THE PROFITABILITY OF INVESTING IN FINE ART: AN ANALYSIS OF RESALE DATA FROM SOTHEBY’S, CHRISTIE’S, AND PHILLIPS

ANNA VASINA, VALERY BUKANOV, VALERIA KOLYCHEVA, ALEXANDER SEMENOV, DMITRY GRIGORIEV

1 Saint-Petersburg State University; Center for Econometrics and Business Analytics; Russia,
2 Herbert Wertheim College of Engineering, University of Florida, USA

E-mail: d.a.grigoriev@spbu.ru

ABSTRACT

The main goal of the presented study was to determine the profitability of investing in unconventional assets, namely, fine art paintings. To do this, we have gathered and analyzed data on the resales of paintings from Sotheby’s, Christie’s, and Phillips from 2003 to 2021. We calculated the annual effective rate of return for each painting and divided the data into profitable and unprofitable investments. We also analyzed the impact of various factors, such as the initial sale price, the length of ownership, and the annual effective rate of return, on the resale price of the paintings. We used correlation analysis to debunk some myths about investing in art, namely myths about shelter-investments and the masterpiece effect. We also compared the annual effective rate of return for paintings with that of traditional financial assets like government bonds and gold. The results showed that investing in paintings can be quite profitable and is a good investment strategy.

Keywords: Auctions, Price Index, Fine Paintings, Resale, Speculative Investments

1. INTRODUCTION

The art market has traditionally been driven by collectors seeking unique paintings, but in recent decades it has also attracted a growing number of investors looking to diversify their investments and make profits. Auction houses, which first emerged in the 17th century, began to thrive in the 18th century as the culture of luxury and collectibles reached new heights. Initially, these houses specialized in selling books, engravings, and minerals, but in the second half of the century, the market for paintings grew significantly, leading to the rise of major auction houses like Sotheby’s [1] and Christie’s [2], which now dominate the global market for art auctions. Later, Phillips, another major auction house, emerged [3] and also became one of the top auction houses in the world. In the late 19th century, the art market saw unprecedented popularity, but it faced a major crisis in the early 20th century as demand for auctions declined [4]. It wasn't until the second half of the century that the art market stabilized and sales volumes in the art industry began to grow again.

In the modern era, the art exchange is rapidly developing and improving, with Sotheby’s, Christie’s, and Phillips all increasing their turnover and opening more branches around the world. These auction houses' websites contain information about all sales of paintings, which enables interdisciplinary research at the intersection of economics and art. While there are methods for determining the current value of a portfolio of paintings or a specific work [5], the factors that influence the value of a unique work of art during repeated sales are not yet fully understood [6].

Classical, unchanging metrics such as the time period, artist, style, and size of a painting [7][8][9][10], its color [11] and even fakes [12] have been studied and described in many works, but other factors that may vary from one sale to the next, such as the time and location of the sale, the auction house conducting the sale, and the estimated value of the lot, have not been as thoroughly explored. In this paper, we propose to study the potential for investing in the art market through the analysis of paintings that have been resold at auctions, as this allows us to determine their realized yield.
In order to analyze the dynamics of the resold painting values, we collected a dataset of paintings that were sold at auctions of one of the three largest auction houses at least twice. We have implemented scripts [13] that scraped information from Sotheby's, Christie's and Phillips sites and stored it in files. To collect data from Sotheby's, we used a pre-defined list of the 5,000 most expensive artists from artceclopedia.com to search for paintings by their names and extracted the necessary set of features (Fig. 1a, Fig.2) from each lot page using the Puppeteer library for Node.js (a tool for collecting dynamically generated Javascript data).

To collect information from the Phillips and Christie's sites, we had to implement another method of extracting information (Fig. 1b), since they do not provide a by-author search. A parser script written in Python using the Selenium library, which can be used to automate and control work with Google Chrome, does most of the work in this case.

