NFT Bedtime Story

Background and Motivation

In a post-pandemic world, online platforms that foster creativity and trade have been more popular than ever before. Non fungible tokens (NFTs) are unique digital assets that are created and traded on blockchain networks, and their value is often subject to uncertainty due to their novelty and lack of established market patterns. By modeling NFTs using probabilistic models, we can better capture the uncertainty and variability in their value and make more informed decisions about their buying, selling, and trading. Moreover, probabilistic models can help identify patterns and trends in NFT markets and provide valuable insights on the consumers and sellers in these marketplaces. This dataset specifically focuses on user adoption for LooksRare and Coinbase. Adoptions are defined in terms of transactions, meaning it could be the purchase or sale of NFTs.

LooksRare and Coinbase are both exchange platforms in the crypto world. While LooksRare solely focuses on NFT exchanges, Coinbase operates in other divisions along with NFTs. LooksRare kicks off the year launching an NFT platform on January 10th, 2022, as an alternative to OpenSea[1]. While still a relatively small platform, as a newcomer, users in this space were curious to see how the company planned to establish itself in the NFT space. Coinbase, an already established player in the crypto space, was looking to pivot to NFTs and launched their platform Coinbase in the week of April 20th, 2022[2].

The story I am trying to tell is that a mix of economic, industry and social influences make up the adoption habits of NFT consumers. More specifically, since we are looking at transactions, I wanted to better gauge metrics that would impact these transactions in different ways. To capture these effects on a macro, meso, and micro level, I have identified covariates to serve as a proxy for these influences that create a good fit for modeling LooksRare and Coinbase adoption.

Model Selection

Individual Distribution

First, establishing the underlying story of the individual level distribution of LooksRare and Coinbase weekly adoption is vital to building the model. Exponential individual level distribution implies a "memoryless" property to the model. This inherently means that individuals purchasing NFTs are not affected by time (ie, not duration dependent). This would not make sense in the context of purchasing NFTs unless they have an equal likelihood of purchasing at any given moment. Customers can be less willing to buy over time if experiencing some kind of market shock, more likely to buy over time if the values appear to be increasing at a promising rate, or have an increasing likelihood of buying to a certain time horizon and then

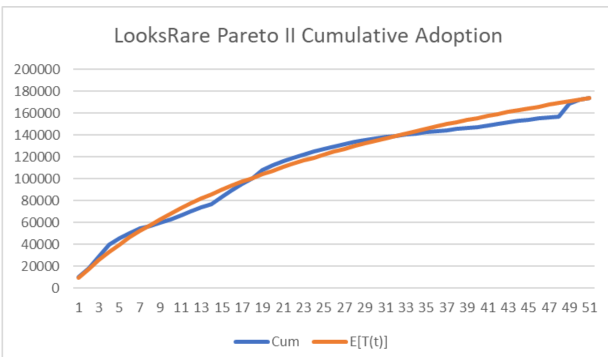
[1] TechCrunch
[2] TechCrunch

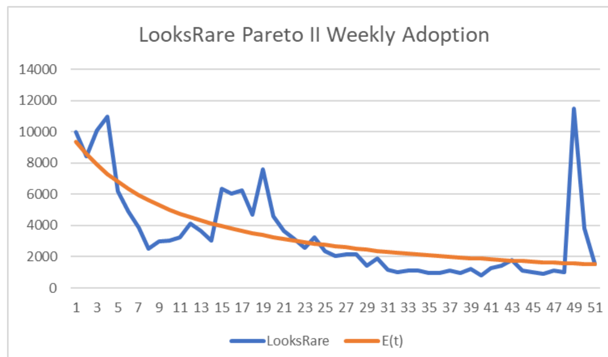
plateauing or dropping off again for potential market circumstances or consumer preferences. In other words, a customer's hazard function of adopting NFTs can be decreasing or increasing depending on their relationship to time.

Pareto II: Exponential Gamma

Motivation: If the consumer population is time agnostic when it comes to NFT purchasing, an exponential gamma would fit our story.

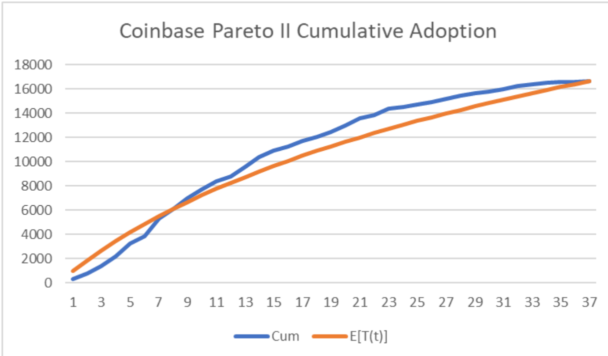
LooksRare Hypothesis: Though LooksRare just launched in January of 2022, it is one of few NFT marketplaces. Since they are very specialized and relatively new, I would expect their r value to be higher, indicating a less heterogeneous population, but still an r value less than 1.

