 30th, 2021. Similarweb.com provides web panel (visits) data. The data exploration and modeling phases are performed using Mathematica (Wolfram Research, 2021) and Datamodeler (Evolved Analytics, 2021). The dataset is available for download at: https://www.autonomous.economymonitor.com/s/NFT-master.csv

Methodology

We start with exploratory data analysis to understand the shape of the data: the main statistical properties and correlations between variables. Then, the modeling stage is done using symbolic regression via genetic programming, a technique that has previously been used to study crypto-economic systems (Venegas, 2021). There are two rounds of modeling: the first round is performed mainly to discover the driver variables (this focus helps the evolutionary algorithm find and develop creative paths, rather than losing time with spurious associations between variables); the second round consists in building groups of models (“ensembles”) with explanatory and predictive power. This workflow is repeated for each of the two target variables: price and popularity. Finally, a multi-target modeling stage is introduced. Here, we compare ensembles of diverse, optimal models (both accurate and simple).

The general idea is straightforward: if people's beliefs in *price appreciation* and *popularity increase* reach consensus, their browsing activity should intensify. Moreover, where there is dissensus and increasing uncertainty, the choice of predictive variables should be revisited.

Results

Descriptive Statistics

The main statistical properties of the dataset are analyzed (Figure 2). We find that all numerical variables are continuous, and observations are uniform (records are complete 92% of the time and above). The values are strictly positive (no zero-crosses).

DataSummaryTable

Col	Label	Type	Uniformity	Class	Unique	Distribution Plot	Zero–Cross	Min	Mean	Median	Max
1	Date	ABC	100%		226	Lots of different values		10/10/2021	9/9/2021	9/9/2021	9/9/2021
2	Arizona	123	98%		220			1.0	30.8	21.7	222.8
3	California	123	100%		217			4.3	18.0	17.4	40.2
4	Florida	123	100%		215			3.0	18.3	16.8	61.1
5	Georgia	123	100%		213			1.0	13.6	12.3	52.8
6	Illinois	123	100%		214			2.9	15.2	12.1	175.6
7	Massachusetts	123	100%		214			1.0	16.6	14.9	76.3
8	Michigan	123	99%		215			1.0	13.8	11.6	58.8
9	New Jersey	123	100%		215			2.5	18.8	17.3	62.1
10	New York	123	100%		212			3.1	15.2	14.8	44.1
11	North Carolina	123	100%		219			1.8	27.8	19.3	181.4
12	Ohio	123	92%		199			1.0	14.2	10.5	69.5
13	Pennsylvania	123	100%		213			2.3	26.2	21.4	119.8
14	Texas	123	100%		212			2.6	14.7	14.1	58.1
15	Virginia	123	100%		212			2.0	13.5	12.4	62.7
16	Washington	123	100%		212			1.2	13.8	12.5	54.9
17	Price	123	100%		225			354.3	863.2	706.2	1643.3
18	Popularity	123	98%		222			61.0	5467.7	3963.7	34845.0

Figure 2. Data summary table with the main statistical properties of the dataset. Own construction using Datamodeler; Source: Similarweb and NFTindex.tech.

