Master of Science in Banking & Finance

2008 – 2010

Master Thesis

The Swiss Art Index

Franziskus Dürr

Coach:

Dr. Marco Rüstmann

Lecturer Centre of Alternative Investments

Submission Date: Jun 21, 2010

Abstract

In this study we analyzed 5231 Swiss Art auction results proprietary gathered from Christie’s and Sotheby’s. The results cover the period between 1993 and 2009. The 2-step hedonic approach developed by Kraeussl and van Elsland (2008) and the classic artist dummy approach have been applied. We could show that the 2-step approach contains two misconceptions which lead to inappropriate or wrong solutions. [1] The solution achieved after 2 steps did not fully converge which leads to biased coefficients and thus a biased index. [2] The estimation of the reputation suggests an exact solution, which leads to a loss of too few degrees of freedom. The results lead to too optimistic standard errors and thus, to a lower uncertainty in the index. Using the artist dummy approach, 75.68% of the variability of the data could be explained. The resulting index performed on average 3.58% over the analyzed period. This is slightly below the bond and fund of hedge funds index. Beside the lower returns the risk of the Swiss Art index was substantially higher than the other two asset classes. The Art market shows only in contrast to the equity index a favourable risk return relationship. Since the correlations with bonds and hedge funds are quite substantial, there is no diversification benefit arising. The broad Swiss Art Index was not an attractive investment during the analyzed period. However, if the sample is reduced to the top 20 artist, the average return increases to 6.1% and for the top 10 even to 7.8%. Thus, the quality of art defines if art is a favourable or poor investment.

Contents

Abstract II
Contents III
List of Figures V
List of Tables VI
List of Formulas VII
List of Abbreviations VIII
1 Introduction 1
2 Problem 3
3 Method 4
3.1 Naive Art Price Indices 4
3.2 Repeat Sales Method 4
3.3 Hedonic Regressions 5
3.4 Discussion 7
4 Possible Difficulties Using Hedonic Regression 8
4.1 Transformation of the Dependent Variable 8
4.2 Dummy Variable Model vs. Single Period Regressions 9
4.3 Sign Restriction of Regression Coefficients 9
4.4 Weighted vs. Unweighted 10
4.5 Outliers Treatment 11
5 2-Step Hedonic Approach 12
5.1 General 12
5.2 Methodology 12
5.2.1 2-Step Hedonic Approach 12
5.2.2 Artist Dummy Approach 15
5.3 Quality Characteristics 15
5.3.1 Available Data 17
5.3.2 Sale Price 18
5.3.3 Sale Date 18
5.3.4 Auction House and Location 18
5.3.5 Surface 18
5.3.6 Technique 19
5.3.7 Living Status 19

5.3.8 Signature 20
5.3.9 Reputation 20
5.3.10 Not Applicable 20
5.3.10.1 Estimate 20
5.3.10.2 Currency 20
5.3.10.3 Nationality 21
5.3.10.4 Place of sale 21
5.3.10.5 Topic 21
5.3.10.6 “Fallen from Fashion” Artist 22
5.3.10.7 “Blue Chips” Artist 23
5.3.11 Overview of Relevant Characteristics 23
6 Data 24
6.1 General 24
6.2 Descriptive Statistics 24
6.2.1 General 24
6.2.2 Evaluation of Estimates 26
7 Results 28
7.1 Derivation of the Reputation 28
7.2 Comparison of 2-Step and Artist-Dummy Approach 29
7.2.1 Misconception of the 2-Step Approach 30
7.3 The Swiss Art Index 33
7.3.1 Index 34
7.3.2 Model Evaluation 36
7.3.3 Comparison of Different Asset Classes 40
8 Conclusion 43
9 Outlook 44
9.1 Data 44
9.2 Model 44
10 Declaration of Authorship 46
A References 47
B Summary of the 2-Step Hedonic Approach Model 49
C Top 50 Swiss Artists 50
D Swiss Art Index Values 52
E R Code (ex.) 53

List of Figures

Figure 1: Nude, Green Leaves and Bust, Pablo Picasso 1
Figure 2: Historical Auction Market Share 2
Figure 3: Number of Observations per Year 25
Figure 4: Dispersion of Hammer Prices Including Buyer's Premium 25
Figure 5: Dispersion of Estimate Deviations 26
Figure 6: Model Evalution Dependent of Iteration Steps 31
Figure 7: Time Dummy Coefficients (ex.) and log(Iteration) 31
Figure 8: Artist Reputation Dependent of the Iteration Step (ex.) 32
Figure 9: Swiss Art Index Returns 34
Figure 10: Swiss Art Index Development Including Confidence Interval 35
Figure 11: Residual Analysis of the Artist Dummy Approach 38
Figure 12: Residual Analysis of some Artist (ex.) 39
Figure 13: Reputation with Reference Artist Ferdinand Hodler 40
Figure 14: Comparison of Different Asset Classes 42

List of Tables

Table 1: Overview of Different Hedonic Models..16
Table 2: Available Data 17
Table 3: Keywords for Topic Variable 22
Table 4: Shares of Estimate Classification 27
Table 5: Top 20 Swiss Artists 33
Table 6: Overview of Past Results...35
Table 7: Hedonic Regression Result with Dummy Approach 37
Table 8: Index Value of Different Asset Classes...41
Table 9: Correlations of the Asset Classes...42
Table 10: Summary of the 2-Step Hedonic Approach Model....................................49
Table 11: The Top 50 Swiss Artists..51
Table 12: Index Values and Return Artist Dummy Approach....................................52

List of Formulas

Formula 1: Hedonic Regression Model 13
Formula 2: Calculation of the Reputation...14
Formula 3: Calculation of the Art Index Value 14
Formula 4: Calculation of the Confidence Interval of the Art Index Value..................15
Formula 5: Separation of Characteristics 28
Formula 6: Subtraction of a Reference Artist from the Analysed Artist......................28
Formula 7: Separation of the Reputation from other Characteristics 28
Formula 8: Exponentiation of Formula 7..28
Formula 9: Reputation Term of one Painting...29
Formula 10: Transformation of the Average Price per Artist Term 29
Formula 11: Combination of Reputation and Coefficient 30

List of Abbreviations

a.k.a.	also known as
CHF	Swiss Francs
CPI	Consumer Price Index
e.g.	for example
ex.	examples
GDP	gross domestic product
mn	Million
p.	Page
q.e.d.	latin: quod erat demonstrandum (that which was to be demonstrated)
U.S.	United States
USD	US-Dollar
vs.	versus

1 Introduction

From time to time art gets attention throughout the world, due to record sales like the sculpture from Alberto Giacommetti called “L'Homme qui marche” which was sold in February 2010 by Sotheby’s in London for about 65 million pounds.[1] However, this record did not last that long. On the 4th of May, Picasso’s “Nude, Green Leaves and Bust” was sold for 106.5 million USD.[2]

Figure 1: Nude, Green Leaves and Bust, Pablo Picasso

Usually, art is recognized as something for enthusiasts who enjoy the aesthetic return or rich people looking for status symbols. Maybe this is also a result of the market characteristics. Unlike other real asset classes art is unique and not standardized at all. Even real estate, widely recognized as illiquid, inefficient, intransparent and with high transaction costs, is much more standardized than art. The “Jones Lang LaSalle Transparency Index” uses five categories to evaluate the transparency of a certain real estate market:[3]

1. Investment Performance Indices
2. Availability of Market Fundamentals Data
3. Listed Vehicles Financials
4. Regulatory and Legal Factors

[1] Sotheby’s: www.sothebys.com, London 03.02.2010 7PM, L10002 Lot 8 [13.02.2010]
[2] Christie’s: www.christies.com, New York 04.05.2010, 2410 Lot 6 [13.05.2010]
[3] Real Estate Transparency Index 2008

5. Professional Standards and Transaction Process

If we try to apply these criteria to the art market [1] we already have problems to find accurate performance indices. This issue will be discussed in detail later. [2] The market is driven by fundamentals only to a small extent, since the supply depends on the people willing to sell their pieces. Business cycles can force people to sell, but at the cost of lower prices, which prevents other people from selling at a heavy discount. Thus, the decreasing supply works like a natural hedge against a deep fall of prices. Most of the time, the individual preferences of a collector or the name of an artist drive the prices. [3] There are only few funds available which normally don't track the market but hold a portfolio of certain styles and buy the pieces for their investors available on the market.

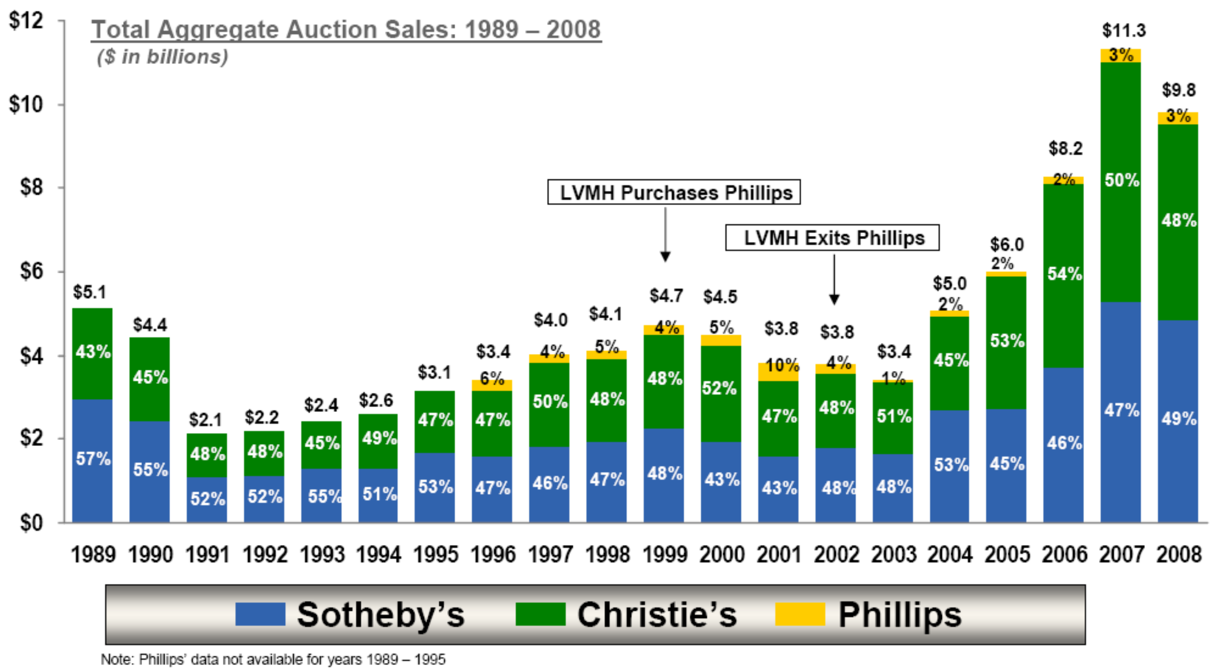

Figure 2: Historical Auction Market Share[4]

[4/5] The market is pretty unregulated and the standards are defined mainly by the two auction houses Christie's and Sotheby's. Figure 2 shows that the two big names are sharing a great proportion of the market. We can conclude that all the above mentioned facts lead to an intransparent and inefficient market. This combined with the high costs for participation and the knowledge needed, leads to a market where only professionals participate successfully. This also shows that there is very much unused potential for growth if art could be available as an investment for a wider range of people. This work should be a small piece in the process to make art as an investment more comprehensible to investors.

[4] Sotheby's Investor Briefing

2 Problem

Investors are seeking for enhanced and uncorrelated returns. Since most of them are holding a core portfolio containing traditional investments, they are always looking for diversification and thus, better risk-return relationships. Alternative investments can possibly provide this diversification by adding untraditional risk exposure to the portfolio. However, to add this risk exposure to the portfolio, it is necessary to have financial instruments tracking the evolution of the prices of non-traditional assets. Since, some of these alternative assets are real assets with heterogeneous characteristics the calculation of a price index is not as straight forward as for traditional assets. Baumol (1986) argued that it is not possible to calculate the true value of art, since art does not pay a dividend that can be discounted. Another issue which needs to be considered is the high risk arising from demolition, theft or fraud which has to be insured. This increases the cost of holding art. In addition the auction houses charge a premium for selling art which leads to high transaction costs and thus high spreads for trades. This results in an underperformance of art in comparison to equity market investments.[5] The risk of an asset class can best be calculated based on a representative index. Here we are back at the problem discussed in section 1. There is no accurate index for art available yet. There are many different calculation methods but all with certain drawbacks as we will see in section 3.

[5] Renneboog et al. (2002)

3 Method

Academic literature proposes three different methods according to Kraeussl and Logher (2008) to evaluate art prices and build indices. It is nearly in every case a trade off between simplicity and accuracy. To increase the accuracy extensive statistic know-how and a broader range of data is needed as we will see in the following lines. We have a more detailed look on the development of the hedonic regressions since this will be the relevant approach as you see later in this thesis.

3.1 Naive Art Price Indices

Naive Price Indices like the average price method or later the representative painting method use average or median auction prices. The assumption of the former is that the distribution of quality of the painting is pretty constant over time. Obviously, this is not necessarily the case which biases the index. Therefore, Renneboog and van Houtte (2002) introduced a new method which is similar to the Consumer Price Index. A basket of representative paintings is selected. There are two methods how to evaluate the prices of the constituting paintings which are not sold in the next period. First, a periodic revaluation of the prices by experts and second, replacing the originally selected painting by a close substitute like a painting of the same artist and the same quality and size. Obviously, the drawback of this method is the subjectivity of choosing representative paintings and substitutes.
Another disadvantage of naive price indices is that outliers can massively distort the calculation. A possible alleviation of this problem could be the use of the median instead of the mean. Nevertheless, the drawbacks make this approach not practicable.

3.2 Repeat Sales Method

This method only uses paintings which are sold at least twice during a defined period. Based on these two prices a return for the time between the two sale dates is calculated. Afterwards, all returns on a certain date of the repeated sales are consolidated to an index value. This approach became popular through the work of Mei and Moses (2002) and is even operationalized.[6] The advantage of this method is that

[6] Art as an Asset: http://www.artasanasset.com/ [21.02.2010]

there are no differences in quality between the various paintings. Renneboog and Spaenjers (2009) highlight the massive drawbacks of this popular method. First, it is not practicable to identify all resales. Chanel et al. (1996) describes this pretty well: *“Using all observations on sales provides many more observations and also avoids the difficult work of searching for paintings which have been sold twice at least. Unless the artwork sold is described by its number in a catalogue raisonné, one can never be sure that it is the same work: the title is often translated into the language of the country where it is sold; many works bear titles which make them undistinguishable (such as Reclining Nude, or Still Life); dimensions "change" because they are sometimes not accurately reported or measured, etc.”*[7]

Second, art is traded rarely. Only considering repeat sales significantly reduce the dataset causing a sample selection bias.[8] Chanel et al. (1996) highlight that it is not possible to exclude the possibility of selection biases since it may be the case that only good works are sold more than once. Additionally, some participants in the market do not even consider resales like many collectors and museums, what further increases the sample selection bias.[9]

3.3 Hedonic Regressions

The most promising and also most complex approach is using a hedonic regression. Chanel and Ginsburgh (1996) described that real assets like paintings depend at least to some extent on their own characteristics. To construct an index for such markets, it is important to account for such non temporal determinants. Hedonic regressions control for such quality changes by breaking down the price to the corresponding attributes. In the perfect world, only the intrinsic value would be left over. Of course in reality, especially for art auctions where irrational behaviour is often observable, there is also a part, which can’t be described by common characteristics. The first use of hedonic price indices goes back to 1939 when Court used this method to account for commodity price changes in the automobile industry due to increasing complexity of commodity products in this sector. Ridker et al. (1967) applied this approach to houses and Griliches (1971) to car prices. Rosen (1974) described the theory of hedonic prices as an economical problem in which the entire set

[7] Chanel, Gérard-Varet and Ginsburgh (1996), p. 2
[8] Zanola (2007)
[9] Anderson (1974)

of implicit prices guides both consumer and producer decisions in characteristics space. For a long time, hedonic methods have been hardly used. In May 1988 Triplett presented a paper on hedonic methods in statistical agency environments on the 50th anniversary of the "Conference on Research in Income and Wealth". He could only name three instances of U.S. statistical agencies which had made use of hedonic methods for calculating price statistics. Namely, [1] the price index of "New One-Family Houses Sold", [2] "BEA-IBM" a computer equipment price index and [3] the adjustment for aging, estimated with a hedonic function, in the housing component of the consumer price index (CPI). This has been a devastating result for over 20 years of hedonic methods research, wherefore, Triplett found the reason in the preconceptions which actually were already solved or disproved by research.[10] However, the picture turned in the ten years after and hedonic methods have been widely accepted and used in a wide range of capabilities. In 2000, 18 percent of the components of US GDP were deflated by hedonic techniques.[11] Moulton (2001) explained the increasing use of hedonic methods by the successful, high profile role of hedonic adjustments of computers and peripheral equipments, first in the national accounts and subsequently in the BLS price programs.