The Christie's website organizes lots by year-month-auction sessions, so the data collection algorithm consists of several steps. First, the page for each month is analyzed to identify the auctions held in that month. In the second step, the parser processes the identified auction pages, collecting information about the location of the auction and a list of links to the pages of the lots being sold. In the third step, the necessary features are extracted from each page (Fig. 2).

Before analyzing the data, it needs to be converted into a structured format and cleaned. This initial dataset contained other types of works of art by the same artists, such as sculptures, medals, and dishes. These were eliminated using regular expressions. Another problem is the variety of currencies in trading operations, as well as the need to compare prices in different time periods. Thus, all prices were normalized and converted into US dollars, taking into account accumulated inflation for 2022. Further comparative analysis will be carried out from the point of view of the real yield of dollar assets.

It is also possible for different paintings by the same artist to have the same names or to be not named at all (Fig. 3). In these cases, it can be difficult to determine if they are the same painting based on additional features, and manually checking whether their images taken from the auction site are the same will require a lot of effort. To automate this process, we used an algorithm called SSIM (structure similarity index) [14] to compare “dispute” images. This method consists in comparing the illumination, contrast and structure of images. Based on these features, the algorithm makes a conclusion about the similarity of the images with some degree of confidence.

First, we created lists of paintings whose author names and titles coincide, and the areas differ slightly (by no more than 1%). Then, for the candidates selected for comparison, images were downloaded from the auction site and SSIM was applied to them. The resulting similar images, for which the level of "confidence" of this method turned out to be quite high, were checked manually. With the use of this procedure, deals on the same paintings identified in the general list of sales.

Only paintings that were resold at least once were left in the final dataset. If the painting has been resold several times, then this sequence is divided into pairs of consecutive sales.
3. PAINTING RESALE ANALYSIS

The final dataset included 612 famous artists (according to artcyclopedia.com), and a total of 1,578 cases of resales were identified from 2003 to 2021. Most resales took place in New York (624). The time of resale of one work ranges from 132 days to 16 years, while the average duration of ownership is estimated at 6 years. The longest-owned painting (16 years) is the ROSES ROUGES DANS UN VASE by Moïse Kisling (Fig. 4a), which was bought in November 2003 for $60,000 and resold in November 2019 for $81,250. The painting with the shortest time interval between purchase and sale of 133 days is Touch by Antony Gormley (Fig. 4b) which was bought for $6,134 and sold at a loss for $4,965.

The most modern painting in our dataset is Robert Ryman's Conversion (2003) (Fig. 5a) and the oldest painting is Joachim Beuckelaer's The Adoration of the Shepherds (1560) (Fig. 5b). The distribution of paintings by the time of creation is as follows (Fig. 6): XV century: 14 paintings, XVI: 19, XVII: 7, XVIII: 179, XIX: 1018, XX: 85, XXI: 1.

Figure 4. Paintings with the longest (a) and shortest (b) ownership periods between resales

Figure 5. The Most Modern (a) and the Oldest (b) Paintings in the Dataset

Figure 6. Distribution of Paintings by the Century of Creation

Figure 7. The most profitable (a) and the most expensive (b) paintings
3.1 Profitability Analysis

Annualized effective rate of return was calculated for each artist as follows:

\[ AERR = \left( \frac{\text{Norm\_price}_2}{\text{Norm\_price}_1} \right)^{\frac{1}{\text{days\_distance}}} \]

where \( \text{Norm\_price}_1 \) is the normalized price in USD of the first sale to the investor, adjusted for inflation and the exchange rate at the time of the transaction. Similarly, \( \text{Norm\_price}_2 \) is the normalized price of the subsequent sale of this painting. The total annual effective rate of return was found to be 6.59%, which is consistent with previous studies [10]. The standard deviation was 8%.