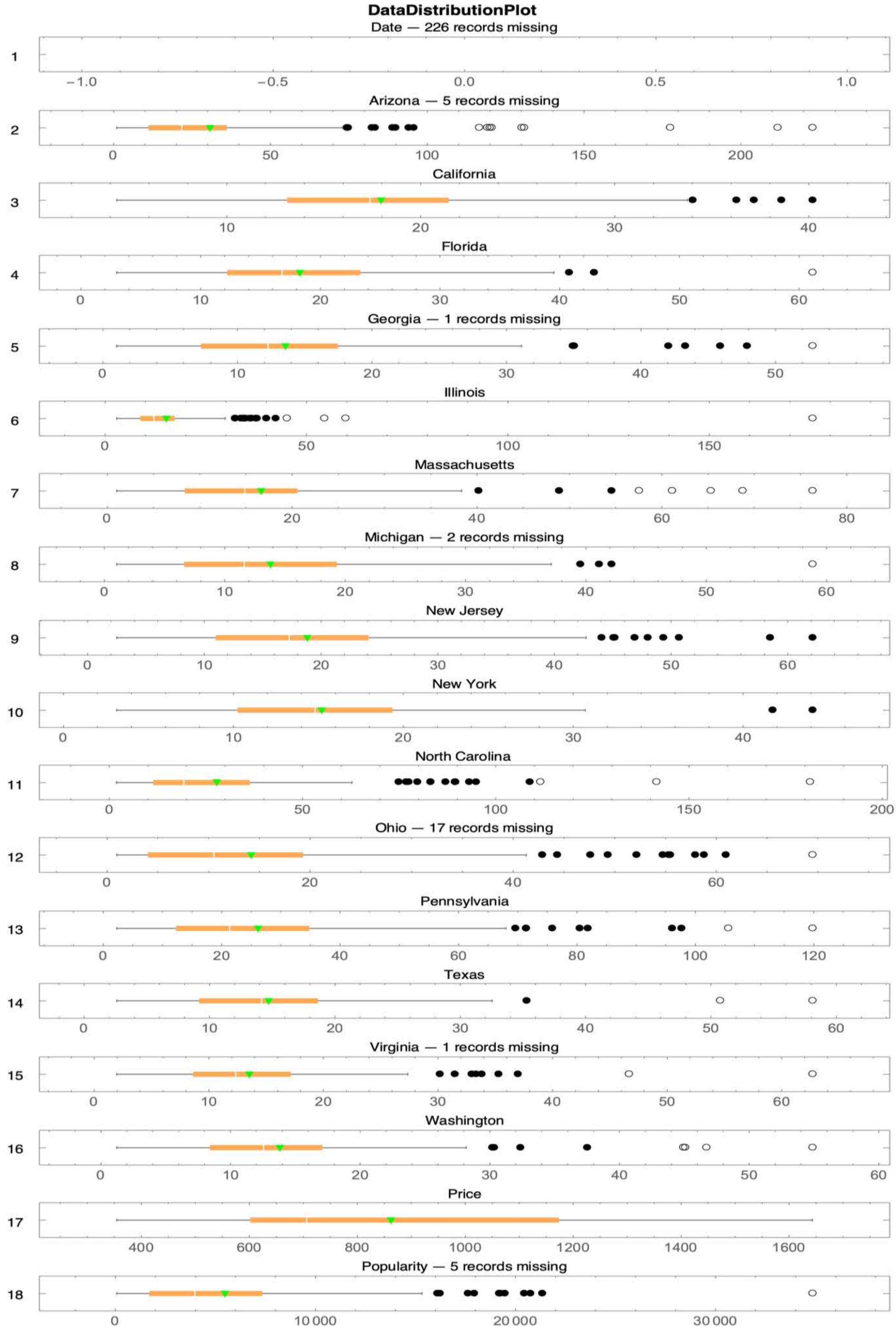

Figure 3. Central values and dispersion for each variable. Own construction using Datamodeler; Source: Similarweb and NFTindex.tech.

Data distribution

The dispersion of the data is different between states. For instance, in Ohio, the 75th percentile is 19.295 with a peak at 69.47, while in Illinois, it is 17.25 with a large outlier at 175.59. This may point to different user behavior across regions in the U.S., or perhaps some data mining experiments being conducted at locations (typically, such experimentation would be conducted using bots, which could generate large volumes of "artificial" activity that may skew the data). The box and whisker plots in Figure 3 depict the spread and locality of data points, with the data within the 25% and 75% percentiles shown in orange, the median as a white dent in the orange bar, and the mean as a green marker; the black dots are outliers, and the white dots are far outliers.

Timeseries

The timeseries plots (Figure 4) provide a sense of the shape of the data. Here we confirm how marketplace usage behavior seems to vary significantly across states. Also, we see the first hints of an apparent decoupling of price and popularity.

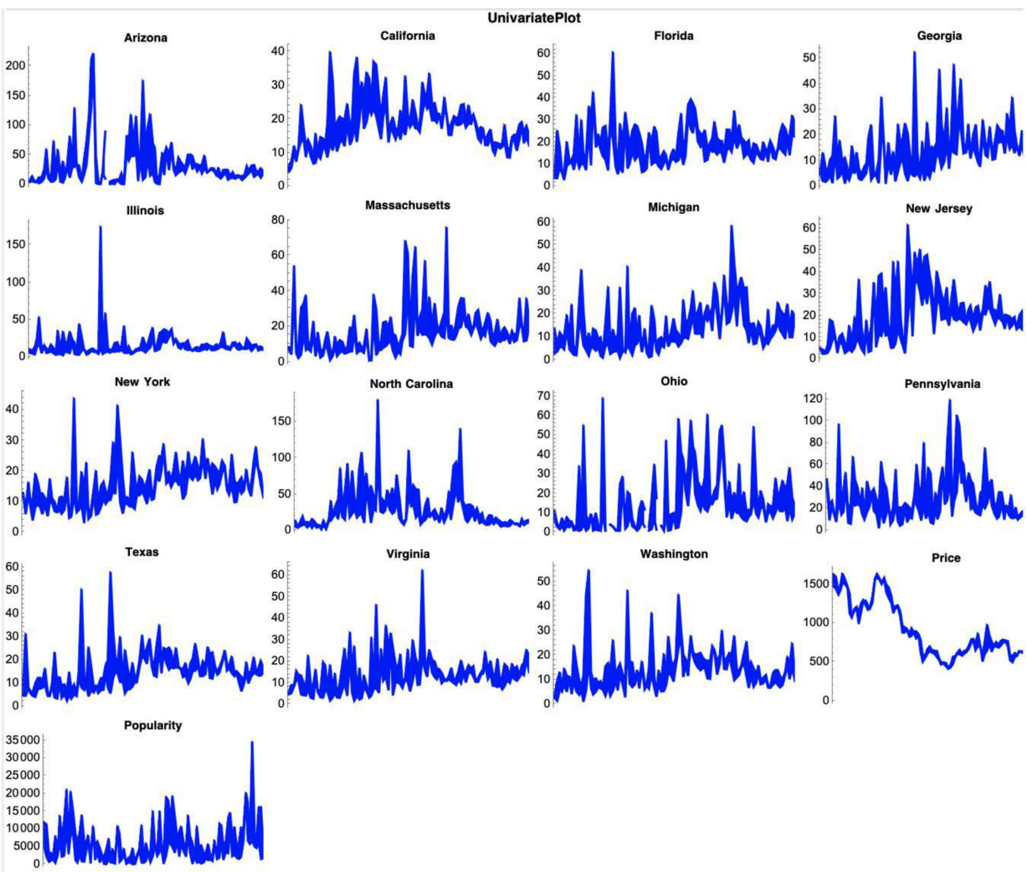

Figure 4. Univariate plots (X axis is time, Y axis is values of each variable). Own construction using Datamodeler; Source: Similarweb and NFTindex.tech.

Correlations

The linear correlations between variables confirm the previous observation regarding the relationship between price and popularity (negative correlation) and between price and the other variables. For instance, we note that at a 0.4 threshold, the price is negatively correlated to site activity in Texas and New York (Figure 5).