The evolution of the hedonic approach in the art sector has been somewhat parallel to the one in other sectors. Anderson (1974) was the first who used hedonic regression on art prices. Buelens & Ginsburgh (1992) used a very small model with only three variables, the year of sale, the school within the country of origin of the painter and a dummy variable indicating whether the painter was still alive or not. This model was applied on the data collected by Reutlinger (1961) and used by Baumol which had not included other attributes. To calculate the price index, they used dummy variables to specify whether the painting is sold in one time period or in another. Later, this approach emerged as the "time dummy variable method". Chanel (1995) extended the hedonic approach by using more data and more characteristics like dimensions, surface, place of sale and a painter dummy. The aim of his paper was to evaluate whether causality exists between financial and art markets or not. He found many indications which make this assumption reasonable. In 1996, Chanel et al. evaluated the power of estimation of hedonic regressions with a bootstrapping approach and found that the standard deviation of the estimator with hedonic regression is four to eight times smaller than with naïve approaches. Renneboog and van

[10] Tripplet (1991)
[11] Landefeld and Grimm (2000)

Houtte (2002) applied the approach of Chanel to the Belgian art market and compared the return with equity market, whereas art underperforms due to high risk, transaction costs, capital gains, resale rights and insurance premium. Until this time, all hedonic regressions were done by using dummy variables for artists. Depending on the number of artists this approach extends the model in an undesirable way and makes the estimation process much more complicated and error-prone. Kraeussl presented an approach to avoid this artist dummy approach by estimating a variable containing the artistic value. In more detail we will have a look at this approach in chapter 5.

The disadvantage of the hedonic regression approach in general is that the choice of relevant characteristics is somewhat arbitrary. However, we focus on auction results which are standardized. These data are also often used in academic literature because they are easily observable and quantifiable.[12] This is also a constraint to later operationalize the index calculation. Thus, we focus only on tertiary market results (auction houses) and let the other markets like direct sales, specialized dealers and galleries aside. Therefore, the index will be only representative for paintings sold by auction houses.

3.4 Discussion

The mentioned drawbacks of the naive price index, namely the not constant quality makes this approach inapplicable for our purpose. Also, the repeat sales methods have a severe problem by wasting a great share of the available results. This index would only be representative if the repeated sales have similar characteristics like the single sales which would be a too strong assumption. The hedonic regressions on the other hand account for the different characteristics and use all information. Therefore, this is the approach we use for this conceptual work and apply it for the first time to Swiss art auction results.

In the next chapter we will take a look into some problems which can arise when hedonic regressions are used.

[12] Renneboog and Spaenjers (2009)

4 Possible Difficulties Using Hedonic Regression

In his paper “Hedonic Regressions: A Review of Some Unresolved Issues” Diewert came up with some important issues that need to be resolved before hedonic regressions can routinely be applied by statistical agencies:[13]

1. Should the dependent variable be transformed or not?
2. Should separate hedonic regressions be run for each of the comparison periods or should we use the dummy variable adjacent year regression technique initially suggested by Court (1939) and used by Berndt et al. (1995) and many others?
3. Should regression coefficients be sign restricted or not?
4. Should the hedonic regressions be weighted or unweighted? If they should be weighted, should quantity or expenditure weights be used?
5. How should outliers in the regressions be treated? Can influence analysis be used?

In the following lines we will discuss the above mentioned issues and also assess the relevance for this work and the application to art auction results.

4.1 Transformation of the Dependent Variable

The transformation of the dependent variable is maybe the most important transformation. Diewert argues that the errors are more likely to be homoskedastic in the model with a transformed dependent variable than in the untransformed model, since models with very large characteristic vectors will have high prices and therefore it is very likely to have relatively large error terms. Thus, it is more probable that the ratio of model price to the mean price is randomly distributed with mean 1 and constant variance instead of a mean 0 and constant variance. This would prefer the logarithmic regression model to the linear counterpart. Due to the property of art auctions, namely the great dispersion of the auction prices, the dependent variable is commonly transformed. If the dependent variable is transformed by the logarithmic transformation Diewert (2003) concludes that there is slight preference to transform the continuous characteristics as well. This transformation is widely accepted and used

[13] Diewert (2003), p. 1

in scientific art literature, e.g. Kraeussl and Lee (2009) transform all continuous variables.

4.2 Dummy Variable Model vs. Single Period Regressions

There are some advantages and disadvantages in preferring a time dummy variable model in contrast to running separate single period regressions. The advantage of the separate approach is that changes in tastes between two periods can already be anticipated in the estimation. However, the results are two separate estimations for two different periods which needs to be arbitrary aggregated to form a single estimate of price change between the periods. Besides the easy solution for the price change between two periods, the dummy variable method also saves degrees of freedom and is less prone to multicollinearity problems.[14]

Another issue which best fits to this question is whether the use of a chain type or a base type index would be more appropriate. The base type approach estimates the price change to an anchor value which was set in the past. In contrast the chain type index estimates the changes between the current and the previous observations. It is reasonable to believe that the spreads of a comparison of two periods t and t-1 is more accurate than one with period t and 0.

4.3 Sign Restriction of Regression Coefficients

Pakes (2001) presented three main ideas concerning hedonic price indices. [1] He suggested that the hedonic function is not only the sum of all producers' marginal costs and thus, he added a market power function which depends on the demand and can vary over time. [2] The price of a characteristic is not linear which means if a certain level of a characteristic is gained the customer is not ready to pay an equal amount for an additional unit. This can lead to wrong signs of the hedonic coefficient. [3] Based on the point presented under [1], two models of two periods can not be compared easily because of the maybe different market power in the two periods. Hulten (2003) sharply summarized the ideas of Pakes (2001):

"Hedonic regressions have been used in research for some time and they are often found to have coefficients which are 'unstable' either over time or across markets,

[14] Griliches (1971), p. 8

and which clash with the naïve intuition that characteristics which are generally thought to be desirable should have positive coefficients. This intuition was formalized in a series of early models whose equilibrium implied that the 'marginal willingness to pay for a characteristic equalled its marginal cost of production'. I hope this discussion has made it amply clear that these models can be very misleading. The derivatives of a hedonic price function should not be interpreted as either willingness to pay derivatives or cost derivatives; rather they are formed from a complex equilibrium process."[15]

The conclusion of this quote is that carefulness is needed by the interpretation of the factors resulting from the hedonic regression. Additionally, the last sentence ("The derivatives … are formed from a complex equilibrium process") indicates that we should distance from the requirement of an economic justification of the factors and get a more statistical view, like the solution from the equilibrium process. This leads us to a more flexible model with as few constraints as possible. Thus, it is suggested to put no restriction on the regression coefficients. In contrast, Diewert (2003) believes that it is reasonable to put a priori sign restrictions on the regression coefficients where one is fairly sure that more of a certain characteristic is better. He argues that statistical agencies are interested in quality adjustments which are consistent with the public's a priori view. However, if the a priori view is correct, the result would be equal as the view. Therefore, by putting no restriction on the model we relinquish to put the model in perhaps unnecessary chains. Another issue brought up before is the possible decreasing marginal utility of an additional unit of a certain characteristic. To anticipate the non linearity of the utility function, it is probably necessary to introduce characteristics which are non linear. One way to introduce such characteristics is to square them.

4.4 Weighted vs. Unweighted

There are many fields where products are standardized but with different properties like computer or car parts. This makes it possible to sell such parts several times. A weighted model would contain some products as many times as they are sold and give them a greater importance in the model. For art this question is irrelevant since

[15] Hulten (2003), p. 7

pictures are always unique. Therefore, a weighted approach is neither necessary nor possible.

4.5 Outliers Treatment

An interesting point is brought up with question 5, namely how outliers should be treated. It is reasonable to assume that especially in art prices large outliers are frequent. Diewert argues that in the unweighted context it should be permitted to delete influential observations, but only observation which affects the estimation of the price change. We will make an influence analysis later to evaluate our model. The goal of our project is to set up a model where the index is calculated automatically. Then, it is no longer possible to make ad hoc outlier analyses. Therefore, standardized criteria need to be found but outlier analysis is to some extent dependent on the judgement of the analyst. A standardized measure can lead to inefficient results in some situations. Thus, this question is very difficult to answer and an optimal solution does not necessarily exist.

Diewert pointed out some issues which need to be considered using hedonic regression. We will later come back to these points after I applied our model.

5 2-Step Hedonic Approach[16]

5.1 General

Beside the observable variables from the auction results, the artist is probably the most important information. As mentioned previously, due to the number of artists available on the market, modelling with artist dummies puts a constraint on the number of artists which can be included in the sample.[17] To bypass this constraint, Kraeussl and van Elsland (2008) developed a 2-step hedonic approach, which first correct the average price per artist for quality and then include it, in a hedonic model which uses nearly all artists instead of only sub-samples of artists. This method allows the use of nearly the entire available data.

The two steps of this approach are:

1. The creation of a new artistic variable, by adjusting the average price per artist for quality
2. Replacement of the artist dummy by the new artistic value variable and the estimation of an index which uses nearly the entire sample

In contrast to the dummy variable approach, the artistic variable adds only one new variable to the regression. This makes the system much more stable to estimate and remove the constraint of the number of variables which need simultaneously be estimated.

5.2 Methodology

5.2.1 2-Step Hedonic Approach

Before the first step can be made, a reference artist needs to be defined. Kraeussl and Lee (2009) ranked in their paper “The Top 500 Artist” all of the 500 artists according of the number of sales. Then they picked the first placed artist as reference. Since our sample does not contain the same amount of transactions it is unsure how practicable this approach is. We will certainly discuss this problem at a later stage

[16] Kraeussl et al. (2008)
[17] Kraeussl et al. (2008), p. 7

again. As mentioned before, the approach contains two steps. After defining the reference artist the following hedonic regression model using a semi-logarithmic[18] function needs to be estimated:

$$\ln P_{it} = \alpha_0 + \sum_{j=1}^{x} \beta_j \cdot X_{ijt} + \sum_{t=1}^{t} \lambda_t \cdot C_t + \varepsilon_{it}$$

Formula 1: Hedonic Regression Model[19]

The dependent variable P_{it} is the auction price including the buyer's premium since this price is available on the homepages of the auction houses and the price which is paid by the buyer. Using the total costs makes sense, since Christie's and Sotheby's do not necessarily charge the same buyers premium. If we assume a buyer is willing to spend e.g. 1 m CHF for one picture, he could place a bid more if the buyer's premium is lower. We should consider whether an increase/decrease of the buyer's premium would or should affect the index. Let us assume that the buyer's premium has increased from one period to another by e.g. 10 percent. If the buyers are willing to pay the same amount as before, they would bid about 10 percent less due to the fee increase. In this case the price including buyer's premium would stick the same and the market price as well. On the other hand, when the buyer still bid the same amount than before, at the end he would pay 10 percent more. Using the price including buyer's premium to calculate the index, would lead to an increase of the market price by 10 percent as well. So considering the buyer's side there is nothing wrong using this measure as dependent variable. Another question is whether this holds for the seller side as well or not. We use the same example like before. An increase of the buyer's premium by 10 percent would lead to substantial lower revenue for the seller under the assumption that the buyer spends the same amount than before. On the other hand, if the buyer pays the same amount for the picture than before, there is no difference for the seller. This lead to the conclusion that the buyer can decide whether his behaviour is or is not elastic to changes in the buyer's premium. The seller on the other hand is exposed to this decision. This needs to be considered at a later stage of this project when derivatives on the index should be accessible. Due to this fact, there might be some exposure for the seller to decisions of the auction houses which can not be hedged. The reason for the logarithmic transformation was already discussed in chapter 4.1. α_0 stands for the intercept of the he-

[18] Bastian et al. (2004), p. 13
[19] Kraeussl and Lee (2009), p. 8

donic regression. β_j indicates the coefficient of the quality characteristic *j* and X_{ijt} reflects the value of the quality characteristic *j* of picture *i* at time *t*. C_t is the time dummy variable which indicates whether a picture is sold in period *t* or not. When the picture is sold in period *t* the dummy takes the value 1 and otherwise 0. Finally, the error term of the regression is reflected by ε_{it}. The quality characteristics (*Xijt*) used in the hedonic regression will be discussed in more detail in chapter 5.3. After estimating the hedonic regression model, the coefficients and the quality characteristics are used to calculate the reputation index ($Reputation_k$) for a certain artist relative to the reference artist.

$$Reputation_k = \frac{\dfrac{\prod_{i=1}^{n} P_{i,k}^{\frac{1}{n}}}{\prod_{i=1}^{m} P_{i,r}^{\frac{1}{m}}}}{e^{\left[\sum_{j=1}^{z} \beta_j \cdot \left(\sum_{i=1}^{n} \frac{X_{ij,k}}{n} - \sum_{i=1}^{m} \frac{X_{ij,r}}{m} \right)\right]}}$$

Formula 2: Calculation of the Reputation[20]

In this equation $P_{i,k}$ stands for the price of painting *i* of artist *k*. Similarly, $P_{i,r}$ reflects the price of painting *i* created by the reference artist *r*. *n* indicates the number of paintings created by artist *k* and *m* the number of the reference artist *r*. As mentioned at the beginning of this subsection. It is unsure how applicable the selection of the reference artist is to our sample. Since our base value is the first year (1993), we can calculate the final price development with the antilog of the corresponding coefficients λ_t of the time dummy variable at time t:

$$Index_t = e^{\lambda_t} \cdot Index_0$$

Formula 3: Calculation of the Art Index Value

This antilog transformation is nothing more than the inversion of the logarithmic transformation of the prices in Formula 1. $Index_0$ is the anchor value at time 0, which is commonly defined as 100 points. Since the values λ_t need to be estimated, there is a certain amount of variability in it. To account for that, we also calculate the confidence interval of the index value to get a feeling about the uncertainty in our estimation. Thus, we extend Formula 3 with the standard errors of the estimation and the factor 1.96 indicating the 95% confidence interval under the normal distribution. This leads us to:

[20] Kraeussl and Lee (2009), p. 10

$$Index_{t,CI} = e^{\lambda_t \pm 1.96 \cdot \sigma_t} \cdot Index_0$$

Formula 4: Calculation of the Confidence Interval of the Art Index Value

5.2.2 Artist Dummy Approach

As a second method, we use the artist dummy approach to evaluate and compare the results of the 2-step approach. For this approach, only the artist variable needs to be added in Formula 1. As mentioned earlier, the disadvantage of this approach is the wasteful dealing with degrees of freedom.

5.3 Quality Characteristics

In the previous subsection we saw how the methodology is applied and how the different steps work. In this part, we will have a closer look at which characteristics are applied. This is maybe a bit redundant with the specific data chapter (chapter 6), but we like to discuss the quality characteristics here and general data issues later. Kraeussl and van Elsland (2008) summarized many studies in a table (see Table 1 on page 16) and structured the characteristics which have been used. Many of these characteristics are the one observable from auction results (which I discuss in more detail in the following subsections) or transformations of them. Often the numeric characteristics are squared to account to some extent for nonlinearity. Nevertheless, there is also information which can not be observed easily like school, art current, publication and number of time exhibited or provenance. Many of these criteria would need a manual intervention and thus judgement of an analyst. In our case this is not desired since the calculation should be strongly data driven and thus straight forward.

Variables	Buelens and Ginsburgh (1993)	De la Barre, Doccio, and Ginsburgh (1994)	Chanel (1995)	Chanel, Gerard-Varet, and Ginsburgh (1996)	Czujack (1997)	Renneboog and van Houtte (2002)	Hodgson and Vorkink (2004)	Biey and Zanola (2005)	Worthington and Higgs (2006)	Kräussl and Schellart (2007)
Year of sale	√ +/-	√ +/-		√ +/-	√ ?	√ +/-	√ +	√ +/-	√ +	√ +/-
Month of the year									√ +/-	
School	√ +/-									
Width		√ +		√ +		√ +	√ +			√ -
Height		√ +		√ +		√ -	√ -			√ +
Width^2				√ -						
Height^2				√ +						
Surface (cm2)		√ -		√ -	√ +	√ +	√ -	√ -	√ +	√ -
Surface (cm2)^2					√ +				√ +	
Technique					√ +/-	√ ?	√ +	√ +/-	√ +/-	√ +
Support		√ +/-			√ +/-		√ +			√ +
Place of sale (country/city)		√			√ +/-	√ +				√ +
Auction house		√		√ +/-	√ +/-	√ +/-	√ +	√ +/-	√ +	√ +
Painter		√		√ +/-		√ ?	√ +/-		√ +/-	
Signed?					√ -	√ +		√ -		√ -
Painter alive?	√ +/-								√ -	√ +
Painter age									√ +	
Painter age^2									√ +	
Painter age^3									√ +	
Painter age^4									√ +	
Works Sold in Calendar Year								√ -	√ +	
Works Sold in Calendar Year^2									√ +	
Art current						√ +/-				
Reputation (average price)										√ +
Publication					√ +					
Number of times exibited					√ +/-					
Working periods					√ +					
Provenance					√ -					
Prior price estimate										√ +
Period	1700 - 1961	1962-1991	1961-1992	1855 - 1970	1963 - 1994	1970 - 1997	1968 - 2001	1988 - 1995	1973 - 2003	1986 - 2006
Sample size	1,111	24,540	25,300	1,972	921	10,598	12,821	1,665	30,227	1,688
R-square	16.3% - 59.3%	81.1%			79.0%	41.5%			67.5%	89.8%
Number of artists		82	82	46	1	71	152	1	50	23

Table 1: Overview of Different Hedonic Models[21]

[21] Kraeussl, van Elsland (2008), p. 35

5.3.1 Available Data

As mentioned earlier we are relying on auction results published by Christie's or Sotheby's. These results contain the following 20 dimensions per sale or picture:

Category	Criterion
Picture	Lot Number
	Title
	Year of Origination
	Price Realized
	Currency
	Estimate Low
	Estimate High
	Size
	Technique
	Study
Artist	Name
	Year of Birth
	Year of Death
	Signature
	Nationality
Auction	Year
	Month
	Day
	Auction House
	Auction Place
	Sale Number

Table 2: Available Data

Anywhere, where dummy variables are introduced the dummy takes the value of one for a certain attribute. However, the number of dummy variables is not equal to the number of attributes of a certain characteristics. If it was, the result would be perfect multicollinearity and none of the coefficients could be estimated. Thus, one attribute is skipped and the estimated coefficients of the others represent the average deviation of the "reference attribute". Often, the reference attribute is chosen in a manner that the highest priced attribute is the reference of a characteristics leading to coefficients with negative signs for the lower priced attributes.

