ERASMUS UNIVERSITEIT ROTTERDAM

ERASMUS SCHOOL OF ECONOMICS

Art Investment: Prices and Returns of Prints

MSc. Candidate: Li Huang

Supervisor: Prof. Dr. Marshall (Xiaoying) Ma

Abstract

This research investigates the pricing model and the investment performance of prints. Our hedonic characteristics, like the size of prints, can explain around 68% of the hammer prices. To derive a precise estimation of price index in the prints market, we utilize two sophisticated methods: hedonic price function and repeat-sale regression, with full considerations. We find that the print returns derived from the hedonic method in the adjacent-period model have a geometric mean of -0.31% and an arithmetic mean of 4.05%, with a Sharpe ratio of 0.0356 during the year 2006-2016. Although the low Sharpe ratio implies the investment in prints is not an ideal choice, the negative correlations with like treasury bonds and commodities provide an opportunity for hedging.

Key words: alternative investment, prints, artwork, hedonic price model, repeat-sale regression

Contents

List of Tables

List of Figures

1 Introduction

As denoted by Wall Street Journal (2010), there are around 6% of total wealth invested in ”passion investment” like artworks, wine, and antiques. As is denoted by Langley (2014), among all such luxuries, the artwork is the most likely to be owned for its investment gains. Therefore, the asset pricing and investment returns of artworks have been attracting volume interest. However, it seems that the majority of art investment studies focus on paintings, while the performance of the investment in prints has been less investigated, especially after the 2008 crisis. To better understand the alternative investment in prints, we collect 280,409 transaction records of prints that occurred during 2006-2016 from 503 different public auction houses around the world.

The asset pricing of real properties often is often based on the Hedonic Pricing Function (HPF). In HPF, the dependent variable is the asset price and the independent variables consist of the asset’s characteristics. Besides asset pricing, we are also interested in investment returns. However, the study of art investment returns is not easy because of two material hurdles: illiquid transactions and heterogeneity of artworks (Mei and Moses, 2002). It is known that tradings of artworks are quite infrequent, only a small fraction of the stock are sold each year. Secondly, due to the variations in artwork qualities, the heterogeneity problem further distorts the calculation of the investment returns.

To circumvent the second problem in the study of investment returns, we need to control for the variances in properties’ characteristics. As denoted by Ginsburgh et al. (2006), there are usually tow approaches in estimating the return of artworks. One of the methods is called the hedonic price model, which incorporates all the measurable hedonic characteristics of the transacted artworks like size, attribution, and transaction date. By controlling for the variances in properties’ attributes, we can obtain a more precise return in the print market. The other approach is named the repeat sales method, which controls for the variance in characteristics through pairing the identical property (Scorcu and Zanola, 2011).

The research question is constructed as

Which variables can influence the sold probability and price of prints, and how well is the investment performance?

Firstly, we are curious that if the sold probability, or say the popularity of a print, could be predicted by hedonic characteristics of the artwork. To examine this hypothesis, we operate a hedonic regression on the dummy variable *sold*, which denotes whether this item is sold (assigned 1) or not (assigned 0) at the time. The first hypothesis is proposed as

Hypothesis 1: Hedonic variables do not influence the sales probability of artworks.

What we concern more about is the explanation power of these hedonic variables. Hence, this research will follow Ma et al. (2018) to utilize the hedonic price function (HPF) in examining the impact of hedonic variables on print prices. We propose the second hypothesis as

Hypothesis 2: Hedonic variables cannot be used to predict the prices of artworks.

Hedonic variables will be regressed on both nominal and real hammer prices, respectively, to inspect how these hedonic variables influence the price as well as construct the pricing model of prints.

After examining the price determinants of prints, we turn to the study of the investment returns in the print market. As aforementioned, there are mainly two methodologies that are utilized in research. However, both methods have their pros and cons. The advantage of the hedonic method is that all of the realized transactions can be employed in the estimation of returns; While the disadvantage is that the measurable hedonic characteristics can hardly capture all of the variances in the quality of artworks. However, Wallace and Meese (1997) denotes that, under the assumption that all omitted variables are orthogonal to those characteristics included in the regression, the coefficient of time dummies can present the constant-quality price fluctuations across years.

For the repeat sales method, the definite advantage is that all of the variances of characteristics can be controlled by tracking the identical artwork. However, there are several problems. First, since the artwork transactions are very infrequently, only incorporating the repeated sales will reduce the sample size dramatically. Furthermore, there are selection biases if constructing the dataset by choosing the art objects that have been sold at least twice.The transaction of the artwork would more likely occur if this artwork is popular, and hence the returns are higher than the average. Besides, as denoted by Renneboog and Spaenjers (2013), the repeat-sale objects are usually sold in anywhere in the world but resoled at arguably more expensive salesrooms. Hence, we conjecture that the returns based on the repeat-sale method are higher than that base on the hedonic method.

Hypothesis 3: The investment returns derived from the repeat sales method are higher than that calculated by the hedonic price model.

As denoted by Mei and Moses (2002), usually, art dealers suggest their clients buy the most expensive artwork they can afford, which presumes that masterpieces have higher expected returns in the market. The segmentation of prints at different price level is supposed to be important. As denoted by Renneboog and Spaenjers (2013), only a small fraction of investors can invest in higher-end works since they are indivisible. Besides, wealthy investors would neither purchase lower-end works since such art cannot signal their social status (Mandel, 2009). Since the more expensive artworks imply more speculation possibility, the distribution of returns is estimated to be skewed to the right. Hence, under such a condition, quantile regression would be very helpful (Zietz et al., 2008). By using quantile regression, the pricing of hedonic variables is allowed to differ across the price level distribution of hammer prices. Hence, we propose the hypothesis as

Hypothesis 4: There is no masterpiece effect in prints investment.

Lastly, we are concerned about whether the returns on prints are as competitive as traditional financial assets, or whether this alternative investment can provide a hedging opportunity for portfolio management. Hence we proposed the last hypothesis as follows

Hypothesis 4: The investment performances of prints are similar to other asset classes.

2 Literature Review

The price determinants of artworks have attracted the interest of researchers more recently. As Anderson (1974)'s study on art investment in paintings, this alternative investment class is often analyzed regarding prices, the investment return and risk, together with portfolio allocation with financial or other real assets.

There several different methodologies to derive art investment returns. For instance, Stein (1977) regards the auctioned artworks each year as random samples of the underlying artworks. Hence, he constructs price indices on the grounds of the yearly average transaction price. Another simple method is calculating the geometric mean return on repeated sales (Baumol, 1986). Repeat sale is the sale whenever the transaction of the identity artwork happens again. However, as denoted by Renneboog and Spaenjers (2013), these simple methods fail to capture the variations in artwork's quality for the construction of artwork price index.

One popular method that enables the controlling for artwork quality for the price index calculation is proposed by Bailey et al. (1963). They denote that, if the calculation of the index is based on sales prices of the same property at different times, the problem of quality differences in estimating price indexes can be solved. The methodology is named the repeat-sale regression (RSR).

Because the RSR method calculates price index based on repeat sales, it specifically controls for the quality differences among artworks. By using the repeat sale pairs, RSR can estimate the average return of a portfolio of properties in every period of time. This method is firstly proposed to solve the price index problem in real estate (Bailey et al., 1963), and then has been widely adopted in later studies in this field like Goetzmann (1992).

RSR also has been widely used in the construction of art price index (See Baumol (1986), Mei and Moses (2002), Renneboog and Spaenjers (2013), and Ma et al. (2018)). In the study of prints investment, Pesando (1993) follows RSR to calculate the price indexes Pesando (1993) for over 35,000 transactions from 1977 to 1992 and find that the nominal semiannual log index ranges from -0.025 to 1.577 (base time: 1977 first half-year), with a spike around 1990.

However, as is denoted by Renneboog and Spaenjers (2013), there are some problems in current RSR studies. First of all, the number of observations in repeat sales datasets is usually small because artwork transactions are infrequent. For instance, Mei and Moses (2002) can only collect 4,896 pairs over more than one hundred years. Wallace and Meese (1997) indicates that such a small dataset would make the estimators of RSR be distorted badly due to influential observations. The second problem is that only incorporating samples from repeat sales can suffer from selection bias, i.e., the sample selection is not random. Besides, in the research of Mei and Moses (2002), the paintings are firstly transacted anywhere, but the resale is at Southey's or Christie's New York, where the sales are among the most expensive transactions in the world (Rosen, 1974). Hence, the price index derived from this method is assumed to be upward biased.

As aforementioned, the other popular method to calculate the price index of property is called the Hedonic method. The hedonic price model is firstly proposed by Lancaster (1966) based on consumer theory. Then Rosen (1974) extends it to the study of real estate and indicates that Hedonic regressions are able to control for attribute variant in the property by ascribing implicit prices to the "utility-bearing characteristics." In the widely applied time-dummy Hedonic regression model, all observations can be pooled by incorporating time dummies. Hence the prices are regressed on a series of both value-determining characteristics and time dummies. If we assume that all omitted variables (OVB) are orthogonal to the variables included in the regression, the coefficients of time dummies present price trends by controlling the quality of the artwork.