Next, we evaluated the resale price influence of initial sale price, period of ownership in years, resale price, and annualized effective rate of return. The results are shown in Tables 1-3. It was necessary to divide the dataset into two groups: profitable and unprofitable resales of the painting [15]. Profitable paintings have a positive AERR, and unprofitable paintings have a negative AERR. In other words, an investment in USD is recognized as profitable only when its profitability has exceeded the inflation rate in the United States. As a result, 1,246 transactions turned out to be profitable (or 80.5% of the total), and 301 (or 19.5%) were unprofitable. Consequently, the vast majority of paintings in our sample make a profit from their resale.

The profitable painting sample exhibits a range of 68 million US dollars, while the duration between resales ranges from 16 years. Among these paintings, the most lucrative one in terms of Annual Effective Rate of Return (AERR) is Love by Robert Indiana (1928-2018) (Fig. 8a). It was acquired on February 11, 2015, for nearly $262,000 and sold just 246 days later on October 15 for $1,118,416. On the other hand, the highest-priced painting among the profitable resales is Bloodline: Big Family No.3 by Zhang Xiaogang (Fig.8b). It was initially purchased for $662,843 in 2008 and resold approximately 6 years later in early 2014 for slightly over 111 million US dollars (AERR 8.39%).

Among the profitable paintings, the smallest profit in terms of the annual effective rate was obtained from Kniender weiblicher Akt (Crouching female nude) by Egon Schiele. It was bought in June 2004 for $173,600 and resold after a decade for $212,500 (AERR 0.05%). The most inexpensive painting among the profitable resales was Mark Tobey's Untitled, which was sold for a mere $3,000.

Moving on to the unprofitable purchases, the most expensive painting among them was Montauk III by Willem De Kooning (Fig. 8a). It was acquired for $9,938,500 in New York in November 2010 and resold for $10,245,000 in May 2014. However, the AERR for this transaction stood at -1.30%. The least profitable purchase in our entire dataset was Untitled by Joan Mitchell (Fig. 8b) which was purchased for $828,800 in 2005 and resold a year later for $228,000 (AERR -73.37%). It is important to highlight that the mean, median, and mode are significantly smaller than the overall range of variation across all the metrics analyzed. In profitable resales, the mode consistently represents the smallest value, whereas in unprofitable resales,
the median is the smallest. In all distributions, the average stands out as the largest indicator, with the exception of the period of ownership of the painting, which warrants a separate discussion.

When examining the period of ownership, we observe that the mode remains the same in both samples, representing the most frequently occurring value for the duration of owning a painting before resale. Furthermore, it is worth noting that the dispersion characteristics in this feature exhibit nearly identical patterns. Upon reviewing the structural characteristics presented in Table 2, it becomes apparent that the initial sale prices are significantly lower. This suggests a risk-oriented investment strategy, whereby investors acquire very affordable paintings with hopes of profitable resale once the artist or style gains popularity.

Among the four considered features, the period of ownership of paintings displays the closest resemblance to a uniform distribution, exhibiting the least amount of asymmetry in both profitable and unprofitable samples. Consequently, we can infer that investors do not demonstrate any particular preferences regarding the duration between purchasing and selling a painting. On the other hand, the initial sale price emerges as the most diverse feature.

Table 2. Summary of Correlations for Repeat Art Sales

<table>
<thead>
<tr>
<th>Coefficient</th>
<th>Profitable</th>
<th>Unprofitable</th>
</tr>
</thead>
<tbody>
<tr>
<td>Initial Sale Price Ownership Period</td>
<td>-0.15</td>
<td>-0.35</td>
</tr>
<tr>
<td>Resale Price Ownership Period</td>
<td>-0.21</td>
<td>-0.17</td>
</tr>
<tr>
<td>AERR Ownership Period</td>
<td>-0.34</td>
<td>-0.12</td>
</tr>
<tr>
<td>Resale Price Initial Sale Price</td>
<td>0.56</td>
<td>0.73</td>
</tr>
</tbody>
</table>

4. ANALYSIS METHODS

Plotting (Fig. 9) the average annual return against the period of ownership of the painting, we noticed that the trend line practically does not deviate from 0. Therefore, it can be said that long-term "buy-and-hold" strategies do not increase the profitability of art investing. It should be noted that this result also does not confirm the basic recommendation of experts on our set of deals on works by famous artists: in order to get the maximum profit, the owner must keep the painting for more than 20 years [17]. However, it may be relevant for investors who invest long-term in a cheap work by a beginner artist, expecting them to "take off" over time.