5.3.2 Sale Price

In subsection 4.1 we already discussed the dependent variable and that it should be transformed using the logarithm. In subsection 5.2 we exercised whether it is important to account for the buyer's premium or not. We concluded that only for the seller side the interpretation is not accurate. Nevertheless, using the hammer price would even be less accurate since the auction houses charge different buyer's premiums. The online availability of the price including buyer's premium supports this decision, since we are looking for plain processes to update the index in a later stage.

5.3.3 Sale Date

The sale date contains year, month and day of the specific auctions. These information are used to introduce the time dummy variable which covers a certain period of time. Since the auctions are not held periodically it is unclear how to define the length of the periods yet. In academic literature often used periods are quarterly or annually.

5.3.4 Auction House and Location

Many academic papers conclude that that the auction house is an important characteristic to define the price. Christie's and Sotheby's are widely associated with high prices.[22] Also a difference for sales at different locations around the world is observable. London and New York are recognized to attract a broader range of seller and thus, the especially valuable pictures are sold there to achieve higher sell prices. Often a combined dummy containing the auction house and the location is generated (e.g. Christie's London, Sotheby's New York). Since we only use the two big names for our index and for this prototype only Swiss sales, the difference is maybe not that big, nevertheless we introduce a dummy variable for the auction house to answer this question.

5.3.5 Surface

Every auction result contains the two dimensions width and height. There are different ways to add these characteristics. Width and height could be two separate vari-

[22] Valsan (2002), Agnello (2002), Renneboog and Spaenjers (2009)

ables, however often a surface variable containing the product of width and height is introduced. Since both variable are expected to be highly correlated, only one of them should be used to avoid multicollinearity. Kraeussl and Lee (2009) argue that the sale price increases with a decreasing effect, due to the fact that large painting are more difficult to display. Thus, they log the surface value.[23] Another way to account for this effect would be to square the surface to account for decreasing prices of big paintings.

5.3.6 Technique

In this section we summarize the technique used for the picture. Two different approaches can be found in academic literature to model these characteristics. Kräussl and Schellart (2007) modeled the medium and the support as two characteristics with individual dummies. The medium contained oil, mixed media and other media whereas the support consisted of canvas, hardboard, cardboard, paper panel and mixed or other forms of support. However, in their recent studies[24] they only use one variable which is combined like oil on canvas, acrylic on canvas, oil on board and so on. Several studies found that paintings with oil on canvas are the most valuable.[25] We stick to the combined variable since otherwise a very complex separation algorithm would be needed which would make the data preparation process far more complex than it already is.

5.3.7 Living Status

The living status indicates whether an artist is still alive or not. The year of death is available for every auction. A rational assumption would be that dead artist can not produce new paintings any longer and thus, a natural shortage arises. However, living artist can improve their reputation. Based on the availability of the year of death we will introduce a dummy indicating whether the artist is dead or alive.

[23] Krauessl, Lee (2009), p. 9
[24] Krauessl, Lee (2009), p. 8
[25] Agnello (2002), Valsan (2002)

5.3.8 Signature

Many studies concluded that investors are willing to pay higher prices if there is prove that the painting can be allocated to a specific artist.[26] This is possible through a signature or a monogram. We classified estate stamps as no signature, since they could be added after the death of an artist.

5.3.9 Reputation

The calculation of the reputation was described in detail in section 5.2 and is part of the “2-Step Hedonic Approach”. It is reasonable to assume that this variable will describe a great share of the total sample variance.

5.3.10 Not Applicable

A kind of “pro memoria” characteristics are listed in the following subsections. They are either not relevant for this work (because we are looking on a clearly defined sub sample of the global market, Swiss paintings) or they are not appropriate to our data. We will discuss each of them and justify why they are not applicable in our model.

5.3.10.1 Estimate

Ashenfelter and Graddy (2003) concluded that the availability of an estimate price results in a higher final price. Since in our data all, beside a handful, observations have a lower and a higher estimate of the price, this characteristic does not make sense for our data. Kraeussl and Lee (2009) mentioned another property of the estimate. They argue that estimates function as equilibrium. Buyers are not willing to substantially overpay a painting. However, the result of the analysis of the estimates showed that only every fourth auction ends within the estimates. If there is an equilibrium function of the estimates it would be of very low power (see also 6.2.2).

5.3.10.2 Currency

For the Swiss Art Index the currency is identical for all results. All are denominated in Swiss Francs. In a later stage when global data is used, the results are in different

[26] Agnello (2002), Renneboog and Spaenjers (2009)

currencies. Thus, it is necessary to convert the result into an identical base currency before calculating an index.

5.3.10.3 Nationality

Like for the currency, the nationality could be an additional characteristic which could describe a part of the volatility of the data. Right now only Swiss painters are used. Thus it is not yet necessary to include this characteristic. However, this information is not contained in the auction result.

5.3.10.4 Place of sale

In section 5.3.4 we already discussed the importance of the place of sale. New York and London attract more sellers and thus, the especially valuable pieces are sold there. Again, for the prototype we only use Swiss auctions which took place in Zurich or rarely in Geneva, therefore we do not need to enter a place of sale variable.

5.3.10.5 Topic[27]

Renneboog and Spaenjers (2009) added an additional variable which categorize the painting in a certain topic like animals or landscapes (see Table 3). They match the title of a certain painting with keywords. In their data base the most titles are in English. Nevertheless, they have accounted to some extent for French titles by also adding French keywords. Another issue which they pointed out is that some keywords are part of others like chat is fully contained in chateau. In this case, they match only words with the same length like the keyword but then also the plural needs to be defined as keyword. Here we see the difficulties. We would need the keywords in all common languages for a global index and all common possibilities which is practically not feasible.

[27] Renneboog and Spaenjers (2009), p. 33

1. ABSTRACT: "abstract", "composition"
2. ANIMALS: "horse", "cheval", "chevaux", "cow_", "cows", "vache", "cattle", "cat_", "cats", "chat_ ", "dog_", "dogs", "chien", "sheep", "mouton", "bird", "oiseau"
3. LANDSCAPE: "landscape", "country landscape", "coastal landscape", "paysage", "seascape", "sea_", "mer_", "mountain", "river", "riviere", "lake", "lac_", "valley", "vallee"
4. NUDE: "nude", "nu_", "nue_"
5. PEOPLE: "people", "personnage", "family", "famille", "boy", "garcon", "girl", "fille", "man_", "men_", "homme", "woman", "women", "femme", "child", "enfant", "couple", "mother", "mere_", "father", "pere_", "lady", "dame"
6. PORTRAIT: "portrait"
7. RELIGION: "jesus", "christ_", "apostle", "ange_", "angel", "saint_", "madonna", "holy_", "mary magdalene", "annunciation", "annonciation", "adoration", "adam and eve", "adam et eve", "crucifixion", "last supper"
8. SELF-PORTRAIT: "self-portrait", "self portrait", "auto-portrait", "autoportrait"
9. STILL_LIFE: "still life", "nature morte", "bouquet"
10. UNTITLED: "untitled", "sans titre"
11. URBAN: "city", "ville", "town", "village", "street", "rue", "market", "marche", "harbour", "port_", "paris", "london", "londres", "new york", "amsterdam", "rome_", "venice", "venise"

Table 3: Keywords for Topic Variable[28]

5.3.10.6 "Fallen from Fashion" Artist[29]

Another criterion which Renneboog and Spaenjers (2009) introduced is the "fallen from fashion" artist. They defined an artist as fallen from fashion according of "Gardner's Art through the Ages". If the artist was included in the 1926, 1959 or 1980 edition but not in the 1996 or 2004 he is fallen from fashion. The disadvantage of this classification is that it is based on the view of renowned academics. Additionally, the definition of the years of the in- or exclusion is arbitrary. There is no explanation why a fallen from fashion artist can not be part of the 1996 edition and be excluded in the one of 2004. This characteristic is dependent on the information of a third party and beside arbitrariness hardly realizable. A feasible way to account for changes in the taste of the buyers could be to calculate rolling windows for different time spans.

[28] Renneboog and Spaenjers (2009), p. 33
[29] Renneboog and Spaenjers (2009), p. 34

5.3.10.7 "Blue Chips" Artist[30]

Similar to the characteristic previously mentioned, the blue chips artist is defined by an information provider. The "grove art online" database is an accumulation of information about paintings. The so called "masterpiece effect" is analysed using the top 5 percent of word counts in this database or if the artist is considered in all five Gardner textbooks. Like for the fallen from angel artists the disadvantages for the "blue chips artists" are similar. This characteristic is dependent of the quality and the assessment of the analysts maintaining the database.

5.3.11 Overview of Relevant Characteristics

We just saw that there are many possibilities to add further characteristics. Often they are dependent on the judgement of experts or analysts. Since our approach should be strongly data driven, we do not want to rely on such arbitrary judgement. Additionally, such expert opinions are hardly realizable and dependent of continuous flow of information. Thus, an operational risk occurs by being dependent on such information. If this information is no longer published a serious problem arises, which put the methodology and the confidence in the index into question. Thus, we keep as few data sources as possible to minimize such risks.

Therefore we stick to the following characteristics:

- Price including buyer's premium
- Date dummy
- Auction house
- Surface
- Technique
- Living status (Alive value 1)
- Signature
- Reputation

[30] Renneboog and Spaenjers (2009), p. 35

6 Data

6.1 General

In this chapter we will have a look at the data which is available from different sources. Some data from Christie's and Sotheby's is available online through their websites. The auction results which are not online can be accessed on-site. This information is documented in printed catalogues and price lists. Since Christie's is part of the KTI project they digitized their results as a preparatory effort. The online available results from Sotheby's had been prepared by students as a part of their work in a research module. After a closer look into the data, it became apparent that a substantial share of the total data was missing, which needed great efforts to be completed. After days of data gathering and completion we have been able to close all gaps and get a full sample of Swiss Art auction results during the period of 1993 and 2009. Note, sales of Swiss artists abroad are not contained in the sample. After removing all kind of works, which do not belong to our data sample (e.g. sculptures, woodcuts, collages, photographs, different printing techniques) 5231 results are left over. From these results all the required information as shown in Table 2 are available. An important information for the calculation of the reputation of a certain artist is his name. Many difficulties arises like spelling errors, different first names (sometimes only one is indicated, sometimes many, sometimes only the first, sometimes the second) and special characters which are sometimes country or language specific. This makes it hard to identify an artist as unique and match all pictures he had painted to him. To show the magnitude of this problem we present two numbers. We started with 906 different artists. After correcting all data errors and other problems we could reduce this number to 606! This is a huge problem which needs to be taken into account.

6.2 Descriptive Statistics

6.2.1 General

To get a better feeling about our data sample, the number of observations per year is shown in Figure 3. The amount of data ranges from about 200 to well over 500 auction results per year. This of course is depending on the number of auctions held by the two houses but also on the number of lots sold in the respective year. Sotheby's normally holds two Swiss Art auctions a year and Christie's one per nine months. This leads to years with one Christie's auction and followed by a year with two.

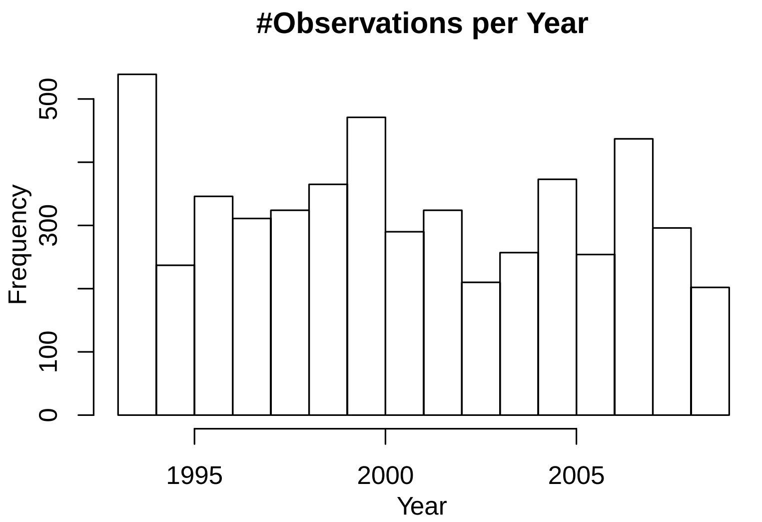

Figure 3: Number of Observations per Year

We can also see the economic cycles which lead to a natural shortage during economic down turn. Sellers which do not need to get liquidity postpone their sales until better times and higher prices are achievable. In Figure 4 the volatility of the final results is shown. In the upper figure, the plot is distorted by the massive outliers, which are depicted with circles.

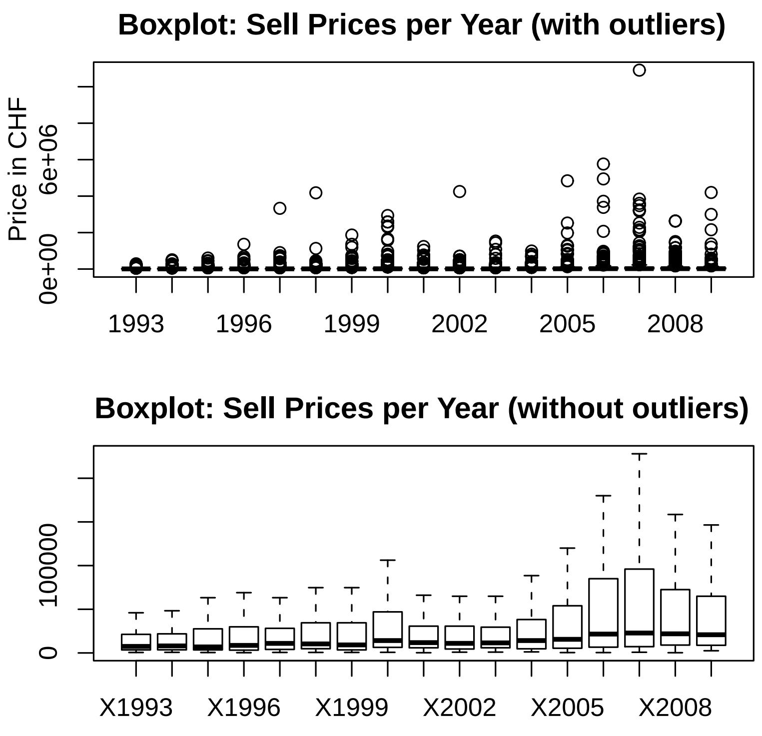

Figure 4: Dispersion of Hammer Prices Including Buyer's Premium

In the lower figure they are dropped which leads to a better view on the majority of the observations. An increase of the volatility at the end of the period is observable which peaked in 2007. During the entire time frame the median (marked as a black bar inside the box) shows an upward trend. The conclusion that the market price increased during this time would be naïve since the characteristics and the quality of the sold paintings do not need to be constant during this period.

6.2.2 Evaluation of Estimates

Since we have a full sample of auction results for the period between 1993 and 2009, we are able to evaluate the estimates defined by art experts of the two auction houses. In the four plots below the dispersion of the deviation between higher or lower estimate and the final price in percent are shown. The two upper plots show the final prices which were below the lower estimate and the two lower plots display the final prices above the higher estimates. On the left hand side all results are displayed. There are many outliers apparent which distort the figure. Thus, on the right side the outliers are removed to get a clearer view of the area where most of the observations are.