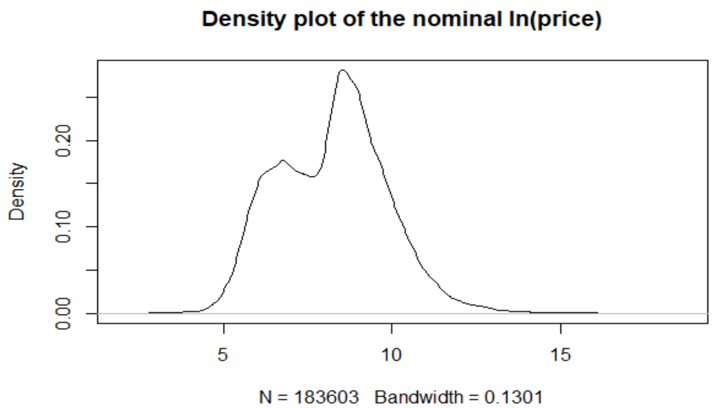
Compared to RSR, because Hedonic estimation incorporates not only repeat sales but also single sales, this method can reduce selection bias and increase the number of observations dramatically. Chanel et al. (1996) argues that considering there is a huge difference between the returns of unobserved private sales (See Guerzoni (1994)) and that of public auction transactions, selection bias can take an important place in computing returns. Hence, it is better to use all observations, including single sales, to calculate price indexes. Additionally, Renneboog and Spaenjers (2013) indicates that, since this Hedonic model can utilize available information more efficiently, it can provide a more reliable price index estimate than RSR.

3 Data

The information of prints transactions collected from 503 different public auction houses worldwide, including the famous Sotheby's and Christie's in London, and other small auction houses. Since there may be replicated entries, we first remove duplicated observations in R (R Core Team, 2019). The time span in this dataset is from 2006 to 2016. The total number of unique transactions amounts to 280,409 pieces in the original dataset.

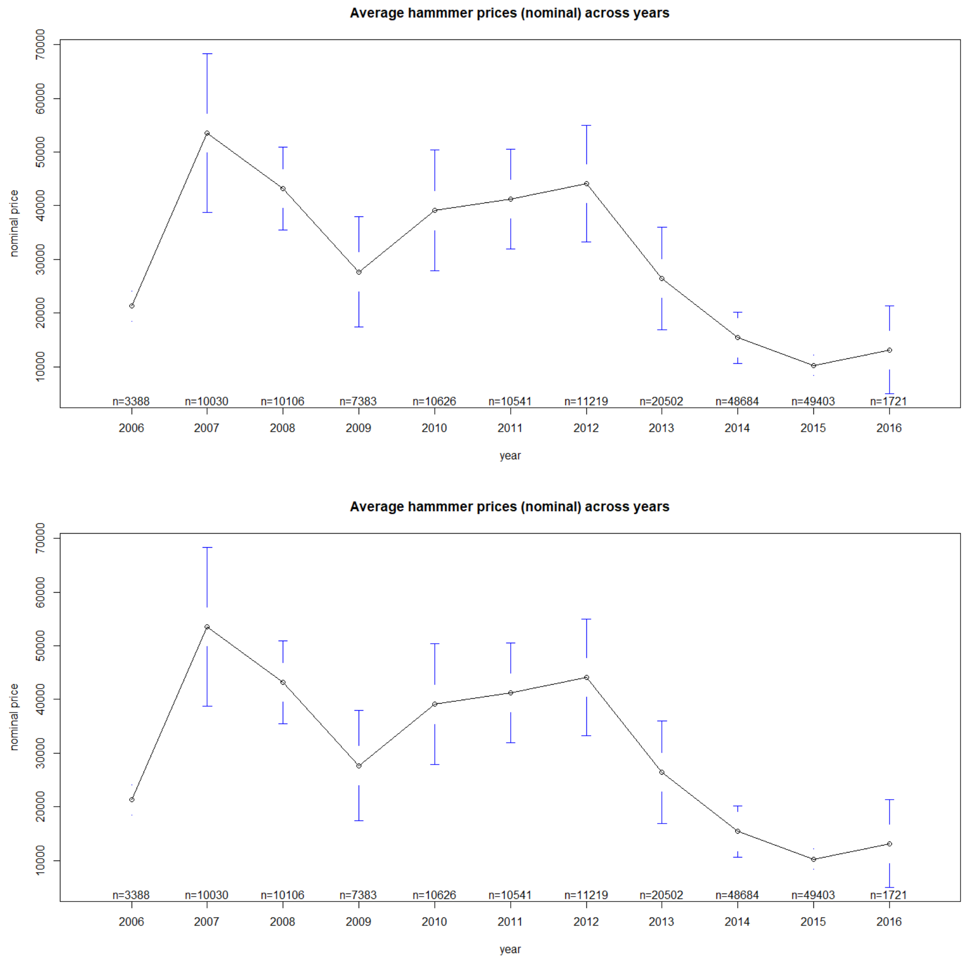
The hammer prices are the sold price, excluding transaction fees charged by the auction house. To uniform the currency of transaction records, we transform all other currencies to US dollars at the exchange rate of that day. Additionally, it provides that the hammer price is the nominal price at the year, we calculate the corresponding real price by Consumer Price Index (CPI) of US at the year, setting the base year as 2006. Moreover, because the prices of prints are skewed to the right greatly, we take the natural logarithms of hammer prices in regressions. The distribution of the nominal hammer price in USD is shown in figure 2. Note that there are 183,603 pieces with sold prices, and the 96,806 are records with no prices (either unsold or missing value).

Figure 1: Density Plot of the Nominal Price



To observe the price trends throughout the year, we plot the average price by years. See the figures below:

Figure 2: Average Prices across the Year



We follow Ma et al. (2018) to construct Hedonic models by incorporating all hedonic pricing variables in regressions. Our hedonic models include a wide range of attributes related to the artist, the artwork, and the sale. Table 1 presents the descriptive statistics of these hedonic variables.

Artist characteristics

The artist fixed effects will be included in hedonic regression, to capture the artist's uniqueness as well as reputation.

Artwork characteristics

We collect a number of price dominant variables as hedonic characteristics that present the attribution, authenticity, size, and topic categories:

Attribution dummies

Renneboog and Spaenjers (2013) denotes that attribution can be a crucial factor affecting the artwork prices. There are six different levels of attribution in the auction world: "attributed (to), studio (of), circle (of), school (of), after, and (in the) style (of)" (Renneboog and Spaenjers, 2013). Bandle (2016) denotes that such formulations enable the auction house to make uncertainty known by distinguishing creatorship based on the degree of certainty in artwork's attribution.

According to Bandle (2016), "attribute" means the auction house specialist doubt the facticity of the designated creator: it is likely as an original work, but it is not guaranteed. "studio" means that the in-house specialist believes that the artwork was made in the studio of the named creator and possibly under their supervision. "circle" indicate that the artwork is believed to be created by another artist, but was made during the lifetime of the designated creator as well as in their manner. "school" can only present that this artwork comes from the same period but not sure about the named creator. The formulation "after" denotes that the in-house specialist believes that the artwork was created by other artist based on the original work by the named artist. While "style" presents the uncertainty of the art work's create year. In our dataset, there are only 0.92% of the observations have such attribution.

Authenticity dummies

The dummy variable "signature" implies the print is signed, dated, or inscribed. In our dataset, around 78.13% of the observations were endowed signature. The variable "numbered" denotes whether the print has a creation number assigned by the creator.

Size

The variables height and width measures the size of prints in centimeters. We also include "height_sq" and "width_sq" as the square of height and width respectively.

Topic dummies

Since the price and returns can depend on the painting's topic, in this research, we also introduce the topic categories for prints. Based on the titles of prints, in total 11 categories are compiled, which are: people, abstract, untitled, urban, portrait, religion, animal, self-portrait, still life, nude, and landscape. The topic classification in text analysis is implemented by the dictionary for paintings in machine learning. Based on the result, there are still some titles that do not belong to any topic above, hence the unknown topic is taken as the omitted benchmark in regressions.

transaction characteristics

Month dummies

As denoted by Renneboog and Spaenjers (2013), the busiest auction months are from May to June and from November to December. Hence, we also include monthly dummies as there are supposed to be seasonal effects in prints auction.

Auction house dummies

For the assumption that the auction house will have an impact on the sale price, by following Ma et al. (2018), we make discrimination between auction houses that were influential throughout the sample period. Regarding the famous Sotheby's and Christie's, we distinguish between their London, New York, and other activities (eg., chr_london, chr_ny, chr_other). These two institutions take up about 21.8% of all observations. For the auction house Bothams and Phillips, we distinguish their London branch and other locations. Additionally, the sales by important European and American auction houses is grouped as *auction_eu* and *auction_us* respectively.