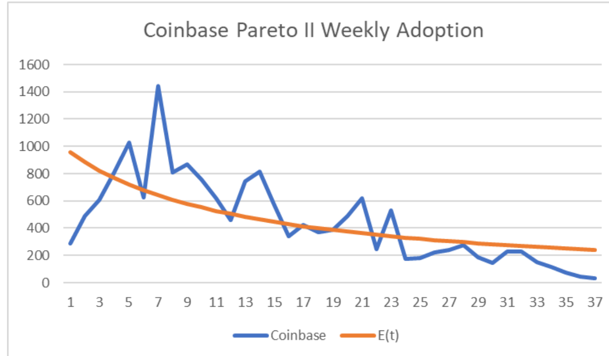




Pareto II LooksRare	
r	0.113
alpha	11.491
LL	-1122850.901
MAPE	49.50%

Coinbase Hypothesis: Since Coinbase is an established player in the crypto space, I would expect their customer base to be more heterogeneous. Coinbase has more offerings than simply an NFT marketplace, and thus would bring in new customers whereas LooksRare, as a leader and specialized space would attract a more specific customer.





Pareto II Coinbase

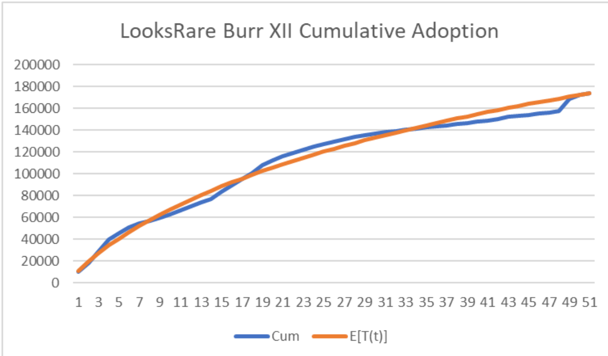
r	0.012
alpha	11.912
LL	-142927.809
MAPE	75.59%

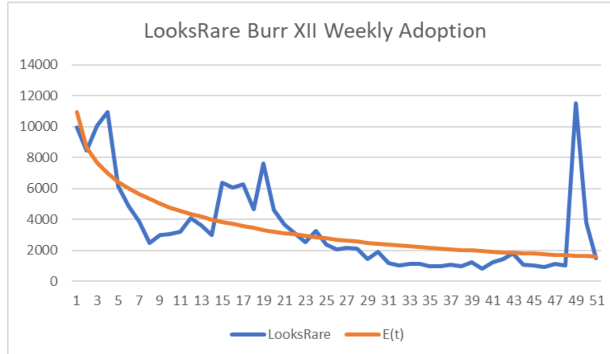
The Pareto II model appears to fit the dataset as heterogeneous for both LooksRare and Coinbase. LooksRare appears to be a little less heterogeneous than Coinbase, which aligns with the original hypothesis. LooksRare The model performs with a MAPE of 49.50% for LooksRare. It is significantly worse for Coinbase at a MAPE of 75.59%. In the incremental tracking plot it follows a similar trajectory of the plot itself, but overestimates in the early week and under estimates for the majority of the data set. In the weekly adoption this becomes even more clear and the plot does not pick up any of the ups and down of the adoption. This indicates that some alterations would be needed to make this model capture more of the nuance of the data.

Burr XII: Duration Dependence

Motivation: Given the original story of duration dependence, that customers would be sensitive to time before making a decision to purchase NFTs, it seems most likely that the Burr XII would best fit the data.

LooksRare Hypothesis: With the volatility of the NFT market, and the relatively new atmosphere, I would expect some positive duration dependence with c being 1< c <2. In terms of the hazard function the more a customer waits, the more likely it would be to adopt. I would expect less duration dependence than an established player such as Coinbase.