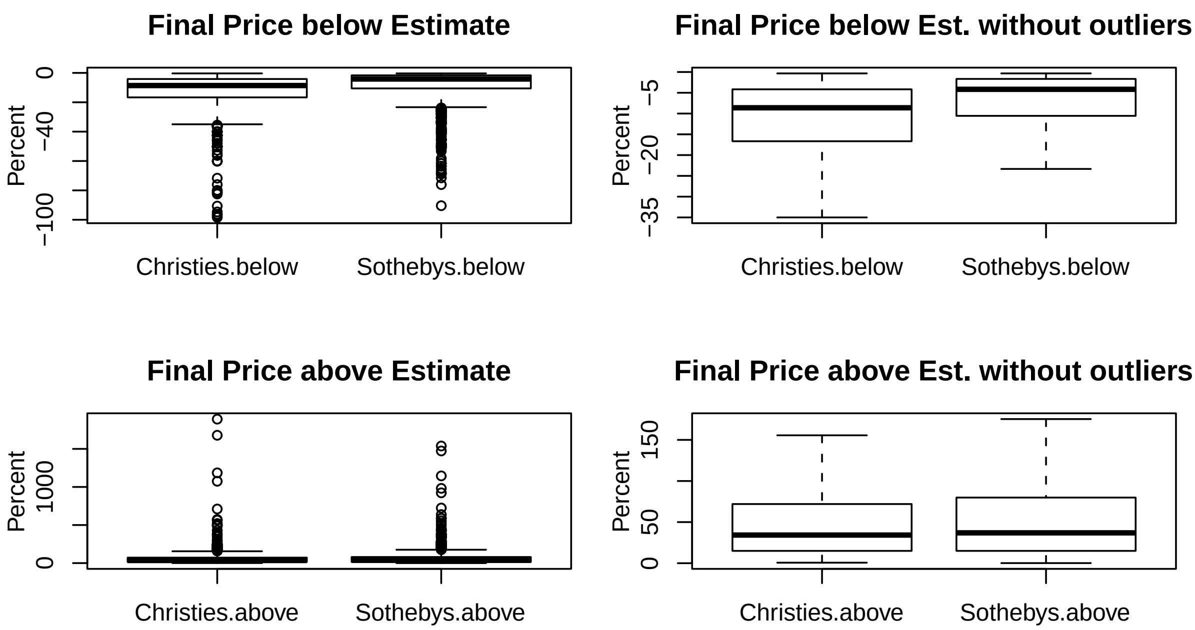

Figure 5: Dispersion of Estimate Deviations

For both auction houses the extreme results are in the same range. There are paintings which achieved a final price of 80% or more below the lower estimate (Plot upper left). On the other hand, we see paintings which where sold for about 15 times the upper estimate (1500%, plot lower left). Without the outliers, it is visible that the median deviation of the lower estimate is around 10% for Christie's and around 5% for Sotheby's results (plot upper right). The upper deviations of both auction houses are similar considering the median deviation of about 40%. However, the upper end of the box indicates that still 25% of the observations, which exceed the upper estimates, are more than 75% higher than the upper estimate. This is a remarkable deviation of a substantial number of observations. Of course the final prices had been adjusted to account for the buyer's premium which is not considered in the estimates. The adjustment factor is 20%. This is a reasonable approximation since Christie's Switzerland[31] has a buyer's premium of exactly 20% with no threshold and Sotheby's Switzerland[32] has 25% until 50'000 CHF, 20% between 50'000 and 1'000'000 CHF and 12% above. By analyzing the plots in Figure 5 we could interpret the dispersion of the observations but not the frequency. This gap is closed by the following table.

	Below	Inside	Above	#Results	Share
Christie's	42.37%	25.19%	32.44%	1973	37.72%
Sotheby's	39.35%	26.18%	34.47%	3258	62.28%

Table 4: Shares of Estimate Classification

In the first three columns the proportion in every interval (below lower estimate, inside range, above upper estimate) are indicated. The fourth column shows the number of results for each of the auction houses. The last column displays the market share of the auction houses. Sotheby's could nearly deliver two thirds of the results in the period between 1993 and 2009. The estimates of both auction houses have similar characteristics. Around 40% of the results are below the lower estimates. On the other side, over 30% are above the higher estimate. Only every fourth result is inside the forecasted range. To evaluate whether the one or the other house make better forecasts, a Chi-Squared Test is used to evaluate the distribution of the results. The null hypothesis for this test is that no difference in the distribution is observable. The p-value of the test (0.09323) indicates that no difference on the five percent confidence level is apparent. Thus, the quality of the forecasts of both auction houses is identical.

[31] Christie's: http://www.christies.com/features/guides/buying/buyers-premium.aspx [05.06.2010]
[32] Sotheby's: http://www.sothebys.com/help/faq/faq_duringauction.html#a03 [05.06.2010]

7 Results

In the previous section we had a look on the property and quality of the data. In this section we go a step further and look into the model and index development. As already mentioned many times we started with the 2-step hedonic approach of Kraeussl. In a first step, we were unable to replicate the model quality he described in his papers. Since, it was not apparent where the problem could be, we started with the verification of his reputation estimator.

7.1 Derivation of the Reputation

To evaluate the correctness of Formula 1 we subtract the intercept and the time dependency of the logarithm of the price.

$$\ln P_{it} - \alpha_0 - \sum_{t=1}^{t} \lambda_t \cdot C_t = \sum_{j=1}^{x} \beta_j \cdot X_{ijt} + \varepsilon_{it}$$

Formula 5: Separation of Characteristics

Now, we find on the right hand side all characteristics of a painting including the reputation of a certain artist. Since the reputation (Formula 2) is calculated in comparison of a reference artist we subtract the formula above from the analyzed artist. We indicate the analyzed artist with *k* and the reference artist with *r* leading to:

$$\ln P_{itk} - \ln P_{itr} - \alpha_{0k} + \alpha_{0r} - \sum_{t=1}^{t} \lambda_{tk} \cdot C_{tk} + \sum_{t=1}^{t} \lambda_{tr} \cdot C_{tr} = \sum_{j=1}^{x} \beta_j \cdot X_{ijtk} - \sum_{j=1}^{x} \beta_j \cdot X_{ijtr} + \varepsilon_{itk} - \varepsilon_{itr}$$

Formula 6: Subtraction of a Reference Artist from the Analysed Artist

Since the intercept α and the error term ε are identical for both paintings, they offset each other. Additionally, we separate the reputation from the other characteristics.

$$\ln P_{itk} - \ln P_{itr} - \sum_{t=1}^{t} \lambda_{tk} \cdot C_{tk} + \sum_{t=1}^{t} \lambda_{tr} \cdot C_{tr} = \sum_{j=1}^{x} \beta_j \cdot X_{ijtk} + Rep_k - \sum_{j=1}^{x} \beta_j \cdot X_{ijtr} - Rep_r$$

Formula 7: Separation of the Reputation from other Characteristics

In a next step, we exponentiate both sides.

$$\frac{P_{itk} \cdot e^{\sum_{t=1}^{t} \lambda_{tr} \cdot C_{tr}}}{P_{itr} \cdot e^{\sum_{t=1}^{t} \lambda_{tk} \cdot C_{tk}}} = \frac{e^{\sum_{j=1}^{x} \beta_j \cdot X_{ijtk} + Rep_k}}{e^{\sum_{j=1}^{x} \beta_j \cdot X_{ijtr} + Rep_r}}$$

Formula 8: Exponentiation of Formula 7

In a last step we bring all terms below the price fraction and regroup them, which leads us to nearly the same solution like Formula 2:

$$\frac{\frac{P_{itk}}{P_{itr}}}{e^{\sum_{j=1}^{x}\beta_j \cdot X_{ijtk} - \sum_{j=1}^{x}\beta_j \cdot X_{ijtr} + \sum_{t=1}^{t}\lambda_{tk} \cdot C_{tk} - \sum_{t=1}^{t}\lambda_{tr} \cdot C_{tr}}} = \frac{e^{Rep_k}}{e^{Rep_r}} = Rep'_k$$

Formula 9: Reputation Term of one Painting

In contrast to Formula 2, Formula 9 is only valid for one painting. The transformation to all paintings is done by a standardization of the characteristics. We need to compare the average painting of artist *k* with the average painting of the reference artist *r*. Thus, the mean of all terms needs to be built. To avoid making the formulas (Formula 5 to Formula 9) more complicated than they already are, we relinquish to insert all these calculations of means. By considering the average painting per artist issue, the only difference left over from the comparison of Formula 2 and Formula 9 is the calculation of the geometric mean per artist. This can be explained by the exponentiation done in the step from Formula 7 to Formula 8. In Formula 7 we would have the sum of the logarithm of the prices of all paintings of artist *k* divided by the amount of paintings build by the respective artist *n*. If we exponentiate this, we find:

$$\exp\left(\frac{1}{n}\sum_{l=1}^{n}\ln P_{l,k}\right) = \exp\left(\sum_{l=1}^{n}\ln\left(P_{l,k}{}^{\frac{1}{n}}\right)\right) = \exp\left(\ln\prod_{l=1}^{n}\left(P_{l,k}{}^{\frac{1}{n}}\right)\right) = \prod_{l=1}^{n}\left(P_{l,k}{}^{\frac{1}{n}}\right) q.e.d.$$

Formula 10: Transformation of the Average Price per Artist Term[33]

So we have proven that the formula to calculate the reputation (Formula 2) is exactly the separation of the unknown reputation in the hedonic regression model (Formula 1).

With the deeper understanding of the concept behind this estimator, we were able to locate a misconception in the 2-step approach as we will see later.

7.2 Comparison of 2-Step and Artist-Dummy Approach

In this chapter we compare the two approaches. To recapitulate we mention the differences and their advantages or disadvantages again. The artist dummy approach is a straight forward approach to estimate the multiple linear regression model. However, this has a major disadvantage, since every artist needs a separate dummy.

[33] Papula (2003)

This leads to an additional coefficient with every additional artist and a model which gains complexity very fast. In our case, we have 606 different artists which need to be covered with 605 dummies. This was the reason why Kraeussl came up with his 2-step approach. His intention was to reduce the number of variables and thus the waste of degrees of freedom. Thus, he calculates in a first step a model without reputation. Calculates the reputation and insert the reputation in the model and calculate it again. In the next subsection we will prove why some of his assumptions are wrong.

7.2.1 Misconception of the 2-Step Approach

In Formula 7 until Formula 9 we showed the extraction of the reputation. This reputation term Rep_t^k correctly contains a beta coefficient (β_k) multiplied with the reputation (Rep_{t-}^{*k}).

$$Rep_t^k = \beta_k \cdot Rep_t^{*k}$$

Formula 11: Combination of Reputation and Coefficient

This coefficient is implicitly contained in the reputation. In the second step we enter this reputation again in the model. Thus, the result of this estimated coefficient should be one, since it was already in the first estimation. However, Kraeussl and I receive a slightly different value. I call this beta the *"reputation coefficient"* a.k.a. *"bias coefficient"*. This *"bias coefficient"* arises because of the separate estimation of the reputation in the first case. We left the reputation aside, which leads to biased coefficients for the other characteristics in the model. Now we go a step further. If we iterate the whole process, the "bias coefficient" should decrease with every step and converge to the exact solution which should be 1. Also, the coefficients of the other characteristics should converge to their exact values which in the end lead to a solution which should be close to the solution with the artist dummy approach. Figure 6 can show all of these ideas. The *"bias coefficient"* (blue line) decreases to one. The convergence is nearly completed after five iteration steps. The green line indicates the model which is used by Kraeussl. We see that the deviation of the *"bias coefficient"* for this model is around 10% of the exact value. The red line and black lines indicate the r-squared of the two approaches.

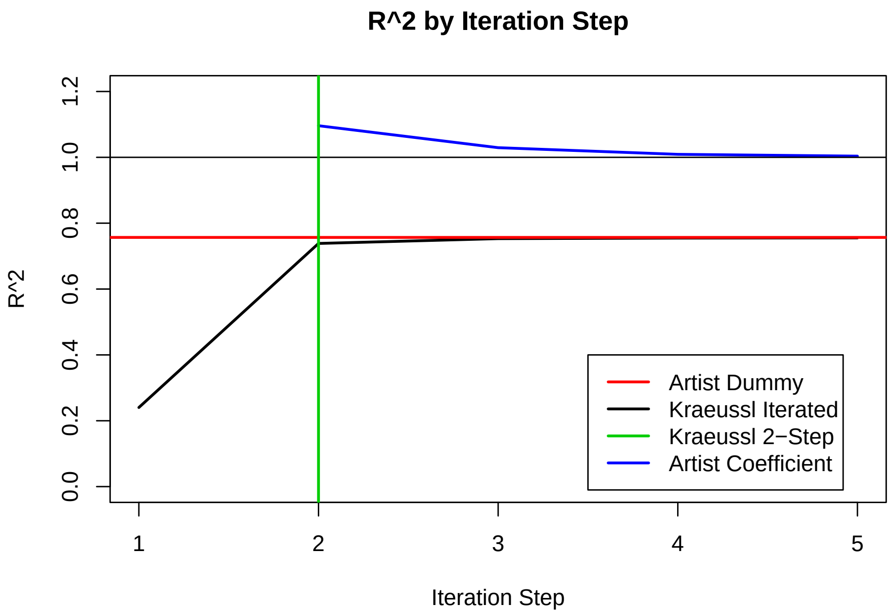

Figure 6: Model Evalution Dependent of Iteration Steps

The artist dummy approach has a constant r-squared of 0.7568. On the other hand the r-squared for the second step of the 2-step approach (green line) is 0.7385. After another step the value increases to 0.7531. After five steps we achieve nearly the identical value as with the artist dummy approach.

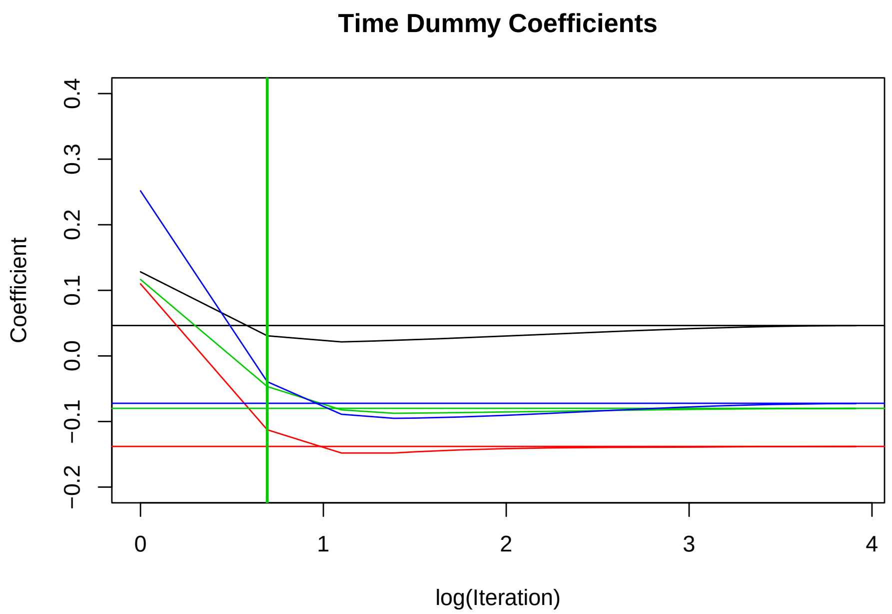

Figure 7: Time Dummy Coefficients (ex.) and log(Iteration)

Worse, than the values of the *"bias coefficient"*, is that the basis of the index is also biased. In Figure 7 the first four time dummy coefficients (year 1994-1997) including their exact solution, according to the artist dummy approach, are displayed. Only four lines are used to avoid confusion. Note the x-axis is log transformed to make the relevant area on the left better visible. Again, the vertical green line indicates the Kraeussl solution. It is obvious that some coefficients are substantially different from the exact solution.

