Hedonic Pricing in the Sneaker Resale Market

Kevin Ma and Matthew C. Treiber

Under the supervision of

Professor Kyle Jurado, Primary Advisor

Professor Michelle P. Connolly, Seminar Advisor

Professor Grace Kim, Seminar Advisor

Honors Thesis submitted in partial fulfillment of the requirements for Graduation with Distinction in Economics in Trinity College of Duke University.

Duke University
Durham, North Carolina
2020

Acknowledgements

We would like to thank our thesis advisors Dr. Kyle Jurado, Dr. Michelle Connolly, and Dr. Grace Kim, all who have given us invaluable advice and support throughout the writing process. Without their comments, ideas, and questions, our thesis would not be where it is today. We are also extremely grateful for our classmates from our thesis workshop classes for asking questions that prompted ideas and offering valuable comments on drafts. We would also like to thank our friends and family for their support throughout the development of our thesis.

Abstract

This paper explores the secondary resale market for high-end and limited-edition sneakers, specifically analyzing the determinants that affect what value sneakers trade for in the secondary market. While it is common knowledge that the sneaker resale market is a thriving and active secondary market, there is little to no empirical research about what exactly causes such sneakers to sell for exorbitant prices in the resale market. The study utilizes a hedonic pricing approach to investigate the determinants of sneaker resale price. We use a dataset of sneaker resale transactions from the online marketplace StockX between the years of 2016 and 2020 as the basis for our research. After analyzing the results, we have determined that the amount of "hype" that surrounds a sneaker as well as supply scarcity are statistically significant factors when determining the resale price premium a particular sneaker commands in the secondary market. This work adds to the sparse literature on the sneaker resale industry and brings an econometrics-approach to determining the price a given pair of sneakers commands in the resale market.

JEL Classification: C20, J19

Keywords: sneakers, resale markets, secondary markets, hedonic analysis

I. Introduction

The Adidas Yeezy Boost 750 "OG Light Brown", one of Kanye West's signature sneakers with Adidas, sells in the resale market for an average price of $1,592 – a premium of over 300% when comparing to the sneaker's original MSRP price tag of $350 (StockX, 2019). The significant markup of coveted athletic sneakers is not unique to Kanye West's signature shoes or to Adidas. The sneaker resale market, the secondary marketplace in which buyers will pay exorbitant amounts of money for coveted limited-edition sneakers, often at large premiums over the sneakers' MSRP, has exploded to a $2 billion USD industry in the US (Jones, 2018).

Sneaker culture has made its way to the forefront of mainstream culture over the last decade – evolving from an underground subculture to a full-blown frenzy that has propelled the sneaker to an iconic fashion staple, collector's item, symbol of status, and even alternative investment vehicle. The rise of "sneaker culture" has led to the growth of the worldwide sneaker market to a value of over $55 billion USD (Weinswig, 2016). The dynamics and evolution of "sneaker culture" have been researched extensively, specifically focusing on branding and marketing strategies of major retailers, as well as the role that different parties, including "sneakerheads"[1], athletes, and celebrities, play in the continuing growth of this phenomenon. Yet, there is little literature about what causes certain sneakers to be priced the way they are in the secondary resale market, and virtually no accepted approach for approximating what a certain sneaker should sell for in the secondary resale market.

This paper aims to research and determine the most significant factors that cause a sneaker to sell above its retail price, and specifically develops a time-dummy hedonic regression model that estimates the contributory value of such factors. In a sneaker market where consumers face complex consumption

[1] Sneakerheads are colloquially defined as those who collect, trade, and or admire sneakers as a hobby.

decisions, the hedonic model offers a way to identify attributes that impact consumers' marginal willingness to pay and to estimate the implicit price of these attributes. The hedonic pricing model is common in real estate markets and uses all available transaction data to estimate a model that prices each property based on its individual attributes (Xu, 2017). In the case of the real estate market, the building can be seen as a "bundle of goods", comprising of its different characteristics, analogous to the way a particular sneaker is differentiated by characteristics like sneaker type and brand (Monson, 2009). The price is then determined by the sum of the implicit value of these characteristics (Rosen, 1974). The hedonic pricing model allows the distinction between price changes arising from individual characteristics such as the number of bedrooms, square footage, construction of new public transport as well as external macroeconomic and policy developments (Xu, 2017).

Data on secondary market sneaker transactions is just becoming more accessible in the past few years. Today, the two primary online platforms that provide authenticity guarantees and streamlined selling / buying processes are GOAT app and StockX.com. Before websites such as these existed, a seller would have to build seller reputation, and provide customer service in the form of product information and guarantees of authenticity. The purchasing experience for consumers was fraught with scamming, counterfeit sneakers, and lengthy shipping times. Examples of online marketplaces like these that provide a selling interface without a middleman verifying for authenticity that are still available today include eBay, Poshmark, Depop, Grailed, Craigslist, and Facebook Marketplace. The online space for selling and buying sneakers is even more fragmented when considering the (relatively) new online storefronts for consignment stores like Flight Club, which was founded in 2005 and went online in 2014, and Stadium Goods, founded as an online-only consignment store in 2015. Simultaneous demand for limited-edition sneakers has been driven by companies like Nike taking advantage of digital marketing and unique release procedures, such as the 'SNKR Cam' that used Augmented Reality (AR) to launch a product like the Nike Pigeon SB Dunk. This particular product release in November 2017 led loyal customers to different locations in New York City to scan a special edition newspaper using a built-in camera in the Nike app for the opportunity to make a purchase. Other unique releases have included: a 'Shock Drop' in which a product is released at a random time and users are notified upon release, and the 'SNKR Stash' in which a product is made available in a specific geofenced location that customers must be inside in order to complete their purchase. These complex digital releases reflect a consumer base that is tech-savvy, and extremely interested in using the internet to gauge popularity, or "hype" of certain sneakers. In a survey of sneaker collectors, it was found that this demographic is largely composed of males ages 18-23, with an average collection size of 14.34 sneakers (Cassidy, 2018). However, the number of women who identify as "sneakerheads" continues to increase, and virtually all sneakers, especially the limited-edition sneakers that we will be focusing on, can be found on the feet of men and women of all ages (StockX, 2019).