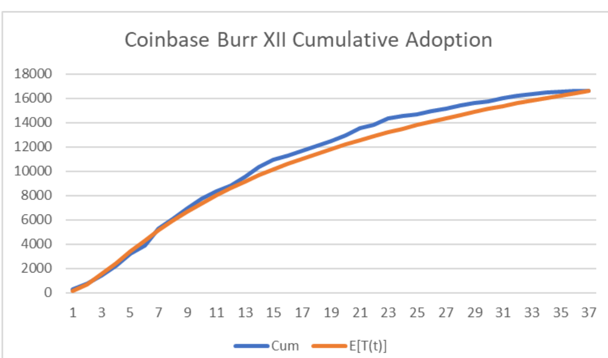


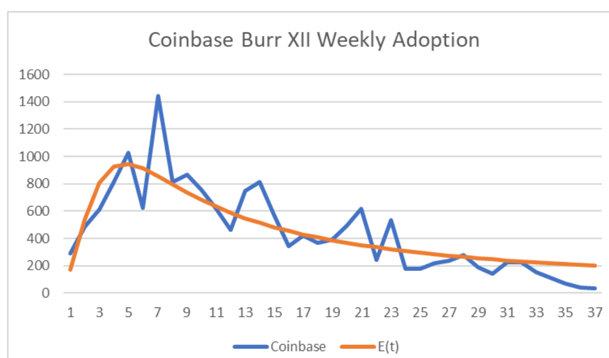


Burr XII LooksRare

r	0.156
alpha	13.752
c	0.890
LL	-1122614.054
MAPE	51.08%

<u>Coinbase Hypothesis:</u> I would expect positive duration dependence, of a c>2 of this kind of platform. As the company itself is pretty well known and is now venturing into a different sector, it would make sense for the likelihood of adoption to increase over time, to compensate for the heterogeneity in adoption over consumers.





Burr XII Coinbase

r	0.003
alpha	19.601
c	2.173

LL	-141974.983
MAPE	54.25%

From this initial analysis, the Burr XII model results confirm the initial assumption of duration dependence. However, it was surprising that the Burr XII model gave a value of c = 0.890, less than 1, indicating negative duration dependence. While this makes sense within the context of the data, where the largest increase hits right after launch of the platform, I would expect an increasing likelihood of adoption as the company continues to develop due to the influence of other covariates. While the consistency of values needs to be done with a more robust model, there is enough evidence to conclude that the Burr XII model is a better overall model for this data set.

Looking at the fit itself, LooksRare does not seem to waver too much in its parameters and error values between Pareto II and Burr XII., with a MAPE of 51.08%. Coinbase appears to be much better fitted through the Burr XII, with a new MAPE of 54.25%. The Burr XII smoothes out the first two peaks as opposed to the gradual decline of the Pareto II on a week by week basis. For the cumulative plot, the LooksRare Burr XII flies right through the middle, while the Coinbase Burr XII seems to underfit after week 11. These can be further adjusted with outside factors that could contribute to the deviation from a pure Weibull function. For these reasons I decided to move forward with the Burr XII as I started to add covariate.

Finding Covariates

For this paper, I decided to include covariates of macroeconomic, industry, and consumer behavior factors. For macroeconomic factors, it is often difficult to have them reflect consumer behavior. Inflationary expectations do just that. Taken from Trading Economics, I disaggregated the monthly rates into weekly metrics[3]. This combines the macroeconomic factor influence of inflation with consumer feelings. While high inflation rate can cause crypto market to fall as the contractionary expectation increases, the expectation of inflation can grow the value of NFTs[4]. Often these factors reflect a self selection bias that makes an industry like this even more volatile. If consumers think there will be a downturn in the economy, they will start to pull out their money from investments, furthering the downturn in the economy.

For industry specific factors, I looked into major Crypto Hack and Crypto Crashes over 2022. There were two main crypto hacks that occurred in March and October. In March, Axie Infinity blockchain gaming platform was the victim of a crypto hack[5]. In October, over US$718 million stolen from decentralized finance sites across 11 different hacks[6]. There were also two main

[3] Trading Economics
[4] Crypto.com
[5] Forkast
[6] Forkast

crypto crashes that occurred in May and November[7]. In May, Terra-LUNA crashed, causing the U.S. peg to waver. November's cryptocrash came from the downfall of FTX where other platforms and crypto currency customers were deeply affected. To account for this, I indicated the hack and crash as throughout the whole month it occurred.

I decided to also include NFT mint and ETH from the given data set. Since the only industry level information I found useful was the addition of hacks and crashes, the mint and ETH data rounds out the variability in the crypto space. This variability influences customer decision making while also being a by-product of said decision making. This made it a viable numeric proxy to my binary indicator variables.