Table 1: Descriptive Statistics of Hedonic Variables

Variable	Obs.	Mean	Std. Dev.	Min	Max
Artwork Characteristics					
Authenticity dummies					
signature	280,409	0.78132	0.4134	0	1
provenance	280,409	0.1307	0.3371	0	1
numbered	280,409	0.3509	0.4773	0	1

Continued on next page

Table 1 – continued from previous page

Variable	Obs.	Mean	Std. Dev.	Min	Max
Attribution dummies					
attributed	280,409	0.0013	0.0364	0	1
studio	280,409	0.0000	0.0033	0	1
circle	280,409	0.0000	0.0050	0	1
school	280,409	0.0000	0.0027	0	1
after	280,409	0.0079	0.0883	0	1
style	280,409	0.0000	0.0027	0	1
Size variables					
height	280,409	58.0099	50.1325	0.64	997
width	280,409	56.3840	51.2396	0.2	995
height_sq	280,409	5878.4110	25438.7100	0.4096	994009
width_sq	280,409	5804.6470	26424.0800	0.04	990025
Topic dummies					
topic_PEOPLE	280,409	0.0443	0.2058	0	1
topic_ABSTRACT	280,409	0.0110	0.1045	0	1
topic_UNTITLED	280,409	0.0527	0.2234	0	1
topic_URBAN	280,409	0.0470	0.2116	0	1
topic_PORTRAIT	280,409	0.0109	0.1038	0	1
topic_RELIGION	280,409	0.0100	0.0993	0	1
topic_ANIMAL	280,409	0.0093	0.0959	0	1
topic_SELF-PORTRAIT	280,409	0.0058	0.0762	0	1
topic_STIL_LIFE	280,409	0.0056	0.0744	0	1
topic_NUDE	280,409	0.0100	0.0993	0	1
topic_LANDSCAPE	280,409	0.0148	0.1209	0	1
Transaction Dummies					
Auction house dummies					

Continued on next page

Table 1 – continued from previous page

Variable	Obs.	Mean	Std. Dev.	Min	Max
auction_eu	280,409	0.0759	0.2649	0	1
auction_us	280,409	0.0912	0.2879	0	1
soth_london	280,409	0.0232	0.1506	0	1
soth_ny	280,409	0.0472	0.2122	0	1
soth_other	280,409	0.0136	0.1157	0	1
chr_london	280,409	0.0280	0.1649	0	1
chr_ny	280,409	0.0657	0.2478	0	1
chr_other	280,409	0.0400	0.1960	0	1
bon_london	280,409	0.0093	0.0961	0	1
bon_other	280,409	0.0329	0.1785	0	1
Month dummies					
January	280,409	0.0328	0.1781	0	1
February	280,409	0.0456	0.2085	0	1
March	280,409	0.0711	0.2570	0	1
April	280,409	0.0914	0.2882	0	1
May	280,409	0.1232	0.3287	0	1
June	280,409	0.1110	0.3142	0	1
July	280,409	0.0287	0.1671	0	1
August	280,409	0.0156	0.1237	0	1
September	280,409	0.0746	0.2627	0	1
October	280,409	0.1293	0.3355	0	1
November	280,409	0.1617	0.3682	0	1
December	280,409	0.1151	0.3192	0	1

Repeat Sales Samples

Calculating the returns of artwork in the repeat sales method requires that the item must be sold at least twice. Since there is no unique id denoting the identity of artwork, we pair the prints of the same artist, title, height, and width. Since there are usually different editions of prints with the same title and size by the same artist, and the sorted results confirm it with some unreasonable returns in very short periods. Based on that, we additionally control for the provenance variables. In this case, the pairs are the prints having the same inner factors (artist and artwork characteristics) but different transaction variables. The number of pairs for repeat sale amounts to 3,361, sorting the holding period as at least 180 days. The operation was done in R Core Team (2019).

4 Methodology

To inspect the asset pricing model and investment returns of prints, we conduct several hedonic regressions. Fixed effects are supposed to be included in regressions, since there may be omitted variables that are unobservable, or because of heterogeneity (for instance, there are many unmeasurable characteristics in artists that may influence the price of their artworks). Incorporating fixed effects can alleviate the problem by allowing different "intercept" for each different group. In this research, we control for the fixed effect in auction house branch level, artist, year, and month, with two-way clustering in the year and auction houses.

Although it usually assumes that the error terms for individual observations are independent and identically distributed (iid), however, there is often autocorrelation of the same observations in different periods or in different auction houses. Hence, the estimate of standard errors should be cluster-robust so as to reduce the standard error. Considering this, we employ a two-way clustering method in the year and auction house in hedonic regressions by using the package of Correia (2016) on STATA (StataCorp., 2015).

4.1 Hedonic Linear Probability Regression

To inspect the sold probability in Hypothesis 1, we regress Hedonic variables on the dummy variable *Sold*, which is a dummy variable indicating whether the item is sold at the time. The function is as follows:

$$Sold_{it} = \alpha + \sum_{m=1}^{M} \beta_m X_{mit} + \sum_{t=1}^{T} \gamma_t D_{it} + \epsilon_{it}, \tag{1}$$

where $Sold_{it}$ equals 1 if the item i is sold at time t, 0 otherwise. X_{mit} denotes the characteristic m of the item i at time t. D_{it} is a time dummy which equals one if the transaction happened at the time t. The marginal effect of characteristic m on the sold probability is estimated as β_m. We adopt linear probability model instead of logit model to get more intuitional explanations of characteristic.

4.2 Hedonic Price Function

By following the Hedonic price function (Rosen, 1974), we regress characteristic variables on the natural logarithm of real and nominal USD hammer prices respectively:

$$ln(P_{it}) = \alpha + \sum_{m=1}^{M} \beta_m X_{mit} + \sum_{t=1}^{T} \gamma_t D_{it} + \epsilon_{it}, \quad (2)$$

where P_{it} is the real (or nominal) hammer price of the print i at time t, while the Hedonic characteristics X and time dummies D on the right side are the same as in Equation (1).

4.3 Investment Returns

As aforementioned, we will adopt the two most popular approaches to analyze the artwork investment returns, i.e., repeat-sales regression (RSR) and Hedonic method.

4.3.1 Investment Returns (Hedonic Method)

In Hedonic method, all of the observations can be incorporated in the regression dataset by controlling for Hedonic variables. Assume that other implied characteristics that are not included, are orthogonal to the current hedonic variables in regression (Wallace and Meese, 1997), the time coefficients can inflect the price level of the year. Following the work of Renneboog and Spaenjers (2013), we firstly construct the art index of prints by taking the time variation of the year. The hedonic index in year t can be defined as

$$\Pi_t = exp(\hat{\gamma_t}) \times 100, \quad (3)$$

with the initial year equaling 0. Thereafter, the estimated annual return of the year t is:

$$r_t = \frac{\Pi_t}{\Pi_{t-1}} - 1. \quad (4)$$

It is notable that since the log transformation is prior to the regression, the index is the geometric mean of prices over time instead of arithmetic mean. According to Silver and Heravi (2007), if the time variations are present in the heterogeneity-controlled dispersion of prices, the transformation bias is considerable. Assuming that the residuals of hedonic regression in Equation (2) are normally distributed in each time t, the transformation bias can be adjusted by constructing the corrected are index as (Triplett, 2004):

$$\Pi_t^* = exp[\hat{\gamma}_t + \frac{1}{2}(\hat{\sigma_t^2} - \hat{\sigma_0^2})] \times 100, \tag{5}$$

where $\hat{\sigma_t^2}$ and $\hat{\sigma_0^2}$ denote the estimated variance of the regression residuals for the transactions in year 0 and t respectively. Hence the corrected estimated return in year t equals to

$$r_t^* = \frac{\Pi_t^*}{\Pi_{t-1}^*} - 1. \tag{6}$$

4.3.2 Investment Returns (RSR method)

The returns of repeat sales can be derived by examining the price changes in two sales of the same real property. Since the idiosyncratic characteristics remain between sales, the differences in price can be explained by time dummies. However, as is discussed by Renneboog and Spaenjers (2013), this method only takes the artworks that have been sold at least twice, resulting in survivorship biases. That is to say, the art index derived from repeat sales samples is assumed to be upward biased. However the advantage of RSR is that the index derived is based on the price changes of the identical prints that the variance in quality of the prints can be controlled(Mei and Moses, 2002).