The only sneaker resale platform that is transparent with all transactions that occur on its marketplace is StockX, and thus is the platform that we will draw our dataset from. StockX brings buyers and sellers together but also serves as a middleman to verify the authenticity of the sneakers being sold. Using StockX resale transactions from 2015 to 2020, we explore a hedonic pricing model for resale transactions of sneakers. We believe that the hedonic model that we develop will not only bolster the existing sparse literature on the sneaker resale market, but also act as a tool to aid in evaluating investment decisions of highly coveted sneakers. We will only be analyzing "deadstock" (a term used by the sneaker community to indicate that a sneaker is brand new) and authentic sneakers, and not pre-owned or counterfeit products. The benefit of scraping data from verified marketplaces such as StockX means that all sneakers are "deadstock" and 100% authentic.

It is worthwhile to give a bit more background about sneaker classification in the secondary market as well as explain how we will frame the concept of *a particular pair of sneakers* throughout the rest of this paper. Let's take the Air Jordan 1 Royal ("Royal 1s") as an example to better illustrate this concept. The model of the Air Jordan 1 "Royal" is the Air Jordan 1, while the colorway is "Royal" (what Jordan Brand has dubbed the specific colorway for this particular sneaker). It is along this classification (the sneaker model along with the colorway) that *a particular pair of sneakers* in the secondary resale market is designated – this is how sneakerheads and sneaker resale platforms alike classify different shoes. There will be separate listings for the Air Jordan 1 "Royal" and the Air Jordan 1 "Reverse Shattered Backboard", for example. Often times, sneakers that share a model but have different colorways will trade in completely different price ranges. That is why it is crucial to understand this concept of classification for a particular pair of shoes. We will refer to this classification as *a particular pair of sneakers/shoes* or a *particular sneaker/shoe* or a sneaker *silhouette*. Figure 1 shows that while the Air Jordan Retro 1 "Royal" and the Air Jordan Retro 1 "Travis Scott" have the same exact sneaker model, they have a different colorway and are thus classified as different *silhouettes*, and trade in different price ranges. To sneakerheads and sneaker resale platforms, **they are considered different sneakers despite having the same model**. Sneakers must have the same model as well as colorway in order to be classified as having that particular *silhouette.*

Figure 1 – Comparison of Different *Sneaker Silhouettes*

		
Name	Air Jordan Retro 1 "Royal"	Air Jordan Retro 1 "Travis Scott"
Common Name	Royal 1s	Travis 1s
Celebrity Signature?	No	Yes
Retail Price	$160	$175
Average Resale Price (last 12 months)	$305	$1,115

It is also important to discuss the timing of events that are involved in the release of a particular sneaker. Figure 2 below provides a timeline of the events involved in a typical sneaker release.

Figure 2 – Illustration of Timing in a Typical Sneaker Release

Before the official announcement of a sneaker release by a manufacturer (Nike, Adidas, New Balance, etc.), images or rumors of the sneaker release will circulate on the internet and garner the attention of the sneaker community. On social media, enthusiasts might discuss how popular the sneaker will be, what the potential resale prices could be, and, of course, how aesthetically pleasing the sneakers are. Afterwards, the manufacturer officially announces the release of a sneaker as well as a release date for that sneaker – they typically announce official releases from the brand's social media accounts, and through the brand's mobile app. For example, Nike may announce the release date for a new pair of shoes on Twitter, along with uploading the product with the official release date on their mobile app, SNKRS. Before the official release date, those who are able to secure the sneakers before release (often people with connections to the brand or local retailers) will command high price premiums in the resale market. Discussion and "hype" for the particular sneaker continues and usually peaks right around the release date. At release date, brick and mortar stores as well as online carriers sell the sneakers for retail price. Those who are able to secure the coveted sneakers for retail price (typically by having connections at brick and mortar retailers, winning online or in-store raffles, or with automated online computer programs called "bots") will often try to sell the sneakers shortly after release to capture a premium above the retail price. After the official release, the resale market is fully active. There is usually a decline in resale price premiums as well as a decline in "hype" in the post-release phase. The last event that *could*, but not always, occurs is a restock. A restock occurs when the manufacturer releases more pairs of a particular sneaker into the primary market for retail price. Hence, there is also an inflow of pairs of a particular sneaker into the secondary resale market. Typically, resale price premiums decline after a restock due to less perceived exclusivity as well as increased supply in the secondary market. Although it is especially uncommon practice to restock more than once, a manufacturer has the ability to restock a particular sneaker as many times as they see fit. For the purposes of our research, a restock will be documented only when a manufacturer makes an official announcement, through their mobile app or social media. A brick and mortar store releasing a limited run of shoes they may have not sold on the official release date, because they found additional pairs in a stockroom, or pairs were unclaimed from a raffle, does not count as a "restock", because these are not additional pairs being produced by the manufacturer in a factory.

The next section provides a review of hedonic methods applied to the art and baseball card resale markets, existing research on the factors that impact the sneaker resale market, and past research studying the effects of Twitter chatter on movie sales. Sections III. and IV. explain the theoretical framework for the model, describe the dataset, and outline the methods used. Section V. presents the findings of the hedonic regressions with accompanying analysis. The conclusion will discuss the possible areas to build upon, implications, as well as limitations of the research.

II. Literature Review

Hedonic Approach in the Baseball Card & Art Resale Markets

Mulligan and Grube (2007) use a simple hedonic approach in pricing baseball cards. Sports memorabilia make a strong candidate for the hedonic pricing approach because of the relative homogeneity of collectibles like baseball cards – cards should have their value determined by characteristics intrinsic to the card, such as the particular player's batting averages, popularity, the player overcoming hardships like injury, etc. Batting average, World Series appearances, and whether or not the player is deceased were generally statistically significant over the different time periods in the data. The researchers also found that there is an aura effect which elevates the value of cards for players who have recently retired and as time passes, the card price declines (or grows less rapidly). Witkowska (2014) applies the hedonic pricing approach to a group of artworks by selected Polish painters auctioned in 2007-2010. Witkowska selects characteristics of the selected artwork such as technique, living status of artist, and surface area of artwork. Witkowska concludes that the specifications of the models are crucial for markets such as art because the specifications impact hedonic quality adjustment, specifically that different hedonic models cause different values for the price index and so it is difficult to determine which indexes describe "true" price impact.

Based on this research, we will specify different models and compare the results as we cannot attribute the impact of a particular attribute based on one model. Both papers utilize models with and without time dummy variables to analyze selected attributes while capturing and not capturing the fixed effects of particular periods. We believe that this will be significant in our own research as trends in the sneaker industry, and in turn, the sneaker resale market, change rapidly (Lux, Mortiz, and Bug, 2018).

Factors that Impact the Sneaker Resale Market

Existing literature have explored certain factors as determinants of price in the resale market, specifically: scarcity, collaboration with certain celebrities and other companies, prices before and after release, as well as overall "hype" or buzz about sneakers on social media and approval throughout the sneakerhead community.

Cassidy (2018) claims that sneaker investors, enthusiasts, and collectors have an overwhelming preference towards supply-scarce sneakers. This preference for supply-scarce sneakers is motivated by the consumers' consumption of sneakers as conspicuous consumption products, in which products demonstrate a signaling effect about the consumer (Cassidy, 2018). For example, Adidas has made the mistake of restocking (releasing more pairs of sneakers that already released to the public at retail price) high-profile sneakers that had sold out and carried high prices in the resale market (Welty, 2017). An extreme example is the debut colorway of the Adidas Yeezy Powerphase, which was selling in the resale market for up to $1,150 shortly after release but was restocked multiple times which led the price to hover around $125 today, only $5 dollars more than the sneaker's MSRP (Dunne, 2017). While restocking high-profile sneakers has netted sportswear companies increased revenue in the short term, these restocks also tarnish the price of many sneakers in the resale market. Collectors and investors alike treat their sneakers as investments and are unlikely to pay for sneakers in the secondary resale market that will increase in supply and, consequently, experience a sharp price drop in the resale market.

Collaboration sneakers between the sportswear company creating the sneaker and other high-profile celebrities and entities have caused such collaborations to sell above their non-collaboration counterparts in the resale market. For example, the average Air Jordan 1, a very popular sneaker in of itself, fetches prices of $300 to $500 in the resale market; however, collaborations with design studios such as OffWhite and Fragment Design demand prices from $1500 to $3000 (Luber, 2018). Kanye

West's collaboration sneakers with Nike still sell for prices up to $6,000 in the resale market and also represented 6 out of the 10 most valuable sneakers on the resale market in 2015 (Adams, 2016).

Khaki compiled sales data from eBay for a popular retro Reebok basketball shoe released in 2013 and found that this shoe had the highest average sale price in the weeks leading up to the release date, followed by a drop in average sale prices immediately after the release. The average sale price eventually leveled out over the observed time period (Khaki, 2013). Khaki focuses on many of the same variables we have identified, including days since release, and price premium above retail price, but his analysis is limited to summary statistics segmented by week. Only focusing on one particular sneaker with 1,500 data points from completed sales on eBay means that these results may not extend to sneakers of a different model, brand, or colorway. However, our approach will keep in mind this interesting relationship between high price premiums before release date ad declining price premiums in the weeks after release.

Lux, Mortiz, and Bug (2018) posit that approval from the sneakerhead community is a crucial aspect in regards to value appreciation in the resale market. Ultimately, the "hype" or buzz around certain sneakers is created by the people, and not by the marketing efforts of the actual companies. Stock and Balachander claim that oftentimes intentional scarcity will feed "hype" or increased desirability of a product especially when a large fraction of consumers is unable to purchase the product. We suspect that characteristics like scarcity are highly correlated to buzz generated by the sneakerhead community or "hype."

Using Twitter to Measure "Hype"

Rui, Lui, and Whinston (2011) found that the number of Tweets as well as the positive Tweets regarding a movie is associated with higher movie sales for a particular title. They also find that the effect of Tweets from users with more Twitter followers is significantly larger than the effect of Tweets from users with comparatively less Twitter followers. This research provides evidence that Twitter can be a valuable platform to evaluate the attention that a certain product is garnering and that this attention is associated with higher sales of the product. We can adopt a similar strategy in using Twitter to measure the “hype” of a specific sneaker silhouette and studying the effect of “hype” on resale price premiums for that sneaker silhouette (not overall sales). Additionally, Rui et. al have shown that not all Tweets are equal – Tweets from users with more Twitter followers have more impact than those from users with less followers. When measuring the effect of each Tweet, we will instead be weighing the effect of each Tweet by the amount of retweets and likes the Tweet generates. Tweets from users with more Twitter followers will inherently generate more likes and retweets when compared to Tweets from users with less followers; therefore, our methodology not only reflects the added impact of a Tweet from a user with a large following base, but also reflects the potential impact of Twitter users who are outside of a user’s following base who may like and retweet a given Tweet. A viral Tweet that generates significant attention can come from a Twitter user that does not have a large following base.

III. Empirical Design

The basic hedonic pricing approach, based on instrumental work by Lancaster (1966) and Rosen (1974), posits that the price of a particular good (P) is a function of its individual attributes. Lancaster concludes that instead of choosing between quantities of products, consumers base their decisions on a good's attributes and their respective intensities. Rosen hypothesizes that hedonic price refers to the implicit price of the good's attributes and is observable through differentiated products.

$$P = f(x_1, x_2, \dots, x_n) \tag{1}$$

Sneakers in the resale market appreciate in value in a fashion similar to sports memorabilia or fine art, though for different reasons. The supply, or scarcity, of sneakers in the sneaker resale market are limited to the amount manufactured and then released by the particular brand; however, often times manufacturers release additional pairs of a particular sneaker after the initial release date (the additional amounts released and time after initial release depend on the specific sneaker). Changes in demand also affect the prices of sneakers in the resale market. Demand for such sneakers increases with increased interest in the particular brand, specific sneaker, signature athlete or celebrity attached to the sneaker, with interest in the sneaker for its own sake, and with increased expected return on them as alternative investment assets. More specifically, the overall "hype" that a sneaker generates among the sneaker community is instrumental in driving its demand, and in turn, price in the secondary resale market. The hedonic price model that we estimate incorporates the aforementioned factors in a manner that is useful for evaluating the value for sneakers in the resale market. The generalized form of the model is:

$$ln(P_{it}) = \lambda_0 + \sum_{j=1}^{k} \beta_j X_{ijt} + \lambda_{17} D_{17} + \cdots + \lambda_{20} D_{20} + E_{it} \tag{2}$$

The model features a dummy variable for each year from 2016 to 2020 belonging to the set of all years in the data $\{2016, \ldots, 2020\}$. Therefore, the linear regression is normalized to the first year, 2016. P_{it} represents the transaction price of sneaker i in the time period t, λ_{year} is the estimated coefficient for each respective time dummy variable D_{year}, and β_j is the vector of estimated coefficients for each of k variables of transaction i in the time period t represented by vector X_{ijt}. The natural logarithm of price is typically utilized in these models to avoid issues with extremely large or small price values; however, we will discuss adopting different functional forms for the dependent variable in the end of this section of the paper.

Supply scarcity and overall "hype" are two crucial variables that are part of vector X_{ijt} which serve as proxies for the true effect of supply scarcity and "hype", respectively. As previously mentioned, there is a preference towards supply-scarce products in the sneaker resale market; therefore, the scarcity is an imperative factor to be captured. Sneaker manufacturers do not publicly release how many pairs of a certain sneaker are released in the primary sneaker market at initial release. For this reason, we will approximate the supply for a particular sneaker by assigning a "score" based on the number of transactions during a 30-day period post release relative to the aggregate number of transactions during a 30-day period post release for all the sneakers in the dataset. We discuss the approach in further detail in Section IV. A limitation of this method is that the same physical until of a particular shoe might be traded in the secondary marketplace more than one time within the 30-day period that we designate (i.e. Person A sells physical unit to Person B one day after release and then Person B sells the same physical unit for a higher price to person C seven days later).

As for "hype", there is no universal way to quantify the amount of attention or buzz surrounding a given phenomenon. For this reason, we approximate the effect of hype through a continuous variable incorporating the number of Tweets associated with the sneaker on Twitter. This continuous variable will track the Tweets associated with the appropriate keywords for a particular *silhouette*. We detail our specific approach in tracking "hype" in Section IV. Using Twitter to quantify a value of "hype" serves as a suitable proxy for several reasons. For one, the sneaker community is active on Twitter and often discusses "hyped" sneakers on the platform; therefore, Twitter is the most appropriate platform that captures the true amount of "hype" for a given sneaker. For the most part, "sneakerheads" are the ones that are buying and selling sneakers in the secondary marketplace. While someone who is not likely to purchase or sell sneakers in the secondary market will Google a sneaker, true movers of the market will Tweet about them. For this reason, we deem capturing the Tweets on Twitter as a more fitting proxy compared to Google search data. Additionally, there is increased desirability for a particular sneaker in the resale market because of an associated network effect on Twitter. Because sneakers in the resale market are often seen as high-fashion items and symbols of status, if sneakerheads observe many mentions of a particular sneaker, especially close to the release date, the value of the sneaker inherently increases to that individual. While "hype" is an abstract factor to measure, it inevitably has a profound effect on the price that a sneaker sells for in the resale market.

It is worth noting a few drawbacks in using the hedonic pricing model. Due to the complicated nature of sneaker resale value, as well as the underlying tastes that motivate purchasing decisions, it is difficult to specify an appropriate functional form for the model *a priori* using a theoretical approach. There are several basic functional forms which include linear, log, and semi-log that can be applied to the hedonic price model. However, an incorrect choice of functional form may result in inconsistent estimates (Bloomquist and Worley, 1981). Despite the widespread use of the hedonic pricing model, the theory of hedonic pricing provides little guidance on the choice of appropriate functional form (Butler, 1982). Thus, the specification of a particular functional form for the model will contain model specification bias (Jiang, Philips, and Yu, 2015). Therefore, we utilize an approach that tests different specifications of the dependent variable to find the most suitable transformation. Furthermore, while the sneaker resale market is efficient at responding to new information, hedonic pricing assumes that the market is perfectly efficient and responds to new information immediately.

There are a few reasons why the hedonic pricing approach is particularly suitable for the secondary sneaker resale market. First of all, the method can be used to estimate values based on actual transactions. This is particularly important in the secondary sneaker resale market as it is difficult to understand the motivation in a change in tastes or sudden virality, for instance. Second, the sneaker resale market is a good indication of value because it is rather efficient in responding to information (such as Nike scaling back supply or popular musicians wearing a particular sneaker on tour). Additionally, Diewert, Heravi, and Silver (2008) claim that hedonic pricing models are apt for "product areas with a high turnover of differentiated models," such as the sneaker resale market. There is a plethora of different brands and types (running, lifestyle, basketball, designer) of sneakers in the resale market. Additionally, new designs are constantly introduced to the primary sneaker market, and in turn, the secondary resale market. On the other hand, the make quality and functionality, are, by design, rather homogenous across different brands and types of sneakers which helps reduce unobserved variation in factors that cannot be captured through data alone. There is also a wealth of resale transactional data on different sneakers which will allow for fine-tuning of the characteristics being investigated.

IV. Data

The data for this paper come largely from StockX.com as well as Twitter. As there are no publicly available datasets that describe the secondary sneaker resale market, we scraped the transaction data of a random sample of 30 different pairs of sneakers from the "Most Popular" sneakers tab on StockX.com. Specifically, the data are scraped from StockX.com using API calls to the Request URL for an individual shoe, which then gives a set of JSON data to be converted into a CSV file. Scraped variables include the time of transaction, initial release date, brand, the transaction price, and the size. In addition to the transactional data from StockX.com, we used publicly available sources such as *Sneakernews.com* and *Complex.com* to determine whether the sneakers were a celebrity, athlete, or brand collaboration as well as the type of sneaker (basketball, running, lifestyle, skateboarding, etc). We scraped Twitter to extract a list of Tweets that mention the particular sneaker using a freely available tool on GitHub developed by Ahmet Taspinar.[2] We will go into the specifics of the variables that will be utilized in our regressions. There are several important caveats and limitations that are also discussed further in this section.