Figure 9. The average annual return based on the length of time the painting was owned

According to the research by Baumol [18], investing in paintings does not yield as high a return as investing in other financial assets that are comparable in terms of risk. In an effort to create an index of prices for art objects, Pesando utilized the least squares method and the method of regression of repeat sales in [19] by regressing changes in the
logarithmic index of the price of each painting on a set of dummy variables. These variables are assigned a value of -1 at the time of the first sale and a value of +1 at the time of the second sale. In [20], this method was improved by dividing the components into time-dependent and non-time-dependent ones, ultimately leading to a reduction in error. Despite this, the method has yet to be applied to the painting market.

Therefore, we aim to use this method on our collected dataset to build a short-term index of painting value. Based on Goetzmann [21], the cost of painting \( i \) at time \( t \) is \( \hat{P}_{i,t} \), and \( \eta \) is a normally distributed random value. Therefore, the final cost equals:

\[
\hat{P}_{i,t} = P_{i,t} \cdot \eta_{i,t}
\]  \( (1) \)

Assume that painting \( i \) was bought at time \( b \) and resold after \( s \) periods. From equation (1), the yield for the time period \( s \) equals:

\[
\frac{P_{i,t}}{P_{i,b}} = \frac{\eta_{i,t}}{\eta_{i,b}} \prod_{t=b+1}^{b+s}(1 + r_t)(1 + e_{i,t}),
\]  \( (2) \)

where \( r_t \) is the component of market yield, and \( e_{i,t} \) is temporal yield. By logarithmizing equation (2), we get the linear model we need to estimate:

\[
p_{i,b+s} - p_{i,b} = \log \left( \frac{\eta_{i,b+s}}{\eta_{i,b}} \right) + \sum_{t=b+1}^{b+s} (Y_t + \epsilon_{i,t})
\]  \( (3) \)

where \( Y_t = \log(1 + r_t), \epsilon_{i,t} = \log(1 + e_{i,t}) \).

The matrix form will be

\[
\hat{\beta} = Z\gamma + \epsilon
\]  \( (4) \)

Here, \( \hat{\beta} \) is a vector containing the logarithmic difference between the price of the painting sold at time \( b+s \) and the price of the painting at time \( b \). \( Z \) is a matrix built from two components, which are -1 for painting \( i \) resold once, and +1 for painting resold several times. \( Y \) is a vector of logarithmic price indices (what we estimate), and \( \epsilon \) is a vector of perturbations (random errors, residuals). In equation (3) the standard assumption [21] is that \( \epsilon \) is independently normally distributed in the range \( \epsilon \sim (0, \sigma^2) \).

In terms of the art market, it seems reasonable to assume that market forces tend to reduce part of systematic error over time, which is consistent with typical first-order autocorrelation. This assumption is discussed in [21]. That is,

\[
\epsilon_{i,t} = \rho \epsilon_{i,t-1} + u_{i,t}, |\rho| < 1
\]

Thus, we get:

\[
\text{Var}[\epsilon_{i,t}] = \frac{\sigma^2}{1 - \rho^2}, \quad \text{Cov}[\epsilon_{i,t}, \epsilon_{i,t-s}] = \frac{\sigma^2 \rho^{|s|}}{1 - \rho^2}
\]

Ultimately, the logarithmic function is as follows:

\[
L = -\frac{N}{2} \ln(2\pi) - \frac{1}{2} \sum_{t=1}^{N} \ln \left( \frac{\sigma^2}{1 - \rho^2} \right) - \frac{1}{2} \sum_{t=1}^{N} \frac{\epsilon_{i,t}^2}{\sigma^2 (1 - \rho^2)}
\]

where \( i = 1, \ldots, N \).