As shown in Goetzmann (1993), the continuously compounded return of the property i in time t can be represented by u_t , the return of the index of the artwork, and $\eta_{i,t}$, an error term which implies the idiosyncratic return. The first stage of RSR is as follows:

$$r_{i,t} = \mu_t + \eta_{i,t}, \tag{7}$$

where μ_t can be regarded as the average return of prints in time t. The art index u over the period $t = 1, ..., T$ is a $(T \times 1)$ vector in which the individual elements are u_t. Let $P_{i,b}$ and $P_{i,s}$ denote the purchase and sale prices of the property i respectively, while the purchase time is b_i and the sales time is s_i. Hence, in the second stage of RSR, the log price relative for the artwork i held between the purchase date b_i and sales date s_i can be present as:

$$\begin{aligned} r_i = ln\left(\frac{P_{i,s}}{P_{i,b}}\right) &= \sum_{t=b_i+1}^{s_i} r_{i,t} \\ &= \sum_{t=b_i+1}^{s_i} \mu_t + \sum_{t=b_i+1}^{s_i} \epsilon_{i,t}. \end{aligned} \tag{8}$$

The last stage of RSR is estimating the vector μ. Assume r is the $(N \times 1)$ vector of the log price relatives for the N pairs of repeated sales. The simplest estimate of the logged return is:

$$\hat{\mu}_{ols} = (X'X)^{-1}X'y. \tag{9}$$

This ordinary least squares (OLS) requires the assumption that the residual $\varepsilon_{i,t} \sim i.i.dN(0, \sigma_i^2)$. However, ignoring the varying variances can lead to the overweight of those observations consist less information about the estimate of μ, and underweight of those contain more information. Furthermore, there are many problems that the simple unweighted regression cannot tackle with. Hence, we require more sophisticated models like weighted repeated sales functions which can be represented as

$$\hat{\mu} = (X'\Omega^{-1}X)^{-1}X'\Omega^{-1}r, \tag{10}$$

where X is a matrix with N rows of observations and T columns with time dummies equal 1 within the holding period. Ω is a weight matrix, where the weights can be based on, or the Generalized Least Squares (GLS) by using the times between the pair sales (Goetzmann, 1993), or residuals from a three-stage Weighted Repeat Sales model (WRS) proposed by Case and Shiller (1987)

As denoted by Webb (1981), the error term $\varepsilon_{i,t}$ is not independent with the length of holding period. Hence, he suggested a GLS model where the weight matrix comprises the time span between purchase and sales time. However, Case and Shiller (1987) indicate that, apart from the error variance will increase with longer holding period, there may also be fixed components of the error variances which are not correlated with the length of holding period. In order to jointly control for the fixed and the varying components of error variance, they proposed a three-stage procedure.

The first step is based on the ordinary least square (OLS) regression as in Equation (9). In the second step, we regress the squared errors derived from the regression above to estimate the effect of error variance compounding linearly through time:

$$\hat{e_i}^2 = \alpha + \theta(s_i - b_i) + \epsilon_i, \tag{11}$$

where α allows the intercept and $(s_i - b_i)$ denotes the time span between the purchase time and sales time. As indicated by Case and Shiller (1987), the slop θ can be explained as the residual risk per period.

The last step is to fit the weight matrix Ω with the residuals ϵ from Equation (14), in the format of

$$\Omega = diag\left(\frac{1}{\hat{e}^2}\right), \tag{12}$$

where Ω is a diagonal matrix with the diagonal elements $1/\hat{e}^2$. And then redo a RSR shown in Equation (10), with the additional weight matrix Ω derived from Equation (12) above.

Similarly, since the logarithmic transformation causes a downward bias of the arithmetic mean, we adjust the bias by incorporating the estimated variance, assuming that the art investment returns are log normally distributed (Goetzmann, 1993):

$$r_t^* = exp(\mu_t + \frac{1}{2}\sigma^2) - 1, \tag{13}$$

where σ_t is the standard error estimated the second stage of RSR (Case and Shiller, 1987).

5 Empirical Results

5.1 Hedonic Characteristics to Sales Probability and Price

To answer the hypothesis 1, we inspect a large number of hedonic characteristics that may influence the sales probability of prints. The results of the hedonic linear probability regression on sales dummy is present in the second column of Table 2.

Based on the results, we can find that the *literature* and *exhibited* dummies have a significant positive impact on the probability of sold at the 5% level, with a slightly more than 3% increase in sold probability respectively. The authenticity dummies *signature*, *numbered* and *provenance* can also significantly increase the probability with 3.15%, 2.37%, and 3.99% respectively. The topic of prints can also influence the sold probability, for instance, if the topic is UNTITLED or PORTRAIT, the chance of being sold will decrease, while if the topic is URBAN or ANIMAL, there is more chance that the print will be sold in the auction house. Furthermore, as compared to the influence of the aforementioned variables, the auction house branch level can influence the sold probability stronger. The coefficients of auction_us, soth_london, soth_ny, soth_other, chr_london, chr_ny, and chr_other are all positive, with larger magnitude, and most of them are significant at the 1% level. It is obvious that the hedonic variables have an influence on the sales probability. Even though the F statistic for the joint hypothesis is not available due to dropped variables as too few clusters, it is extremely likely to reject the first hypothesis due to the significant variables.

Table 2: Hedonic Regressions on Sold Probability and Prices

VARIABLES	**Sold**	**Nominal**	**Real**
signature	0.0315***	0.1929***	0.1929***
	(0.0086)	(0.0411)	(0.0411)
numbered	0.0237**	0.0353	0.0353
	(0.0082)	(0.0395)	(0.0395)
provenance	0.0399***	0.3652***	0.3652***
	(0.0110)	(0.0505)	(0.0505)

Continued on next page

Table 2 – continued from previous page

VARIABLES	Sold	Nominal	Real
literature	0.0362**	0.3078***	0.3078***
	(0.0155)	(0.0528)	(0.0528)
exhibited	0.0307**	0.7205***	0.7205***
	(0.0099)	(0.0905)	(0.0905)
attributed	-0.0642*	-0.5682***	-0.5682***
	(0.0295)	(0.1221)	(0.1221)
studio	0.5194*	0.2942	0.2942
	(0.2520)	(0.3333)	(0.3333)
circle	-0.2396	-0.9614	-0.9614
	(0.2631)	(0.7041)	(0.7041)
school	0.3721***	1.7603***	1.7603***
	(0.0875)	(0.4335)	(0.4335)
after	-0.0358	0.2498**	0.2498**
	(0.0307)	(0.0981)	(0.0981)
style	0.1640***	-0.8429***	-0.8429***
	(0.0373)	(0.1393)	(0.1393)
height	0.0001	0.0049***	0.0049***
	(0.0001)	(0.0008)	(0.0008)
height_sq	-0.0000**	-0.0000***	-0.0000***
	(0.0000)	(0.0000)	(0.0000)
width	0.0002**	0.0064***	0.0064***
	(0.0001)	(0.0010)	(0.0010)
width_sq	-0.0000**	-0.0000***	-0.0000***
	(0.0000)	(0.0000)	(0.0000)
topic_PEOPLE	0.0006	0.0648**	0.0648**
	(0.0065)	(0.0233)	(0.0233)
topic_ABSTRACT	-0.0199	-0.2337***	-0.2337***
	(0.0163)	(0.0564)	(0.0564)
topic_UNTITLED	-0.0321***	-0.1247***	-0.1247***
	(0.0088)	(0.0179)	(0.0179)

Continued on next page

Table 2 – continued from previous page

VARIABLES	Sold	Nominal	Real
topic_URBAN	0.0180*	0.0970***	0.0970***
	(0.0094)	(0.0261)	(0.0261)
topic_PORTRAIT	-0.0262**	0.1080*	0.1080*
	(0.0107)	(0.0563)	(0.0563)
topic_RELIGION	-0.0029	0.0082	0.0082
	(0.0141)	(0.0486)	(0.0486)
topic_ANIMAL	0.0228**	-0.0859	-0.0859
	(0.0093)	(0.0734)	(0.0734)
topic_SELF-PORTRAIT	0.0145	0.2140***	0.2140***
	(0.0154)	(0.0614)	(0.0614)
topic_STILL_LIFE	0.0189	0.0639*	0.0639*
	(0.0154)	(0.0345)	(0.0345)
topic_NUDE	0.0082	0.0396	0.0396
	(0.0112)	(0.0224)	(0.0224)
topic_LANDSCAPE	0.0023	0.0179	0.0179
	(0.0110)	(0.0299)	(0.0299)
auction_eu	-0.0075	0.1903**	0.1903**
	(0.0220)	(0.0776)	(0.0776)
auction_us	0.0832**	0.1948	0.1948
	(0.0263)	(0.1123)	(0.1123)
soth_london	0.0756***	0.8808***	0.8808***
	(0.0157)	(0.1369)	(0.1369)
soth_ny	0.1055***	0.7583***	0.7583***
	(0.0190)	(0.1463)	(0.1463)
soth_other	0.0646**	0.6150***	0.6150***
	(0.0281)	(0.1235)	(0.1235)
chr_london	0.1139***	0.8147***	0.8147***
	(0.0318)	(0.1346)	(0.1346)
chr_ny	0.1521***	0.5161***	0.5161***
	(0.0260)	(0.1196)	(0.1196)