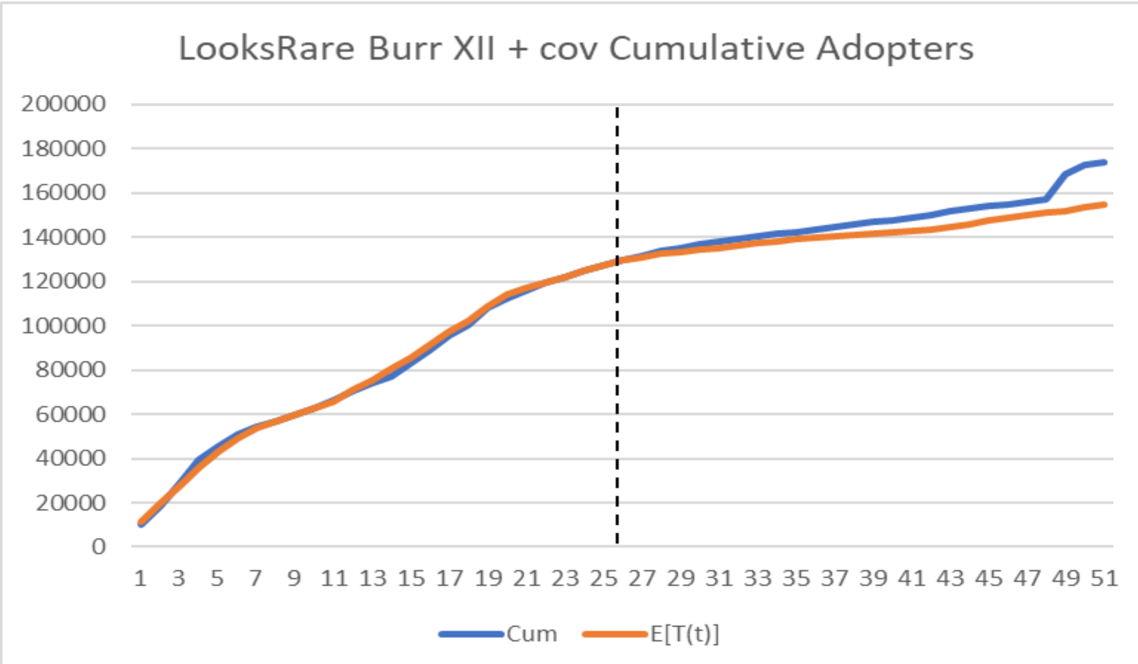
The best proxy I could find for consumer curiosity would be Google trends. I included Google trends of both "LooksRare NFT" and "Coinbase NFT" on a week by week basis to gather when more people were searching about each respective company. This, much like inflationary expectations, is a proactive metric, and can better guide what future decisions might look like.

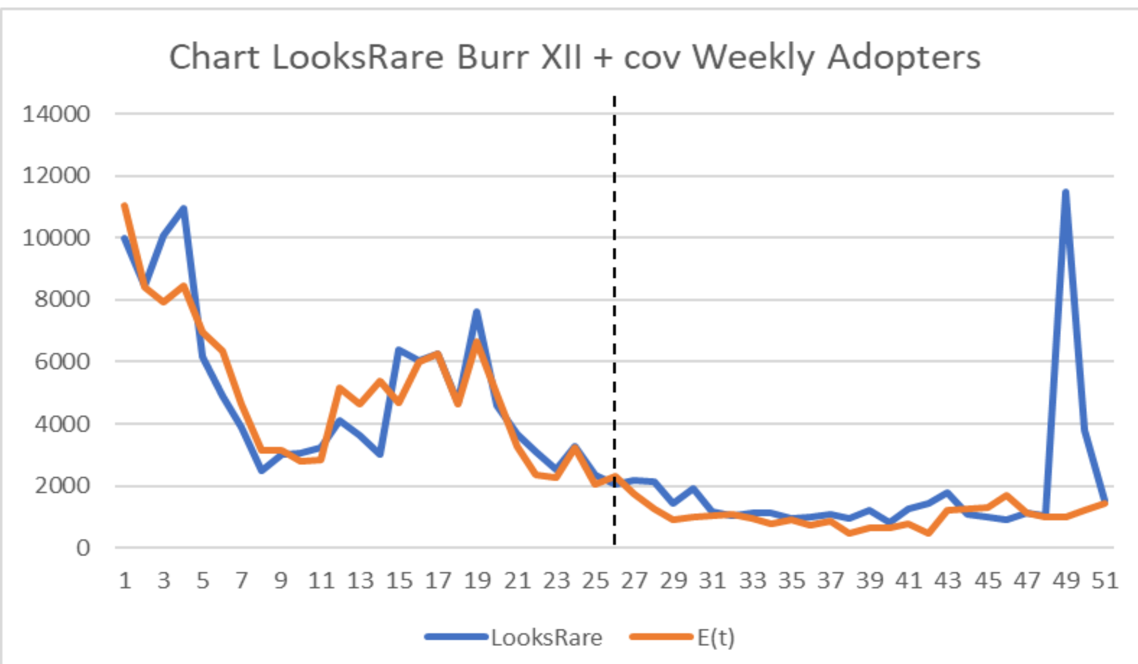
Given the different ranges of these covariates, the best way to compare the effects of these covariates was to standardize them. Aside from the binary variables, I standardized the variables to better compare its impacts on the model.