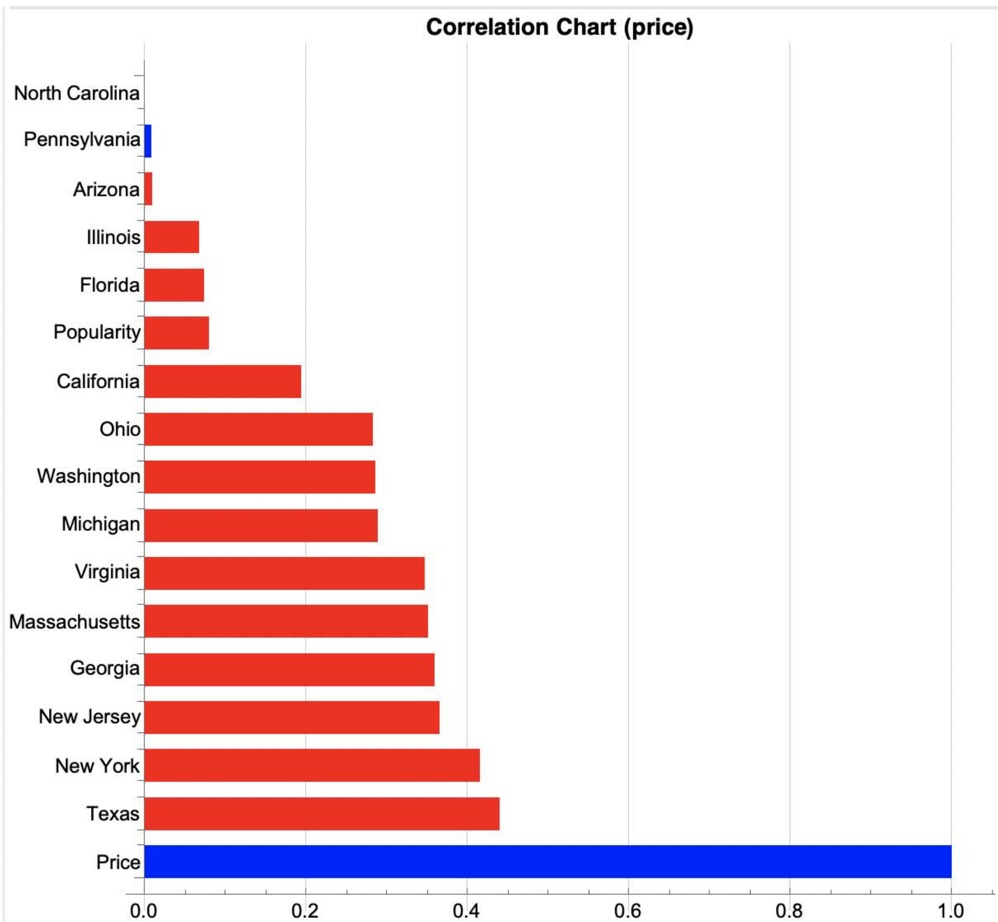

Figure 5. Absolute correlations with price. Positive correlations in blue; negative correlations in red. Own construction using Datamodeler; Source data: Similarweb and NFTindex.tech.

Modeling

<u>First round.</u> During the first modeling round, we developed 1,138 models for price and 971 for popularity. We discover the more prevalent variables across models in each case (depicted in Figure 6 and 7 as predominantly solid red bars): for the price, marketplace site activity in New Jersey, New York, Ohio, and Texas; for popularity, New Jersey, North Carolina, and Ohio. The interpretation of the figures is as follow: in each model (X axis) a variable will

be present (or not)—the variables that present higher visual density are likely to be more important.

<u>Second round.</u> For the second modeling round, we create a subset with the variables in common, Ohio and New Jersey. We then run the second round of modeling using the same modeling parameters (i.e., types of mathematical blocks) but only that new subset of variables. As we will see below, the number of total models generated in this case is different: the reason is that once the evolutionary algorithm is constrained to work with fewer variables, it can itself specialize and develop new populations of models.

Price

Figure 8 shows the 11 models included in the ensemble, out of 1,099 models developed in the second round. The red dots in the "knee" of the pareto denote models that have a good balance between complexity and accuracy.

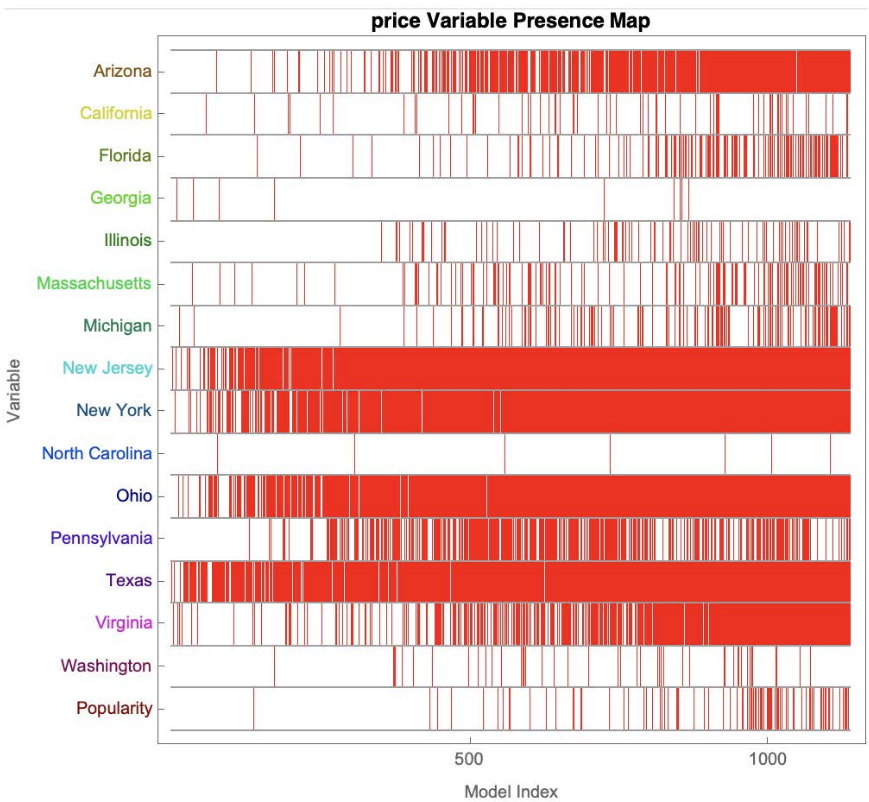

Figure 6. Variable presence across models (target: price).
Own construction using Datamodeler.