<u>Dependent Variables</u>

<u>Price, Retail Price, and Price Premium</u>

The price of the specific transaction is simply the amount that the buyer paid for a pair of sneakers for resale transaction *i* before shipping and transaction costs (StockX displays past transactions without the shipping and transaction costs). Unlike the transaction price, the retail price is the fixed price that the sneaker manufacturer charges in the primary market for a certain sneaker and hence it is constant throughout. Different resale transactions of a particular kind of sneaker will have the same retail price. We calculate, for a given sneaker, the absolute price premium as well as the price premium

[2] https://github.com/taspinar/twitterscraper

as a percentage of retail price as follows:

$$Absolute\ Price\ Premium = Transaction\ Price - Retail\ Price \tag{3}$$

$$Price\ Premium\ as\ \%\ of\ Retail\ Price = \frac{Absolute\ Price\ Premium}{Retail\ Price} \tag{4}$$

Table 1 - Summary of Dependent Variables

Variable	Mean	Standard Deviation	Min	Max
Transaction Price	433.2914	288.0153	25	3750
Retail Price	217.0529	59.89552	110	395
Absolute Price Premium	216.2384	289.9628	-175	3575
Price Premium as % of Retail	1.102813	1.589247	-0.875	20.42857

<u>Independent Variables</u>

<u>Shoe Size (US), Brand, and Type Dummy Variables</u>

Each of the above independent variables are represented with a series of dummy variables. While shoe size is a numerical measure, in our dataset the shoe size typically ranges from 3 to 18 with half-sizes for all sizes except for sizes 15 and above and thus will be coded as a series of dummy variables. The most common shoe size in our dataset is size 10 with 38,895 observations and the least popular shoe size is size 3 with 32 observations. In our regressions, these dummy variables for the shoe sizes will serve as fixed effects for the shoe size of the transaction. The brand is simply the sneaker manufacturer for that particular sneaker (Nike, Jordan, Adidas, New Balance, Asics, etc). The most common brand in our dataset is Adidas with 170,266 observations. The least common brand is Asics, with 260 observations. The type of sneaker describes the category of use that the sneaker falls under (basketball, lifestyle, athletic, skateboarding, etc). The most common type of sneaker in our dataset is lifestyle (casual) with 184,276 observations, while the least common type of sneaker is skateboarding, with 8286 observations.

<u>Brand Collaboration, Celebrity or Athlete Signature, and Brand/Celebrity/Athlete Twitter Followers</u>

As Lux et. al (2018) have claimed in their research, brand collaboration as well as celebrity and athlete signature shoes are important for the value that sneakers resell for in the secondary market. Brand Collaboration, Celebrity and Athlete Signature Shoe are three dummy variables that take on values 0 or 1 (1 if the shoe is part of a brand collaboration or is a celebrity/athlete signature shoe and 0 if not). We utilize the number of Twitter followers that the particular brand/celebrity/athlete had at time *t* to weigh the impact that these collaborations have on our dependent variables. We approximate Twitter followers at time *t* by using the Wayback Machine to view celebrity, athlete, and brand Twitter accounts at previous time periods starting before the first transaction for a particular sneaker, and at yearly intervals until the last observation for that sneaker. Then, we assume linear growth in Twitter followers between each of these time periods and assign each transaction with the approximate number of Twitter followers that celebrity, athlete, or brand had at the time. We found this to be a reasonable way to capture Twitter follower growth considering Twitter's API does not allow for scraping historical Twitter follower numbers. Additionally, by getting a snapshot of a particular celebrity or athlete's Twitter followers every year, we were able to capture unusual spikes in Twitter following, for instance, Kawhi Leonard's rise in followers after his 2019 NBA Championship. Of course, if the shoe does not feature a celebrity, athlete, or brand collaboration, all of these weighted terms will be 0.

“Hype” Twitter Value

As discussed previously, the sneaker community builds up and maintains the “hype” for a particular sneaker before and after a sneaker release, especially on Twitter. To approximate the amount of “hype” that surrounds a certain sneaker at time *t* for transaction *i*, we utilize a metric that we have defined as “Tweet Value Stock”. In equations (5) and (6), we establish the basis for our methodology, mimicking a capital accumulation approach with a depreciation factor. “Tweet Value” assigns a value for each Tweet scraped regarding a pair of sneakers, with Tweets that get more exposure (retweets and likes) being assigned a higher value. “Tweet Value Stock” at time *t* reflects the accumulation of Tweet Value over time, factoring depreciation, and is the ultimate variable to be used in our regressions.

$$Tweet\ Value_t = 1 + \#\ of\ Retweets_t + \#\ of\ Likes_t \tag{5}$$

$$Stock\ Tweet\ Value_t = (1-\delta) Stock\ Tweet\ Value_{t-1} + Tweet\ Value_t \tag{6}$$

We start collecting Tweets for a sneaker 3 months before the first available transaction on StockX. We start collecting Tweets based on the first available transaction as opposed to the initial release date for two reasons. First of all, there are some insiders who resell the sneakers in the secondary resale market days before the official release date, and so we use the first transaction as our baseline instead of the release date. Second, sneaker manufacturers often surface images and announce potential releases of limited-edition sneakers multiple months in advance (this can vary between sneaker release), and thus, the sneaker community will start discussing the announcement months ahead of the initial release. This phase is very important and where much of the “hype” around a sneaker is built up.

With this method, the $Tweet\ Value\ Stock$ at time *t* is the depreciated stock from *t-1* plus $Tweet\ Value_t$, the new Tweet value for time *t*. As opposed to simply utilizing the raw number of Tweets as our approximation for “hype”, the above method allows us to weigh the value that each Tweet has a function of likes and retweets while also incorporating the diminishing value of old Tweet value. We discuss the different rates for our depreciation δ in Section V.