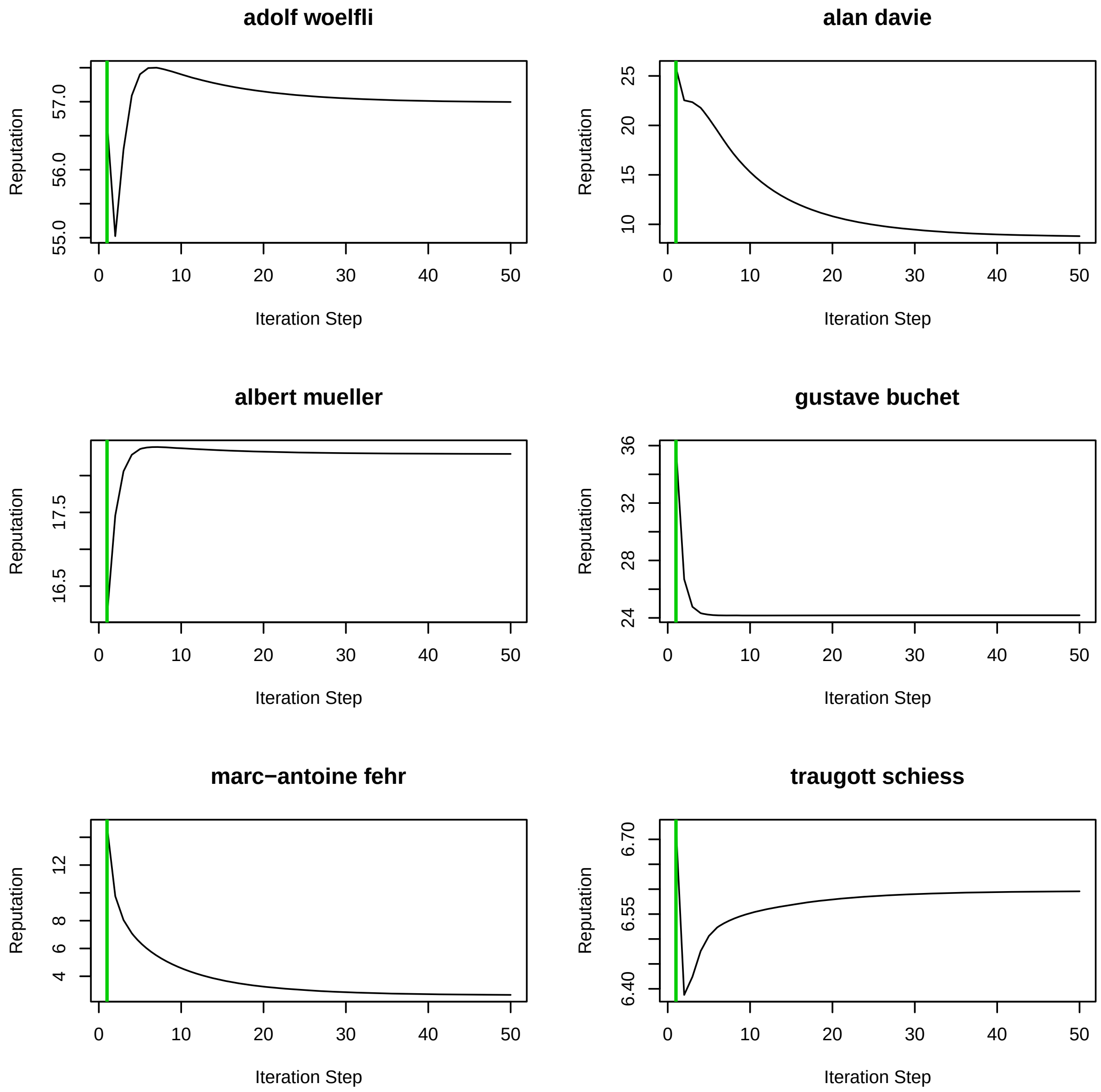

Figure 8: Artist Reputation Dependent of the Iteration Step (ex.)

This problem is supported by the development of the reputation. A few examples are shown in Figure 8. Some converge pretty fast, some rather slow. Some even show sharp bends. They could be triggered by substantial differences in the coefficients or in the reputation of related artists. This could lead to such severe differences from one step to another. As in the previous plots, the green line indicates the solution of Kraeussl which can be very misleading. This clearly shows that the 2-step approach leads to very inefficient estimations of some artists and thus wrong index values.

The summary of the problems of the Kraeussl 2-step approach:

- The solution is not fully converged, thus the coefficients are biased
- The estimation of the reputation suggests an exact solution, this leads to a loss of too few degrees of freedom, which leads in the end to too optimistic standard errors and thus lower uncertainty in the index

7.3 The Swiss Art Index

In the previous section we saw that the 2-step approach is a misconception, thus we further focus on the classic artist dummy approach. A possible problem brought up earlier is the choice of the reference artist. For the artist dummy approach this would be the dummy variable which is skipped. Diewert mentioned that “The problem with arbitrary choices is that the end results may not be invariant to these choices.”[34]

	Number of Sales	**Sum of Sales**	**Av. Sale Price**
ferdinand hodler	272	116'013'576	426'521
giovanni giacometti	162	56'511'630	348'837
félix vallotton	201	35'581'840	177'024
albert anker	134	29'712'672	221'736
cuno amiet	254	24'454'385	96'277
augusto giacometti	125	20'188'665	161'509
adolf dietrich	110	11'861'229	107'829
ernest biéler	147	10'706'392	72'833
gottardo segantini	64	7'384'020	115'375
robert zuend	56	4'967'014	88'697
giovanni segantini	39	4'929'910	126'408
wolfgang-adam toepffer	63	4'148'461	65'849
willy guggenheim	49	3'150'210	64'290
gustave buchet	65	2'320'860	35'706
albert mueller	36	2'297'270	63'813
arnold boecklin	12	2'290'830	190'903
max gubler	61	2'271'826	37'243
alice bailly	32	2'260'180	70'631
hermann scherer	12	2'211'046	184'254
johann heinrich fuessli	10	1'866'400	186'640

Table 5: Top 20 Swiss Artists

However, since the artist dummy approach is a closed system, a different reference artist would be compensated in the coefficients of the artist dummies which lead to the same solution for the time dummies and thus the index values. We used Ferdi-

[34] Diewert (2003), p. 32

nand Hodler as reference artist which could achieve the highest sum of sales during the period (see Table 5).

7.3.1 Index

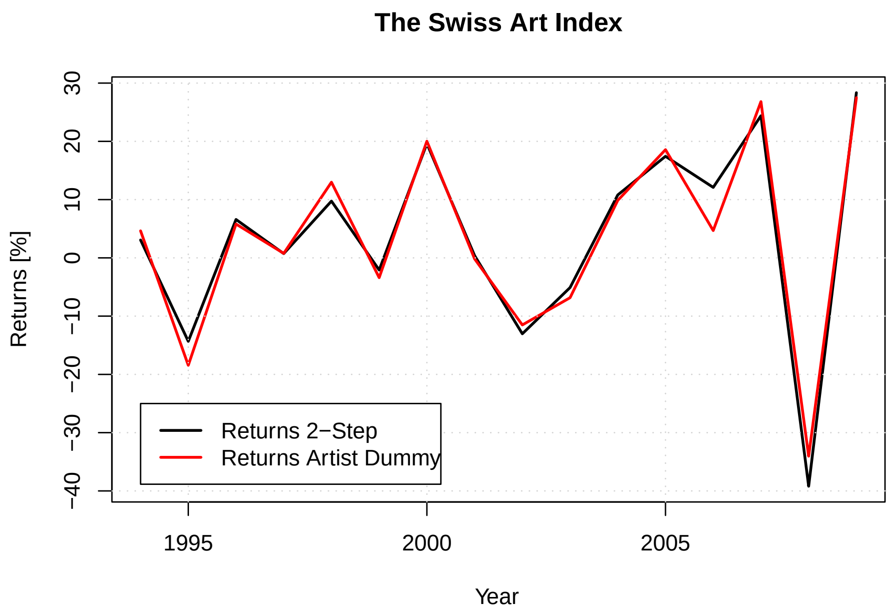

Figure 9: Swiss Art Index Returns

Formula 3 is used to calculate the Swiss Art Index. Table 12 (see Appendix D) shows the index value and the return of each year. The Swiss Art market performed on average 3.58% per year during the period of 1993 and 2009. The cumulated return is 57.26%. However, the volatility is rather high. The differences of the two approaches are rather small except in the year 2006 where the overestimation of the 2-step approach is nearly 10%.

Study	Year	Sample	Period	N	Nominal return	Real return	Data
1. Average per year							
Stein	1977	Pre-WW II paintings (in US)	1946-1968	8,950	10.47%		Art Prices Current
		Pre-WW II paintings (in UK)	1946-1968	35,823	13.12%		Art Prices Current
Worthington and Higgs	2004	Paintings	1976-2001	94,514	2.54%		Art Market Research
2. Geometric mean estimator							
Baumol	1986	Paintings	1652-1961	640		0.55%	Reitlinger
Frey and Pommerehne	1989	Paintings	1635-1987	1,198		1.5%	Reitlinger
		Paintings	1950-1987			1.7%	Reitlinger
3. Hedonic regression							
Anderson	1974	Paintings	1800-1970	> 13,000	3.3%		Reitlinger and Mayer
Buelens and Ginsburgh	1993	Paintings	1750-1961	ca. 5,900		0.91%	Reitlinger
Chanel et al.	1996	Paintings	1855-1969	1,972		4.9%	Reitlinger
Agnello and Pierce	1996	American paintings	1971-1992	15,216	9.3%		Art Sales Index
Renneboog and Van Houtte	2002	Belgian paintings	1970-1997	10,598	5.6%		Art Sales Index
Higgs and Worthington	2005	Australian paintings	1973-2003	37,605	6.96%		Austr. Art Auction Records
4. Repeat-sales regression							
Goetzmann	1993	Paintings	1716-1986	3,329		2.0%	Reitlinger and Mayer
		Paintings	1900-1986			13.3%	Reitlinger and Mayer
Pesando and Shum	1999	Picasso prints	1977-1996	8,257		1.48%	Gordon's Print Price Annual
Mei and Moses	2002	Paintings	1875-1999	4,896		4.9%	Mei & Moses
		Paintings	1950-1999			8.2%	Mei & Moses
Pesando and Shum	2008	Modern prints	1977-2004	80,214		1.51%	Gordon's Print Price Annual

Table 6: Overview of Past Results[35]

Renneboog and Spaenjers (2009) arranged and overview of different research papers. However, the comparison is very difficult since the analyzed periods, the index calculation approaches, the sample and the sample sizes are very inhomogeneous. Our result of 3.58% is in line with the range of other results achieved with hedonic regressions (Table 6, 3. paragraph).

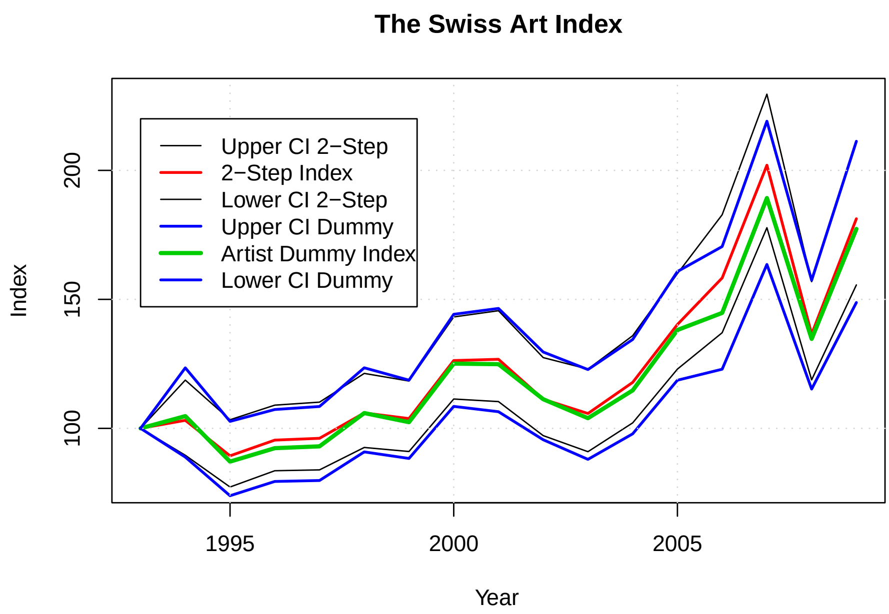

Figure 10: Swiss Art Index Development Including Confidence Interval

[35] Renneboog and Spaenjers (2009), p. 43

As already seen in the return plot (Figure 9) the market decreased sharply in 2008 and increased in a similar way the year before and after. This occurrence is like the results published by Kraeussl and Lee (2009). They found a similar decrease for 2008 and the top 500 Artists. Since the index values needs to be estimated with a hedonic regression model the result is not exact. The uncertainty is defined through the standard errors in the estimation of the coefficients. Thus, the confidence interval needs to be considered (blue lines). This uncertainty can easily be a deviation of 20% or even more. Therefore, based on the nature of the index the solution is not exact like other indices with observable market prices (e.g. stock market indices). As mentioned at the end of chapter 7.2.1, by pretending the exact calculation of the reputation not enough degrees of freedom are lost. As a result, the standard errors of the index are smaller as they should be. This is observable by considering the black lines in comparison to the correct solution (blue lines).

7.3.2 Model Evaluation

Above we had a look at the final result of the index. Now we look in more detail into the accuracy of the model. Note that all 605 artist dummies have been skipped in Table 7 to have a better overview. In every line we find the estimate of the coefficient (β and λ in Formula 1), the corresponding standard error which indicates the uncertainty of the coefficient and the t value of the corresponding test whether the estimate could be zero or not. The last column indicates with stars the significance of the test on the corresponding confidence level.

	Estimate	Std. Error	t value	Pr(>|t|)	
(Intercept)	7.1008	0.1501	47.3230	< 2e-16	***
Date1994	0.0464	0.0822	0.5640	0.5728	
Date1995	-0.1380	0.0823	-1.6760	0.0938	.
Date1996	-0.0800	0.0752	-1.0630	0.2879	
Date1997	-0.0723	0.0768	-0.9410	0.3469	
Date1998	0.0576	0.0766	0.7510	0.4524	
Date1999	0.0236	0.0738	0.3200	0.7489	
Date2000	0.2239	0.0711	3.1480	0.0017	**
Date2001	0.2222	0.0798	2.7830	0.0054	**
Date2002	0.1070	0.0761	1.4070	0.1596	
Date2003	0.0386	0.0833	0.4640	0.6429	
Date2004	0.1373	0.0794	1.7300	0.0837	.
Date2005	0.3230	0.0760	4.2530	0.0000	***
Date2006	0.3699	0.0817	4.5250	0.0000	***
Date2007	0.6380	0.0731	8.7310	< 2e-16	***
Date2008	0.2977	0.0780	3.8170	0.0001	***

Date2009	0.5726	0.0877	6.5300	0.0000	***
Auction.HouseSothebys	-0.0533	0.0277	-1.9280	0.0539	.
lnSurface	0.5647	0.0155	36.3740	< 2e-16	***
Techniqueoil on board	-0.0871	0.0566	-1.5370	0.1242	
Techniqueaquarell on paper	-0.8717	0.0705	-12.3750	< 2e-16	***
Techniquepencil on paper	-2.1535	0.0664	-32.4240	< 2e-16	***
Techniqueoil on wood	-0.0734	0.0738	-0.9940	0.3203	
Techniquepencil and aquarell on paper	-0.9173	0.0772	-11.8850	< 2e-16	***
Techniqueindian ink on paper	-1.3453	0.1008	-13.3490	< 2e-16	***
Techniquegouache on paper	-0.8472	0.1113	-7.6100	0.0000	***
Techniqueother media	-0.7805	0.0388	-20.1240	< 2e-16	***
AliveFALSE	0.1214	0.1791	0.6780	0.4977	
Signatureyes	0.2715	0.0374	7.2610	0.0000	***

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.8316 on 4586 degrees of freedom
Multiple R-squared: 0.7568,   Adjusted R-squared: 0.7232
F-statistic: 22.55 on 633 and 4586 DF, p-value: < 2.2e-16

Table 7: Hedonic Regression Result with Dummy Approach

The value of the R-squared is with 0.7568 pretty high. This indicates that three fourth of the total sample variance can be absorbed with our model, which is quite a good result. However, many of the time dummies are not significant. This means that the null hypothesis that the coefficient and thus the index values are different from zero could not be rejected. Therefore, we can find in Figure 10 the values of the lower confidence level below 100. A look at the next variable indicates that the sale prices in auctions held by Sotheby’s are slightly lower than the one of Christie’s. The surface coefficient is significant and positive which means that bigger paintings are worth more. On the other hand, the technique coefficients are all negative. This makes sense since we used oil on canvas as the reference dummy which is known as the worthiest technique. The most negative coefficients we find for pencil on paper and indian ink on paper. Commonly, these techniques are used to make drawings which normally are quite artless. Also the last two variables, whether a signature is on the picture or the painter is still alive are as expected. A signature increases the value of a picture since the artists is unambiguous allocatable. Also the death of an artist increases the value of his paintings since a natural shortage arises. However, this variable was not significant and thus has no impact on the explanation of the data. In a further step this variable should be dropped to increase the degrees of freedom and thus lowering the standard errors for the other coefficients.

After analyzing the model output, we also need to check the assumptions for the residuals.