[7] Forkast

Final Model

<u>LooksRare</u>





r	0.073
alpha	184.232
c	1.065
B_NFT	-0.159
B_ETH	-0.569
B_inf	4.045
B_google	0.247
B_Hack	-0.864
B_Crash	0.515
LL	-794340.807
in sample MAPE	16.22%
out sample MAPE	32.80%
BIC	1588805.95

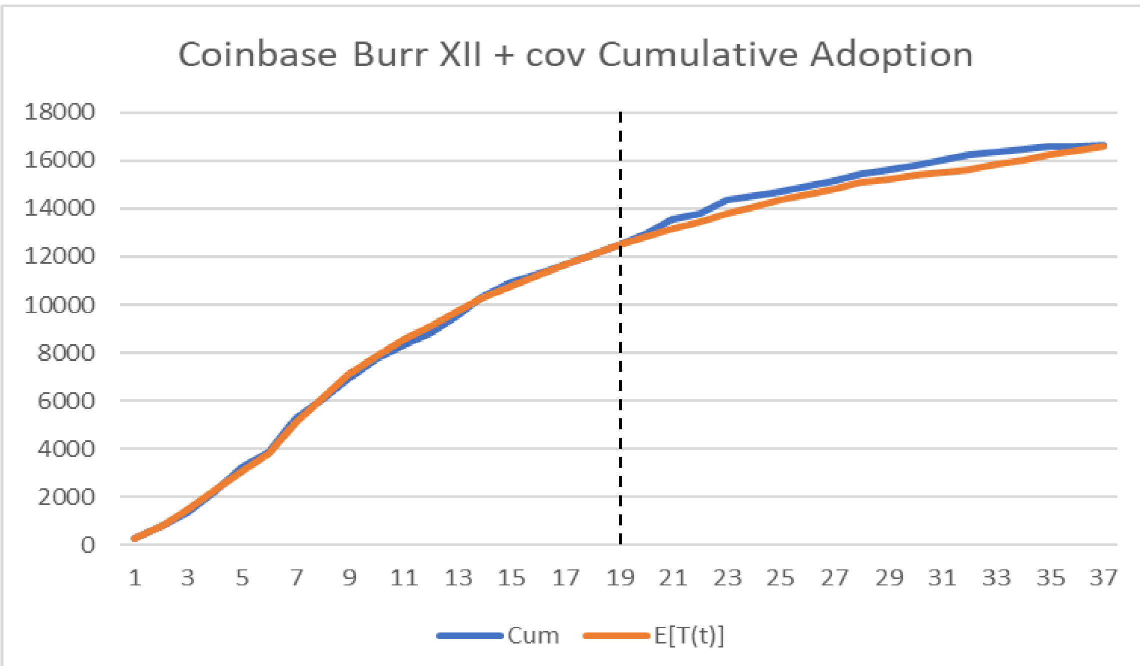
This model has a low r-value of 0.073 and a c-value of 1.065 indicating a heterogeneous population and positive duration dependence. This means that the NFT adoption pool is pretty heterogeneous and that there is some duration dependence for those who adopt LooksRare. The covariates with large influence are Ethereum volume, inflation, hacks and crashes. Ethereal volume and NFT mint would make some intuitive sense given that adoption includes both selling and buying. An increase in transaction volume and mints may increase the number of buys but also impact the number of sales for a particular individual.

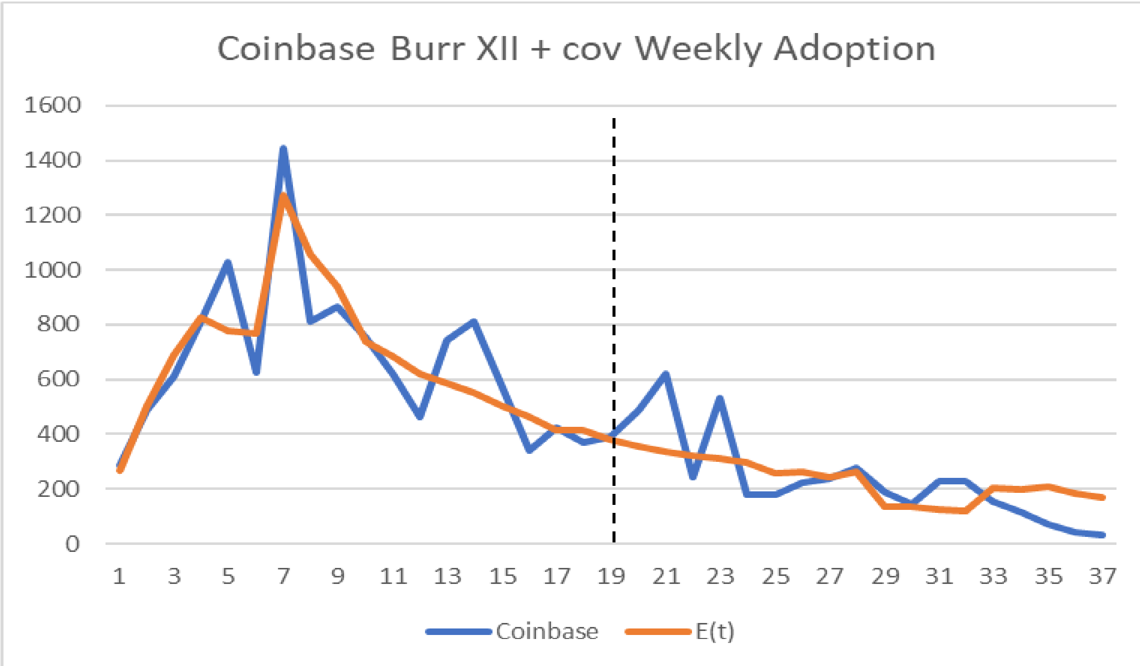
Inflationary expectations and Google trends having a positive beta value also makes some intuitive sense. If we believe inflation will increase, we are more likely to purchase in the short run. Similarly, Google trends can be a good indicator, both on the seller and buyer side of activity that is increasing on the LooksRare platform. More conversation and curiosity of NFTs found on LooksRare website can translate to more sellers or buyers.