Continued on next page

Table 2 – continued from previous page

VARIABLES	Sold	Nominal	Real
chr_other	0.0929***	0.4392**	0.4392**
	(0.0266)	(0.1559)	(0.1559)
bon_london	0.0011	0.4497**	0.4497**
	(0.0213)	(0.1630)	(0.1630)
bon_other	0.0035	0.0469	0.0469
	(0.0273)	(0.1158)	(0.1158)
2007.year	-0.1257***	0.1964***	0.1685***
	(0.0143)	(0.0358)	(0.0358)
2008.year	-0.2937***	-0.0128	-0.0785
	(0.0365)	(0.0497)	(0.0497)
2009.year	-0.2995***	-0.0558	-0.1180**
	(0.0159)	(0.0426)	(0.0426)
2010.year	-0.2611***	0.0229	-0.0556
	(0.0121)	(0.0391)	(0.0391)
2011.year	-0.2537***	0.1066**	-0.0030
	(0.0138)	(0.0426)	(0.0426)
2012.year	-0.2854***	0.0469	-0.0831**
	(0.0140)	(0.0346)	(0.0346)
2013.year	-0.2549***	-0.2384***	-0.3829***
	(0.0069)	(0.0611)	(0.0611)
2014.year	-0.2587***	-0.6423***	-0.8029***
	(0.0068)	(0.0945)	(0.0945)
2015.year	-0.2750***	-0.7423***	-0.9042***
	(0.0071)	(0.0912)	(0.0912)
2016.year	-0.1063***	-0.8160***	-0.9903***
	(0.0300)	(0.1335)	(0.1335)
February	0.0368	-0.0395	-0.0395
	(0.0601)	(0.0961)	(0.0961)
March	-0.0109	-0.0483	-0.0483
	(0.0427)	(0.0752)	(0.0752)

Continued on next page

Table 2 – continued from previous page

VARIABLES	Sold	Nominal	Real
April	-0.0011	0.0975	0.0975
	(0.0447)	(0.0921)	(0.0921)
May	-0.0054	0.1193	0.1193
	(0.0393)	(0.0714)	(0.0714)
June	-0.0203	0.1838	0.1838
	(0.0459)	(0.1020)	(0.1020)
July	-0.0307	-0.2025	-0.2025
	(0.0398)	(0.1145)	(0.1145)
August	-0.0159	-0.1913	-0.1913
	(0.0398)	(0.1172)	(0.1172)
September	-0.0226	-0.1228	-0.1228
	(0.0427)	(0.1113)	(0.1113)
October	-0.0292	0.0178	0.0178
	(0.0419)	(0.0994)	(0.0994)
November	-0.0174	0.0384	0.0384
	(0.0416)	(0.1204)	(0.1204)
December	-0.0677	-0.0498	-0.0498
	(0.0376)	(0.1246)	(0.1246)
Constant	0.8617***	7.5799***	7.5799***
	(0.0412)	(0.2198)	(0.2198)
Observations	271,063	176,482	176,482
R-squared	0.1585	0.6781	0.6863
Artist FE	YES	YES	YES
Auction House FE	YES	YES	YES
Year FE	YES	YES	YES
Month FE	YES	YES	YES

Robust standard errors in parentheses

*** p<0.01, ** p<0.05, * p<0.1

The asset pricing model of prints incorporates all the aforementioned hedonic variables to get a better explanation of the asset price. The results of HPF on nominal price and real price are shown in the last two columns of Table 2, respectively. There are no differences in the coefficients, including both the magnitude and the significance but the slight disparity in R-square. It is obvious that since the real prices are adjusted by the yearly CPI, this change can be mostly incorporated in the year fixed effects (Year FE).

From the table, we can find that the majority of the intrinsic factors have a significant impact on the prices of prints. Among the authenticity dummies, all but "numbered" are significantly positive at the 1% level. It is reasonable that authentic artworks are more expensive, ceteris paribus. The size dummies show that the price will increase with both height and width, but with a slight diminishing effect since the coefficient of squared height and squared width are negative. The topic of prints can also influence the prices; some topics like PEOPLE, URBAN, SELF-PORTRAIT have significantly higher prices holding other factors constant; while topics like ABSTRACT, and UNTITLED correspond to lower prices. The auction place can also affect the price: if the prints are sold in famous auction houses like Sotheby's and Christie's, the prices are usually higher, ceteris paribus. For instance, if the print was sold at Sotheby's London branch, the estimated sold price would be 141.48 USD higher. Similarly, the second hypothesis is rejected despite the missing F statistic.

5.2 Investment Return - Hedonic Price Model

Table 4 shows the annualized investment return of prints calculated based on the regressions in Table 3. γ_t denotes the coefficient of year t in the regression above, Π_t and r_t are the uncorrected price index and yearly return in year t; while Π_t^* and r_t^* are the corrected price index and yearly return in year t adjusted based on the Equation (6) and Equation (7) respectively. The arithmetic mean of nominal returns is -0.39%, as compared to -2.08% of real returns. The geometric annual return from 2006 to 2016 is -5.1% and -6.75% for nominal return and real return, respectively, during the sample period.

Table 3: Price Indices and Returns - Hedonic Price Model

a. Annualized Nominal Return of Prints

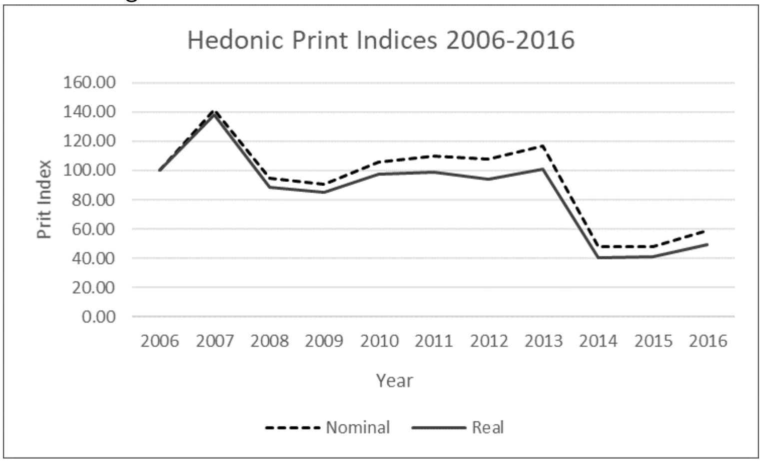
	γ	SD	σ^2	Π_t	$r_t(\%)$	Π_t^*	$r_t^*(\%)$
2006	0.0000	0.9328	0.8702	100.00	NA	100.00	NA
2007	0.1964	1.0850	1.1772	121.70	21.70	141.89	41.89
2008	-0.0128	1.0478	1.0978	98.73	-18.88	94.89	-33.12
2009	-0.0558	1.0081	1.0162	94.57	-4.21	90.79	-4.32
2010	0.0229	1.0387	1.0789	102.31	8.19	105.57	16.29
2011	0.1066	1.0303	1.0615	111.24	8.73	110.28	4.46
2012	0.0469	1.0553	1.1136	104.80	-5.79	107.57	-2.46
2013	-0.2384	1.3799	1.9041	78.79	-24.82	116.99	8.75
2014	-0.6423	1.3102	1.7165	52.61	-33.23	47.90	-59.05
2015	-0.7423	1.3176	1.7361	47.60	-9.53	48.07	0.35
2016	-0.8160	1.5238	2.3219	44.22	-7.10	59.27	23.30
							-0.39

b. Annualized Real Return of Prints

	γ	SD	σ^2	Π_t	$r_t(\%)$	Π_t^*	$r_t^*(\%)$
2006	0.0000	0.9332	0.8709	100	NA	100	NA
2007	0.1685	1.0850	1.1773	118.36	18.36	137.95	37.95
2008	-0.0785	1.0480	1.0983	92.45	-21.89	88.87	-35.58
2009	-0.1181	1.0083	1.0166	88.86	-3.88	85.30	-4.01
2010	-0.0556	1.0388	1.0791	94.59	6.44	97.59	14.40
2011	-0.0028	1.0305	1.0619	99.72	5.42	98.86	1.30
2012	-0.0832	1.0554	1.1138	92.02	-7.72	94.44	-4.47
2013	-0.3829	1.3798	1.9037	68.18	-25.90	101.21	7.17
2014	-0.8057	1.3077	1.7100	44.68	-34.47	40.56	-59.93
2015	-0.9041	1.3175	1.7357	40.49	-9.37	41.02	1.13
2016	-0.9918	1.5237	2.3217	37.09	-8.40	49.72	21.22
							-2.08

The line chart of the adjusted nominal and real print indices is shown in Figure 3 below. It is obvious that the nominal returns are higher than the real returns due to inflation.

Figure 3: Hedonic Print Price Indices 2006-2016



5.3 Investment Return - Repeat Sales Method

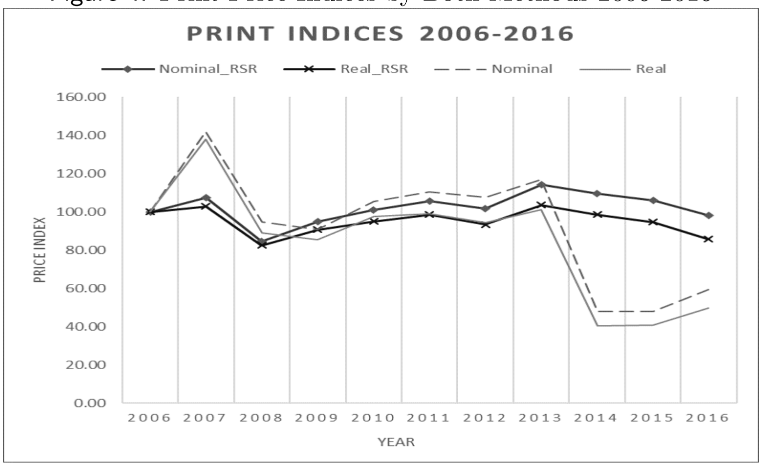
The total amount of collected pairs of repeat sales is 3,361, of which the holding period is greater than 180 days. We calculate the price indices in the weighted repeat sales model (WRS) proposed by Case and Shiller (1987) as aforementioned. The indices and returns of the print market by year is shown in Table 4 below.

