The "Right" Price for Collectible Sneakers: A Predictive Analysis on the Sneaker Resale Market

Roger Wu

Advisor: Professor Eric Auerbach

Northwestern University: Department of Mathematical Methods in the Social Sciences*

June 5, 2020

Abstract

The market of collectible sneakers has grown dramatically with the digitalization of American society. Sneakers that are designed by world-famous celebrities and limited in quantity are highly sought after by fervid collectors, thus have become investment opportunities for many individuals. This paper examines the sneaker resale market based on data available on StockX, one of the biggest sneaker reselling platforms, and aims to identify models that predict the rational reselling price for collectible sneakers from an investor's perspective.

The predictive models are created based on multivariate hedonic regressions that establish a relationship between various characteristics of a pair of sneakers, such as the retail price, the brand, and the release time, and the average resale price. Various models are created to capture different perspectives on the sneaker resale market, and are put into real-life tests to evaluate their respective predictive power. The results are eye-opening, as models that incorporate a more quantitative view of the sneaker resell market and capture a large amount of variation in the average resale price are almost as predictive as models that are more qualitative from an investor's point of view. Moreover, the results indicate that the best investment strategy is to exploit the changes in sneaker value over time, as older sneakers are likely to be more valuable on the resale market.

This paper aims to contribute to the already rich literature on pricing and dynamics of the secondhand market of collectible goods. Even though sneakers as collectibles are relatively new and more popular within the younger population. this research, by deploying a traditional logic of investment and returns, seeks to apply the insights developed from examining collectible sneakers into broader contexts.

1. Introduction

1.1 Collectible sneakers and the culture of "reselling"

Sneakers play a role in everybody's life. We need them to carry out daily activities and to explore the world. When most people talk about sneakers, they are referring to a certain pair of athletic shoes they wear or use for sports. Different from other types of shoes, sneakers are meant for comfort and physical activities, and are more casual than most other types of shoes. In general, sneakers are considered to be a necessity, and are valued for their functions.

However, not all sneakers are created equal. While some are created simply for use and consumption, others are carefully designed, purposefully branded, and enthusiastically marketed as products that are more than sneakers themselves. Collectible sneakers are similar to other collectibles such as artworks and stamps – they are limited in quantity and highly sought after by collectors, who, referred to as "sneakerheads" by the public, value these sneakers for the meaning they carry rather than their functionality. As a result, a secondhand exists for these collectible sneakers, and sneakerheads actively trade with each other on various platforms to obtain the specific pair they want at a fair price.

For example, Nike generally releases collectible sneakers in limited quantities on Saturdays at a set retail price[1]. The retail price is an arbitrary price set by a pair of sneakers' producer or brand, such as Nike and Adidas. Because these sneakers are rare in quantities, not everyone who wants them is able to buy them at retail, so sneakerheads will attempt to buy them at a higher price on the resell market, and this price is called resale price[2]. The range of resale price is wide, as a pair of sneakers that retails for, for example, $220, may be resold for more

---

[1] It is important to note that whether a pair of sneakers is considered as a collectible item is not solely determined by its producer (Nike, for instance). It should be considered more as a result of the market rather than by creation.

[2] One of the biggest differences between retail and resale price is that retail price is set for the first-hand market while the resale price is set by the secondhand market.

than $1,000. The main reason for this phenomenon is that sneakerheads usually attach unique meanings to specific sneakers. For example, a pair of Air Jordan 12 "Flu Game" is significant because it commemorates the game in which Michael Jordan played with a flu and delivered an unbelievable performance, and sneakers enthusiasts as well as loyal Jordan fans are willing to pay a lot to have that specific representation of their memories in their collection. Moreover, in the current age of global digitalization, celebrity culture has reached its peak, which results in tons of people fervidly following their idols. That explains why a pair of Adidas Yeezy's, designed by Kanye West, widely regarded as one of the most iconic figures in the sneaker community, sometimes resale for over $2000 despite retailing at only $200.

As a result, sneakers, just like other collectibles, become investments for both collectors and pure resellers who simply want to make profits from selling. Theoretically, the resale price for a specific pair of sneakers is determined in a free market by all participants. Buyers and sellers need to coordinate in order for transactions to happen, and thus there are market dynamics that we can study to better understand how sneakers can be treated as investments. This research will focus on exploring how sneakers are to be valued as investments to generate returns.