What surprised me was the opposing signs of hacks and crashes. I think both covariates are trying to capture some sort of potential to buy or sell and thus have opposite effects on each other. Moreover, the time lag that might be created from a crash or hack is difficult to capture, and this serves as a proxy for when something like that occurs.

To create the in-sample model, I used the first 26 weeks in the in sample and the last 25 weeks in the out of sample. The in-sample MAPE is 16.22% which is low enough to catch some of the peaks and troughs in the first half that are captured visually pretty nicely. The out of sample MAPE is 32.80% which is fairly high, but could potentially be attributed to the larger spike at the end. While I could have tried to capture it with covariates, it would have been an intentional spike that might have led to an overfitting of the data, rather than capturing random from the overall dataset. The BIC is 1588805.95, which was the smallest of the other combinations of covariates I tried, despite a larger number of parameters.

Coinbase





r	0.003
alpha	49.964
c	2.157
B_NFT	0.060
B_ETH	0.124
B_inf	0.025
B_google	0.048
B_Hack	0.000
B_Crash	-0.618
LL	-102926.927
in sample MAPE	14.99%
out sample MAPE	57.55%
BIC	205978.19

New users for Coinbase also have a low r, once again a lower r than LooksRare. This confirms our original understanding of the two platforms, and shows the relative heterogeneity Coinbase has over

LooksRare.The c value is also much higher than LooksRare, with a c value of 2.157 or greater than 2. We can expect as time goes on for adoption to get faster. With the integration of its platform on other marketplaces, this makes intuitive sense.

Given the larger heterogeneity of Coinbase relative to LooksRare, it makes sense that the magnitude of impact for the covariates is much smaller. All of the covariates seem to have a positive beta except for crypto crashes, which have a negative beta, and crypto hacks, which had a magnitude of 0. Given that these covariates had a larger impact on LooksRare, I thought it would be fair to keep them in the model for both. This could potentially be due to the specialization of LooksRare in just the NFT market, Coinbase may have not felt as deep impacts of crypto crashes. Once again we are looking at transactions, and in a potentially volatile environment, uncertainty can cause buyers and sellers to pull out.

I split the dataset in half at Week 19 to perform in sample and out of sample fit metrics. The in-sample MAPE is reliable at 14.99%. However, the out of sample MAPE at 57.55%. At week 21 and 23 are big spikes preceding the inevitable crypto crash in November. The inability to capture this nuance in the in sample, made the fit for the out of sample much worse. Still the general trends allow for a good fit for this model. The BIC is 205978.19, which was the smallest of the other combinations of covariates I tried, despite a larger number of parameters.

Limitations

Model

The anticipated worry with many covariates is always overfitting. I was careful to not include too many variables that were correlated with each other, and that were over emphasizing certain peaks and troughs in the model. Still, it would be interesting to try this model in the future with other covariates and see how it compares to the interactions of these variables

Heterogeneity

I decided against incorporating heterogeneity directly into my models after trying the "hard-core-never-triers" model as well as a latent class model. For the "never-triers" model, the value of pi was either very close to 1 or 0 suggesting very little about the customers itself (See Appendix). Even with lower adoption rates in the first year, it is difficult to interpret this as a large portion of the population refusing to never buy.

Similarly, with the Latent Class 2 segment Weibull with covariates, I learned very little about the differences covariates caused between the two segments from running the model, and pi remained very close to 1. Without latent class, it appeared that my model was able to capture the nuances of adoption over time. What my covariates may lack is representing characteristics of a particular customer group. Given more time, I would retry the latent class model with a different mix of covariates to see if more segments emerge.