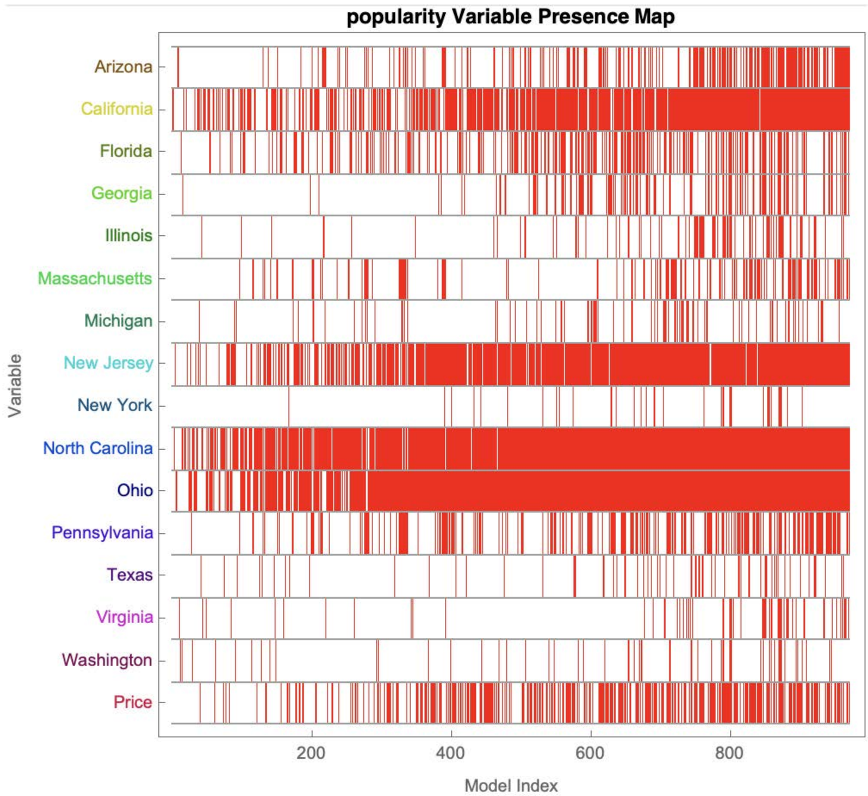

Figure 7. Variable presence across models (target: popularity). Own construction using Datamodeler.

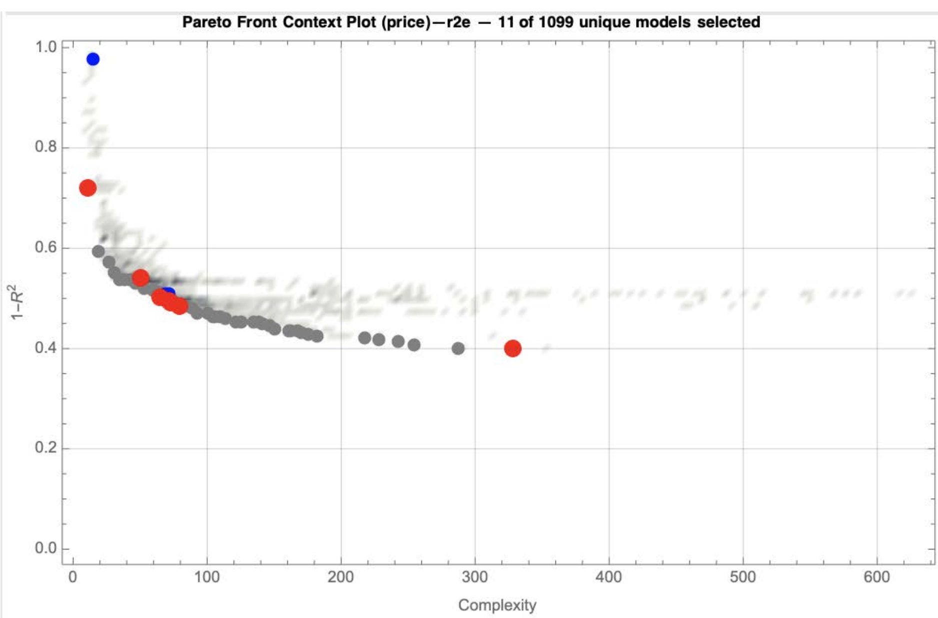

Figure 8. Accuracy – Complexity trade-off (price). Gray dots are suboptimal models, not included in the ensemble.

In Figure 9, the models are ranked by complexity and error.