Figure 3 – Twitter Hype over Time for Air Jordan Retro 1 "Travis Scott"

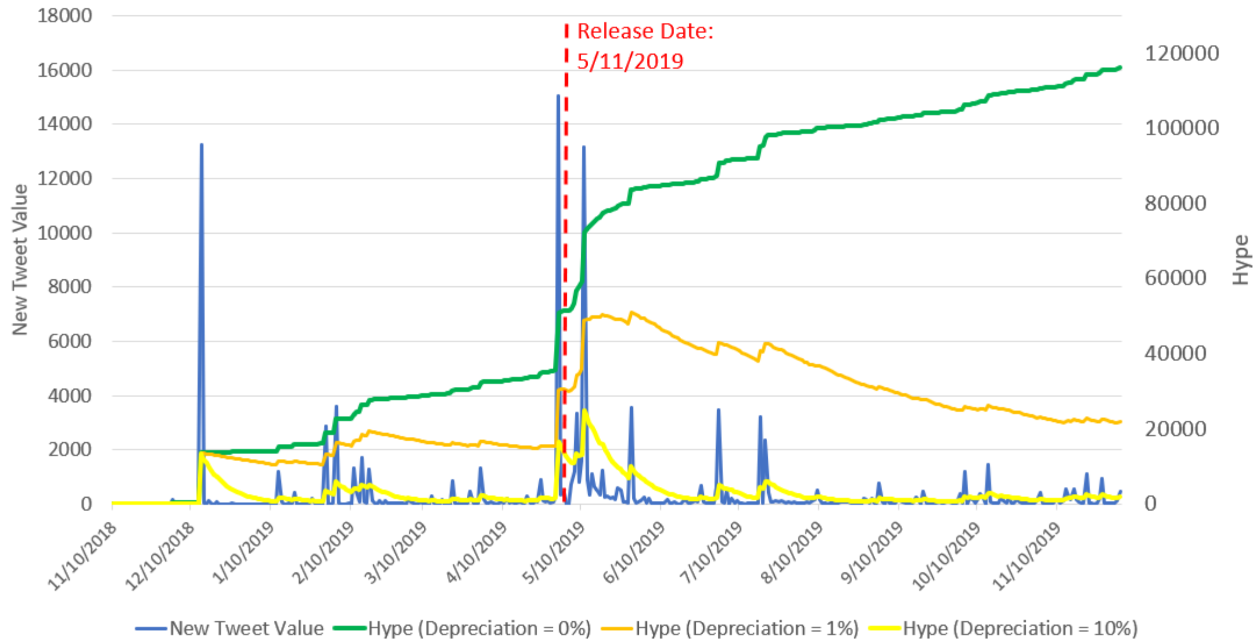

Figure 3 above displays Twitter "hype" over time for the Air Jordan Retro 1 "Travis Scott" as an example. The value of $Tweet\ Value_t$ is represented by the blue line, while $Tweet\ Value\ Stock_t$ is represented by the green, orange, and yellow lines (0%, 1%, and 10% depreciation, respectively). As the figure shows, $Tweet\ Value_t$ and $Tweet\ Value\ Stock_t$ (with 1% and 10% depreciation, respectively) reach all-time highs around the time of the release of the sneaker, which was on May 11th, 2019.

A few limitations occur in our approach of approximating "hype", specifically in scraping Twitter. First of all, we only have access to public Tweets because we cannot access Tweets of private accounts. Second, in scraping Twitter, there are many automated bot accounts that do not contribute to the sentiment of the sneaker community towards a shoe; however, using the described method of assigning a value to each Tweet will mitigate the impact that these bot accounts have on our "hype" measure. Nonetheless, we are simply approximating the amount of "hype" as there is no standard or universal way to quantify the true value.

<u>Supply Scarcity Proxy</u>

As discussed in Section III., it is imperative that we obtain some measure of the scarcity or supply for a given silhouette. To approximate this metric, we first take the number of transactions for a silhouette 30 days after initial release for all of the 30 sneakers in our dataset and then add these figures together and divide by the number of sneakers in our dataset, 30, to obtain an average 30-day post release number of transactions. We then divide the individual figures for every given silhouette by the average 30-day post release number of transactions figure. Using this method, we have created a relative "score" that proxies the initial supply for a given sneaker silhouette.

<u>Days Since Release</u>

This numerical variable is the number of days that have elapsed for transaction *i* compared to the initial release date of the sneaker (when the sneaker manufacturer releases the sneakers to the public in-store and online). The resale transaction date can occur before, at, and most often after the initial release date that the sneaker manufacturer sets.

<u>Restock</u>

This variable is a dummy variable that signifies whether a particular sneaker has experienced an officially announced restock, a phenomenon we describe in Section I. We make sure to base this variable on the announcement of a restock by the manufacturer, either on its official social media page or mobile app. When a restock of a limited-edition pair of sneakers occurs, sneaker news outlets such as *Sneakernews.com* and *Complex.com* report on this announcement, making it straightforward to find restock information about the sneakers in our dataset. Because sneaker resale markets are somewhat efficient, as soon as there is an official announcement for a restock, the resale market should quickly price in this information into trading prices.

Year Time Dummy Variables

As the transactions span between 2016 and 2020, our regressions have fixed effects dummy variables to absorb any macro level shocks or changes that might occur in the sneaker resale market in any given year, which is especially relevant in controlling for growth in StockX users over time. This allows us to consider the increase in StockX user growth that may have an impact on the number of transactions that occur in a 30-day period after a particular sneaker's release and compare sneakers that were released in different years in our proxy for supply.