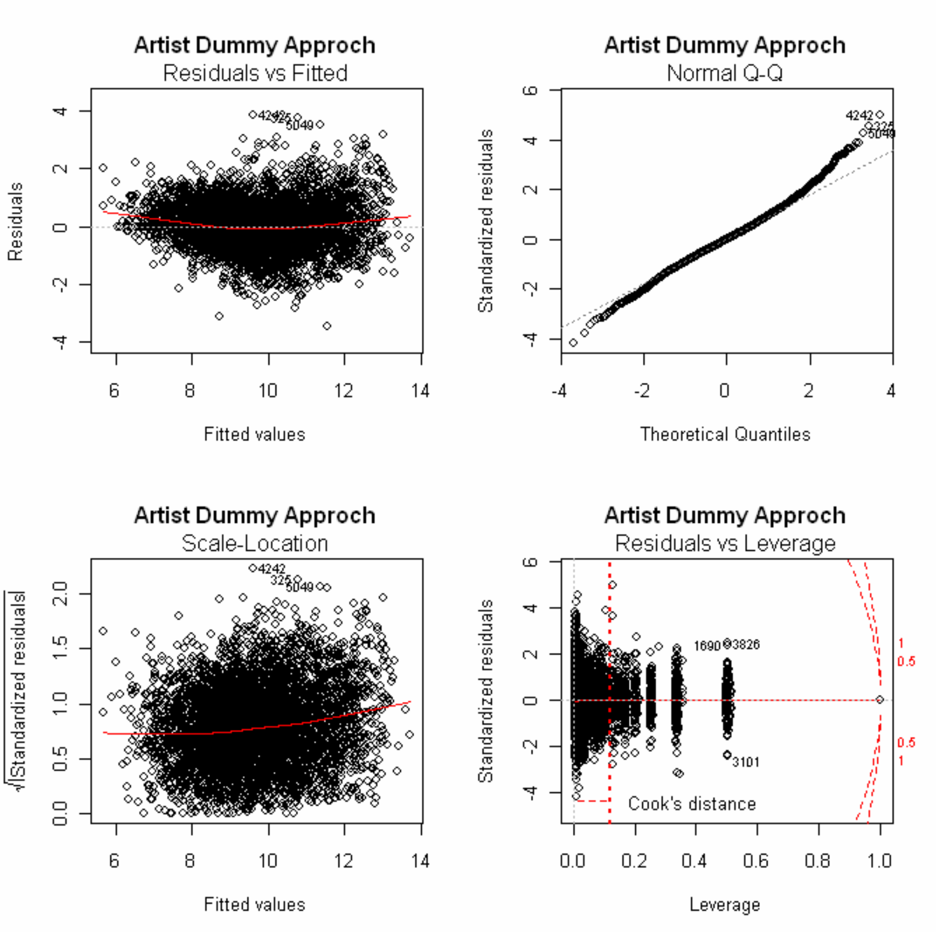

Figure 11: Residual Analysis of the Artist Dummy Approach

The Tukey-Anscombe Plot[36] on the upper left shows the residuals plotted against the fitted value. The corresponding assumption would be a constant standard deviation and symmetric residuals. The red line indicates that the mean is not constant and on both side positive. Q-Q plot (upper right) fits pretty well in the middle. However, both ends of the distribution show tendencies of heavy tails. The lower left figure shows a similar thing like the Tukey. On the lower right we see the residuals plotted against a leverage factor. The leverage is high if the observation has a big influence on the model. A bad combination is a big influence and a big residual. This leads to an observation with big influence on the final regression which is not desirable. Results on the right side of the critical lines (red) are critical.

[36] Anscombe and Tukey (1963)

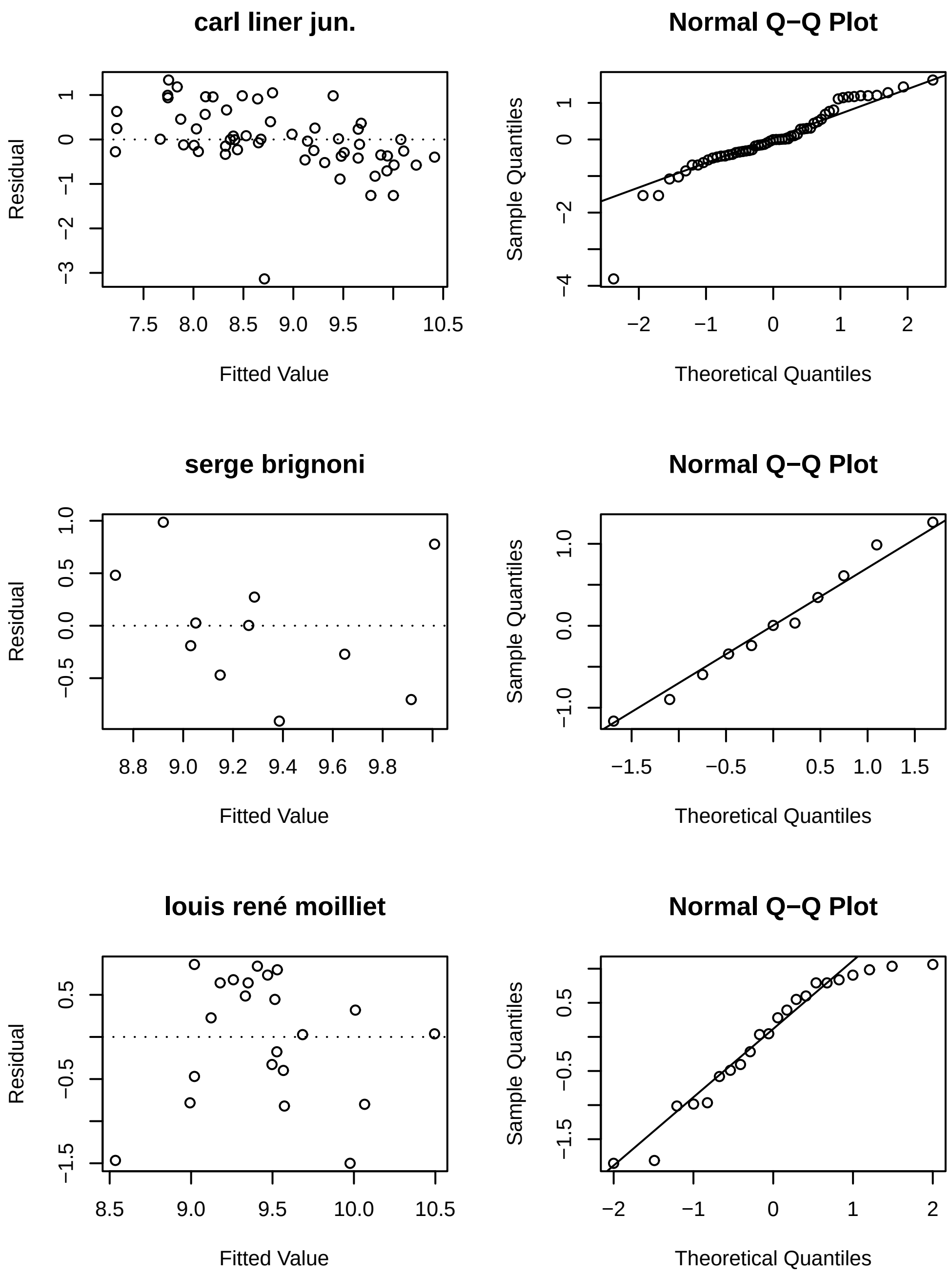

Figure 12: Residual Analysis of some Artist (ex.)

In Figure 12 we show a few examples of artists and the corresponding residual analysis. The two plots per line and artist are like the upper plots in Figure 11. Thus the interpretation is analogue. For Carl Liner jun. we see one outlier and maybe some tendencies that the mean of the residual is decreasing at the right side. The same outlier who is observable in the left plot is also visible in the Q-Q Plot on the very left. The rest of the observations fit pretty well. A very nice example is in the middle. The residuals are symmetric, the mean is more or less constant and thus, the residuals fulfil the normal distribution assumption. Finally, Louis René Moilliet shows some exposure to negative residuals. Also the Q-Q Plot indicates some problems on the right hand side of the distribution which is in this case light tailed. Since all continuous variables are already transformed, this measure, to make the residuals more like they should be according the assumptions, is already assigned. Another option would be using robust estimation methods. This should be improved anyway since our goal is to make the calculation applicable and then no residual analysis can be performed anymore. Based on the artist dummy coefficient we can calculate the reputation in a similar way like the Swiss Art Index.

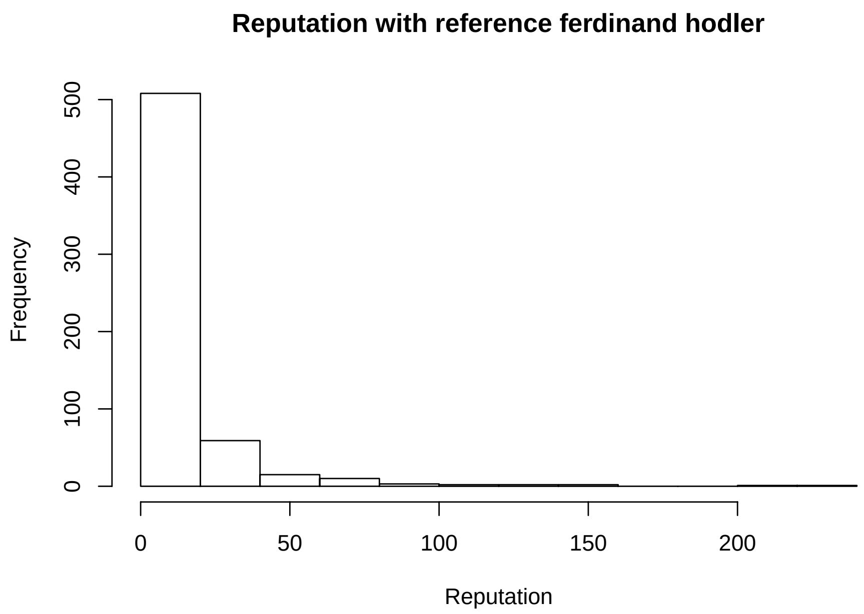

Figure 13: Reputation with Reference Artist Ferdinand Hodler

The Reputation of the reference artist Ferdinand Hodler is 100. We have a few other artists with a higher reputation but the majority is substantially lower. More detailed information about the top 50 artist can be found in the appendix (see appendix C). The grey marked artists have a higher reputation than Hodler. Some of them are quite famous but there are also a few names which do not sound familiar at all. Often the “unknown“ artists are sold very infrequently. This leads to only few results. If these paintings were sold with special characteristics or achieved higher prices than estimated according to the model, this would be compensated by a higher reputation.

7.3.3 Comparison of Different Asset Classes

In this chapter we compare the performance of Art Investments, with funds of hedge funds (HFRI FoF Composite), equities (MSCI World) and bonds (JPM GBI). The average logarithmic mean return for the period between 1993 and 2009 are beside the one of the equity market quite close. They range between 3.99% for bonds to 3.58% for art. However, the risk according to the standard deviation is for art (16.43%) much higher than for bonds (7.99%). Fund of hedge funds lie between with 10.09%.

Year	Swiss Art	HFRI FoF Comp	MSCI	JPM GBI
1993	100.00	100.00	100.00	100.00
1994	104.75	96.33	103.16	101.08
1995	87.11	104.08	119.07	117.28
1996	92.31	115.14	128.65	118.40
1997	93.03	128.65	141.24	115.45
1998	105.93	117.29	166.62	127.92
1999	102.39	142.64	197.98	116.77
2000	125.09	143.49	164.48	115.50
2001	124.88	146.02	133.79	113.43
2002	111.30	146.57	104.96	134.56
2003	103.94	162.15	136.09	152.72
2004	114.72	171.45	151.93	166.37
2005	138.13	179.75	159.38	151.66
2006	144.76	191.45	181.37	155.01
2007	189.28	205.26	188.87	167.02
2008	134.67	160.42	108.75	185.97
2009	177.29	178.56	137.89	189.23
Mean 1y Ret	3.58%	3.62%	2.01%	3.99%
SD 1y	16.43%	10.09%	21.35%	7.99%
Cum Ret	57.26%	57.97%	32.13%	63.78%

Table 8: Index Value of Different Asset Classes

Equities have a lower average return (2.01%) and a much higher risk (21.35%) than the others which makes them a very unattractive investment during the analyzed period. Of course the picture of the cumulated return needs to be similar than the average return. Art achieved 57.26% during the 16 year, whereas fund of hedge funds gained 57.97% and bonds slightly more with 63.78%. The global equity market had two bad periods in this phase, namely the dotcom bubble and the global financial crisis which ruined the favourable performance during the previous years making them the worst performing asset classes in the overview during the analysed time period.

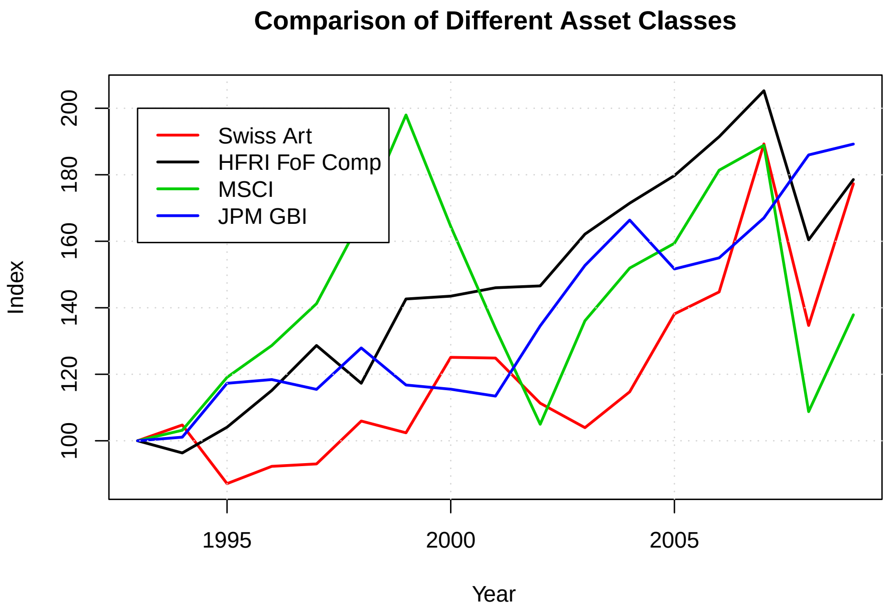

Figure 14: Comparison of Different Asset Classes

This is also observable in the figure above. The MSCI World index (green line) performed remarkably until 1999 exceeding all other asset classes by a factor two. In the three following years all the gains had been lost and the index started nearly at the same level like in 1993. This formation nearly copied in the period after 2002. During the first ten years Swiss Art moved around the anchor value of 100 points. After 2003 the prices increased sharply to peak at nearly 190 points. Bonds and hedge funds have more or less achieved a steady growth with rather small drops. Beside the risk return relationship, the correlations to other asset classes are also important to assess the potential of diversification.

	Swiss Art	HFRI FoF Comp	MSCI	JPM GBI
Swiss Art	1.00	0.81	0.39	0.71
HFRI FoF Comp	0.81	1.00	0.56	0.81
MSCI	0.39	0.56	1.00	0.20
JPM GBI	0.71	0.81	0.20	1.00

Table 9: Correlations of the Asset Classes

Swiss Art has a rather high correlation to bonds and hedge funds but a lower correlation to stocks. However, this is not in a range where the art investment becomes much more attractive. Based on this number it is possible to get similar exposures using bonds or hedge funds but with a much lower illiquidity risk.

The conclusion of these facts is that based on our calculation of the Swiss Art index the art market is a rather unattractive investment.

8 Conclusion

In this master thesis we applied two different approaches to construct an art index for the Swiss Art market. Namely, the 2-step hedonic approach and the classic artist dummy approach. However, we could show that the 2-step hedonic approach developed by Kraeussl is based on a misconception. The approach contains two major problems. [1] The solution achieved after 2 steps is not fully converged which leads to biased coefficients and thus a biased index. [2] The estimation of the reputation suggests an exact solution, which leads to a loss of too few degrees of freedom. The results lead to too optimistic standard errors and thus to a lower uncertainty in the index. In a next step the 2-step hedonic approach was extended to a x-step approach. The solution of the converged x-step approach was identical with the classic artist dummy approach.

The estimates and final prices of the 5231 auctions results gathered from Christie’s and Sotheby’s are analyzed to asses the accuracy of the judgement of the analysts. We could show that only every fourth result lies within the estimates and that for both auction houses the accuracy is similar.

75.68% of the total variance could be absorbed by our model. We could show that paintings sold at Christie’s achieved on average higher prices than sales at Sotheby’s. Other important factors are the surface, the technique, whether the artist is still alive or not and if a signature is on the painting. The results of the analyzed reputation were reasonable in comparison to our reference artist Ferdinand Hodler.

The Swiss Art Market performed during the period of 1993 to 2009 on average 3.58% which leads to a cumulative return of 57.26%. However these returns are below the ones of the bond and fund of hedge fund market. Also the correlation to these two markets is not favourable which leads to very low potential of diversification. This makes Swiss Art an unattractive investment during the analyzed period.

9 Outlook

In this chapter we want to bring up some unresolved issues which are important for the future project development. For a better overview we made two subsections.

9.1 Data

As mentioned during this report, the data gathering and preparation is one of the most crucial point for an art index. Thus, it is inevitable that these processes are standardized. Nevertheless, there will be some additional effort left which needs to be done manually by an analyst.

The techniques have been classified based on their frequency. The eight most common techniques build an own class whereas all others have been dropped into a class other media. However, the last class build a very inhomogeneous group which is not a desired solution. A classification by analysts could improve the quality of this variable.

9.2 Model

We used a set of variables based on the available data. It is possible to generate derivatives of such variables like to square the surface or other continuous variables to account for non linearity. Which additional variable should be added and which not needs further research.

For the Swiss Art index we assumed constant coefficients during the whole period. This applies for the coefficients as well as for the artist dummies. However, the reputation and the taste of the buyers can vary over time. To absorb such effects, maybe a smaller regression window is appropriate. The choice of the length of such a window is rather arbitrary. It will need more analysis work to find an adequate solution. This can be grouped to the second issue of Diewert (see chapter 4.2) whether a base or chain type approach should be used.

We assessed that the assumptions to the residuals are in general quite well fulfilled. However, there are some situations which are not optimal. For those it maybe makes sense to apply robust estimation.

Another issue which needs to be considered is the future update of the index. If we only expand the window of data, all past coefficients including the date dummy and thus the index would be modified. A possible solution could be to fix the history and only calculate the return based on the recent two dummies. With this return and the latest index value it is possible to calculate the new index value.

It is widely known that only few artists absorb a great part of the market. A brief analysis showed that the index returns for the top 10 or top 20 Swiss artists are significantly higher than the one of the broader market. The 20 index performed on average 6.1% and the top 20 even 7.8%. The critical step will be to find reasonable thresholds which data should be included in an index and which not.

In this thesis we were only able to calculate one index value a year. To build financial products a higher frequency is necessary. With more data a quarterly index should be feasible.

10 Declaration of Authorship

I hereby declare that this thesis is my own work, that it has been generated by me without the help of others and that all sources are clearly referenced.

I understand that the Zurich University of Applied Sciences (ZHAW) reserves the right to use plagiarism detection software to make sure that the content of this thesis is completely original. I hereby agree for this thesis, naming me as its author, to be sent abroad for this purpose, where it will be kept in a database to which this university has sole access. I also understand that I will be entitled at any time to ask for my name and any other personal data to be deleted.

Furthermore, I understand that pursuant to § 16 (1) in connection with § 22 (2) of the federal law on universities of applied sciences (FaHG) all rights to this thesis are assigned to the ZHAW, except for the right to be identified as its author.

Franziskus Dürr

………………………………………..

A References

Agnello R. J. (2002): Investment Returns and Risk for Art: Evidence from Auctions of American Paintings, Estern Economic Journal, 28(4), 443-464

Anderson R. (1974): Paintings as an investment, Economic Inquiry, Vol. 12, 13-26

Anscombe F. J., Tukey J. (1963): The Examination and Analysis of Residuals, Technometrics, 5, 2, 141-160

Ashenfelter O., Graddy K. (2003): Auctions and the Price of Art, The Journal of Economic Literature, 41, 763-787

Bastian Ch., Lange Y., Schulze P. (2004): Hedonische Preisindizes – Überblick und Anwendung auf Personalcomputer, Arbeitspapier Nr. 25, Mai 2004

Baumol W. J. (1985): Unnatural Value: Or Art Investment as Floating Crap Game, http://econ.as.nyu.edu/docs/IO/9398/RR85-25.pdf [13.02.2010]

Berndt E. R., Griliches Z., Rappaport N. J. (1995): Econometric Estimates of Price Indexes for Personal Computers in the 1990's, Journal of Econometrics, 68, 243-268.

Biey M., Zanola R. (2005): The Market for Picasso Prints: A Hybrid Model Approach, Journal of Cultural Economics, 29(2), 127-136

Buelens N., Ginsburgh V. (1993): Revisiting Baumol's "Art as Floating Crap Game", European Economic Review, 37, 1351-1371

Chanel O. (1995): Is art market behaviour predictable?, European Economic Review, 39, 519-527

Chanel O., Gérard-Varet L., Ginsburgh V. (1996): The Relevance of Hedonic Price Indices, Journal of Cultural Economics, March 1996, 20, 1

Court A. T. (1939): Hedonic price indexes with automotive examples: The Dynamics of Automobile Demand, The Genearl Motors Corporation, P. 99-117 http://www.rpi.edu/~simonk/pdf/gm1939.pdf [28.03.2010]

Czujack C. (1997): Picasso Paintings at Auction, 1963-1994, Journal of Cultural Economics, 21(3), 229-247

De la Barre M., Doccio S., Ginsburgh V. (1994): Returns on Impressionist, Modern and Contemporary European paintings, 1962-1991, Annales d'Economie et Statistique, 33, 143-181

Diewert E. (2003): Hedonic Regressions: A Review of Some Unresolved Issues, http://insee.fr/en/insee-statistique-publique/colloques/ottawa/pdf/paper_diewert.pdf [28.03.2010]

Griliches, Z. (1971): Hedonic Price Indexes for Automobiles: An Econometric Analysis of Quality Change: in Price Indexes and Quality Change: Studies in New Methods of Measurement, Harvard University Press

Henning T., Ridker R. (1967): The Determinants of Residential Property Values with Special Reference to Air Pollution, Review of Economics and Statistics, 44, 246-257

Hodgson D., Vorkink K. (2004): Asset Pricing Theory and the Valuation of Canadian Paintings, Canadian Journal of Economics, 37(3), 629-655

Hulten C. (2003): Price Hedonics: A Critical Review, Economic Policy Review, Volume 9, Number 3, http://www.newyorkfed.org/research/epr/03v09n3/0309hult.pdf [18.04.2010]

Kraeussl r., Lee J. (2009): Art as an Investment: The Top 500 Artists, http://papers.ssrn.com/sol3/papers.cfm?abstract_id=1536047 [10.04.2010]

Kraeussl R., Logher R. (2008): Emerging Art Markets, http://papers.ssrn.com/sol3/papers.cfm?abstract_id=1304856&download=yes [21.02.2010]

Kräussl R., Schellart E. (2007): Hedonic Pricing of Artworks: Evidence from German Paintings, http://papers.ssrn.com/sol3/papers.cfm?abstract_id=968198 [02.05.2010]

Kraeussl R., van Elsland N. (2008): Constructing the True Art Market Index – A Novel 2-Step Hedonic Approach and its Application to the German Art Market, http://papers.ssrn.com/sol3/papers.cfm?abstract_id=1104667 [24.02.2010]

Landefeld St., Grimm B. (2000): A Note on the Impact of Hedonics and Computers on Real GDP, Survey of Current Business, http://www.bea.gov/scb/pdf/beawide/2000/1200hm.pdf [28.03.2010]

Mei J., Moses M. (2002): Art as an Investment and the Underperformance of Masterpieces, http://papers.ssrn.com/sol3/papers.cfm?abstract_id=311701&download=yes [21.02.2010]

Moulton B. (2001): The Expanding Role of Hedonic Methods in the Official Statistics of the United States, http://www.bea.gov/about/pdf/expand3.pdf [29.03.2010]

Pakes, A. (2002): A Reconsideration of Hedonic Price Indices with an Application to PC's, Harvard University, http://papers.ssrn.com/sol3/papers.cfm?abstract_id=373101 [18.04.2010]

Papula, L. (2003): Mathematische Formelsammlung für Ingenieure und Naturwissenschaftler, 8. Auflage, Vieweg, Wiesbaden

Real Estate Transparency Index 2008: Jones Lang LaSalle, http://www.joneslanglasalle.com [13.02.2010]

Renneboog L., Spaenjers C. (2009): Buying Beauty: On Prices and Returns in the Art Market, http://arno.uvt.nl/show.cgi?fid=93224 [21.02.2010]

Renneboog L., Van Houtte T. (2002): The monetary appreciation of paintings: From realism to Magritte, Cambridge Journal of Economics; May 2002; 26, 3

Reitlinger G. (1961): The Economics of Taste: The Rise and Fall of the Picture Market 1760-1960, http://www.questia.com/PM.qst?a=o&d=54330719 [04.04.2010]

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

Sotheby's Investor Briefing (May 2009): Sotheby's, http://files.shareholder.com/downloads/BID/844559602x0x297908/a9f77a70-50d3-4736-a55e-868acf139baa/FY08%20IR%20slides.pdf [13.02.2010]

Tripplet J. (1991): Hedonic Methods in Statistical Agency Environments: An Intellectual Biopsy: in Fifty Years of Economic Measurement: The Jubilee of the Conference on Research in Income and-Wealth, 207-233, NBER Studies in Income and Wealth, vol. 54, University of Chicago Press.

Valsan C. (2002): Canadian versus American Art: What Pays Off and Why, Journal of Cultural Economics, 26, 203-216

Worthington A., Higgs H. (2005): Financial Returns and Price Determinants in the Australian Art Market, 1973-2003, Economic Record, 88, 113-123

Zanola R. (2007): The Dynamics of Art Prices: The Selection Corrected Repeat-Sales Index, http://polis.unipmn.it/pubbl/RePEc/uca/ucapdv/zanola85.pdf [21.02.2010]

B Summary of the 2-Step Hedonic Approach Model

	Estimate	Std. Error	t value	Pr(>|t|)	
(Intercept)	2.4440	0.1252	19.5250	< 2e-16	***
Date1994	0.0307	0.0704	0.4360	0.6632	
Date1995	-0.1126	0.0727	-1.5490	0.1215	
Date1996	-0.0467	0.0664	-0.7030	0.4820	
Date1997	-0.0394	0.0681	-0.5780	0.5631	
Date1998	0.0581	0.0675	0.8600	0.3896	
Date1999	0.0372	0.0656	0.5670	0.5706	
Date2000	0.2334	0.0628	3.7150	0.0002	***
Date2001	0.2372	0.0693	3.4250	0.0006	***
Date2002	0.1069	0.0677	1.5780	0.1146	
Date2003	0.0559	0.0754	0.7420	0.4581	
Date2004	0.1640	0.0715	2.2940	0.0218	*
Date2005	0.3383	0.0658	5.1430	0.0000	***
Date2006	0.4593	0.0720	6.3810	0.0000	***
Date2007	0.7032	0.0639	11.0060	< 2e-16	***
Date2008	0.3112	0.0695	4.4810	0.0000	***
Date2009	0.5947	0.0763	7.7950	0.0000	***
Auction.HouseSothebys	-0.1351	0.0240	-5.6200	0.0000	***
lnSurface	0.5002	0.0126	39.7340	< 2e-16	***
Techniqueoil on board	-0.1470	0.0474	-3.1040	0.0019	**
Techniqueaquarell on paper	-0.7874	0.0566	-13.9160	< 2e-16	***
Techniquepencil on paper	-1.9899	0.0581	-34.2600	< 2e-16	***
Techniqueoil on wood	-0.1528	0.0627	-2.4360	0.0149	*
Techniquepencil and aquarell on paper	-0.8189	0.0690	-11.8710	< 2e-16	***
Techniqueindian ink on paper	-1.0661	0.0759	-14.0380	< 2e-16	***
Techniquegouache on paper	-0.8840	0.0768	-11.5070	< 2e-16	***
Techniqueother media	-0.7066	0.0303	-23.2840	< 2e-16	***
AliveFALSE	-1.0665	0.0542	-19.6650	< 2e-16	***
Signatureyes	0.1533	0.0305	5.0210	0.0000	***
lnReputation	1.0961	0.0110	99.4200	< 2e-16	***

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.8107 on 5188 degrees of freedom
(2 observations deleted due to missingness)
Multiple R-squared: 0.7385, Adjusted R-squared: 0.737
F-statistic: 505.2 on 29 and 5188 DF, p-value: < 2.2e-16

Table 10: Summary of the 2-Step Hedonic Approach Model

C Top 50 Swiss Artists

Artist	Coefficient	Reputation
paul klee	1.4388	421.5550
sophie taueber-arp	1.1957	330.5871
jean-etienne liotard	0.8431	232.3652
alberto giacometti	0.7460	210.8444
johann heinrich fuessli	0.3687	144.5825
albert anker	0.3369	140.0543
giovanni giacometti	0.2554	129.1004
otto meyer-amden	0.2459	127.8810
angelika kauffmann	0.1048	111.0433
augusto giacometti	0.0662	106.8408
Ferdinand Hodler	0.0000	100.0000
giovanni segantini	-0.1015	90.3454
ernst ludwig kirchner	-0.1520	85.8963
johann jakob heuscher	-0.1937	82.3938
sam francis	-0.2874	75.0227
adolf dietrich	-0.3267	72.1315
arnold boecklin	-0.3336	71.6347
julius bissier	-0.3552	70.1040
marc-théodore bourrit	-0.4361	64.6534
félix vallotton	-0.4454	64.0555
johannes stauffacher	-0.4609	63.0735
anna barbara aemisegger-giezendanner	-0.4741	62.2439
charles-édouard jeanneret	-0.4879	61.3902
gottfried mind	-0.4943	60.9992
johann ludwig aberli	-0.5192	59.5008
andré derain	-0.5551	57.4009
adolf woelfli	-0.5623	56.9874
marcel broodthaers	-0.5623	56.9874
louis soutter	-0.5728	56.3961
gottardo segantini	-0.6409	52.6823
johannes mueller	-0.6479	52.3128
maria elena vieira da silva	-0.7016	49.5816
hermann hesse	-0.7151	48.9153
johann jakob biedermann	-0.7491	47.2778
robert zuend	-0.8580	42.4022
johann-baptist zeller	-0.8814	41.4211
ernest biéler	-0.9019	40.5790
cuno amiet	-0.9070	40.3734

hans purrmann	-0.9073	40.3633
serge poliakoff	-0.9267	39.5862
louis rené moilliet	-0.9269	39.5763
raphy dallèves	-0.9519	38.6003
hans arp	-0.9772	37.6360
adrian schiess	-0.9918	37.0901
peter robert berri	-1.0051	36.6026
ulrich-johann rutz	-1.0053	36.5942
richard paul lohse	-1.0117	36.3608
wolfgang-adam toepffer	-1.0407	35.3214
hermann scherer	-1.0446	35.1847
willy guggenheim	-1.0551	34.8147

Table 11: The Top 50 Swiss Artists

D Swiss Art Index Values

	Index	1y Return
1993	100.00	NA
1994	104.75	4.64%
1995	87.11	-18.44%
1996	92.31	5.81%
1997	93.03	0.77%
1998	105.93	12.98%
1999	102.39	-3.40%
2000	125.09	20.02%
2001	124.88	-0.17%
2002	111.30	-11.51%
2003	103.94	-6.84%
2004	114.72	9.87%
2005	138.13	18.57%
2006	144.76	4.69%
2007	189.28	26.82%
2008	134.67	-34.04%
2009	177.29	27.50%
Mean 1y Ret		3.58%
Cum Ret		57.26%