Model Selection Report (price)

	Complexity	$1-R^2$	Function
1	11	0.720	$653.23 + \frac{2510.90}{\text{New Jersey}}$
2	15	0.979	$889.59 - (1.74\times10^{-3})\,\text{New Jersey}^3$
3	50	0.540	$879.34 + \frac{9783.15}{7.64+\text{New Jersey}} - 247.59\sqrt{\text{Ohio}} + 20.40\,\text{Ohio} + 0.25\,\text{New Jersey Ohio}$
4	65	0.503	$459.72 + \frac{14207.95}{9.27+\text{New Jersey}} - 50.78\,\text{Ohio} + 1.00\,\text{New Jersey Ohio} + 0.88\,\text{Ohio}^2 - (3.60\times10^{-4})\,\text{New Jersey Ohio}^3$
5	69	0.509	$3323.25 - 203.11\sqrt{\text{New Jersey}} - \frac{724.93}{\text{Ohio}} - 911.14\sqrt{\text{Ohio}} + 94.07\,\text{Ohio} + 1.84\,\text{New Jersey Ohio} - (3.49\times10^{-2})\,\text{New Jersey Ohio}^2$
6	71	0.507	$937.45 - \frac{35559.62}{\text{New Jersey}^4} + \frac{3874.54}{\text{New Jersey}} + 3.45\,\text{New Jersey} - \frac{7.84}{-9.10+\text{New Jersey}} - 235.55\sqrt{\text{Ohio}} + 24.13\,\text{Ohio}$
7	72	0.497	$2584.80 - 1187.19\,\text{New Jersey}^{1/3} + 49.07\,\text{New Jersey} - \frac{70.01\,\text{New Jersey}}{\text{Ohio}} + 6.42\,\text{Ohio} + \frac{4271.79}{\frac{21}{\text{New Jersey}}+\text{Ohio}}$
8	73	0.491	$1549.98 + \frac{14902.05}{9.73+\text{New Jersey}} - \frac{546.08}{\text{Ohio}} - 725.95\sqrt{\text{Ohio}} + 76.34\,\text{Ohio} + 1.33\,\text{New Jersey Ohio} - (2.62\times10^{-2})\,\text{New Jersey Ohio}^2$
9	73	0.492	$1581.21 + \frac{11770.18}{7.64+\text{New Jersey}} - \frac{519.54}{\text{Ohio}} - 700.49\sqrt{\text{Ohio}} + 74.00\,\text{Ohio} + 1.24\,\text{New Jersey Ohio} - (2.49\times10^{-2})\,\text{New Jersey Ohio}^2$
10	80	0.485	$2736.75 - \frac{419.14}{\text{Ohio}} + 2.49\,\text{New Jersey Ohio} + 0.40\,\text{Ohio}^2 - (3.64\times10^{-2})\,\text{New Jersey Ohio}^2 - 660.68\,(\text{New Jersey Ohio})^{1/4}$
11	328	0.399	$-96944.59 - 585.05\sqrt{\text{New Jersey}} + 28.35\,\text{New Jersey} + \frac{4766.58}{42.60+2\,\text{New Jersey}+\frac{\text{New Jersey}^2}{-6+\text{Ohio}}+3\,\text{New Jersey}\sqrt{\text{Ohio}}} - 894.54\sqrt{\text{Ohio}} + 8.65\,\text{New Jersey}\sqrt{\text{Ohio}} + 112.27\,\text{Ohio} - \frac{468.31\,\text{Ohio}}{27.53+\frac{\text{New Jersey}^2}{-6+\text{Ohio}}} - (2.18\times10^{-2})\,\text{New Jersey Ohio}^2 + \frac{7856043.60}{\left(8.82+\frac{1}{9.07+\text{New Jersey}+3\,\text{New Jersey}\sqrt{\text{Ohio}}+\text{Ohio}^3}\right)^2}$

Figure 9. Ensemble constituent models (price) from lower to higher complexity.

In Figure 10, the range of embedded model predictions is shown in blue, while the modeling outliers (most difficult data records to model) are shown in red. We note that the prediction performance is more challenging at higher price levels.

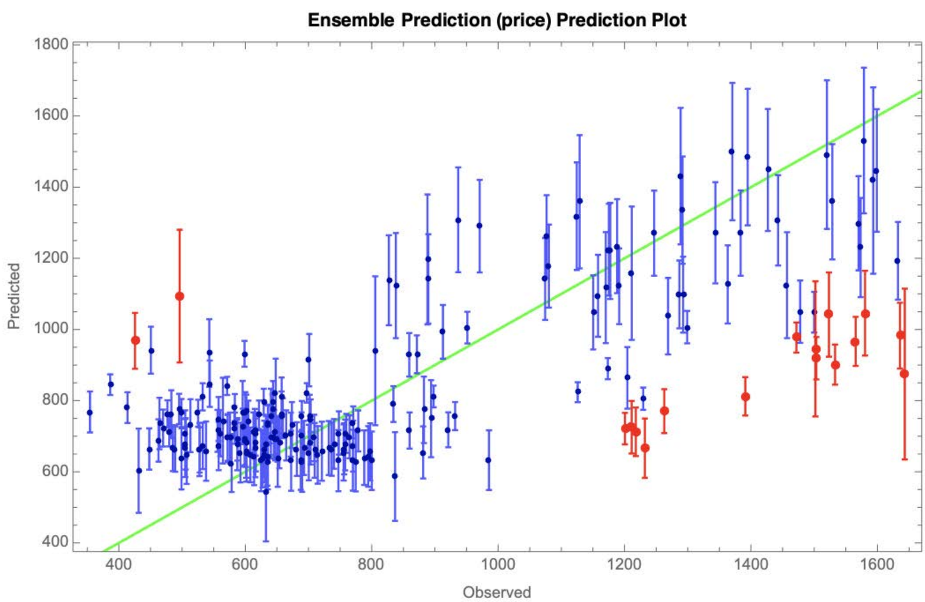

Figure 10. Ensemble performance (price). The range of computed values is represented by the length of the lines.

Popularity

In the second round, we developed 1000 models for popularity; Figure 11 shows the 11 models selected for the ensemble.