Data Limitations

The data is over a short time horizon and thus it is difficult not to overstate the robustness of the model and its covariates. In particular, inflationary expectations was a variable reported on a monthly basis, which I turned into weekly for this analysis. There could be some data nuance lost in the extrapolation of those rates. Without direct access to consumer preferences, it was difficult to create a story around their transaction history. Lastly, having a better representation of the adoption population could have lended better to a hard core never trier model, and also customer segmentation. Adoption for Coinbase specifically is relatively low, but can create misinterpretations of who could become a potential adopter of an NFT platform.

Social Contagion

I had sourced data on what I called "celebrity induced spikes"" (Donald Trump NFT, Elon Musk, etc). I did not include this variable in my model, partially due to the lack of explainability and also to maintain parsimony of variables (and thus reduce overfitting).

Conclusion

Through this analysis, I have understood the importance of covariates in a very volatile market like NFTs. I would advise LooksRare and Coinbase to specifically try to identify proactive measures and indicators that allow them to get ahead of market trends. Specifically inflationary expectations, and other macroeconomic factors that might indicate earlier on that customers are thinking about the changes in the market. Similarly Google Trends of company activity can be another way to preempt any changes in consumer behavior.

Consumer pool is heterogeneous, new adopters trying to segment early on may not prove to be efficient. Consumer demographic data in a heterogeneous market like this may not develop insights that help segment the market. Potentially segmenting on other factors would better enhance the company's understanding of its customers and their adoption of the platforms. While market crashes and hacks are not necessarily predictable, taking security measures to protect platforms can be a good step in avoiding large fluctuations in these periods. Focusing on these factors relative to customer segment characteristics could better tell a story of what kind of adoption exists on these platforms and how to better model them for the future.

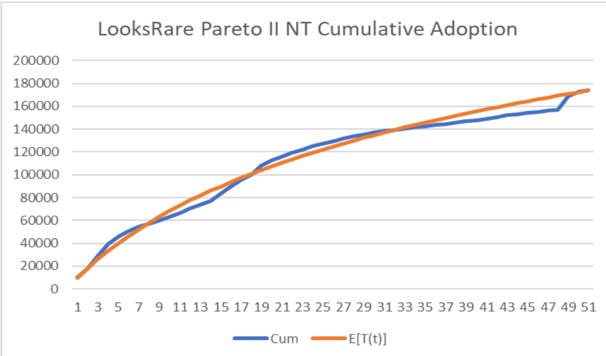
Appendix

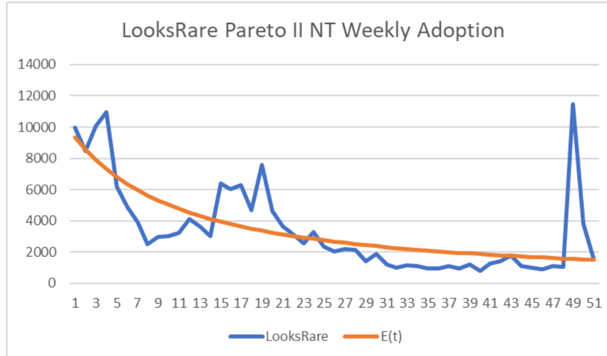
Model Summaries

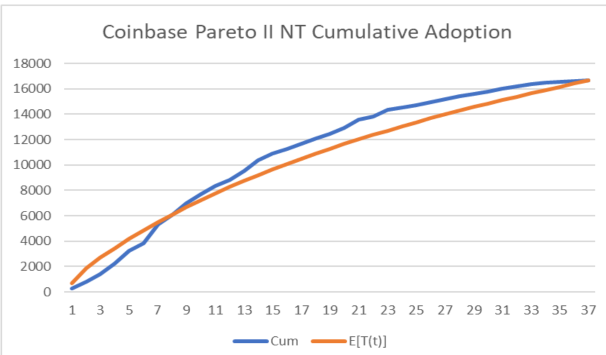
Pareto II LooksRare		Pareto II LooksRare NT		Burr XII LooksRare		Burr XII LooksRare NT		Burr XII COV LooksRare		Burr XII COV LooksRare LC	
r	0.113	r	0.113	r	0.156	r	0.157	r	0.073	lambda1	0.948
alpha	11.491	alpha	11.491	alpha	13.752	alpha	13.803	alpha	184.232	c1	0.948
		pi	1.000	c	0.890	pi	1.000	c	1.065	B_NFT	0.210
LL	-1122850.901	LL	-1122850.903	LL	-1122614.054	c	0.887	B_NFT	-0.159	B_ETH	-0.745
MAPE	49.50%	MAPE	49.50%	MAPE	51.08%	LL	-1122613.323	B_ETH	-0.569	B_inf	2.669
						MAPE	51.28%	B_inf	4.045	B_google	0.508
								B_google	0.247	B_Hack	-0.647
								B_Hack	-0.864	B_Crash	0.188
								B_Crash	0.515	lambda2	1.000
								LL	-794340.807	c2	0.942
								in sample MAPE	16.22%	B_NFT	0.210
								out sample MAPE	32.80%	B_ETH	-0.745
								BIC	1588805.95	B_inf	2.669
										B_google	0.508
										B_Hack	-0.647
										B_Crash	0.188
										pi	1.000
										LL	-794861.3368
										in sample MAPE	19.40%
										out sample MAPE	99.28%
										BIC	1589957.54