V. Results

Table 2 – OLS Regressions on Log of *Price Premium as % Retail*

		Regression 1 Log Price Premium %		Regression 2 Log Price Premium %
	Number of Observations	349,556	Number of Observations	349,556
	Prob > chi2	0.0000	Prob > chi2	0.0000
	R-Squared	0.6489	R-Squared	0.6494
	Adjusted R-Squared	0.6489	Adjusted R-Squared	0.6493
	Root MSE	0.64852	Root MSE	0.64807
	Twitter Hype (δ=5%)	0.0000041***	Log Twitter Hype (δ=5%)	0.0309168***
		(0.0000002)		(0.0009084)
	Supply Proxy	-0.1839263***	Supply Proxy	-0.1940782***
		(0.0015293)		(0.0015645)
	Restock	-0.9500945***	Restock	-0.9235048***
		(0.0047547)		(0.0047662)
Brand Dummies	Asics	0.5535577***	Asics	0.6318008***
		(0.0408412)		(0.0409358)
	Converse	1.8525329***	Converse	1.8207559***
		(0.0114081)		(0.0114667)
	Jordan	2.1760778***	Jordan	2.2407729***
		(0.0136992)		(0.0135123)
	New Balance	1.8725931***	New Balance	1.8947488***
		(0.0415927)		(0.0415526)
	Nike	0.9283579***	Nike	0.9622157***
		(0.0096019)		(0.0093410)
	Adidas	1.1251170***	Adidas	1.0996207***
		(0.0055892)		(0.0053508)
Type Dummies	Basketball	-0.8677774***	Basketball	-0.8696927***
		(0.0144276)		(0.0144146)
	Lifestyle	-1.5298751***	Lifestyle	-1.4809308***
		(0.0124051)		(0.0122005)
	Skateboarding	0.4302635***	Skateboarding	0.4163602***
		(0.0109249)		(0.0109355)
Influence Dummies	Brand Collaboration	1.5371633***	Brand Collaboration	1.5002193***
		(0.0058529)		(0.0059167)
	Celebrity Signature	2.2221654***	Celebrity Signature	2.2125778***
		(0.0057334)		(0.0056718)
	Athlete Signature	-0.8720630***	Athlete Signature	-0.8341601***
		(0.0100178)		(0.0100038)
	Days Since Relese	-0.0000871***	Days Since Relese	-0.0000851***
		(0.0000078)		(0.0000077)
	Constant	-1.2959547***	Constant	-1.6129390***
		(0.0195678)		(0.0211759)
	Shoe Size Fixed Effects	YES	Shoe Size Fixed Effects	YES
	Year Fixed Effects	YES	Year Fixed Effects	YES

Standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Regarding the rate of depreciation for our Twitter Hype variable, we ultimately selected a depreciation rate of 5% after experimentation with depreciation rates ranging from 1% to 50%. A depreciation rate of 5% signifies that approximately all of the "hype" associated with one Tweet decays through depreciation after 90 days. We found that depreciation rates higher than 10% depreciated too quickly; consequently, the distributions for Twitter Hype with depreciation rates higher than 10% were extremely right-skewed.

In Regression 1, we see that all our independent variables are statistically significant at a significance level of 0.01, showing that sneaker resale market prices incorporate these variables in resale prices. This regression model produces an R-Squared value of 0.6489 (adjusted R-Squared = 0.6489), suggesting that the model is able to explain over half of the observed variation in sneaker resale premiums. We find that our initial hypotheses regarding Twitter Hype ($\delta = 5\%$), Supply Proxy, and Restock are correct based upon the signs of the coefficients for these variables: the more "hype" a silhouette garners, the higher the resale premiums; the lower the supply (higher the exclusivity) for a silhouette, the higher the resale premiums; a silhouette that has been restocked suffers a decrease in resale premiums. In Regression 1, the magnitude of Twitter seems small. This is because we are evaluating the impact of one Tweet depreciated over time. Additionally, we observe that sneakers that are part of a brand collaboration and sneakers that feature a celebrity signature have a positive effect on our dependent variable. This is unsurprising as silhouettes that feature a brand collaboration or celebrity signature are typically the most sought-after sneakers in the resale market. The sign of Days Since Release is also consistent with our initial hypotheses: the resale market is biased towards more recent releases, and thus the sign of Days Since Release is negative.

One of our variables, Athlete Signature, demonstrates a different result than we expected: the coefficient for Athlete Signature is negative. We believe that this is due to the fact that athletes who have contracts with sneaker manufacturers release a new signature shoe model every year (each model has various different colorways, as discussed in Section I.). For example, LeBron James has released 17 different signature shoe models over the course of his career, one for each year he has played in the NBA. It is likely the case that sneakerheads understand that new LeBron James signature shoes will continue to be released year after year, lowering their desirability and exclusivity; however, this is not the case for celebrity signature sneakers, which do not have consistent yearly releases. Another explanation for this unexpected result, which likely overlaps with the aforementioned explanation, could be that recently, since athlete signature sneakers are typically designed for performance and not for fashion, they are also less desirable as status symbols and fashion accessories.

In Regression 2, we use a log transformation on Twitter Hype so that we can interpret the coefficient as an elasticity of the resale price premium. Additionally, taking the log of Twitter Hype creates a more normal-like distribution. This regression model produces an R-Squared value of 0.6494 (adjusted R-Squared = 0.6493), again suggesting that the model is able to explain more than half of the observed variation in sneaker resale prices. The slight increase in the R-Squared figures is likely due to the more appropriate log transformation for Twitter Hype. The coefficient of 0.0309168 on Twitter Hype suggests that a 10% increase in Twitter Hype is associated with 0.3% increase in the *Price Premium as % of Retail*, approximately. The coefficients for the rest of the terms remain consistent with our results in Regression 1 discussed in the preceding paragraph.

Table 3 – OLS Regressions on Log of *Price Premium as % Retail* with Interaction Effects

		Regression 3 Log Price Premium %
	Number of Observations	349,556
	Prob > chi2	0.0000
	R-Squared	0.65
	Adjusted R-Squared	0.65
	Root MSE	0.64746
	Log Twitter Hype (δ=5%)	0.0284385***
		(0.0011142)
	Supply Proxy	-0.1894349***
		(0.0015780)
	Restock	-0.9353819***
		(0.0049130)
Brand Dummies	Asics	0.6520137***
		(0.0409051)
	Converse	1.8001337***
		(0.0115023)
	Jordan	2.1799127***
		(0.0140451)
	New Balance	1.7991134***
		(0.0416881)
	Nike	0.9075482***
		(0.0098399)
	Adidas	1.1314544***
		(0.0060169)
Type Dummies	Basketball	-0.8819429***
		(0.0144546)
	Lifestyle	-1.5441032***
		(0.0124771)
	Skateboarding	0.4157730***
		(0.0109403)
Influence Dummies	Brand Collaboration	1.6040309***
		(0.0074004)
	Celebrity Signature	2.1774753***
		(0.0063996)
	Athlete Signature	-0.7827758***
		(0.0105119)
	Days Since Relese	-0.0000771***
		(0.0000078)
Interactions	Twitter Hype (δ=5%) * Brand Collaboration	-0.0000138***
		(0.0000007)
	Twitter Hype (δ=5%) * Celebrity Signature	0.0000027***
		(0.0000002)
	Constant	-1.5099496***
		(0.0232050)
	Shoe Size Fixed Effects	YES
	Year Fixed Effects	YES

Standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

In Table 3, we explore the effect of two interaction effects: Twitter Value * Brand Collaboration and Twitter Value * Celebrity Signature. The coefficient on the former is negative, suggesting that Tweets about a silhouette that feature a Brand Collaboration have a dampened effect on resale price premiums. This could be because sneakerheads understand that brand collaborations are typically very exclusive and coveted despite how much "hype" is building up over Twitter over this particular sneaker. On the other hand, the coefficient of the latter is positive, suggesting that Tweets about a silhouette that feature a Celebrity Signature have an amplified impact on resale price premiums. This might occur because fans of celebrities building up "hype" can be more passionate about a sneaker that features their favorite celebrities.