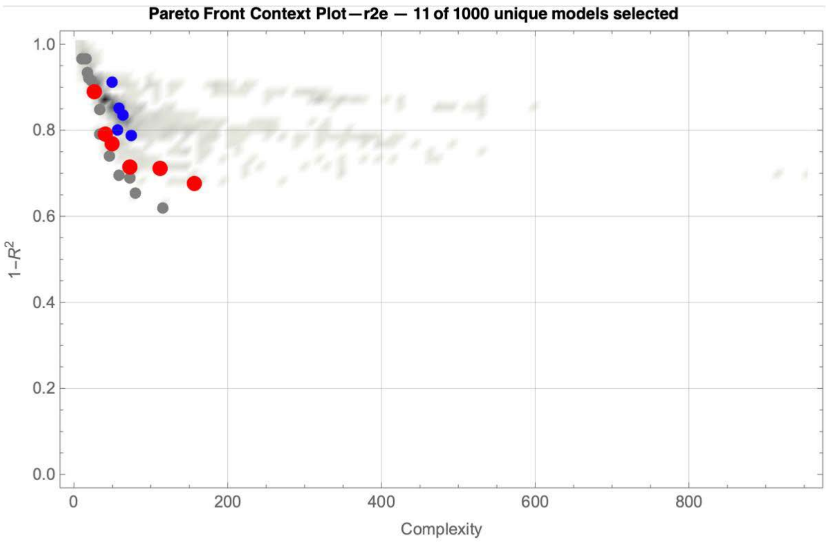

Figure 11. Accuracy-complexity trade-off (popularity). Gray dots are suboptimal models, not included in the ensemble.

The models are again ranked by complexity and error (Figure 12). We note some peculiarities: how the minimum complexity for the popularity models (27) is higher than in the case of the price models (11) and how the minimum error (0.889 vs. 0.720) is also higher.

Model Selection Report (popularity)

	Complexity	$1-R^2$	Function
1	27	0.889	$7534.11-109.90\,\text{New Jersey}+\frac{925.92}{-10.44+\text{Ohio}}$
2	41	0.792	$7519.61-114.81\,\text{New Jersey}-\frac{769.51}{7.22+\text{New Jersey}-2\,\text{Ohio}}$
3	49	0.767	$6860.81-125.75\,\text{New Jersey}-\frac{766.52}{7.22+\text{New Jersey}-2\,\text{Ohio}}+60.57\,\text{Ohio}$
4	49	0.912	$4381.26-217.07\,\text{New Jersey}+807.80\left(\text{New Jersey}^2\right)^{1/4}+551.61\sqrt{\text{Ohio}}$
5	57	0.800	$4532.09-\frac{678.23}{7.22+\text{New Jersey}-2\,\text{Ohio}}+\frac{1057.97}{14.08-\text{Ohio}}+52.59\,\text{Ohio}$
6	59	0.850	$5036.89-\frac{399.99}{10-\text{Ohio}}+37.74\,\text{Ohio}-\frac{153.92}{-0.96+\text{Ohio}}+\frac{926.17}{-10.44+\text{Ohio}}$
7	63	0.835	$6605.89-109.52\,\text{New Jersey}-\frac{1553.36}{-8.27+\text{New Jersey}}-\frac{1898.19}{\text{Ohio}}+406.82\sqrt{\text{Ohio}}-\frac{68.77}{-0.95+\text{Ohio}}$
8	73	0.714	$7560.14-122.85\,\text{New Jersey}-\frac{756.92}{7.22+\text{New Jersey}-2\,\text{Ohio}}-\frac{16.17}{0.95-\text{Ohio}}-\frac{2673.49}{\text{Ohio}}+39.54\,\text{Ohio}$
9	75	0.788	$9856.54-167.29\,\text{New Jersey}-\frac{422.62}{10-\text{Ohio}}-\frac{143.55}{-0.96+\text{Ohio}}+\frac{877.00}{-10.44+\text{Ohio}}-\frac{25912.89}{\text{New Jersey}+\text{Ohio}}$
10	112	0.712	$-14020.79+\frac{206646.27}{8.61-\frac{1}{8.61-\text{New Jersey}}}-\frac{15793.38}{\text{New Jersey}}-173.69\,\text{New Jersey}-\frac{597.47}{7.08+\text{New Jersey}+\frac{2}{\text{Ohio}}-2\,\text{Ohio}}+\frac{945.45}{14.08-\text{Ohio}}$
11	157	0.676	$-8731.16+\frac{161521.57}{8.61-\frac{1}{8.61-\text{New Jersey}}}-154.53\,\text{New Jersey}-\frac{649.52}{7.08+\text{New Jersey}+\frac{2}{\text{Ohio}}-2\,\text{Ohio}}-\frac{2543.36}{\text{Ohio}}-\frac{47656.99}{\text{New Jersey}+3\,\text{Ohio}+\frac{1}{-\sqrt{\text{Ohio}}+\text{Ohio}^2}}$

Figure 12. Ensemble constituent models (popularity) from lower to higher complexity.

The ensemble prediction tends to have a better performance at lower popularity levels (Figure 13), with a large dispersion at the peak point—where a single traffic spike occurred, as seen previously in Fig 5.