1.2   Goal of research

The goal of this paper is to identify potential investment opportunities in the sneaker resale space from analyzing the existing data on past transactions to generate a predictive model that one can use to estimate the rational resale price for a pair of collectible sneakers. By doing so, an investor and exploit market inefficiency by purchasing sneakers that are listed under their rational resale price and selling them at that price afterwards. Assuming that general characteristics and public reception of sneakers have not drastically changed over time, past

transactions can provide valuable insights on how a specific pair of sneakers is supposed to be valued. Hedonic regression models will be the essence of such testing and the basis for prediction in the analysis.

Data on past transactions are available at StockX, one of the largest platforms on which people resell sneakers. StockX, corresponding to its unique name, treats sneakers as stocks and closely monitors the trend of every pair published on the website. It contains data such as the release date and the retail price for a specific pair of sneakers, the number of transactions occurred, the 52-week high and low of its resale price, and the average resale price across the last 12 months. Moreover, as mentioned above, StockX is also a marketplace for trades, as buyers and sellers coordinate through the website's auction-style listings. Every pair published features a "lowest ask", at which a buyer can immediately purchase the sneakers, and a "highest bid", at which a seller can immediately sell the sneakers. Of course, buyers and sellers can repetitively adjust their ideal price so that transactions occur when the ask matches the bid.

The mathematical model is based on a multivariate model to show what may affect the resale price of collectible sneakers:

$$y = b_0 + b_1 x_1 + b_2 x_2 + \dots + b_{k-1} x_{k-1} + b_k x_k + \varepsilon$$

with y being the average resale price and the x's in the function being features of sneakers that are selected through rounds of model testing. Because the goal of this research is to identify the optimal predictive model that one can use to trade and make profits from collectible sneakers, we want a model that has the strongest predictive power, and will do so by applying the few models created to actual cases. The application will compare the predicted price with the current lowest ask from the buyer's perspective to determine whether investment opportunities exist for a specific pair of sneakers and if yes, how much the theoretical profit would be.

However, before we dive deeper into the specifics of data collection and analysis, it would be helpful to examine the existing literature on collectible goods, as they provide valuable insights regarding what is to be accomplished from this research.

2. Literature review

2.1   Relevant literature on collectible goods

Most of the existing research on the pricing of collectibles employ multivariate regression models that are similar with the one proposed above. Moreover, the focus of most of the literature is on the pricing in the secondhand market, which is in correspondence with my interest in determining a proper resale price for collectible sneakers.

Scorcu and Zenola (2011) study the dynamics of collectible artworks in the auction-style secondhand market. The research aims to show which characteristics affect the auction price of art collectibles, and uses a dataset of past auctions of Picasso paintings. Final auction price is regressed on variables such as dimension, style, media, salerooms, and year of sale to set up a model that represents the predicted price of artworks in auctions. Qualitative characteristics of artworks are translated into dummy variables to be incorporated. Moreover, the regression model is separated into quantiles to show the effect of variables on different ends of the market. The research found the existence of significant differences both in the way prices respond to characteristics (particularly in the media variables) and in the rates of return from an investment in Picasso paintings across different price ranges. In general, this paper resembles the most with my research on the collectible sneakers market as it features a similar model as well as particular interests in treating collectibles as investments.

Dickie, Delorme, Jr. and Humphreys (1994) also study the proper pricing in the collectibles industry. Different from the previous paper, this research takes interest in studying the price of collectible coins, which is particularly popular in the United States. In the regression, independent variables include denomination, vintage, type, the mint, the value of the silver and the condition of the coins, while those that are qualitative are represented through dummy variables. Result showed that prices of rare coins are predictably related to characteristics of coins valued by collectors, consistent with the attainment of equilibrium prices for coins.

Slightly different from the previous two groups of researchers, Ashenfelter (2008) incorporates more real-world applications into his research. The aim of Ashenfelter's work is to study the effect of weather on the quality and price of Bordeaux wines. He first employs background research to show that the vintage and the vineyard are two natural dimensions for consideration directly related to the quality and price of wines, and then analyzes how vintage is affected by weather variability and how vineyard is affected by climate variability so that weather and climate are translated into reliable independent variables. The paper features a hedonic regression on data from the 1960s and shows that weather is a significant determinant of wine quality and price, and that wines that are produced during hot, dry summers are of the highest quality, thus the highest price as well. The predictive model is then used to match wine prices for 1980s and 1990s and fits well with the actual sale records. As a result, the model was later employed by many other researchers to predict wine quality. Besides its powerful insights on how to establish a predictive model in the collectibles industry, this paper also features a discussion on the real-world applications from its results. For example, Ashenfelter talks about the indicated market inefficiency in that wines that are priced unfairly high in early years, as well as related findings to how climate change and global warming can affect the wine industry.