Through our regressions, we are able to better address several of our questions. First of all, we find that the secondary sneaker resale market incorporates the determinants that we selected into resale prices, and hence, price premiums. Second, we find that increased discussion of a particular pair of sneakers on Twitter is associated with higher price premiums. Despite the fact that we are not sure whether the sentiment of such Tweets is negative or positive, it is safe to say that more discussion and attention on Twitter, more specifically, increased "hype" will result in sneakers commanding higher price premiums in the resale market.

VI. Conclusion

This paper builds a hedonic regression model for the high-volume sneaker resale market. The selected determinants are able to account for a majority of the observed variation in sneaker resale premiums. Of course, due to the subjective nature of sneaker purchasing, there are some unobservable characteristics such as aesthetic beauty that undoubtedly impact resale premiums.

The current model could be potentially strengthened by a sentiment analysis of the Tweets scraped. We ignored the contents of the Tweet themselves (other than the presence of name of the given silhouette) and focused on the amount of traction the Tweet was generating via likes and retweets; however, a sentiment analysis of the Tweets could potentially mitigate the inability to quantify characteristics such as aesthetic beauty. Additionally, it would be interesting to design specific models for certain categories of popular sneakers in the secondary market (Yeezys (Kanye West's signature shoe), Jordans). Taking a deeper dive into such categories and specifying more specific models could yield better results that our approach was not able to capture. There are many different types of Yeezy sneakers – and investigating how Kanye West's album sales, for instance, impact resale prices within the context of a Yeezy model could be worthwhile.

To our knowledge, this paper is the first investigation of sneaker resale prices using rigorous econometric and statistical methods. The insights gained can be beneficial for those attempting to make a profit reselling or investing in sneakers in the secondary resale market. A quick Twitter keyword search can aid in gauging the amount of "hype" a particular sneaker is generating. Celebrity and Brand Collaborations tend to command high resale premiums in the secondary market as well.

References

Adams, Dexter (2016) “The Yeezy Effect: A Look at Adidas Before and After Kanye West”, Retrieved from https://www.kubashi.com/kicks/adidas-kanye-west

Bloomquist, G. & Worley, L. (1981). Hedonic prices, demands for urban housing attributes and benefit estimates, Journal of Urban Economics, vol. 9, pp. 212-221.

Butler, R. V. (1982). The specification of hedonic indexes for urban housing, Land Economics, vol. 58, pp. 94-108.

Cassidy, Nicholas George (2018). “The Effect of Scarcity Types on Consumer Preference in the High-End Sneaker Market,” 47.

Diewert, E., Heravi, S., Silver, M., 2008. Hedonic imputation versus time dummy hedonic indexes. National Bureau of Economic Research.

Dunne, Brendan (2017). “The Yeezy Crash of 2017”, Retrieved from https://solecollector.com/news/2017/06/adidas-yeezy-powerphase-calabasas-resell-price

Jiang, L., Phillips, P.C.B., Yu, J., 2015. New methodology for constructing real estate price indices applied to the Singapore residential market. J. Bank Financ. 61, S121–S131. doi:10.1016/j.jbankfin.2015.08.026

Jones, R. (2018, February 8). GOAT and Flight Club Merge to Become ‘World’s Largest Sneaker Marketplace’. Retrieved from http://footwearnews.com/2018/focus/athletic-outdoor/goat-flight-clubmerge-make-worlds-largest-sneaker-marketplace-492858/

Khaki, Sherief A., (2013). “Maximizing Revenue Selling Retro Sneakers on eBay: A Data Driven Approach,” 39.

Lancaster, K. (1966). A New Approach to Consumer Theory. Journal of Political Economy, 74(2), 132-157. Retrieved April 20, 2020, from www.jstor.org/stable/1828835

Luber, Josh (2016) “The History of Adidas Resell”, Retrieved from https://stockx.com/news/the-history-of-adidas-resell/

Lux, Moritz, Student Reutlingen-University De, and Dr Peter Bug., 2018 “Sole Value - The Sneaker Resale Market,” 41.

Monson, M., 2009. Valuation using hedonic pricing models. Cornell Real Estate Review 7, 10.

Mulligan, R.F., & Grube, A. (2007). MODELING MARKETS FOR SPORTS MEMORABILIA.

Rosen, S., 1974. Hedonic Prices and Implicit Markets: Product Differentiation in Pure Competition. Journal of Political Economy 82, 34–55. doi:10.1086/260169

Rui, Huaxia & Liu, Yizao & Whinston, Andrew. (2011). Whose and What Chatter Matters? The Impact of Tweets on Movie Sales. Decision Support Systems. 55. 10.2139/ssrn.1958068.

StockX. (2019). THE STOCK MARKET FOR THINGS. Retrieved from https://stockx.com/how-itworks

Weinswig, D. (2016, March 18). Sneaker Culture Fuels $1 Billion Secondary Market. Retrieved from https://www.forbes.com/sites/deborahweinswig/2016/03/18/sneaker-culture-fuels-1-billion-secondarymarket/#45eb10a77911

Welty, Matt (2017) “The Sneaker Industry’s Take on the Current Highs and Lows of Jordan Brand” Retrieved from http://www.complex.com/sneakers/2017/08/the-sneaker-industrys-take-on-the-current-highsand-lows-jordan-brand

Witkowska, D. Int Adv Econ Res (2014) 20: 281. https://doi.org/10.1007/s11294-014-9468-x

Xu, Jiakun (2017). Hedonic Modeling of Singapore's Resale Public Housing Market. Honors thesis, Duke University. Retrieved from https://hdl.handle.net/10161/14268.