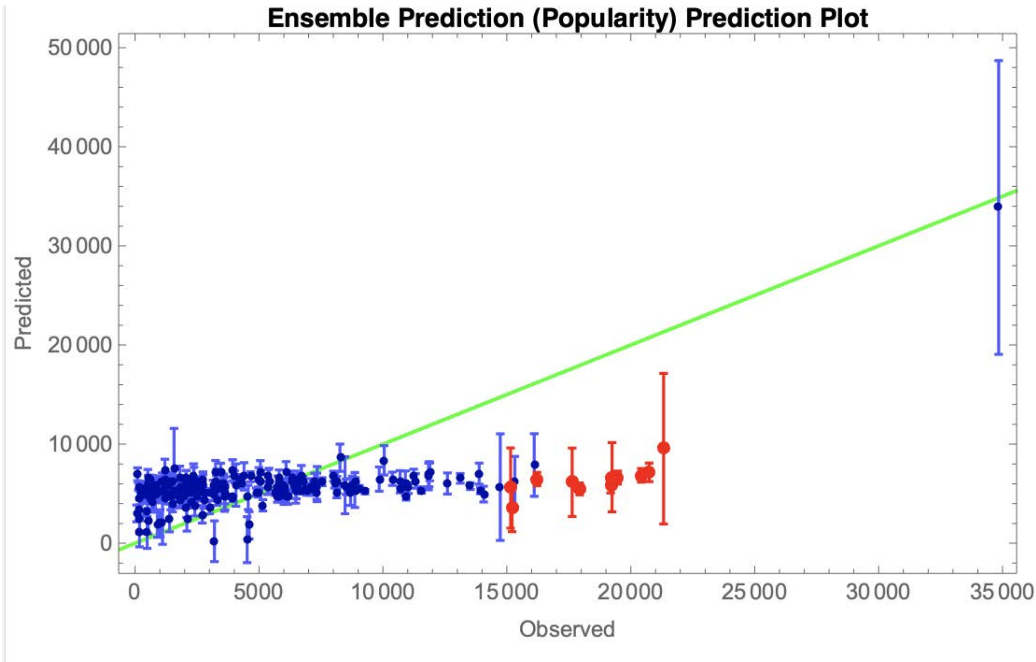

Figure 13. Ensemble performance (popularity). The range of computed values is represented by the length of the lines.

Multi-target response

Finally, we analyze the responses for the variables that are common to both model ensembles (New Jersey and Ohio). The explorer in Figure 14 allows us to see the effect of changing parameter values simultaneously on multiple target behaviors.

The trade-off curves move from red (minimum) to green (maximum) values of daily visits in each state. Normally, we would expect a curve with no loops or discontinuities; however, in this case we notice possible pathologies in the models (specially for popularity).

A closer inspection of the individual price (C.1 and C.2) and popularity (D.1 and D.2) response plots shows how the dissensus of models is predictable around mean values (blue line) with bounds defined by an envelope (yellow ribbon) and constituent models (ensemble submodels) depicted as gray lines inside the ribbon. In the case of price, we see a few instances when the submodels go outside of the boundaries, but mostly they remain within the ribbon. Popularity, however, is harder to model—there are several input values for which an output prediction value would be undefined (seen as pronounced spikes in the ribbon).

This difficulty in modeling the system when optimizing for popularity is consistent with the choice of variables: we should expect dissociation of behavior since the users are specialized in different blockchains (Ethereum and Binance Blockchain) which are largely mutually exclusive in terms of technology and user community. To be sure, modeling price is not without challenge: the size of the ribbon widens even for small values of pages visited (e.g., 20) which indicates an increase in uncertainty.

Revisiting the site categories of Figure 1, but this time at the local level (as shown in Fig 15 and 16), we find that the distribution of categories for the sites visited by the audience of the marketplace differ: in Ohio, it is predominantly technology and social network sites (like Discord, Twitter, Medium and Reddit), while in New Jersey it is mainly Other types of Financial sites (like crypto finance, such as Coinbase and Coingecko). This makes sense considering that the users in New Jersey are closer to the large financial center of New York City, and likely many of them work in the financial industry or are connected with people who do. On the other hand, users in Ohio might be predominantly in the early stages of their exposure to NFT markets (in the phase of discovery via social networks) – and, with less exposure to the financial industry, their level of interest in the financial aspects of NFT assets might be lower than in the case of users from New Jersey.

By the fact that C.1 and C.2 have conventional shapes (models diverge around a mean across different values of the parameter space), while D.1 and D.2 contain singularities, we confirm that the intensity of site activity (as expressed by pages per visit) is a better predictor of price than a predictor of popularity. The uncertainty of the prediction itself changes, with minimal uncertainty (narrower ribbons) around the value of 40 pages per visit in New Jersey and 55 in Ohio.

This behavior is precisely what we should expect given the selection of target variables: NFT investors in Opensea mainly use the Ethereum blockchain, and users of AirNFTs mainly use the Binance Blockchain – in other words, they belong to separate technology networks and effectively different communities. While both A and B show abnormal shapes (typically the relationship between two variables that are prevalent across models, when a third dimension is encoded in color, would be more similar to an arc), the discontinuity in B confirms again that price and popularity are disassociated when the groups of users are different, especially in Ohio.

The results suggest that price and popularity can be modeled meaningfully only within the same community (users of the same technology or NFT-issuing blockchain network). It also shows that the prediction accuracy varies with changing levels of activity in different geographical locations, which suggests that NFT investors' behavior is modulated by cyberspace and physical space factors.