Besides, we plot both the hedonic and repeat-sales price indices for comparison in Figure 4. From the plot, we find that the repeat-sale indices are flatter than hedonic indices, specifically in 2007 as well as during 2014-2016. From 2008 to 2013, the price indices derived from both methods are quite similar, especially in real price indices where there are even overlap between *Real_RSR* and *Real* in the figure below.

Table 4: Price Indices and Return by Repeat Sales Method

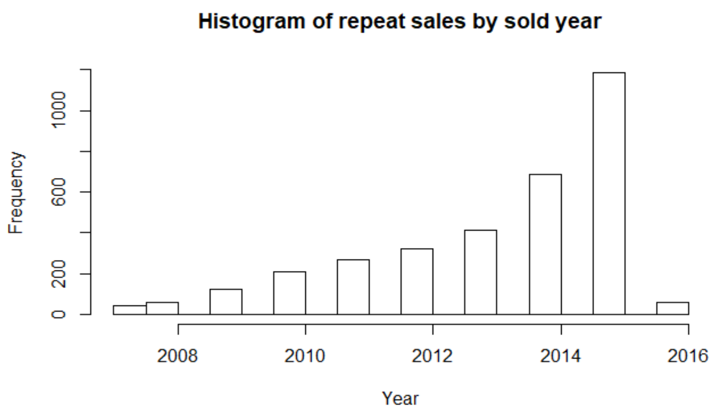
	Nominal			Real		
	$\hat{\mu_t}$	r_t	r_t^*	$\hat{\mu_t}$	r_t	r_t^*
2006	19.7%	21.8%	24.3%	23.3%	26.3%	28.5%
2007	4.9%	5.1%	7.3%	1.0%	1.0%	2.8%
2008	-25.9%	-22.8%	-21.2%	-23.8%	-21.2%	-19.8%
2009	9.3%	9.8%	12.1%	7.9%	8.3%	10.2%
2010	4.3%	4.4%	6.6%	2.9%	2.9%	4.7%
2011	2.5%	2.5%	4.7%	1.9%	2.0%	3.7%
2012	-6.0%	-5.8%	-3.8%	-7.1%	-6.9%	-5.3%
2013	9.4%	9.8%	12.1%	8.6%	9.0%	10.9%
2014	-6.2%	-6.0%	-4.0%	-6.7%	-6.5%	-4.8%
2015	-5.4%	-5.3%	-3.3%	-5.8%	-5.6%	-4.0%
2016	-9.8%	-9.4%	-7.4%	-11.8%	-11.1%	-9.6%
			2.5%			1.6%

Figure 4: Print Price Indices by Both Methods 2006-2016



The reason why the RSR indices are lower is most likely because of the methodology. Because the sample period is from 2006 to 2016, there are fewer resales in earlier years, which can be confirmed from the histogram of repeat sales by year in Figure (xxx) below. Hence, unless the sample selection includes more years before 2007, the RSR index for 2007 cannot capture the majority of the repeat sales data but those prints bought in 2006 or early 2007. Therefore, we can say that the RSR index in 2007 is probably not accurate due to the severer distortion in sample selection as aforementioned.

The RSR indices for 2014-2016 can better capture the repeat sales market because there are more "first-time" sales available in previous years within our sample window. However, because the survivorship bias may become stronger when the prints market is bear in this period. Investors may hold the "ought to have been depreciated" artworks to wait for better prices due to loss aversion. It is also very possible that the hedonic indices cannot precisely measure the investment gains: we can observe that the total transaction volume increased along 2012-2016, but the average price decreased. It may be due to some changes in the overall average quality of prints sold in the whole market that can not be measured, and such changes can negatively influence the hammer prices. For instance, the auction houses start to record all transactions, or they start to collect some prints of lower quality than before, which are not included in our hedonic variables. In this case, the RSR indices during 2014-2016 can better present the price index of the print market. It is notable that the results for the year 2016 are neither representative because the dataset only consists of the transactions that occurred in the first quarter of 2016.



5.4 Robustness Checks

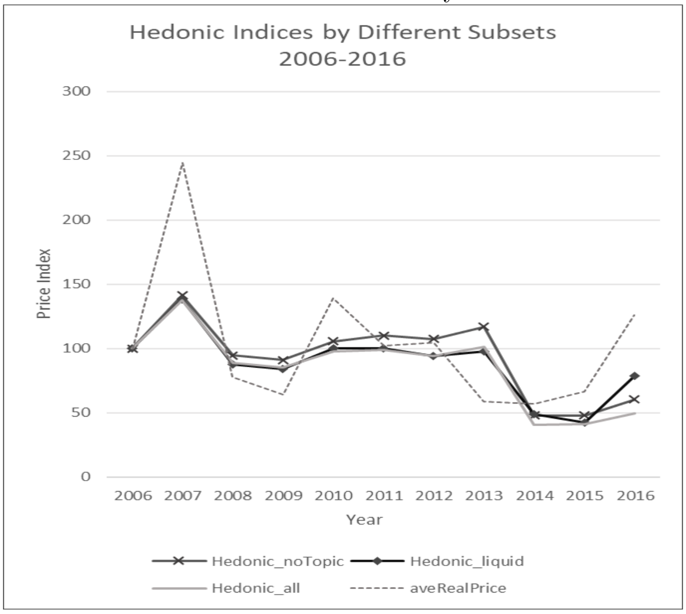
To ensure the robustness of the results, we first conduct hedonic regressions based on two other different setups:
1. dropping the topic dummies because the text analysis based on the dictionary may not be very accurate (reduce the number of variables).
2. excluding the artists whose transaction records are less than 100 pieces since these artists are less liquid (reduce the number of observations).

Besides, the regressions above assume that the coefficient of hedonic variables hold constant across the entire sample window (2006-2016), which may be too strong because the shadow prices of prints' attributes (in other words, the tastes) may vary from time to time. Hence, to alleviate this problem, adjacent-period models can be applied Triplett (2004): by separating the sample in sub-periods, the coefficients of hedonic variables are allowed to fluctuate. We follow Renneboog and Spaenjers (2013) to operate the divided HPF for every two coherent years within the sample window and then join together two print indices by rescaling as upon that in 2006.

5.4.1 Sample Setups

The hedonic price function of the different setups (excluding topic dummies and excluding illiquid artworks) can be referred to in Table 8 and 9 of the Appendix. Here we mainly discuss the investment returns and price indices based on different sample setups. The results are shown in Figure 5 below.

Figure 5: Robustness Check: Hedonic Indices by Different Subsets 2006-2016



The dash line presents each mean of real price across the year, and the rest three solid line are all hedonic indices. As present in Figure 5, these three different setups have similar results. In the last year 2016, the price increased a lot while only one hedonic index increases a lot. This index, *Hedonic_liquid*, is derived from the work of the artists whose prints are sold at least 100 times. It may imply that the higher price level in 2016 is mainly driven by the higher prices of "liquid" artworks. On the other side, the deviation of *Hedonic_noTopic* shows that failing to include omitted variables (OVB) can distort the calculation of price index: the assumption that all other omitted characteristics are orthogonal to the included characteristics is probably too strong.

5.4.2 Adjacent-Period Model

As aforementioned, the adjacent-period model allows the fluctuation of coefficients across the year. In this way, the estimation of a hedonic price index would be more accurate. The results are shown in Table 5 below. The arithmetic average return amounts to 4.05%, and the geometric annualized return is -0.308%.