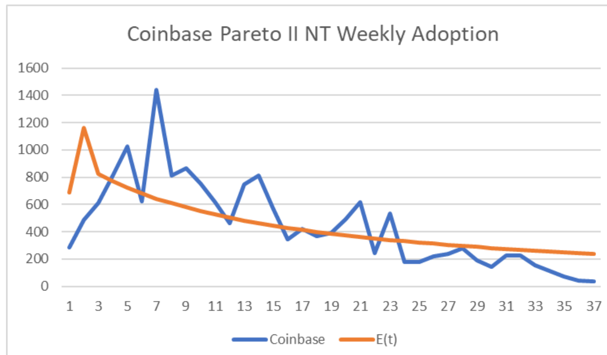
Pareto II Coinbase		Pareto II Coinbase NT		Burr XII Coinbase		Burr XII Coinbase NT		Burr XII COV Coinbase		Burr XII COV Coinbase LC	
r	0.012	r	0.012	r	0.003	r	1.808	r	0.003	lambda1	0.00
alpha	11.912	alpha	11.815	alpha	19.601	alpha	119.716	alpha	49.964	c1	0.00
		pi	0.714	c	2.173	pi	0.018	c	2.157	B_NFT	0.21
LL	-142927.809	LL	-142893.263	LL	-141974.983	c	1.607	B_NFT	0.060	B_ETH	-0.75
MAPE	75.59%	MAPE	74.39%	MAPE	54.25%	LL	-141639.118	B_ETH	0.124	B_inf	2.67
						MAPE	31.53%	B_inf	0.025	B_google	0.51
								B_google	0.048	B_Hack	-0.65
								B_Hack	0.000	B_Crash	0.19
								B_Crash	-0.618	lambda2	1.00
								LL	-102926.927	c2	1.00
								in sample MAPE	14.99%	B_NFT	0.21
								out sample MAPE	57.55%	B_ETH	-0.75
								BIC	205978.19	B_inf	2.67
										B_google	0.51
										B_Hack	-0.65
										B_Crash	0.19
										pi	1.00
										LL	-142052.75
										in sample MAPE	25.40%
										out sample MAPE	54.56%
										BIC	284340.37

Never Triers Models:

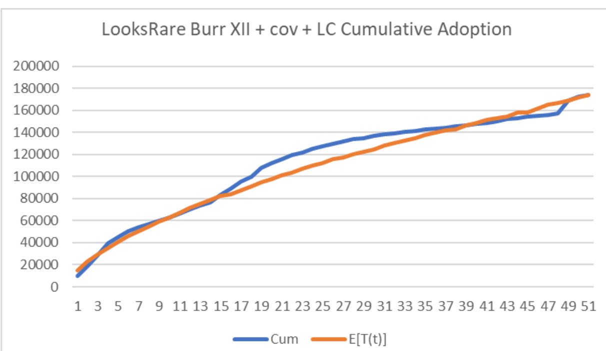


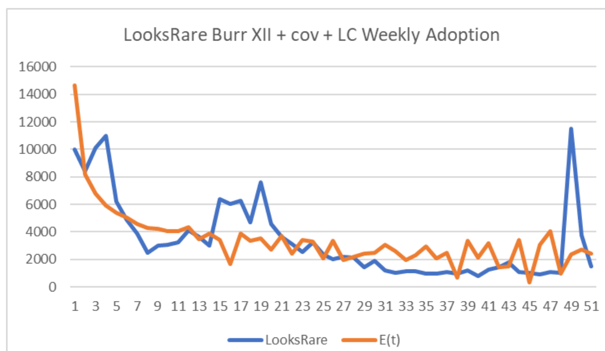






Latent Class 2 segment Weibull:





Burr XII COV LooksRare LC	
lambda1	1.000
c1	0.015
B_NFT	0.210
B_ETH	-0.745
B_inf	2.669

B_google	0.508
B_Hack	-0.647
B_Crash	0.188
lambda2	1.454
c2	1.917
B_NFT	0.210
B_ETH	-0.745
B_inf	2.669
B_google	0.508
B_Hack	-0.647
B_Crash	0.188
pi	1.000
LL	-11505908.842
MAPE	273.39%
BIC	23011872.95