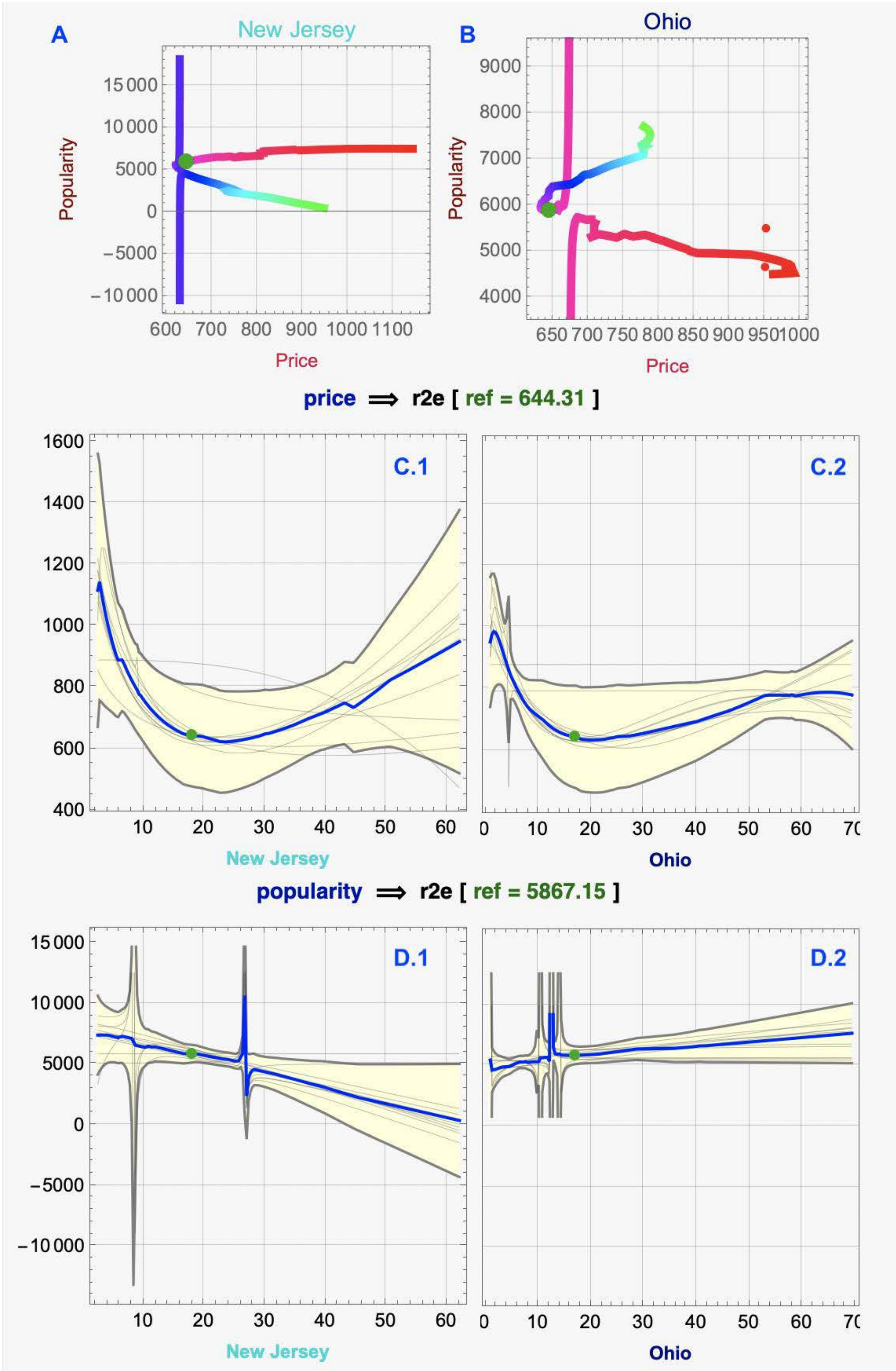

Figure 14. Multi-target response explorer. A) Trade-off curve for variable Ohio; A) Trade-off curve for variable New Jersey; C.1, C.2) Response plots for target: price; D.1, D.2) Response plots for target: popularity.

Figure 15. Audience preferences in Ohio, Opensea.io Dec 2020-Nov 2021.

Figure 16. Audience preferences in New Jersey, Opensea.io Dec 2020-Nov 2021.

Conclusions

We found that NFT market participant behavior differs across different states in the U.S. Market beliefs on price and popularity (expressed as changes in activity in the off-chain marketplace) are indicators of interest and a strong precursor of economic activity.

A definitive classification of motivations among possible extremes, i.e., purely speculative or community-driven, will require the analysis of other facets besides grouping at the regional level. However, we confirmed that expertise resides both in the expert and a social system: users in different states also have different priorities regarding topics of interest related to financial literacy. A future study may cover the relationship between the general popularity of collectives and the degree of specialization of marketplace users in financial topics.

References

Bauer, R., & Smeets, P. (2015). Social identification and investment decisions. *Journal of Economic Behavior & Organization, 117*(C), 121-134, https://EconPapers.repec.org/RePEc:eee:jeborg:v:117:y:2015:i:c:p:121-134.

Ericsson, K. A., Krampe, R. T., & Tesch-Römer, C. (1993). The role of deliberate practice in the acquisition of expert performance. *Psychological Review, 100*, 363-406.

Evolved Analytics LLC. (2021). *DataModeler, Version 9.5*.

Gobet, F. R. (2016). *Understanding expertise: A multidisciplinary approach*. Palgrave.

Nadini, M., Alessandretti, L., Di Giacinto, F., Martino, M., Aiello, L.M., & Baronchelli, A. (2021) Mapping the NFT revolution: market trends, trade networks, and visual features. *Scientific Reports, 11*, 20902. https://doi.org/10.1038/s41598-021-00053-8

Peters, K. (July 08, 2021). *Binance Exchange*. Received from https://www.investopedia.com/terms/b/binance-exchange.asp

Stein, E.W. (1997). A look at expertise from a social perspective. In P.J. Feltovich, K.M. Ford, & R.R. Hoffman (Eds.), *Expertise in Context: Human and Machine* (pp. 181–194). MIT Press.

Venegas, P. (2021). Trustable risk scoring for non-bank gateways in blockchain and DLT financial networks. In D. Braha, M.A.M. de Aguiar, C. Gershenson, A.J. Morales, L. Kaufman, E.N. Naumova, A.A. Minai, & Y. Bar-Yam (Eds.). *Unifying Themes in Complex Systems X* (pp. 165-178). ICCS 2020.

Wolfram Research, Inc. (2021). *Mathematica, Version 12.2.0*