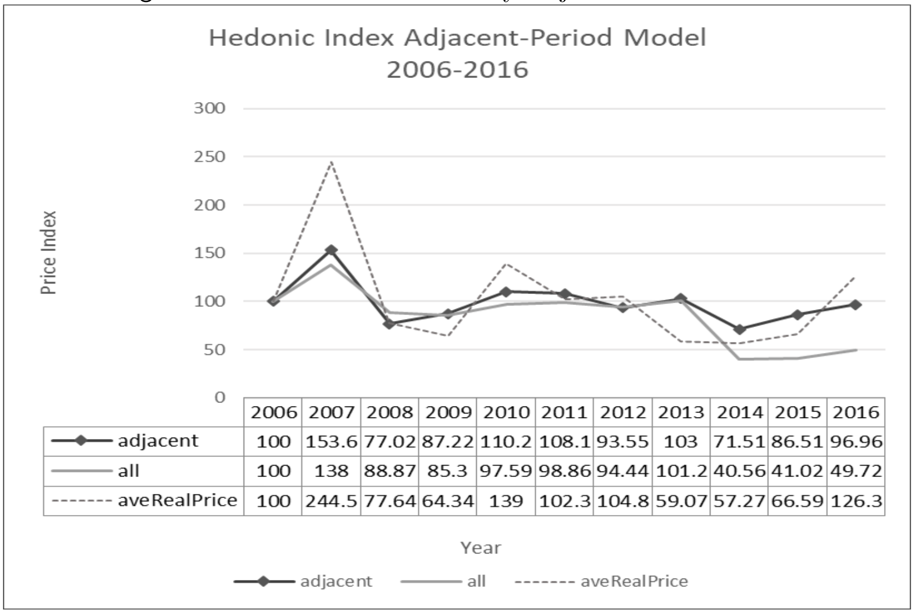
To better compare the differences with the previous result based on the constant-

Table 5: Price Indices and Returns by Adjacent-Period Model

year	γ	SD	variance	Π_t	$r_t(\%)$	Π_t^*	$r\_t^*(\%)$
2006	0.0000	0.8874	0.7876	100.00		100.00	
2007	0.3111	1.0116	1.0233	136.49	36.49	153.56	53.56
2008	-0.2766	1.0268	1.0542	75.83	-44.44	77.02	-49.85
2009	-0.0938	0.9841	0.9684	91.04	20.06	87.22	13.24
2010	0.0790	1.0023	1.0047	108.22	18.87	110.20	26.35
2011	0.0516	1.0285	1.0577	105.29	-2.70	108.12	-1.89
2012	-0.0829	1.0441	1.0902	92.04	-12.58	93.55	-13.47
2013	-0.2484	1.2829	1.6459	78.01	-15.25	102.99	10.09
2014	-0.3802	1.3174	1.7355	68.37	-12.35	71.51	-30.57
2015	-0.0976	1.2810	1.6408	90.70	32.66	86.51	20.98
2016	-0.0526	1.2978	1.6843	94.88	4.60	96.96	12.08
							4.05

coefficient model in the repeat sales method, we plot the graph as below:

Figure 6: Hedonic Price Indices by Adjacent-Period Model



From Figure 6, we can find that the previous result assuming constant tastes for hedonic characteristics is quite different from the price indices calculated by assuming varying tastes for hedonic characteristics. Hence, the hedonic indices are not robust in this research.

5.5 Quantile Regression

In previous researches, Mei and Moses (2002) find masterpieces significantly underperform the middle-level and low-level paintings. Pesando (1993) also derives the similar conclusion that the masterpiece of prints usually has lower returns. Hence, we would also examine the "masterpiece effect" by the quantile regression within the top 25%, bottom 25%, and the group in between. We continue to utilize the adjacent-period model for quantile regressions. The results can be referred to the Table 10-12 in the Appendix. The line chart of low-end, middle, and high-end prints price indices is shown below:

Figure 7: Quantile Regressions in Adjacent-period Model

From the price indices present in Figure 7, we can find that the high-end prints have the highest fluctuation in price index cross years, which implies higher risks in the masterpiece investment. The standard deviations of returns are 0.15, 0.09, and 0.28, from low-end to high-end portfolios. We can find that the increase in the high-end price index in 2007 can explain why the average price in 2007 drives so high.

The arithmetic average returns of the low-end, middle, and high-end prints are 3.56%, -0.29%, and 3.97%, respectively. Based on the Sharpe ratio which measures both returns and volatility, we suggest an investment strategy that focuses on the low-end market in prints market.

5.6 Comparisons and Correlations with Other Assets

To better evaluate the investment performance in the prints market, we compare its arithmetic mean return, geometric mean return, and Sharp ratio with other assets. Considering that the number of repeat sale pairs are extremely small in earlier years, the RSR print price indices in our research are obviously more likely to be distorted, let alone the intrinsic disadvantage of selection bias. Hence, we opt for the hedonic index, specifically the hedonic index calculated in the adjacent-period model, which is supposed to be more precise. We compare the real returns of prints with that of S&P 500 stocks, 10-year treasury bond, Baa corporate bonds, gold, and commodities (S&P GSCI).

Table 6 shows the performance of these assets. We can find that even the print market has reasonable arithmetic mean return, its low geometric mean return, high risk (standard deviation), and low Sharpe ratio denotes that the sole investment in prints is not a good choice. However, the correlation matrix in Table 7 reveals that the print market is negatively correlated with the 10-year treasury bonds (p-value 0.177) and commodities (p-value 0.005). It is interesting to find that the magnitude of the correlation with commodities is considerable, and significant at the 1% level. Hence the investment in prints may provide an excellent hedging opportunity.

Table 6: Investment Performance 2006-2016

ASSET	AMean(%)	GMean(%)	SD(%)	Sharpe
Prints (Hedonic)	4.05	-0.31	26.10	0.0356
S&P 500	6.91	5.04	19.01	0.1992
10 yr T-Bonds	3.12	2.77	8.87	0.0000
Baa Corp Bonds	5.34	5.01	8.85	0.2507
Gold	7.64	6.12	18.37	0.2459
Commodities	10.71	6.11	35.94	0.2112

a. AMean stands for the arithmetic annual return of the asset class.

b. GMean stands for the geometric annual return of the asset class.

c. SD is the standard deviation of the returns.

d. Sharpe is the arithmetic Sharpe ratio calculated by the arithmetic excess return dived by SD.

Table 7: Correlation Matrix between Assets

	Prints	SP500	T-Bonds	Corp Bonds	Gold	Commodity
Prints	1					
SP500	0.5320	1				
T-Bonds	-0.4637	-0.8041***	1			
Corp Bonds	0.2426	0.5904*	-0.4153	1		
Gold	0.3177	-0.0725	0.2510	0.4440	1	
Commodity	-0.8093***	-0.6945**	0.4937	-0.5642	-0.4889	1

6 Discussion and Conclusion

In this research, we explore an asset pricing model of prints by hedonic price function. The HPF can explain more than 60% of the hammer price of prints. Apart from the prices, we also examined the returns of prints investment by utilizing both the hedonic method and repeat-sales regression (RSR). The results show that the market returns calculated by repeat-sales regression are higher than that derived from the hedonic method. It confirms our assumption of selection bias in the sample of RSR.

From Figure 4, we can find that the indices derived from the RSR method fluctuate much less than that from the hedonic method. In the earlier years like 2007, because the sample window starts from 2006, a lot of repeat-sale pairs sold in 2007 are left out: only the prints that are bought in 2006 or earlier 2007 and sold in 2007 can be successfully included in the repeat-sale samples of the year 2007. Hence, the price indices and returns computed by RSR method can hardly capture the whole repeat-sale market, let alone the whole market.

However, there are also some crucial advantages of RSR compared with the hedonic method. For instance, in the RSR method, all variance in quality of prints can be eliminated by pairing the identical prints. While hedonic method can certainly not control for all of the variances in quality.

The most important advantage of the hedonic price model is that, this methodology can relieve the problem of selection bias by utilizing all observations. However, the accuracy depends on whether and to what extent the omitted variables can affect the variables included in the regression. In this research, we execute the robustness check by removing the topic dummies in the hedonic regression. It then shows that the exclusion of such variables can distort the resulting price indices and market returns. Hence, we would argue that the selection of hedonic characteristics is very important for the accuracy of results. Furthermore, the assumption that the omitted characteristics are orthogonal to the included variables is probably too strong.

Furthermore, to alleviate the assumption of constant-coefficient across the whole sample window, we adopt an adjacent-period model to make the regression on samples in every two adjacent years. As shown in Figure (xxx), the market return derived from the adjacent-period model differs a lot in the last three years (2014-2016) as compared to the simple HPF. Hence, we suggest using the adjacent-period model for the calculation of price indices and returns in the hedonic method.

After implementing the study of the price index and return of prints, we further the research by testing the masterpiece effect. As suggested before, we would adopt the adjacent-period model to calculate the returns by hedonic regressions. The resulting returns show that the low-end prints have the highest Sharp ratio. Even though the return of high-end prints have slightly higher returns than low-end works, the much higher standard deviation (a proxy of risk) reduces the Sharpe ratio. Hence, we would like to recommend an investment strategy for low-end prints. However, in contrast to the conclusion from Mei and Moses (2002) and Goetzmann (1993), we find that the investment in the masterpiece is still ponderable if compared to the middle works. This difference can result from the differences in prints and other artwork classes like paintings.

In the last step, we investigate the investment performance of print with several assets, together with their correlations. In Table (4), we find that the arithmetic mean of prints is only slightly greater than the 10-year treasury bonds, which is often regarded as the risk-free rate. And obviously, the volatility of prints is much higher than the T-bonds. By observing the Sharpe ratio, we find that the investment in only prints is not a good choice. Since the returns are negatively correlated with like 10-year treasury bonds and commodities, it has the potential to be combined with other assets in an investment portfolio for more desired results.

This research provides information on the investment of prints in recent years. Besides, we examine the problems and advantages of the two most popular methods in calculating the market returns by empirical analysis. We hope that the in-depth investigation could provide more insights into the hedonic price model and repeat sales method for further study in the investment of real property.

7 References

References

Anderson, R. C. (1974). Paintings as an investment. *Economic Inquiry*, 12(1):13–26.

Bailey, M. J., Muth, R. F., and Nourse, H. O. (1963). A regression method for real estate price index construction. *Journal of the American Statistical Association*, 58(304):933–942.