2.2 Relatedness of research to existing literature

The previous researches are helpful in shaping my own, as their methods all focus on the collectibles industry and their results are indicative of what hypothesis are appropriate from my investigation of the collectible sneakers market. The hedonic regression model is used in all of the studies, with the common goal to determine the theoretical price of certain collectible items. While the first two papers demonstrate a good practice of how to utilize hedonic regression to generate insights into important characteristics of collectible items that are valued, the third paper by Ashenfelter promptly features a predictive analysis of the proper price for wines. The method of evaluating the effectiveness of a particular predictive model has provided me with helpful instructions on how to design my research.

Conversely, my research contributes to the current literature regarding pricing of collectible items. My study is to consider sneakers as another form of collectibles different from traditional ones such as artworks, coins and wine, so that our study of the collectibles market can be expanded to meet real-world applications. Moreover, sneaker resale features a direct online bidding model that is different from traditional auctions as well. My research evaluates sneakers as a potential investment to show the market's financial implications, and aims to generate a larger discussion on how collectibles are valued.

3. Data collection

3.1 Web scraping from StockX

StockX contains a rich database on transactions of collectible sneakers. There are thousands of collectible sneakers published, and market data from various perspectives are available for almost all of them. Important data include release date, retail price, average resale

price, price range, etc. of a specific pair of sneakers, and the trend of transaction prices is pictured as a stock graph for every pair that is available on the website. For my research, I created a web scrawler to collect data for a diverse range of collectible sneakers.

In order to select a sample that is as random as possible, I used the "featured" page on StockX. Even though there are certainly underlying reasons why StockX sorts sneakers in such a way[3], the randomness of sneakers are preserved through the various brands they are from, as well as a diverse range of popularity across them, with some generating over 20,000 sales while others are transacted for fewer than 10 times. The first 25 pages of the "featured" sneakers are scraped to obtain the web link to each pair of sneakers. Then, every link obtained from the previous process is used to extract the specific data needed for analysis.

3.2   Data characteristics

The scraping obtained links to 975 pairs of sneakers from StockX, and these links provided more than 10,000 data entries. Specifically, data scraped are on important characteristics for each pair of sneakers such as its release date, retail price, brand, model, colorway, as well as useful resale price records such as 52-week high and low, volatility, average price in the last 12 months, etc..

Some characteristics are qualitative while others are quantitative. For example, brand is a qualitative variable, and all the sneakers are categorized by StockX as from either Nike, Adidas, Air Jordan or other brands. Therefore, these three brands are going to be the focus of analysis. The release dates of sneakers scraped range from 2020 to as far as 1985 to represent a wide spectrum of sneakers to be incorporated into the analysis.

---

[3] For example, those that are popular and so may generate more sales are pushed to the top of the list.

Most qualitative characteristics are transformed into dummy variables so that they can be properly incorporated into the regression model. Brands are categorized by assigning an 1 if a pair of sneakers is of a certain brand, and a 0 otherwise. Release dates are sorted into 6 groups: 2020, 2019, 2018, 2017, 2016, and before 2016, and 1 and 0 are assigned respectively for all sneakers.[4] A new dummy variable, collaboration, is added to represent whether a pair of sneakers is considered a collaboration between its producer, such as Nike, and another brand or a celebrity, because it is hypothesized that such collaborations are significant in raising a pair of sneakers' resale value.[5] Moreover, dummy variables on the interaction between collaboration and brands are also created to be used in regression models. Below is a table of all the variables collected and created, while Table 1 to 3 are descriptive statistics on the data scraped from StockX to establish models.

Variables collected							
Quantitave	Retail price	52-week high 52-week low	12-month high 12-month low	Volatility	Volume	Price premium	Average resale price
Qualitative	Sneaker name	Brand	Model	Picture	Release date		
Variables created							
Dummy	Collaboration	Adidas Air Jordan Nike Other brands	2020 release 2019 release 2018 release 2017 release 2016 release Before 2016 release				

---

[4] 2016 is regarded as a proper threshold because sneakers that are released more than 4 years ago are generally considered to be old releases in conversations among sneakerheads. Moreover, it is interesting to note that StockX features more recent releases than older releases.

[5] Specifically, only sneakers that are collaborated with other brands or influencers outside of sportswear companies are regarded as collaborations. For example, sneakers designed by both Nike and famous hip-hop artist, Travis Scott, are considered collaboration works. Sneakers that feature athletes and designers within sportswear companies are not considered as collaborations. For example, a pair of Kobe Bryant sneakers released by Nike is not considered a collaboration because Bryant is signed to Nike.

Table 1: Descriptive statistics on numeric variables

Variable	obs	mean	std.dev	min	max
Average sale price	967	494.2968	1486.222	28000	56
Retail price	930	177.5419	91.85556	1195	60
52 week high price	967	878.7849	1878.398	28000	105
52 week low price	947	325.0338	1231.412	28000	25
Volatility	958	0.1683925	0.1363355	2.57	0.01
Number of transactions	967	2826.741	5936.861	87651	1
Price premium	914	1.308162	4.480896	66.5	-0.71

Table 2: Descriptive statistics on dummy variables

Variables	Value	Frequency	Percentage
Collaboration	0	580	59.49
	1	395	40.51
	Total	975	100
Brands	Air Jordan	309	31.69
	Adidas	183	18.77
	Nike	402	41.23
	Other brands	81	8.31
	Total	975	100
Release date	2020	60	6.15
	2019	479	49.13
	2018	153	15.69
	2017	98	10.05
	2016	60	6.15
	Before 2016	125	12.82
	Total	975	100

Table 3: Descriptive statistics on interaction between collaboration and brand

	Air Jordan		Adidas		Nike		Other brands	
Collaboration	0	1	0	1	0	1	0	1
0	318	262	521	59	345	235	556	24
1	348	47	271	124	228	167	338	57
Total	666	309	792	183	573	402	894	81

4. Analysis and results

4.1   Model creation and explanation

The research is based on a multivariate regression model:

$$y = b_0 + b_1x_1 + b_2x_2 + \ldots + b_{k-1}x_{k-1} + b_kx_k + \varepsilon$$

with the dependent variable being the average resale price of sneakers and the independent variables being various characteristics that are considered relevant. There are many different combinations of these characteristics as independent variables. In order to select the most relevant ones to be tested as potential investment strategies, it's beneficial to lay out a few hypotheses that are based on specific knowledge of the sneaker resale market so that the ideal models incorporate only the relevant variables.

For example, collaboration is often considered as an important determinant of a pair of sneakers' value in the sneaker community, and so is the age of the sneakers, as older sneakers tend to be valued more than recent releases. Moreover, rarity is believed to be another key determinant of the resale value of sneakers, and can be represented by the number of transactions[6]. Various quantitative variables, such as the retail price of the sneakers and the 52-week high and low, are believed to be directly associated with the resale value as well. We want to create both quantitative and qualitative models so that both investors who do not know much specifics about sneakers and those who are sneaker experts can find ways to invest. On the other hand, variables such as color are left out because they are not believed to be related to the resale value of sneakers[7], and are not significant while tested by running regressions on them.

---

[6] It is assumed that if a pair of sneakers is transacted more often than another pair, its quantity is less limited. Such an assumption is generally true based on the fact that the sneaker resale market is free and open.

[7] It is also true that several combinations of colors tend to be popular when it comes to sneakers. For example, the combination of black and red on Air Jordan's seems to be a fan-favorite. However, compared with other determinants such as collaboration, the color of sneakers are not considered significant by most sneakerheads, as we have also seen many sneakers with popular colors lose value on the resale market.

Therefore, the following three models are believed to be relevant:

$$Average\ resale\ price = b_0 + b_1*52\text{-}week\ high + b_2*52\text{-}week\ low + \varepsilon \qquad (1)$$

$Average\ resale\ price = b_0 + b_1*collaboration + b_2*retail\ price + b_3*2020\ release + b_4*2019$

$release + b_5*2018\ release + b_6*2017\ release + b_7*2016\ release + b_8*number\ of\ transactions +$

$\varepsilon \qquad (2)$

$Average\ resale\ price = b_0 + b_1*collaboration + b_2*retail\ price + b_3*2020\ release + b_4*2019$

$release + b_5*2018\ release + b_6*2017\ release + b_7*2016\ release + b_8*number\ of\ transactions +$

$b_9*Air\ Jordan + b_{10}*Adidas + b_{11}*Nike + b_{12}*(Air\ Jordan*collaboration) +$