Table 12: Index Values and Return Artist Dummy Approach

E R Code (ex.)

```
# Module 12, The Swiss Art Index, Master Thesis
# MSc Banking and Finance
# Author: dufr
############################################################################
####


rm(list=ls())

setwd("M:/private/Master/4. Semester/Modul 12/")
library(xlsReadWrite)
library(MASS)
library(car)
library(fSeries)
library(xtable)

#Test whether choice of reference artist is relevant
refartist <- "hodler"
#refartist <- "anker"
#refartist <- "giovanni giacometti"

############################################################################
#
############################################################################
#
#Import adjusted data
SwissArt8 <- read.csv("Data/SwissArt8.csv",sep=",")

############################################################################
#
############################################################################
#
#Data preparation

dimnames(SwissArt8)[[2]]
#X"
# [2] "Jahr"
# [3] "Month"
# [4] "Day"
# [5] "Auction.House"
# [6] "Auction.Place"
# [7] "Sale.Number"
# [8] "Lot.Number"
# [9] "Bildtitel"
#[10] "Kuenstler"
#[11] "Nationality"
#[12] "Enstehungsjahr"
#[13] "Geburtsjahr.Künstler"
#[14] "Todesjahr.Künstler"
#[15] "Verkaufspreis.mit.Buyers.Premium..Hammer."
#[16] "Schätzpreis.tief.Sfr."
#[17] "Schätzpreis.hoch.Sfr."
#[18] "Grösse.cm"
#[19] "Signatur..yes.no."
#[20] "Style"
#[21] "Dim.1"
#[22] "Dim.2"
#[23] "Studie"

#remove Studys
```

```
#SwissArt8 <- SwissArt8[-which(SwissArt8[,23]!=""),]

#prepare data set for regression
regdata <- matrix(rep(NA,dim(SwissArt8)[[1]]*9),ncol=9)
dimnames(regdata)[[2]] <- c("lnPrice", "Date","Auction Hou-
se","lnSurface","Technique","Alive","Signature","Artist","lnReputation")
regdata <- data.frame(regdata)

regdata[,1] <- log(SwissArt8[,15])

regdata[,c(3,8)] <- SwissArt8[,c(5,10)]
regdata[,2] <- as.factor(SwissArt8[,2])
regdata[,4] <- log(SwissArt8[,21]*SwissArt8[,22])

write.csv(data.frame(table(SwissArt8[,20])),"Results/Reputation/style.csv")

#Technique
regdata[,5] <- as.character(SwissArt8[,20])
techniques <- c("oil on canvas","oil on board","aquarell on paper","pencil
on paper",
            "oil on wood","pencil and aquarell on paper","indian ink on pa-
per","gouache on paper")


for (i in 1:dim(regdata)[[1]]){
      if (length(grep(regdata[i,5],techniques))==0){
            regdata[i,5] <- "other media"
      }
      if (regdata[i,5]==""){
            regdata[i,5] <- "other media"
      }
      if (regdata[i,5]=="ink on paper"){
            regdata[i,5] <- "other media"
      }
}
data.frame(table(regdata[,5]))

regdata[,5] <- factor(regdata[,5],levels=c("oil on canvas","oil on
board","aquarell on paper","pencil on paper",
                        "oil on wood","pencil and aquarell on pa-
per","indian ink on paper","gouache on paper","other media"))


#Living Status death=True
regdata[,6] <- !is.na(SwissArt8[,14])
regdata[,6] <- factor(regdata[,6],levels=c("TRUE","FALSE"))

#Signature
regdata[,7] <- SwissArt8[,19]
which(SwissArt8[,19]==" ")

#Define refartist as first level
refid <- grep(refartist,levels(regdata[,8]))

#artist
refartid <- grep(refartist,levels(regdata[,8]))
regdata[,8] <- fac-
tor(regdata[,8],levels=c(levels(regdata[,8])[refartid],levels(regdata[,8])[
-refartid]))


#search whether all results are complete
apply(regdata,2,function(x) any(is.na(x)))
```

```
#remove sales with no dimensions
regdata <- regdata[-which(is.na(regdata[,4])),]

dimnames(regdata)[[2]]
#[1] "lnPrice"       "Date"          "Auction.House" "lnSurface"
#[5] "Technique"     "Living.status" "Signature"     "Artist"
#[9] "Reputation"

dim(regdata)

#necessary
options(contrasts=c("contr.treatment","contr.poly"))

apply(regdata,2,function(x) is.factor(x))

############################################################################
######
############################################################################
######
#Step 1: Estimate model without reputation

years <- sort(unique(regdata$"Date"))

step1.lm <- lm(regdata[,1]~.,data=regdata[,-c(1,8,9)])
summary.step1 <- summary(step1.lm)
summary(step1.lm)

#step1.lm.aic <- stepAIC(step1.lm, scope=list(upper = ~ ., lower = ~ Date),
direction="both")
#
#summaries.aic <- summary(step1.lm.aic)

#stepAIC(step1.lm,direction="both")

step1.lm.infl <-lm.influence(step1.lm)

par(mfrow=c(2,2))
plot(step1.lm,main="Step1")
#Huber line
abline(v=2*step1.lm$rank/dim(regdata)[[1]], lty=3,col=2,lwd=2)
savePlot(filename ="Results/Reputation/ModelanalysisStep1.eps",type =
c("eps"))
dev.off()

write.csv(summary.step1$"coeff","Results/Reputation/step1.lm.csv")


############################################################################
######
############################################################################
######
#Step 2: Estimate reputation

dimnames(regdata)[[2]]
#[1] "lnPrice"       "Date"          "Auction.House" "lnSurface"
#[5] "Technique"     "Living.status" "Signature"     "Artist"
#[9] "Reputation"

artists <- unique(regdata[,8])
reference <- unique(regdata[,8])[grep(refartist,unique(regdata[,8]))]

repabove <- rep(NA,length(artists))
```

```
repbelow <- rep(NA,length(artists))
reputation <- rep(NA,length(artists))

artistscontr <- rep(NA,length(artists))

for (j in 1:length(artists)){

            refid <- grep(as.character(reference),regdata[,8])
            artid <- grep(as.character(artists[j]),regdata[,8])

            if (length(artid)==0){
                  print("No result for this artist in period")
            }else{
            repabove[j] <-
(prod(exp(regdata[artid,1])^(1/length(artid))))/(prod(exp(regdata[refid,1])
^(1/length(refid))))

            delta <- list()
            dataid <- c(2,3,4,5,6,7)
            #dataid <- c(2,3,4,5,6,7,8)
            for (k in dataid){
                  if (k==4){
                  delta[[k-1]] <- sum(regdata[artid,k])/length(artid) -
sum(regdata[refid,k])/length(refid)


                  }else{
                  delta[[k-1]] <- (table(regdata[artid,k])/length(artid) -
table(regdata[refid,k])/length(refid))[-1]

                  }

            }
            if (length(unlist(delta))!=length(summary.step1$coeff[-1,1])){
                  print("Different lengths in calculation of the reputa-
tion")
            }else{
            repbelow[j] <- exp(sum(unlist(delta)*summary.step1$coeff[-
1,1],na.rm=T))


            reputation[j] <- repabove[j]/repbelow[j]
            regdata[artid,9] <- log(reputation[j]*100)
            #regdata[artid,9] <- repabove[j]*100
            }
      }

      artistscontr[j] <- artists[j]

      print(j)
}

sum(reputation*100>100,na.rm=T)

#cbind(levels(artists)[artists],levels(artists)[artistscontr])

wri-
te.csv(cbind(levels(artists)[artists],levels(artists)[artistscontr]),"Resul
ts/Reputation/Reputation_Contr.csv")

wri-
te.csv(cbind(levels(artists)[artists],reputation*100),"Results/Reputation/R
eputation.csv")
```

```
windows(width = 7, height = 5)
hist(exp(regdata[artists,9]),main=paste("Reputation with refer-
ence",reference,sep=" "),xlab="Reputation")
savePlot(filename ="Results/Reputation/Reputation.eps",type = c("eps"))
dev.off()

###########################################################################
######
###########################################################################
######
#Step 3: Estimate model including reputation

step3.lm <- lm(regdata[,1]~.,data=regdata[,-c(1,8)])
summary.step3 <- summary(step3.lm)
summary(step3.lm)


step3.lm.aic <- stepAIC(step3.lm, scope=list(upper = ~ ., lower = ~ Date),
direction="both")

summaries.aic <- summary(step3.lm.aic)

#stepAIC(step3.lm,direction="both")

step3.lm.infl <-lm.influence(step3.lm)

par(mfrow=c(2,2))
plot(step3.lm,main="Step3")
#Huber line
abline(v=2*step3.lm$rank/dim(regdata)[[1]], lty=3,col=2,lwd=2)
savePlot(filename ="Results/Reputation/ModelanalysisStep3.eps",type =
c("eps"))
dev.off()

write.csv(summary.step3$"coeff","Results/Reputation/step3.lm.csv")


pdf("Results/Reputation/Modelanalysis_per_Artist_2Step.pdf")

artists2 <- levels(artists)[artists]
for (i in 1:length(artists2)){
      artid <- which(artists2[i]==regdata[,8])
      par(mfrow=c(2,2))
      plot(fitted(step3.lm)[artid],resid(step3.lm)[artid],main=artists2[i])
      abline(h=0,lty=3)

      qqnorm(stdres(step3.lm)[artid])  # QQ-Plot of the residuals
      qqline(stdres(step3.lm)[artid])

      plot(fitted(step3.lm)[artid],sqrt(abs(stdres(step3.lm)[artid])))
      abline(h=0)


      plot(step3.lm.infl$hat[artid], stdres(step3.lm)[artid],
xlim=range(c(0,step3.lm.infl$hat[artid]),na.rm=T)) # standart. Res.
      abline(h=0,lty=3)
      abline(v=2*step3.lm$rank/dim(regdata[artid,])[[1]],
lty=3,col=2,lwd=2)

#     par(mfrow=c(2,2))
#     plot(step3.lm)
      print(i)
```

```
}
dev.off()


##########################################################################
######
##########################################################################
######
#Step 4: Calculate index

swissartindex <- matrix(rep(NA,17*4),ncol=4)
#cum-
prod(exp(summary.step3$"coeff"[2:17,1])/exp(c(0,summary.step3$"coeff"[2:16,
1])))*100
swissartindex[1,1:3] <- 100
dimnames(swissartindex)[[1]] <- c(1993:2009)
dimnames(swissartindex)[[2]] <- c("Index 2-Step","Lower CI","Upper CI","1y
Returns 2-Step")

coefficient <- rbind(c(0,0),summary.step3$"coeff"[2:17,1:2])

#exp(coefficient[,1])
#exp(coefficient[,1]-2*coefficient[,2])
#exp(coefficient[,1]+2*coefficient[,2])

for (i in 1:16){
      swissartindex[i+1,1] <-
exp(coefficient[i+1,1])/exp(coefficient[i,1])*swissartindex[i,1]
      swissartindex[i+1,2] <- exp(coefficient[i+1,1]-
2*coefficient[i+1,2])/exp(coefficient[i,1]-
2*coefficient[i,2])*swissartindex[i,2]
      swissartindex[i+1,3] <-
exp(coefficient[i+1,1]+2*coefficient[i+1,2])/exp(coefficient[i,1]+2*coeffic
ient[i,2])*swissartindex[i,3]

}

##########################################################################
#####
#VIF
vif(step3.lm)
#GVIF Df GVIF^(1/2Df)
#Date          1.181321 16     1.005221
#Auction.House 1.076162  1     1.037382
#lnSurface     1.467062  1     1.211223
#Technique     1.643021  8     1.031520
#Alive         1.073271  1     1.035988
#Signature     1.077862  1     1.038201
#lnReputation  1.096039  1     1.046919



##########################################################################
######
##########################################################################
######
#Estimate model with artist dummy

artistdummy.lm <- lm(regdata[,1]~.,data=regdata[,-c(1,9)])
summary.artistdummy.lm  <- summary(artistdummy.lm )
```

```
summary(artistdummy.lm)
dummy.coef(artistdummy.lm)[-8]

artistdummy.lm.infl <-lm.influence(artistdummy.lm)

par(mfrow=c(2,2))
plot(artistdummy.lm,main="Artist Dummy Approch")
#Huber line
abline(v=artistdummy.lm$rank/dim(regdata)[[1]], lty=3,col=2,lwd=2)
savePlot(filename
="Results/Reputation/residual_analysis_artist_dummy.bmp",type = c("bmp"))
dev.off()

write.csv(summary.artistdummy.lm,"Results/Reputation/artist.dummy.lm.csv")
write.csv(artistdummy.lm,"Results/Reputation/artist.dummy.lm.csv")
xtable(artistdummy.lm,)
xtable.summary.lm(summary.artistdummy.lm)#,"Results/Reputation/artist.dummy
.lm.tex")



pdf("Results/Reputation/Modelanalysis_per_Artist_Artist_Dummy_selection.pdf
")

artists2 <- levels(artists)[artists]

artists2 <- artists2[c(57,64,68)]

windows(width = 9, height = 12)
par(mfrow=c(3,2))
for (i in 1:length(artists2)){
      artid <- which(artists2[i]==regdata[,8])

      plot(fitted(artistdummy.lm)[artid],resid(artistdummy.lm)[artid],main=
artists2[i],xlab="Fitted Value",ylab="Residual")
      abline(h=0,lty=3)

      if (length(artid)!=1){
      qqnorm(stdres(artistdummy.lm)[artid])  # QQ-Plot of the residuals
      qqline(stdres(artistdummy.lm)[artid])

#
      plot(fitted(artistdummy.lm)[artid],sqrt(abs(stdres(artistdummy.lm)[ar
tid])))
#     abline(h=0)
#
#
#     plot(artistdummy.lm.infl$hat[artid], stdres(artistdummy.lm)[artid],
xlim=range(c(0,artistdummy.lm.infl$hat[artid]),na.rm=T)) # standart. Res.
#     abline(h=0,lty=3)
#     abline(v=2*artistdummy.lm$rank/dim(regdata[artid,])[[1]],
lty=3,col=2,lwd=2)
#
#     par(mfrow=c(2,2))
#     plot(artistdummy.lm)
      }
      print(i)

}
savePlot(filename
="Results/Reputation/resid_analysis_artist_sample.eps",type = c("eps"))
dev.off()
```

```
sum(100*exp(summary(artistdummy.lm)$"coef"[-c(1:29),1])>100,na.rm=T)

#plot reputation
windows(width = 7, height = 5)
hist(100*exp(summary(artistdummy.lm)$"coef"[-
c(1:29,523,587),1]),main=paste("Reputation with reference",reference,sep="
"),xlab="Reputation")
savePlot(filename ="Results/Reputation/Reputation.eps",type = c("eps"))
dev.off()
############################################################################
##########
#Calculate Artist Dummy index

swissartindexdummy <- matrix(rep(NA,17*4),ncol=4)
#cum-
prod(exp(summary.step3$"coeff"[2:17,1])/exp(c(0,summary.step3$"coeff"[2:16,
1])))*100
swissartindexdummy[1,1:3] <- 100
dimnames(swissartindexdummy)[[1]] <- c(1993:2009)
dimnames(swissartindexdummy)[[2]] <- c("Index Artist Dummy","Lower
CI","Upper CI","Return Artist Dummy")

coefficientdummy <- rbind(c(0,0),summary.artistdummy.lm $"coeff"[2:17,1:2])

#exp(coefficientdummy[,1])
#exp(coefficientdummy[,1]-2*coefficientdummy[,2])
#exp(coefficientdummy[,1]+2*coefficientdummy[,2])

for (i in 1:16){
      swissartindexdummy[i+1,1] <-
exp(coefficientdummy[i+1,1])/exp(coefficientdummy[i,1])*swissartindexdummy[
i,1]
      swissartindexdummy[i+1,2] <- exp(coefficientdummy[i+1,1]-
2*coefficientdummy[i+1,2])/exp(coefficientdummy[i,1]-
2*coefficientdummy[i,2])*swissartindexdummy[i,2]
      swissartindexdummy[i+1,3] <-
exp(coefficientdummy[i+1,1]+2*coefficientdummy[i+1,2])/exp(coefficientdummy
[i,1]+2*coefficientdummy[i,2])*swissartindexdummy[i,3]

}


swissartindexdummy[,4]<- returns(swissartindexdummy[,1])

summary(swissartindexdummy[,4])

############################################################################
#########
#figure Index

windows(width = 7, height = 5)
plot(1993:2009,swissartindex[,1],type="l",xlab="Year",ylim=c(min(apply(swis
sartindex[,-4],1,min,na.rm=T)),max(apply(swissartindex[,-
4],1,max,na.rm=T))),ylab="Index",main="The Swiss Art Index",col=2,lwd=2)
lines(1993:2009,swissartindex[,2],col=1)
lines(1993:2009,swissartindex[,3],col=1)
lines(1993:2009,swissartindexdummy[,1],col=3,lwd=3)
lines(1993:2009,swissartindexdummy[,2],col=4,lwd=2)
lines(1993:2009,swissartindexdummy[,3],col=4,lwd=2)

grid()
```

```
legend(1993,220,c("Upper CI 2-Step","2-Step Index","Lower CI 2-Step","Upper
CI Dummy","Artist Dummy Index","Lower CI Dum-
my"),col=c(1,2,1,4,3,4),lwd=c(1,2,1,2,3,2),bg="white")
savePlot(filename ="Results/Reputation/TheSwissArtIndex.eps",type =
c("eps"))
dev.off()

swissartindex[,4]<- returns(swissartindex[,1])



wri-
te.csv(cbind(swissartindex,swissartindexdummy),"Results/Reputation/SwissArt
Index.csv")

##########################################################################
############
#Returns Index
windows(width = 7, height = 5)
plot(1994:2009,swissartindex[-
1,4]*100,type="l",xlab="Year",ylim=c(min(100*swissartindex[,4],na.rm=T),max
(100*swissartindex[,4],na.rm=T)),ylab="Returns [%]",main="The Swiss Art In-
dex",col=1,lwd=2)
#abline(h=0,lty=2)
lines(1994:2009,100*swissartindexdummy[-1,4],col=2,lwd=2)
grid()
legend(1994,-25,c("Returns 2-Step","Returns Artist Dum-
my"),col=c(1,2),lwd=c(2,2),bg="white")
savePlot(filename ="Results/Reputation/TheSwissArtIndex_Returns.eps",type =
c("eps"))
dev.off()

##########################################################################
############
#Comparison with other asset classes
indices <- read.csv("Results/Reputation/Index_ALL.csv",sep=";")
dimnames(indices)[[2]] <- c("Year","Swiss Art","HFRI FoF Comp","MSCI","JPM
GBI")

for (i in 2:5){
      indices[,i] <- indices[,i]/indices[1,i]*100
}


windows(width = 7, height = 5)
plot(indices[,1],indices[,2],type="l",col=2,lwd=2,ylim=c(min(apply(indices[
,2:5],1,min,na.rm=T)),max(apply(indices[,2:5],1,max,na.rm=T))),xlab="Year",
ylab="Index",main="Comparison of Different Asset Classes")
lines(indices[,1],indices[,3],col=1,lwd=2)
lines(indices[,1],indices[,4],col=3,lwd=2)
lines(indices[,1],indices[,5],col=4,lwd=2)
grid()
legend(1993,200,c("Swiss Art","HFRI FoF Comp","MSCI","JPM
GBI"),col=c(2,1,3,4),lwd=c(2,2,2,2),bg="white")
savePlot(filename ="Results/Reputation/Index_Comparison.eps",type =
c("eps"))
dev.off()


write.csv(indices,"Results/Reputation/Indices_ALL.csv")

write.csv(cor(indices[,c(2:5)]),"Results/Reputation/Indices_COR.csv")
```