Bandle, A. L. (2016). *The sale of misattributed artworks and antiques at auction*. Edward Elgar Publishing.

Baumol, W. J. (1986). Unnatural value: or art investment as floating crap game. *The American Economic Review*, 76(2):10–14.

Case, K. E. and Shiller, R. J. (1987). Prices of single family homes since 1970: New indexes for four cities. Technical report, National Bureau of Economic Research.

Chanel, O., Gérard-Varet, L.-A., and Ginsburgh, V. (1996). The relevance of hedonic price indices. *Journal of Cultural Economics*, 20(1):1–24.

Correia, S. (2016). Linear models with high-dimensional fixed effects: An efficient and feasible estimator. Technical report. Working Paper.

Ginsburgh, V., Mei, J., and Moses, M. (2006). The computation of prices indices. *Handbook of the Economics of Art and Culture*, 1:947–979.

Goetzmann, W. N. (1992). The accuracy of real estate indices: Repeat sale estimators. *The Journal of Real Estate Finance and Economics*, 5(1):5–53.

Goetzmann, W. N. (1993). Accounting for taste: Art and the financial markets over three centuries. *The American Economic Review*, 83(5):1370–1376.

Guerzoni, G. (1994). Testing reitlinger's sample reliability. In *8th Conference on Cultural Economics, Witten, August*.

Lancaster, K. J. (1966). A new approach to consumer theory. *Journal of political economy*, 74(2):132–157.

Langley, J. (2014). Capgemini.(2010). *Fifteenth Annual Third-Party Logistics Study*.

Ma, M. X., Li, Y., and Renneboog, L. (2018). Trust in art markets.

Mandel, B. R. (2009). Art as an investment and conspicuous consumption good. *American Economic Review*, 99(4):1653–63.

Mei, J. and Moses, M. (2002). Art as an investment and the underperformance of masterpieces. *American Economic Review*, 92(5):1656–1668.

Pesando, J. E. (1993). Art as an investment: The market for modern prints. *The American Economic Review*, pages 1075–1089.

R Core Team (2019). *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria.

Renneboog, L. and Spaenjers, C. (2013). Buying beauty: On prices and returns in the art market. *Management Science*, 59(1):36–53.

Rosen, S. (1974). Hedonic prices and implicit markets: product differentiation in pure competition. *Journal of political economy*, 82(1):34–55.

Scorcu, A. E. and 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.

Silver, M. and Heravi, S. (2007). The difference between hedonic imputation indexes and time dummy hedonic indexes. *Journal of Business & Economic Statistics*, 25(2):239–246.

StataCorp. (2015). *Stata Statistical Software: Release 14*. College Station, TX: StataCorp LP.

Stein, J. P. (1977). The monetary appreciation of paintings. *Journal of political Economy*, 85(5):1021–1035.

Triplett, J. (2004). Handbook on hedonic indexes and quality adjustments in price indexes.

Wallace, N. E. and Meese, R. A. (1997). The construction of residential housing price indices: a comparison of repeat-sales, hedonic-regression, and hybrid approaches. *The Journal of Real Estate Finance and Economics*, 14(1-2):51–73.

Webb, C. (1981). The expected accuracy of a real estate price index. *Wor# king paper, Department of Mathematics, Chicago State University.*

Zietz, J., Zietz, E. N., and Sirmans, G. S. (2008). Determinants of house prices: a quantile regression approach. *The Journal of Real Estate Finance and Economics*, 37(4):317–333.

8 Appendix

Table 8: Hedonic indices and returns excluding topic dummies

year	γ	SD	variance	Π_t	$r_t(\%)$	Π_t^*	$r_t^*(\%)$
2006	0.0000	0.9367	0.8774	100.00		100.00	
2007	0.1953	1.0865	1.1806	121.57	21.57	141.47	41.47
2008	-0.0126	1.0477	1.0977	98.75	-18.77	94.74	-33.03
2009	-0.0554	1.0100	1.0201	94.61	-4.19	91.01	-3.93
2010	0.0229	1.0420	1.0858	102.32	8.14	105.73	16.17
2011	0.1064	1.0327	1.0665	111.23	8.71	110.16	4.19
2012	0.0464	1.0584	1.1202	104.75	-5.82	107.60	-2.32
2013	-0.2386	1.3819	1.9097	78.77	-24.80	116.90	8.64
2014	-0.6437	1.3131	1.7241	52.53	-33.31	47.88	-59.04
2015	-0.7436	1.3205	1.7436	47.54	-9.51	48.01	0.26
2016	-0.8169	1.5399	2.3714	44.18	-7.07	60.47	25.97

Table 9: Hedonic indices and returns based on liquid artworks

year	γ	SD	variance	Π_t	$r_t(\%)$	Π_t^*	$r_t^*(\%)$
2006	0.0000	0.9269	0.8591	100.00		100.00	
2007	0.1788	1.0724	1.1500	119.58	19.58	138.30	38.30
2008	-0.0893	1.0340	1.0692	91.46	-23.52	87.84	-36.49
2009	-0.1024	0.9613	0.9240	90.27	-1.30	83.95	-4.43
2010	-0.0345	1.0008	1.0016	96.61	7.03	100.43	19.63
2011	0.0182	0.9847	0.9696	101.84	5.41	100.22	-0.21
2012	-0.0590	0.9832	0.9668	94.27	-7.43	94.14	-6.07
2013	-0.3191	1.2501	1.5627	72.68	-22.90	97.91	4.00
2014	-0.7650	1.2926	1.6708	46.53	-35.98	49.12	-49.83
2015	-0.8656	1.2947	1.6762	42.08	-9.57	42.19	-14.10
2016	-0.9033	1.7346	3.0089	40.52	-3.70	78.90	87.01

Table 10: Price Indices and Returns of Low-End Print Market

year	γ	SD	variance	Π_t	$r_t(\%)$	Π_t^*	$r_t^*(\%)$
2006	0.0000	0.2795	0.0781	100		100	
2007	-0.1047	0.4644	0.2156	90.06	-9.94	96.47	-3.53
2008	-0.0869	0.4476	0.2003	91.68	1.80	90.98	-5.70
2009	-0.0307	0.4816	0.2319	96.98	5.78	98.53	8.29
2010	0.0006	0.5339	0.2851	100.06	3.18	102.76	4.29
2011	-0.0076	0.5151	0.2653	99.24	-0.82	98.27	-4.37
2012	0.0003	0.5927	0.3513	100.03	0.79	104.42	6.26
2013	-0.3155	0.7300	0.5329	72.94	-27.08	79.88	-23.50
2014	-0.0109	0.4701	0.2210	98.92	35.61	84.64	5.96
2015	-0.0543	0.4683	0.2193	94.71	-4.25	94.63	11.81
2016	-0.0563	0.9151	0.8375	94.52	-0.20	128.76	36.06
							3.56

Table 11: Price Indices and Returns of Middle Print Market

year	γ	SD	variance	Π_t	$r_t(\%)$	Π_t^*	$r_t^*(\%)$
2006	0.0000	0.3767	0.1419	100		100	
2007	0.1157	0.3724	0.1387	112.26	12.26	112.08	12.08
2008	-0.0580	0.3757	0.1412	94.37	-15.94	94.48	-15.70
2009	-0.0435	0.3550	0.1260	95.75	1.46	95.02	0.57
2010	0.0212	0.3613	0.1305	102.14	6.68	102.37	7.73
2011	0.0042	0.3743	0.1401	100.42	-1.69	100.90	-1.44
2012	0.0023	0.3770	0.1421	100.23	-0.19	100.33	-0.57
2013	-0.0323	0.4518	0.2041	96.82	-3.39	99.87	-0.45
2014	-0.2263	0.6183	0.3823	79.75	-17.64	87.18	-12.71
2015	-0.0180	0.5651	0.3194	98.21	23.15	95.17	9.17
2016	-0.0567	0.5491	0.3015	94.48	-3.80	93.65	-1.60
							-0.29

Table 12: Price Indices and Returns of High-End Print Market

year	γ	SD	variance	Π_t	$r_t(\%)$	Π_t^*	$r_t^*(\%)$
2006	0.0000	0.7637	0.5832	100		100	
2007	0.1946	0.8936	0.7985	121.48	21.48	135.28	35.28
2008	-0.0984	0.8857	0.7845	90.63	-25.40	89.99	-33.48
2009	-0.0938	0.9633	0.9279	91.04	0.46	97.81	8.69
2010	0.0216	0.6388	0.4081	102.19	12.24	78.80	-19.44
2011	0.0254	0.6366	0.4053	102.57	0.38	102.43	29.98
2012	-0.0394	0.6995	0.4893	96.14	-6.27	100.26	-2.11
2013	-0.2866	1.0710	1.1469	75.08	-21.90	104.31	4.04
2014	-0.2740	0.9392	0.8820	76.03	1.27	66.60	-36.15
2015	-0.0600	0.8841	0.7816	94.18	23.87	89.57	34.49
2016	-0.1505	1.0885	1.1848	86.03	-8.66	105.24	17.50
							3.88