$b_{13}*(Adidas*collaboration) + b_{14}*(Nike*collaboration) + \varepsilon \qquad (3)$

with the first model featuring a more quantitative perspective while the second and the third model exploring the relationship between resale value and more qualitative characteristics. Moreover, the third model is an improvement of the second model, as it tries to take brands into consideration. When brands, as a variable, is regressed in a simple regression with the resale price, it is not significant, but exploring the interaction between collaboration and brands may generate possible insights. In general, the first model should be attractive to investors who are unfamiliar with specific knowledge of the sneaker industry, while the second and the third model will help sneaker enthusiasts leverage their expertise in the industry.

4.2   Regression results

Table 4 details the regression results of the previous three models.

Table 4: Regression results

	Average sale price		
52 week high	0.394***		
	(-0.00654)		
52 week low	0.636***		
	(-0.0101)		
Collaboration		240.7***	62.75
		(-43.47)	(-168.6)
Retail price		1.651***	1.836***
		(-0.232)	(-0.232)
2020 release		-797.0***	-807.4***
		(-104.5)	(-104.1)
2019 release		-831.6***	-851.4***
		(-69.96)	(-70.11)
2018 release		-792.7***	-806.9***
		(-80.65)	(-79.96)
2017 release		-597.3***	-599.9***
		(-89.49)	(-90)
2016 release		-575.5***	-492.6***
		(-105.1)	(-105.2)
Number of transactions		-0.000807	0.002
		(-0.00354)	(-0.00366)
Air Jordan			117.8
			(-147.4)
Adidas			85.73
			(-163.9)
Nike			167.6
			(-148.5)
Adidas & Collaboration			56.83
			(-196.4)
Air Jordan & Collaboration			215.5
			(-197.2)
Nike & Collaboration			316.6
			(-180.5)
Constant	-56.16***	683.5***	526.5***
	(-6.261)	(-75.48)	(-157.9)
N	947	924	924
R-sq	0.988	0.236	0.263

Compared with the second and the third model, the first model returns a way higher R^2, thus explaining more variation in the average resale price. An R^2 near 99% shows that the average resale price of a pair of sneakers can almost be completely explained by its 52-week high and low, which should be expected. The coefficient is larger for 52-week low than 52-week high, showing that the sneakers' resale value is generally more spurious on the higher end and closer to the lower end. A high R^2, of course, is significant in determining an investor's strategy because it confirms the validity of the independent variables. An investor can use 52-week high and low as reliable approximators of a pair of sneakers' resale value without worrying about neglecting other characteristics that may also explain this value.

The second and third model, even though returning a lower R^2, are valuable because they answer many of the hypotheses mentioned above. Both model 2 and 3 confirm the significance of collaboration as a determinant of the sneakers' resale value, while model 3 further develops that a collaboration project between Nike and another brand or celebrity is the most valuable[8]. It follows our expectation that if a pair of sneakers is a collaboration project, its resale value tend to be higher than that of those that are not collaborated with other brands or celebrities. However, the difference in coefficients among the three brands in model 3, even though not significant, raises doubts on why the marginal increase in a pair of sneakers' value is larger for Nike than Air Jordan and Adidas, which contradicts our original expectation that these three brands have almost equal effects in contributing to their sneakers' resale value. Nevertheless, an investor who is informed about the sneaker market can still leverage his specific knowledge on whether a pair of sneakers is a collaboration and which brand it comes from to estimate its fair resale value.

---

[8] This is generally true based on the recent trends in the sneaker resale market, as collaborations between Nike and celebrities such as Travis Scott (famous hip-hop artist) and Fear of God (popular high-end streetwear brand) are the most popular sneakers on the market.

Moreover, model 2 and 3 confirm the hypothesis that older sneakers tend to have higher resale values, as the coefficients for older releases are generally larger than those for more recent ones. Even though the coefficient for 2020 release is higher than that for 2019 release, it can be explained by that (1) this dataset is scraped at the beginning of 2020 and so the sample representing 2020 releases is partial, and that (2) sneakers generally go up in value immediately after their release but lose value after a couple months, and eventually regain some of their value after years.

Model 2 and 3 also reject the hypothesis that a pair of sneakers' resale value is negatively related with its number of transactions in the past, as the relationship is almost negligible in regression results. However, based on the assumption that the sneaker resale market is free and demand-driven, a pair's rarity should affect its resale value. A possible explanation of the above result is that the number of transactions is not a good representation of the sneakers' rarity. Instead of focusing on the amount of transactions occurred, we may need to gather more information regarding how many pairs are released at retail for each pair of sneakers, which may be unavailable for the general public.

4.3   Testing of predictive power

The goal of this research is to test the predictive power of the relevant models so that they can present valuable strategic insights for investors who aim to earn profits from the resale market. Therefore, the above models are going to be fitted with another set of data scraped from

StockX[9] to calculate the theoretically "proper" resale value for sneakers in the dataset[10]. A rational investor will handle his investment in the following way: if the theoretical resale value for a pair of sneakers fitted from the models is higher than its lowest ask, the investor is going to invest and earn the theoretical profit of the difference between the two[11]. Otherwise, the investor is not going to invest.

The above strategy is applied to all the sneakers in the dataset, and a decision of whether to invest is assigned to each pair. For those that are considered an "invest", the difference between the expected resale value and the lowest ask is calculated. The theoretical profit is then the sum of all such differences. The process is repeated for all three models, and the results are represented in Table 5 below.

Table 5: Hypothetical profit based on models

Variable	obs	mean	std.dev	min	max	value
Profit from Model 1	139	141.4796	180.4855	0.8723412	991.5629	19665.664
Profit from Model 2	116	176.8088	128.4272	2.294258	738.2519	20509.821
Profit from Model 3	119	167.7658	139.9173	1.423828	735.871	19964.13

The result is insightful from various perspectives. First, the theoretical profits from the three models are almost the same, showing that these models, despite their considerable difference in explaining the variation in the average resale price, have similar predictive powers from an investor's point of view. Therefore, an investor interested in the sneaker resale market can make profits by exploring his opportunities either quantitatively or qualitatively. It was

---

[9] To be more specific, the data collected from the first round of scraping and used for setting up the models are based on StockX as of January 25th, and the data collected from the second round of scraping and used to fitting the models are based on StockX as of April 30th. It is reasonable to assume that the sneaker resale market has not drastically changed during the time interval. However, it is also true that the estimation for the 2020 releases may be a little bit off to be applied to the second dataset based on the first dataset.

[10] The descriptive statistics for the dataset from the second round of scraping are included in the Appendix.

[11] This profit is theoretical because we did not take some of the realities, such as StockX's transaction fee, into consideration. This theoretical amount is believed to be larger than the actual possible profit, but should not be a problem as our goal is to compare models.

hypothesized that Model 1, boasting a near 99% R^2, should generate more accurate approximations of a pair of sneakers' resale value; however, the above result shows that it is less powerful than Model 2 and 3 in helping the investor earn profits. Among the three models, Model 2 turns out to be the most predictive one, having the highest theoretical profit, the highest mean profit, and the lowest standard deviation. The success of Model 2, more importantly, indicates that the best investment strategy for investors interested in the collectible sneakers market is to exploit the change of sneaker value over time, as the age of sneakers is the focus of Model 2. Because sneakers that are older tend to have higher resale values, an investor can possibly exploit this by buying a pair of sneakers immediately after its release, holding on to it, and eventually selling it at a higher price months or even years later.

Model 1 presents an easy and straightforward strategy for investors by advising them to make investment decisions based on historical data, and it may be especially attractive to investors who do not know much about the sneaker resale market and simply want to make profits out of it. Its failure in surpassing more qualitative models such as Model 2 and 3 reflects the usefulness of specialized knowledge in the sneakers industry. Those investor who are interested in and know about various characteristics of sneakers are able to create their own investment strategies based on such knowledge without having to collect historic data on sneakers in order to make profits. It should be encouraging for sneakerheads to realize that knowing whether a pair of sneakers is a collaboration and when it was released can contribute as much as, or even more than researching about specific quantitative data in shaping an effective investment strategy.

Moreover, Model 3, regarded as an improvement of Model 2, does not perform better than Model 2 in generating theoretical profit for investors even though it takes more

characteristics into consideration. There have always been debates on whether the sneakers'
brand is directly related to their resale value, as brands such as Nike, Adidas and Air Jordan are
similar in popularity among sneakerheads. It is also true that a collaboration sneaker from Nike is
generally more valuable than a non-collaboration sneaker from Adidas, while a collaboration
sneaker from Adidas is also generally more valuable than a non-collaboration from Nike.
Therefore, the interaction between collaboration and brands may not be more effective in
explaining a pair of sneaker's resale value than collaboration itself. The irrelevance of brand is a
possible explanation for why Model 3 is less satisfying than Model 2 according to the above
result. As mentioned above, Model 2's success signifies the importance of age in determining a
pair of sneakers' fair resale value, which should inform investors to take advantage of time
change in regards to their investments.

There is also an important flaw we can observe from the above result. The max values for
all three models seems too high based on my knowledge of the sneaker resale market. Even
though it is not uncommon to see sneakerheads earn over $1000 from selling a pair of sneakers
they bought at *retail*, because the investor's strategy is to buy sneakers at a certain *resale* price
and then resell at a higher price, such situations are very unlikely to occur because if there is
actually so much profit to be made, why would other resellers sell at a much lower price?
Therefore, such an observation indicates that all of the three models may have overestimated the
resale value for at least a number of sneakers in the sample.

4.4   Drawbacks and further opportunities

There are, of course, various drawbacks regarding the above research design, as many
important characteristics within the sneaker resale market are still unaccounted for. However,

these drawbacks also provide opportunities for further research to keep exploring the collectible

sneakers market, which, similar to the collectible industry as a whole, tends to be relatively

stable so as to provide rich investment opportunities.

Sizing of the sneakers is another key characteristic that is expected to affect a pair of

sneakers' resale value, as brands such as Nike and Adidas release sneakers in different quantities

across sizes[12]. For example, it is often observed that collectible sneakers in girl's sizes generally

have a higher resale value than those in bigger sizes. However, StockX does not provide

complete data on the average resale price of sneakers in a specific size, and so we did not take

sizing into the above analysis, which can lead to several drawbacks. First, the models we have

created, especially the qualitative ones, are less ideal in explaining the variance in resale value

than the imaginary models that take sizing into consideration. Second, the investor's strategy,

along with the calculation of hypothetical profit, is flawed in the sense that the lowest ask for a

pair of sneakers is for a specific size, and the resale value of the pair in that specific size may be

higher or lower than that of the pair across all sizes. Such drawbacks are unfortunate and are

awaiting further research to potentially address and to provide new insights.

Moreover, because Model 2's success above indicates the importance of age in

determining a pair of sneakers' fair resale value and the possible exploitation of time as an

investment strategy, further research can focus on locating the time threshold, at which a pair of

sneakers is likely to start appreciating in value. As we have observed, sneakers often lose value

on the resale market immediately after their release, but slowly regain their value and even start

to appreciate after a period of time. If we can locate a timespan during which a pair of sneakers is

---

[12] Besides rarity, other factors, such as unbalanced demand, also make sizing a key determinant of the resale value of sneakers.

likely to lose value or the turning point at which its value starts to rise, we can create an investment strategy by investing at that turning point to maximize returns on time.

Besides sizing and specifics on age, there are also other improvements that can be made on the research. For example, even though we have found the number of transactions of a pair of sneakers to be insignificant in determining its resale value, we still believe that rarity plays a key role in the sneaker resale market. It would be beneficial to identify a good proxy for rarity, or use the data on the number of transactions in a different way to reflect rarity. Moreover, we also believe that the sneaker resale market is highly affected by key events around the world. For example, the passing of Kobe Bryant earlier this year have dramatically raised the resale value of Bryant's signature sneakers from Nike, and the recently released documentary of Michael Jordan, *The Last Dance*, have also helped Air Jordan sneakers gain more exposure and popularity. Further researches can incorporate variables that represent such events to evaluate how they play a role in determining a pair of sneakers' resale value.

5. Conclusion

While the sneaker resale market, just like any other secondhand collectible resale market, presents investment opportunities for investors interested in the products, this paper identifies various models that an investor can use to evaluate a pair of sneakers' resale value based on a hedonic multivariate regression design. The resale value for sneakers is affected by a number of characteristics, and knowledge of the sneaker market, as well as individual sneakers themselves, can provide investors with insights into the true value of their investments.

An investor can approach such evaluations either quantitatively or qualitatively. For someone who does not have a lot knowledge of sneakers but still aims to make profits by

investing in the sneaker resale market, a quantitative model featuring historic data on 52-week high and low prices is easier to be applied. The model accounts for a large amount of variance in the dependent variable, but does not generate the optimal amount of profits. On the other hand, a sneaker enthusiastic who knows a consideration amount of the sneaker industry can leverage their knowledge by following models that use key qualitative characteristics such as a pair of sneakers' release time, collaboration status, and rarity to approximate a pair's resale value. An informed investor can also exploit change in sneaker value over time to make profits, as sneakers that are older tend to be more valued on the resale market. These models, even though explains only a part of the variance in the average resale price, are still reliable in generating positive returns. Despite several drawbacks in the design, the research provides results that confirm the attractiveness of the sneaker resale market as an investment opportunity, and also insights into how the collectible industry as a whole is evolving while staying stable and consistent for most investors.

I hope this research serves as an inspiration for future efforts to engage in not only the sneaker resale market, but also the new collectible industry, as the definition of collectible items keeps evolving with rapid modernization and digitalization. While stamps and coins are still valued by their respective collectors, consumer products such as sneakers and streetwear items and accessories from highly sought-after brands are presenting both new investment opportunities and interesting insights into how the market is everchanging. Research into these markets will go a long way in explaining market dynamics and consumer behavior.

Reference

Ashenfelter, O. (2010). Predicting the Quality and Prices of Bordeaux Wine. *Journal of Wine Economics*, *5*(1), 40–52. doi: 10.1017/s193143610000136x

Bapna, R., Jank, W., & Shmueli, G. (2004). Price Formation and its Dynamics in Online Auctions. *SSRN Electronic Journal*. doi: 10.2139/ssrn.902887

Boland, M., & Thornton, D. B. (2014). Pricing in Auction Markets for Collectibles: Theory and Experimental Evidence. *SSRN Electronic Journal*. doi: 10.2139/ssrn.2506350

Burton, B. J., & Jacobsen, J. P. (1999). Measuring Returns on Investments in Collectibles. *Journal of Economic Perspectives*, *13*(4), 193–212. doi: 10.1257/jep.13.4.193

Dickie, M., Delorme, C. D., & Humphreys, J. M. (1994). Price Determination for a Collectible Good: The Case of Rare U. S. Coins. *Southern Economic Journal*, *61*(1), 40. doi: 10.2307/1060128

Kauffman, R. J., & Wood, C. A. (2007). Market depth and efficiency in online auctions. *Proceedings of the Ninth International Conference on Electronic Commerce - ICEC 07*. doi: 10.1145/1282100.1282154

Lynn, M., & Bogert, P. (1996). The Effect of Scarcity on Anticipated Price Appreciation. *Journal of Applied Social Psychology*, *26*(22), 1978–1984. doi: 10.1111/j.1559-1816.1996.tb01783.x

Perrigne, I., & Vuong, Q. (1999). Structural Econometrics of First-price Auctions: A Survey of Methods. *Canadian Journal of Agricultural Economics/Revue Canadienne d&Apos;Agroeconomie*, *47*(3), 203–223. doi: 10.1111/j.1744-7976.1999.tb00222.x

Scorcu, A. E., & Zanola, R. (2011). The "Right" Price for Art Collectibles:A Quantile Hedonic Regression Investigation of Picasso Paintings. *The Journal of Alternative Investments*, *14*(2), 89–99. doi: 10.3905/jai.2011.14.2.089

Appendix

Table 6: Descriptive statistics on numeric variables

Variable	obs	mean	std.dev	min	max
Average sale price	198	308.6313	192.3228	85	1335
Retail price	197	186.9645	50.44985	90	395
52 week high price	198	759.8889	564.7797	204	3999
52 week low price	191	172.2356	118.7593	27	815
Volatility	198	0.1267677	0.0734892	0.02	0.54
Number of transactions	198	8388.662	9160.669	198	57795
Price premium	192	0.8268229	1.185509	-0.53	7.9

Table 7: Descriptive statistics on dummy variables

Variables	Value	Frequency	Percentage
Collaboration	0	100	50.51
	1	98	49.49
	Total	198	100
Brands	Air Jordan	89	44.95
	Adidas	55	27.78
	Nike	54	27.27
	Other brands	0	0
	Total	198	100
Release date	2020	42	21.21
	2019	102	51.52
	2018	36	18.18
	2017	13	6.57
	2016	3	1.52
	Before 2016	2	1.01
	Total	198	100

Table 8: Descriptive statistics on interaction between collaboration and brand

	Air Jordan		Adidas		Nike		Other brands	
Collaboration	0	1	0	1	0	1	0	1
0	21	79	99	1	80	20	100	0
1	88	10	44	54	64	34	98	0
Total	109	89	143	55	144	54	198	0