Now, to get the maximum likelihood estimates for the parameters, we need to maximize this function. To obtain the dummy matrix \( Z \), formed by two different components, we need to assign -1 to the \( i \)-th painting column in case of its sale, and +1 in case of its repeated resale.

5. RESULTS

5.1 Systematic Error Reduction Hypothesis

The hypothesis proposed in [21] is that the art market is partially able to correct systematic pricing errors over time, which is consistent with a pattern of first-order autocorrelation. In other words, the longer a painting is owned, the closer its value is to its fair market value and the less it is affected by any anomalies. Table 4 shows the results of these calculations. To test for the presence of first-order regression perturbation, we used the Darbin-Watson test, which had a value of 1.875 for the return equation. The critical values at the percentage level are 1.576 and 1.967. Since the sample value falls within an uncertain range, we applied an iterative MLE estimate to calculate the model, the results of which are shown in column 3. However, the sample value still falls within an uncertain range, so it is not possible to confirm this hypothesis based on our data. Therefore, long-term ownership of a painting by a famous artist will not necessarily result in the maximum profit.

<table>
<thead>
<tr>
<th>Table 4. Results</th>
<th>OLS</th>
<th>MLE</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Coef</td>
<td>Stdev</td>
</tr>
<tr>
<td>( Z_{2004} )</td>
<td>-0.013</td>
<td>0.112</td>
</tr>
<tr>
<td>( Z_{2005} )</td>
<td>-0.002</td>
<td>0.111</td>
</tr>
</tbody>
</table>
5.2 Masterpiece Effect

Other studies have found evidence of a "masterpiece effect" in the contemporary art market. However, based on our dataset, we can make the same conclusion from Table 5 that Mei and Moses [6] have already come to in their studies: the masterpiece effect does not hold. The coefficient at $\ln(P)$ is too small, that is, the masterpiece does not have a confirmed advantage in resale relative to investing in several paintings of the same total value.

Table 5: Annual Percentage Returns Of Various Assets

<table>
<thead>
<tr>
<th>Year</th>
<th>Index</th>
<th>Government Bonds</th>
<th>Gold</th>
</tr>
</thead>
<tbody>
<tr>
<td>2003</td>
<td>1</td>
<td>-2.06</td>
<td>4.76</td>
</tr>
<tr>
<td>2004</td>
<td>2</td>
<td>3.8</td>
<td>-3.06</td>
</tr>
<tr>
<td>2005</td>
<td>2.3</td>
<td>0.65</td>
<td>4.55</td>
</tr>
<tr>
<td>2006</td>
<td>2.45</td>
<td>-0.28</td>
<td>-1.81</td>
</tr>
<tr>
<td>2007</td>
<td>3.89</td>
<td>3.52</td>
<td>6.74</td>
</tr>
<tr>
<td>2008</td>
<td>8</td>
<td>21.76</td>
<td>8.26</td>
</tr>
<tr>
<td>2009</td>
<td>3.25</td>
<td>4.01</td>
<td>-7.27</td>
</tr>
<tr>
<td>2010</td>
<td>1.89</td>
<td>2.46</td>
<td>2.61</td>
</tr>
<tr>
<td>2011</td>
<td>2.12</td>
<td>10.8</td>
<td>-10.3</td>
</tr>
<tr>
<td>2012</td>
<td>3.15</td>
<td>6.32</td>
<td>-2.11</td>
</tr>
</tbody>
</table>

5.3 Relative Price Change Index

Our dataset enables the calculation of the relative price change index. Let $I_t$ be the price change index in year $t$ relative to the price in 2003. We will calculate the index relative to 2003, taking the 2003 value as 1. The model construction is outlined in the article by Biey and Zanola [21]. They also considered resales over a specific period, and the approach applied embodies the least squares method of repeated sales regression, which has been expanded to accommodate risk. In this way, we can assess the influence of factors on the painting's value after its sale, by isolating shocks that could cause a deviation from the expected profit.

$$I_{t+1} = I_t (1 + e^{y_s y_t - 1}); t = 1, 2, \ldots, n; I_{2003} = 1$$

where $e^{y_s y_t - 1}$ is the percentage change in value over the period $s$ and $y$ is the vector of logarithmic prices from the matrix form of model (4).

We also examined the change in the US government bond price indices over this period, as well as the price of gold, i.e., financial assets comparable in risk to the art market [22].

Plotting this (Fig. 10), it is noticeable that until 2021, the index of painting resale prices slowly increased, except for the years 2008-2009, when the global crisis hit, and the value of government bonds and gold also plummeted. It's also worth noting that from 2016 to 2018, investments in paintings yielded a relatively stable profit compared to gold and US government bonds. On the contrary, in 2011-2012, it was more profitable to invest in the bond market.

D.W. 1.874
6. DISCUSSION AND CONCLUSIONS

This research scrutinizes the profitability of investing in artwork, specifically paintings, from 2003 through 2021. To facilitate this, we compiled a rich dataset from the top three auction houses, preparing and organizing it for comprehensive analysis. Following rigorous data cleaning and preprocessing, we narrowed down our sample to 1,578 paintings that had been re-auctioned at least once. We then dissected the influence of diverse features on the appreciation of the paintings, examining distribution and structural parameters related to these features.

Our findings indicate a feeble correlation between the initial selling price and the duration of possession, the resale price and the length of possession, and the annual effective rate of return and the duration of ownership. Conversely, there's a significant association between the initial selling price and the resale price. In practical terms, the study yields the following key insights:

1. Broadly speaking, the resale market for paintings by renowned artists tends to yield profits, regardless of the time interval.
2. Some investors strategize by purchasing affordable art pieces, betting on the artist's eventual rise to prominence.
3. Prolonged possession of artwork by famed artists doesn't necessarily result in maximizing profits for the owner.
4. Investing in high-end masterpieces doesn't offer a distinct advantage over buying a collection of less expensive paintings that add up to the cost of a piece by a famous artist.
5. The profitability of the art market experiences dips during certain periods, paralleling the downturn in other risk-prone assets.

The next phase of this study served to dispel the notion of the "masterpiece effect." This was achieved by constructing a model to track short-term fluctuations in the price index, using an adapted repeated-sales method [20]. Our primary interest lay in speculative investments in artwork, defined as purchasing a painting with the intent to resell it in the near future.

Our findings revealed a period of profitability in painting investments from 2003 through 2008. However, in 2008, a noticeable downturn occurred, which is largely attributable to the global financial crisis. Post this decline, the index experienced a steady climb and has maintained an upward trajectory until 2021.

Yet, there remain certain aspects that would benefit from further exploration. Specifically, the repeated-sales method zeros in on paintings purely as financial assets, overlooking the costs associated with their storage and completely disregarding their aesthetic appeal to collectors. Additionally, we have not delved into the resale of art collections. It's an established fact [17] that the acquisition cost of an additional painting to supplement a collection isn't the overriding factor, and a completed collection significantly surpasses the cumulative worth of its individual components when sold.

7. ACKNOWLEDGMENTS

Valery Bukanov and Dmitry Grigoriev research for this paper was supported by a grant from the Russian Science Foundation (Project No. 22-18-00588). The authors are grateful to participants at the Center for Econometrics and Business Analytics (ceba-lab.org, CEBA) seminar series for helpful comments and suggestions.

REFERENCES:


