Tilburg University

In Art We Trust

Li, Yuexin; Ma, X.; Renneboog, Luc

Publication date:
2021

Document Version
Early version, also known as pre-print

IN ART WE TRUST

By

Yuexin Li, Marshall X. Ma,
Luc Renneboog

TILBURG ◆ UNIVERSITY

In Art We Trust

Yuexin Li, Xiaoyin Ma, and Luc Renneboog[1]

ABSTRACT

While trust is the cornerstone of any market's functioning, it is of particular importance in markets that are unregulated, illiquid, and opaque, such as the art market. This study examines the role of authenticity, as captured by provenance information in auction catalogs, on the probability of auctioned oil paintings, watercolors, and prints being sold, their price formation, and returns. Auction catalogs include four authenticity dimensions: pedigree (ownership "blockchain," descendance information, type of past owners, such as renowned collectors, and past sales records); exhibition history (e.g., in famous museums or galleries); literature coverage (e.g. in *catalogues raisonnés* or authoritative press), and certification (e.g. artist's physical testimonial, experts' opinions). We find that trust, proxied by provenance information, increases the probability of a work being sold by up to 4%, leads to hammer price premiums up to 54%, and increases annualized returns by 5% to 16%. To address potential endogeneity problems between the provision of provenance, and past prices/price expectations, we perform quasi-natural experiments in difference-in-differences settings on auction houses' provenance policy changes following authenticity litigation, and on a contamination effect of the discovery of fakes and forgeries on the oeuvre of forged artists. We also test transactions less affected by past prices, such as estate sales following the death of a collector. The findings on the relation between provenance and prices are robust to artist reputation, artistic style, auction house reputation, art market liquidity, and artist career timing.

Keywords: auction; hedonic pricing; art investment; art returns; auction house

JEL Codes: D44, G20, G11, Z11

In Art We Trust

ABSTRACT

While trust is the cornerstone of any market's functioning, it is of particular importance in markets that are unregulated, illiquid, and opaque, such as the art market. This study examines the role of authenticity, as captured by provenance information in auction catalogs, on the probability of auctioned oil paintings, watercolors, and prints being sold, their price formation, and returns. Auction catalogs include four authenticity dimensions: pedigree (ownership "blockchain," descendance information, type of past owners, such as renowned collectors, and past sales records); exhibition history (e.g., in famous museums or galleries); literature coverage (e.g. in *catalogues raisonnés* or authoritative press), and certification (e.g. artist's physical testimonial, experts' opinions). We find that trust, proxied by provenance information, increases the probability of a work being sold by up to 4%, leads to hammer price premiums up to 54%, and increases annualized returns by 5% to 16%. To address potential endogeneity problems between the provision of provenance, and past prices/price expectations, we perform quasi-natural experiments in difference-in-differences settings on auction houses' provenance policy changes following authenticity litigation, and on a contamination effect of the discovery of fakes and forgeries on the oeuvre of forged artists. We also test transactions less affected by past prices, such as estate sales following the death of a collector. The findings on the relation between provenance and prices are robust to artist reputation, artistic style, auction house reputation, art market liquidity, and artist career timing.

1. Introduction

Acquiring an asset not only requires an assessment of the risk–return trade-off, it also involves an act of trust in the soundness of information reflecting fundamental value and in the fairness of the overall (trading) system (Guiso, Sapienza, and Zingales 2008). Trustworthy and reliable information reduces the information asymmetry between trading parties. Equity and bond investors depend on annual reports, independent audit reports, credit ratings, analysts' forecasts, and information from financial intermediaries, among others. Reliable information is even more important for markets for alternative investments, such as art markets, which differ from financial markets in several aspects.

First, except for the market for prints, art objects offered for sale are *unique* pieces, the value of which is determined by the characteristics of the art object (e.g., topic, medium, and artist reputation) but also by subjective non-monetary ownership utility (Renneboog and Spaenjers 2013), and by the resale option value that is affected by demand factors (e.g., wealth concentration, equity market evolution, income inequality, or changes in art-collecting audiences) (Goetzman, Renneboog and

Spaenjers 2011; Lovo and Spaenjers 2018; Pénasse, Renneboog and Scheinkman 2021). This implies that pinpointing the "fundamental" value of an art object is much more difficult than for standardized financial assets. Second, in contrast to the considerable amount of information generated about corporations with traded equity and debt, information in the art market is *opaque*, since for the majority of auctioned art, only sparse information is available, such as the artist's name, title, topic, medium, measurements, auction house, date of sale, and lot number, which are the variables traditionally included in hedonic pricing regressions. Prices may also be slow to reflect changing valuations, as the *lack of short selling* induces delays in the incorporation of negative information in prices. Third, the art market is *illiquid*. A decision to sell a painting could take long to be executed, as the right type of auction might not be available for a specific type of art. Fourth, the secondary art market comprises hundreds of intermediaries (auction houses or dealers) around the world and is *not regulated*. A problem in the art world is that fakes and forgeries are occasionally discovered, which may erode trust, reduce market participation, and lead to lower valuations. It is rumored that up to 40 to 50% of the high-end modern art market consists of forgeries (Thompson 2010)[2].

The art auction market is an important alternative financial market. High-net-worth individuals (HNWIs) hold on average 9% of their investment portfolios in art and other types of collectibles (e.g., Bordeaux wines, classic cars, and collector watches). The total value of collectibles held by HNWIs is estimated at more than USD 4 trillion (Deloitte Luxembourg and ArtTactic 2016). Art sales through auction houses and internet auctions have grown rapidly over the past two decades (Deloitte Luxembourg and ArtTactic 2014, 2016) and global art sales exceeded USD 40 billion in 2015 and 2016 (Pownall 2017). The finance and economics literature has focused on the risk–return relationship of art (Mei and Moses 2002; Renneboog and Spaenjers 2013; Korteweg, Kräussl, and Verwijmeren 2016; Lovo and Spaenjers 2018; Li, Ma, and Renneboog 2022); its macro-economic market drivers (Goetzmann et al. 2011); sentiment and hype (Pénasse, Renneboog, and Spaenjers 2014); supply shocks (Pénasse, Renneboog, and Scheinkman 2021); and behavioral anomalies, such as anchoring (Beggs and Graddy 2009; Graddy et al. 2015), gender bias (Adams et al. 2021; Cameron, Goetzmann, and Nozari 2019), or over-extrapolation (Pénasse and Renneboog 2022). However, this broad literature has little to say on the fundamental question of art trade, that is, how information disclosure and trust affect the market.

In the art market, information on provenance is especially important, as it may increase confidence in the authenticity of the art object offered for sale. This is particularly true if the provenance information offers details about an object's pedigree (e.g., can the ownership chain uninterruptedly be traced back to the artist?), the literature (e.g., is the art object referenced in art history books or

---

[2] One of the art market's greatest challenges for paintings by dead artists is to verify their authenticity and provenance. The Fine Arts Expert Institute (FAEI) in Geneva claims that over 50% of the artworks circulating on the market are either forged or cannot be attributed to the correct artist. Source: https://news.artnet.com/market/over-50-percent-of-art-is-fake-130821

catalogues raisonnés?), exhibitions (has the object been exhibited by museums or galleries?), and certification (is there any evidence in the form of old photographs of the art object, certificates by the artist, certified witness statements, etc.?). Pedigree refers to past ownership, which ideally can be traced back to the primary market with sales by the artists themselves or their galleries. Pedigree also comprises inheritance information within the artists' families and about collectors' families. It can be regarded as the biography of a piece of art that lived through the decades, possibly even centuries, and records the owners down to the present day. Traceability may translate into trustworthiness. There is a general upward trend of tracking ownership and providing more detailed provenance. In the current digital age, pedigree can be recorded by means of blockchain technology, which could revolutionize the logging of ownership, as records cannot be forged. Companies such as Verisart provide digital certification of art transactions, which can be applied to both physical transactions as well as digitalized art by means of NFTs (Chohan 2021; Fairfield 2022). The blockchain initiative related to physical art transactions is still in infancy, and detailed analyses of provenance (or "blockchain" by means of textual analysis on auction catalogs), as undertaken in the present study, will remain important for many decades. Our study of the art auction market comprises paintings created in the pre-blockchain area; we consider an epoque of artistic painting spanning more than 500 years (from medieval old masters to contemporary art).

The pedigree dimension also includes comprehensive lists of specific types of people appearing in the ownership chain, such as nobility and royalty, wealthy business professionals, influential politicians, celebrities, sportspeople, and other prominent individuals. The second provenance category, exhibition history, documents previous exhibitions in museums, galleries, and art fairs. Prominent exhibitions can serve as a filter for authenticity and a quality indicator, since exhibitions are curated and often accompanied by an exhibition catalog, which may include new research on the exhibited art, its artist, or style. Third, literature coverage may be important; this includes books, catalogs, and scholarly articles covering the art piece. The most important reference work is the *catalogue raisonné*, which includes all the known artworks by an artist. The fourth category, certification, provides physical or non-physical proof of authenticity issued by the artist, the artist's close family, and art experts. While all four provenance dimensions and their constituting elements undoubtedly enhance trust in art markets, some provenance variables may also embed the quality of the art object (the highest quality works may be exhibited more often or appear more in scientific art history books). In addition, prices of paintings may be inextricably related to the status or glamor of previous owners if a buyer is willing to pay a premium for a painting that was owned by, for example, royalty or celebrities (but the price premium may also reflect that such people may be wealthy enough to have bought in expert advice for their acquisitions and paid for technical analyses to verify authenticity).

We demonstrate that the probability of being sold, price levels, and returns are all significantly affected by the various dimensions of provenance. The probability of being sold increases by about 2% when the catalog documents information on pedigree, by 4% for information on the exhibition history,

and by 3% when the painting is mentioned in the literature. The effect on price is even more pronounced: pedigree, exhibition background, references in the art literature, and the presence of certification drive prices up by 21%, 42%, 54%, and 14%, respectively. The annualized repeat sales returns increase by respectively 11, 16, and 5 percentage points when the catalog provides information on exhibitions, literature, and certification, respectively. In the models yielding these results, we control for artwork characteristics (e.g., topic, measurements, or medium), as well as artist, year, seasonality, and auction house branch fixed effects.

As the estimated relationship between provenance and probability of being sold, hammer prices, and returns may be affected by endogeneity (i.e., the decision to offer a painting for sale and the decision to provide detailed provenance information may depend on expected prices), we conduct four additional analyses to alleviate such concerns. First, we examine the provenance effects on prices of paintings by artists with some (attributed) works that have been discovered to be fakes and forgeries in a difference-in-differences (DiD) setting. We study whether the provision of provenance stems from a possible contagion effect from a forgery to all of an artist's artworks (in terms of buy-in probability, prices, and returns) (Subsection 3.3.1). We find that the impact of provenance increases substantially in a context of increased uncertainty about authenticity. Second, we exploit Christie's provenance policy change in 2012 (following a litigation case) relative to Sotheby's, also in a DiD setting (Subsection 3.3.2). Christie's increase in the quality of its provenance led to a price premium of 37% relative to Sotheby's transaction prices. Third, to address the effect of past prices on provenance provision, we examine the impact of incremental provenance information on prices, while controlling for initial provenance at the first sale and for the previous hammer price, and conclude that incremental provenance is also priced. We report that provenance remains strongly related to prices (Subsection 3.3.3). Fourth, as a seller's decision to offer a painting to an auction house as well as the auction house's decision to provide provenance information may be affected by recent price evolutions (for similar paintings, e.g., by the same artist or school, or for the art market as a whole), we examine the impact of provenance on the probability of being sold and the hammer prices for the subsample of sales for which we expect the above decisions to have been taken exogenously. For instance, such auction sales may result from estate sales after an owner's death. Also in this case, in which market timing would play little role, we find price premiums for provenance (Subsection 3.3.4). All endogeneity tests corroborate our baseline results.

As our models are estimated with high dimensional fixed effects models, which could result in a decrease in predictive power, we perform least absolute shrinkage and selection operator estimations (LASSO), and the findings confirm earlier results. A possible issue is that provenance provision could be related to liquidity of an artist's oeuvre and liquidity could in turn proxy for quality and require a premium. We study whether the provenance price premium coincides with a liquidity premium, but demonstrate that this is not the case. While we control for auction house branch fixed effects in all regressions, we still perform a separate analysis on auction houses and provenance in order to study

whether a substitution effect exists between auction house types and provenance. We document that for art of different price ranges by auction house type (large international, medium-sized, and small auction houses), provenance generates a price premium. Thus, there is no substitution effect nor are provenance premiums limited to art in the highest price quantiles or offered by specific types of auction houses. Moreover, an artist's reputation is not a substitute for provenance provision; on the contrary, established artists' work may be more subject to being forged.

The rest of the paper proceeds as follows. Section 2 describes the methodology and data. Section 3 documents the empirical results and endogeneity tests. Section 4 reports extensions and robustness tests. Section 5 concludes.

2. Data and Methodology

2.1 Data and Variables

While approximately half the global art market is private and comprises a primary market (galleries selling the oeuvre of living artists) as well as part of secondary markets (private sales among collectors or organized by dealers), we can study only the public segment of the secondary market, namely, the auction market organized by hundreds of auction houses around the world. This public market comprises more than half the global market and also leads in terms of price setting by providing publicly available price benchmarks for privately sold art. We focus on the market for oil paintings, watercolors, and drawings, which comprises the largest proportion of the fine art auction market. From the online database *Blouin Art Sales Index,* we collect all sales of paintings from 2007 to 2016. Our sample starts in 2007, as information on provenance and buy-ins (i.e., items that do not reach the undisclosed reserve price and remain unsold) is of poor quality in earlier years. Our dataset includes 1,812,807 transactions of which 1,195,640 objects (66%) were sold at auction by 608 auction houses (branches) all over the world. The paintings and drawings were created by about 150,000 artists. The average (median) hammer price is USD 53,142 (USD 3,400) with a standard deviation of USD 638,181—all real terms (2007 deflated). For each transaction, we collect all the artist, artwork, and transaction characteristics, such as artist name, title of the art object, medium (oil/acryl, watercolor, print), measurements (height and width), attribution, creation year, whether signed and/or dated by the artist, sold or unsold at auction, hammer price, auction lot number, low and high price estimates, auction date, auction house (branch), and detailed provenance information. We apply textual analysis to the provenance text in the auction catalogs to obtain 40 characteristics (see Subsection 2.1.2), which we categorize into four dimensions: pedigree, exhibition, literature, and certification.

2.1.1 Traditional Hedonic Variables

We follow Renneboog and Spaenjers (2013) by including the traditional hedonic pricing variables as control variables in our regressions.[3]

---

[3] The descriptive statistics for the hedonic variables are presented in Online Appendix Table A.I.

Artist characteristics. We include artist fixed effects and a dead artist dummy (Deceased). The former captures each artist's uniqueness and reputation; the latter captures that prices may increase after the death of an artist due to a supply shock (Pénasse et al. 2021). In our dataset, 76.7% of the auctioned paintings were from deceased artists.

Artwork characteristics. We consider the following wide range of price-determining variables that capture the attribution, signature, medium, measurements, and topic of the work of art.

- Attribution: We consider six levels of attribution that capture various degrees of uncertainty/closeness to a specific artist: *Attributed* (to), *Studio* (of), *Circle* (of), *School* (of), *After*, and (in the) *Style* (of) an artist. About 3.4% of the observations in our sample carry such an attribution.

- Signature: We include *Signed, Dated*, and *Inscribed* variables; 80.4% of artworks are signed, about 36.3% are dated, and 11.4% are inscribed.

- Medium: The indicator variables *Oil*, *Watercolor*, and *Drawing* represent the mediums used. About 68.1% of the transactions are oil paintings, 20.5% are watercolors, and 11.4% are drawings.

- Measurements: *Height* and *Width* are included in centimeters (in addition to the squared values *Height_2* and *Width_2*).

- Topic: As the aesthetic and financial appreciation can depend on a painting's topic, we categorize the paintings based on the keyword analysis of the titles. We search for keywords in the seven languages most used in the art auction world (and its catalogs): Dutch, English, French, German, Italian, Portuguese, and Spanish. We partition the paintings in the following categories: *Abstract*, *Animals*, *Landscape*, *Seascape*, *Cityscape*, *Nude*, *People*, *Self-Portrait*, *Portrait*, *Religion*, *Still Life*, *Study*, and *Other Topics*. *Untitled* is used as the omitted benchmark in our regressions.

Transaction characteristics. We include indicator variables capturing the timing of the sale, and the reputation and location of the auction house:

- Year and month: We control for year effects as well as seasonality as the most important auction seasons are in spring (May and June) and fall (November and December).

  Auction houses: We distinguish between different fine art auction houses based on reputation/size. In the case of Sotheby's and Christie's, we introduce dummy variables for their London, New York, and other branches (e.g., *Sotheby's London*, *Sotheby's New York*, and *Sotheby's Other Branches*). For two other important British auction houses, Bonhams and Phillips, we distinguish between their London sales rooms and other branches (e.g., *Bonhams London* and *Bonhams Other Branches*). We also create two dummies to account for the sales by important (large or middle sized) European and American auction houses (*Auction European* and *Auction American*).

2.1.2 Provenance Variables

About 14.2% of the observations in our database provide pedigree information; the proportions for exhibitions, literature, and certification amount to 3.7%, 4.0%, and 3.9%, respectively, and are presented in Table 1.

Pedigree considers the history of past owners and thus, refers to ownership chains. For example, a painting might have been in the collection of the artist's family, prominent collectors, royal and noble families, wealthy families, CEOs, and celebrities. If an uninterrupted ownership chain between the artist and the current owner can be traced, the artwork has a higher probability of being authentic, and thus, there is a possibility of a premium at auction. It is also possible that a "glamour" premium is paid for a painting once owned by a famous individual.[4] As such, "ennobling" provenance can turn an ordinary object into an extraordinary one.

Furthermore, the ownership chain may reveal whether the painting was acquired directly from the artist, from the artist's family, or from the sitter (the person depicted in the painting), or has uninterrupted information on the descent of the painting through the generations. All of the abovementioned factors indicate the painting's authenticity. Obviously, falsified provenance can never be excluded, although one would expect that auction houses would carefully verify provenance information in order to avoid lawsuits and loss of reputation from the auction of forged paintings. We also check whether paintings were sold throughout each work's history by one or more prominent auction houses or established dealers, as it is likely that they more carefully collect and verify the provenance. In this respect, we consider both historic auction houses and dealers that no longer exist following mergers or termination, as well as contemporary ones.

To apply textual analysis in the pedigree dimension, we develop a name list based on more than 150 databases.[5] The variables incorporated in the *Pedigree* are as follows.

- Past ownership: *Prominent Collectors*, *Royalty/Nobility*, *Wealthy Families*, *CEOs*, *Influential People (Time 100)*, *Celebrities*, *Famous Sportspeople*, *Corporate Collection*, and *Private Collection (Anonymous)*, all from around the world. For instance, for *Royalty* and *Nobility* we search for nobility titles in seven languages (English, Latin, Dutch, French, German, Italian, and Spanish). For the sportspeople, we collect the names of the best-paid ones as well as the world champions and superstars (in boxing, golf, basketball, tennis, soccer, football, baseball, racing, motorcycle, cricket, track, auto racing, mixed martial arts, motorsport, and hockey).

---

[4] It is difficult to separate the effects of glamor and artistic quality. On the one hand, ownership by a celebrity may induce a sales premium, but the art object itself might be of high quality in that the celebrity might have been well advised (by art experts) upon original purchase of the work of art. For example, the fact that a painting has been owned by Oprah Winfrey or Elton John at one point in the painting's history might be a selling point to potential buyers who feel sympathetic toward a "star." However, it is unclear whether, if a premium is paid, the premium reflects stardom in past ownership or the possibility that a star is able to pick quality art with high value growth potential, as they can obtain advice from the best art consultants.

[5] The string searches and sources are reported in Appendix B.

- Descent: purchased *Directly from Artist*,[6] *From Artist's Family*,[7] or *From Sitter*.[8] We also report *Other Descent Information*.[9]
- Past sale channel: *Sold by Sotheby's/Christie's*,[10] *Sold by Bonhams/Phillips*,[11] *Sold by Historic Auction Houses* (e.g., Dorotheum, Dowell's, Hôtel Drouot), *Sold by Other Important Auction Houses*,[12] and *Sold by Prominent Dealers*.[13]
- Other collections: *Other Pedigree Information* (indicates that other pedigree information is available not falling in any of the abovementioned categories).

Our first provenance dimension is *Pedigree*. For 256,560 paintings, *pedigree* information is available in auction catalogs. The average pedigree text length amounts to 104 characters. Of all the paintings with pedigree information, 2.1% were once part of a prominent collection, 2.4% were held by royal or noble collectors, 0.7% by (other) wealthy families, 0.06% by collectors who are also CEOs, 0.08% by "influential people," and 0.3% by celebrities and sportspeople.[14] The credibility of the authenticity increases when the ownership can be traced back to the artist or people close to the origin. We document that there is evidence for 10.5% of the paintings that an earlier owner purchased the painting directly from the artist; 5.3% of the paintings were acquired from the artist's family, and 0.33% from the sitter. For 9.8% of artworks, the pedigree text gives additional information about ownership of descendants (not included in the abovementioned categories). When studying previous sales records, about 15.0% of the paintings with pedigree information were sold by Sotheby's and Christie's at one point in the painting's history, 0.9% by Bonhams and Phillips, 1.3% by historically important auction houses, and 5.8% by prominent dealers.

Our second provenance dimension is labeled *Exhibition*, which embeds information about the number and importance of exhibitions (by museums, at art fairs, in museums, by galleries, and in culturally important cities). Past exhibitions may vet the painting, because an exhibited painting is then often examined by experts and curators, who reflect in an exhibition catalog on the position of the painting within the total oeuvre of an artist or within an artistic school or era. Therefore, an often-

---

[6] Examples are as follows: "acquired directly from the artist by the present owner in 2002," "courtesy of the artist," and "gift from the artist."

[7] Examples are as follows: "by descent in the family of the artist until the late 1980s," "descended within the family of the artist," and "purchased from the artist's family."

[8] Examples are as follows: "from the sitter, by descent to the present owner," "by descent in the sitter's family until 2010,"
and "by descent through the sitter's grandson, Montague Peregrine Albemarle, 12th Earl of Lindsey (1861–1938) to his daughter, the late Lady Muriel Barclay-Harvey (1893–1980)."

[9] Examples are as follows: "by family descent for three generations," and "by descent in the family."

[10] Examples are as follows: "Christie's, London, 25 January 1991, lot 20," and "Sotheby's London, Russian Pictures, Icons and Russian Works of Art, 15 February 1984, Lot 106."

[11] Examples are as follows: "Bonhams, London, 14 June 2005, lot 109," and "Phillips, London, 14 June 2000, lot 60."

[12] See examples in Appendix B.

[13] Examples are Georges Petit, Gagosian Gallery, Pierre Matisse, Sidney Janis, and Leo Castelli.

[14] Precise definitions of these ownership categories (e.g., for nobility, celebrities, and influential people) can be found in Appendix B.

exhibited painting may yield a premium at auctions. Our sample comprises 67,713 paintings with exhibition information. On average, a painting was exhibited twice. Among all observations with exhibition information, about 6.2% were exhibited at least once at prominent exhibitions, 0.4% at prominent art fairs, 17.2% in prominent museums, 29.9% at lesser-known museums, 74.1% in cultural cities, and 14.9% in galleries.[15]

Our third provenance dimension is *Literature*. We consider whether artworks are included in *catalogues raisonnés*, which offer a comprehensive listing of all known artworks by the artists, are illustrated on the cover page of art books, are included in art books published by an authoritative press (e.g., a university press[16]), or in any other publication. Our sample comprises 72,906 paintings with literature information with an average text length of 242 characters and with on average 1.5 literature-related references of the painting. Among all paintings with literature information, about 15.7% are mentioned in the *catalogue raisonné*, 1.7% are even on the cover page of art books, 45.9% are illustrated in art books, and 1.2% are mentioned in art books published by an authoritative press.

In the fourth dimension, *Certification*, we search for two aspects: (i) the person or agency who has issued the certification (this can be artists themselves, their family members, associations,[17] experts, or other parties); and (ii) the form of the certification (physical certificate[18] vs. non-physical confirmation, e.g., an oral statement by the artist about the painting's authenticity[19]). For 70,556 paintings certification information is mentioned in the provenance text. Among the paintings with certification information, 31.7% have physical certification issued by the artist, 6.2% by the artist's family, 15.1% by the artist's association, 2.7% by experts, and 27.6% by other parties. In addition, about 5.3% of the observations are presented with non-physical certification by the artist, 2.2 % by the artist's family, 4.5% by their association, 2.9% by experts, and 6.2% by other parties. The general *Certification* dummy has a correlation close to 0 with the other main provenance dimensions (*Pedigree*, *Exhibition*, and *Literature*), which in turn exhibit moderate positive correlations between 0.35 and 0.45. Delving deeper into the more detailed variables within each of the four main dimensions, we find very low correlations, which suggests that all the detailed variables together need to be considered to obtain a reliable picture of a painting's authenticity.

[Insert Table 1 about here]

---

[15] Prominent exhibitions include retrospectives, anniversary exhibitions (birth/death year of artists), biannual, and triannual exhibitions. Prominent art fairs include TEFAF, Art Basel, Art Miami, Biennale, and Frieze; while prominent museums include Getty, Louvre, and Museo del Prado. Cultural cities include New York City, Paris, and London. The detailed keywords and sources are given in the Appendix B.

[16] Examples are Oxford University Press and Cambridge University Press. The sources are in the Appendix B.

[17] Examples are the artist's foundation and registry. For detailed definitions, see Appendix B.

[18] Examples are "a certification by the Picasso Administration will be given to the buyer," "accompanied by a certificate of authenticity from the artist," "accompanied with Solomon Gallery exhibition catalogue as well as letter written by the artist in 2002 following the purchase of the work," and "Registered in the artist's archive in Paris, no, 86170 SWF."

[19] Examples are "verbally authenticated by Dr. Paul Vogt, Essen," and "verbally authenticated by the Ludwig von Hofmann Archive, Zurich."

2.2 Methodology

2.2.1 Hedonic Linear Probability Regression

To investigate the provenance effects, we estimate linear probability of being sold regressions with the dependent variable $Sold_{i,t}$ indicating whether an art object was sold (vs. bought in), while controlling for a wide range of hedonic characteristics:

$$Sold_{i,t} = \alpha_{i,t} + \sum_{p=1}^{P} \beta_p Pedigree_{p,i,t} + \sum_{l=1}^{L} \beta_l Literature_{l,i,t} + \sum_{e=1}^{E} \beta_e Exhibition_{e,i,t}$$
$$+ \sum_{c=1}^{C} \beta_c Certification_{c,i,t} + \sum_{m=1}^{M} \beta_m X_{m,i,t} + \sum_{t=1}^{T} \gamma_t D_{i,t} + \varepsilon_{i,t} \tag{1}$$

where $Sold_{i,t}$ equals 1 if art object i at time t is sold and 0 when bought in. The four provenance dimensions are represented by $Pedigree_{p,i,t}$ (with characteristic p of item i at time t), $Literature_{l,i,t}$ (with literature characteristic l), $Exhibition_{e,i,t}$ (with exhibition characteristic e), and $Certification_{c,i,t}$ (with certification characteristic c). The provenance dimensions are presented as dummy variables capturing whether the artwork's catalog comprises any information on these dimensions, or the natural log of the number of characters used in the provenance text for each of these dimensions. $X_{m,i,t}$ is the value of characteristic m of item i at time t. $D_{i,t}$ is a time indicator variable for the year in which the art object is offered for sale in an auction. The coefficients, β_p, β_l, β_e, and β_c capture the relationships of the provenance dimensions for the probability of being sold. The coefficients β_m reflect the shadow price of each of the m characteristics, and the coefficients γ_t reflect the time trend, which can be used to construct an art price index.

2.2.2 Hedonic Pricing Regression

To measure the impact of provenance information on hammer prices, we resort to a hedonic pricing model, which has the advantage (relative to a repeat sales approach) of including all observed transactions.[20] We regress the natural logarithms of hammer prices (in 2007 real USD) on provenance, while controlling for a broad range of hedonic characteristics:

$$Ln(P_{i,t}) = \alpha_{i,t} + \sum_{p=1}^{P} \beta_p Pedigree_{p,i,t} + \sum_{l=1}^{L} \beta_l Literature_{l,i,t} + \sum_{e=1}^{E} \beta_e Exhibition_{e,i,t}$$
$$+ \sum_{c=1}^{C} \beta_c Certification_{c,i,t} + \sum_{m=1}^{M} \beta_m X_{m,i,t} + \sum_{t=1}^{T} \gamma_t D_{i,t} + \varepsilon_{i,t} \tag{2}$$

where $P_{i,t}$ represents the hammer price of art object i at time t, and the rest of the control variables are defined in Subsection 2.1.

2.2.3 Repeat Sales Regression

As we also intend to study art returns, we turn to a repeat sales analysis. Here, we analyze whether incremental changes of provenance occurring between two sales impact returns. The dependent variable $Return_{i,(1,2)}$ in Equation (3) is the geometrically annualized return of the painting in the repeat sales transaction. We regress this return on (i) the changes in the provenance variables between the first and second transactions, $\Delta Provenance_{a,i,(1-\delta,2-\delta)}$ (with provenance dimension a of item i between two

---

[20] Even over a long time period spanning decades, the number of repeat sales is limited to about merely 3–5% of the total sales, as the average holding periods are extensive (e.g., by collectors) and some art objects never return to the market (e.g., art bought by museums or private collectors) (Renneboog and Spaenjers 2013).

transactions; *a* represents the various characteristics of each of the four provenance dimensions, *Pedigree*, *Exhibition*, *Literature,* and *Certification*) while also controlling for the initial provenance predating the first sale *Provenance*$_{a,i,1-\delta}$ (with provenance characteristic *a* of item *i* at time *1-δ*); (ii) possible changes in sales channel (changes in auction house (branch)) (i.e., *Auction House Upgrade* is a dummy variable equal to 1 if the second sale auction house has a better reputation than the first one[21]); and (iii) the other hedonic variables and fixed effects as defined in Subsection 2.1. In Equation (3), we include the same hedonic control variables $X_{m,i,2}$ and time control variables $D_{i,2}$ as in Equations (1) and (2). The subscripts *1*, *2*, (*1,2*), and *δ* refer to the time of the first sale, the second sale, the holding period, and the time lag, respectively. The coefficients of interest are β_a, which reflect the incremental provenance effects on the returns:

$$Return_{i,(1,2)} = \alpha_{i,(1,2)} + \sum_{a=1}^{A} \beta_a \Delta Provenance_{a,i,(1-\delta,2-\delta)} + \sum_{a=1}^{A} \beta_b Provenance_{a,i,1-\delta}$$
$$+ \beta_u Auction\ House\ Upgrade_{i,(1,2)} + \sum_{m=1}^{M} \beta_m X_{m,i,2} + \sum_{t=1}^{T} \gamma_t D_{i,2} + \varepsilon_{i,(1,2)} \quad (3)$$

The provenance information (auction catalog) is available 4–6 weeks prior to the auction, which is highlighted by means of *δ* in Equation (3).[22]

3. Empirical Results

This section discusses the empirical results of the correlation between provenance information, on the one hand, and the probability of being sold, the hammer price (for our full sample), and the returns (for the repeat sample), on the other hand. We also provide robustness tests on subsamples and perform analyses by means of the least absolute shrinkage and selection operator (LASSO) estimations to alleviate estimation concerns induced by high dimensional fixed effects.

3.1 Provenance Effects and Probability of Being Sold

Is the probability that a painting offered for sale is actually sold (vs. bought in) correlated with the provision of provenance? As the auction house provides a price estimate in the catalog, which can affect the sales outcome, we include the lower estimate (which is assumed to be at or close to the reserve price) in the linear probability Models (1) and (2) of Panel A of Table 2. A low-price estimate that is set high increases the probability that the painting will not be sold (which occurs when the highest bid does not exceed the reserve price). We also control for a large set of hedonic variables detailed in Subsection 2.1.1, as well as for the following fixed effects: Artist, Year, Month, and Auction House Branch Level.[23] Our sample comprises 1,707,136 observations with full hedonic information available as well as low price estimates.

---

[21] See Appendix A for details.

[22] In equations (1) and (2), the δ lag also applies in that all provenance information is available prior to the auction. By inserting δ in Equation (3), we emphasize that the change information affects changes in prices, and hence, returns.

[23] The parameter estimates of the full regression are provided in Online Appendix Table A.II.

[Insert Table 2 about here]

The aggregated provenance dimensions *Pedigree, Exhibition, Literature,* and *Certification* are represented by either (i) an indicator variable capturing the provision of this type of information or (ii) the natural logarithm of the catalog's text length of each of these dimensions. The provision of pedigree information goes hand in hand with a higher probability of being sold, which increases by 1.7% (Column (1) in Panel A of Table 2). Likewise, exhibition history and the appearance of a painting in the art (history) literature increase the probability of being sold by 3.8% and 2.5%, respectively. By contrast, the presence of a certificate does not seem to affect the probability of being sold (but affects prices; see Subsection 3.2). As expected, the negative coefficient of the low-price estimate indicates that artworks with higher reserve prices are less likely to be sold. The specification that uses the provenance dimensions' text length (Column (2) in Panel A of Table 2) shows consistent and similar patterns with exception of a positive correlation for certification. It is important to note that these regressions should not be interpreted as causal relations. In fact, endogeneity (in the form of reverse causality) could very well be substantial since auction houses offer provenance information, especially for paintings with ex ante high probability of being sold or with the highest price potential (which could partially but not entirely be addressed by the inclusion of the price estimates). We undertake a set of endogeneity tests in Subsection 3.3, but first undertake a correlation analysis based on a more granular approach to the provenance information with Models (1) and (2) in Panel B of Table 2.

In the pedigree dimension, past ownership by prominent collectors, royalty/nobility, wealthy families, celebrities, and famous sportspeople are all correlated to increase the probability of being sold by 5.2%, 6.6%, 8.6%, 3.7%, and 9.5%, respectively (Column (1) in Panel B of Table 2). If the painting was in the past part of a corporate collection, the current probability of being sold is 9.8% higher.

With regard to information on descent, when a painting was purchased directly from the artist, the probability of being sold increases by 1.5%.[24] We also examine whether the reputation of past sales channels has an impact on current sales: if the painting was sold by a prominent dealer, the current probability of being sold is 3.8% higher, but when past sales were at one point executed by Sotheby's or Christie's, there is no positive effect on current probability of being sold (possibly because those paintings may at the time already have been sold at higher price levels). A past auction at Bonhams and Phillips (or other historically important auction houses) does not have any positive impact on prices.

When a painting was part of prominent exhibitions, it will be sold more easily (the probability increases by 2.6%; Column (1) of Panel B in Table 2). If a painting was exhibited in museums in the past, it is now sold more easily; exhibitions in prominent museums increase this probability by 5.4% and by 1.8% in the case less prestigious museums. If the artwork was exhibited in cultural cities where

---

[24] If the painting was at one point directly purchased from the sitter or the sitter's family, the probability of being sold is 8.9% lower, which may be affected by the scarcity of such paintings in auctions. Furthermore, such paintings (usually commissioned by the sitter) are of personal value for the sitter or their family and may be less appealing to people not related to the sitter's family.

possibly a more cultured audience of art lovers can be reached, we note an augmentation in the probability of being sold (by 1.9%). Past exhibitions by galleries facilitate a sale (by 4.3%); but displaying the painting at prominent art fairs does not seem to matter.

When a painting is mentioned in the art (history) literature, selling is expectedly easier: a mention in a *catalogue raisonné*, the depiction on the cover page of art books, and inclusion of an illustration in an art book augments the probability of being sold by 2.4%, 5.1%, and 2.1%, respectively. Physical certification by artists or their associations or foundations (e.g., the Andy Warhol Foundation and the Keith Haring Foundation) is important and affects sales: the physical certificate issued by the artist increases the probability (by 6.0%) as does certification by an artist's association (by 3.5%). Interestingly, if the painting is accompanied by non-physical certification (by the artist's family, experts, etc.) the painting does not sell more easily, which suggests that non-physical certification does not remove possible doubts regarding the authenticity.

3.2 Provenance Effects and Hammer Prices

We turn to the relationship between provenance and hammer prices in Panels A and B of Table 2, in the same vein as for the probability of being sold. We first estimate Equation (2) for 1,111,220 auction transactions for which we have complete information on all the hedonic characteristics presented in Subsection 2.1.[25] The provenance variables in Columns (3) and (5) of Panel A of Table 2 are dummy and textual length variables (as defined in Subsection 3.1), respectively. After controlling for all the traditional hedonic variables and fixed effects, we find that the presence of provenance information goes hand in hand with higher price levels (Column (3)).[26] If information is made available on the painting's pedigree, the price is 20.7% higher (*exp(0.1885)-1*); with an exhibition history, the price augments by 41.9%; literature drives the price up by 53.5%, and with certification the price increases by 13.9%. Thus, art (history) literature sources and an exhibition history affect prices most. Models including the provenance text length yield qualitatively similar results (Column (5) in Panel A). As mentioned in Subsection 3.1, the caveat on endogeneity applies, which we address in Subsection 3.3.

We dissect the four main dimensions of provenance into their detailed constituting elements and show the results in Panel B of Table 2. For the pedigree components, we observe that past ownership

---

[25] The sample size used to estimate the models of Columns (3) and (5) of Table 2 is smaller than that of Columns (1) and (2), because the sample for the latter models includes all paintings offered for sale (including those who were bought in).

[26] The models presented in Panels A and B of Table 2 show only the parameter estimates of the variables capturing provenance details. The coefficients of the hedonic variables and fixed effects (included in Equation (2)) are given in Online Appendix Table A.III. Those results are in line with past research (e.g., Renneboog and Spaenjers 2013). Artworks with the attributions "style," "after," "school," "circle," "studio," and "attributed" are priced at large discounts (relative to the price of an authenticated artist), while signed, dated, or inscribed works tend to have higher prices. Oil paintings and watercolors command higher prices than drawings. Furthermore, prices increase with size (measurements), up to the point that the work becomes too large, which is captured by the negative coefficients on the squared terms. In addition, portraits and studies are sold at a discount. Unsurprisingly, Sotheby's London and Christie's London sell artworks with the highest prices, ceteris paribus.

by prominent collectors, royalty/nobility, wealthy families, and famous sportspeople has a strong impact on current price levels. Paintings reach prices that are on average 24.4% (31.3%) higher if they had been held by prominent collectors (royalty/nobility). Past ownership by wealthy families or famous sportspeople correlates with current price premiums of 42.5% and 50.1%, respectively. When the auction catalog reports that the painting was at one point purchased directly from the artist, the current hammer price is 13.8% higher; a past purchase directly from a sitter now augments the price by 11.3%; and any descent information (not included in the above categories) increases prices by 23.6%. Thus, information related to the past ownership chain (purchases from the artist or his family), prominent collectors, or famous owners) enhances the marketability of paintings and their hammer prices.

In addition, past sales channels also affect auction prices: the fact that a past auction sale was made by Sotheby's or Christie's, Bonhams or Phillips, or by historically important auction houses (e.g., Dorotheum, or Hôtel Drouot), may be perceived as (past) recognition of quality and authenticity which is mirrored in prices that are higher by 23.6%, 7.1%, and 10.0%, respectively. The current price premium may thus reflect the (past) notoriety of these auction houses. Similarly, past sales by prominent dealers, are correlated with auction prices, which are on average 32.9% higher.

Details about the second pillar–past exhibitions are also strongly related to price levels, as they may reflect the intrinsic quality of paintings. For instance, if the painting had been included in prominent exhibitions, the auction price is 26.2% higher. Likewise, exhibitions organized by prominent museums or in important cultural cities enhance a painting's exposure to a larger art-loving audience and is related to higher prices (by 58.9% and 24.2%, respectively).

Likewise, it is important that a painting is named or depicted in the art literature: all aspects within the Literature dimension correlate with prices: the inclusion of the painting in the *catalogue raisonné*, being depicted on the cover page of or in art books, discussed in books published by an authoritative (university) press are related with higher prices (with significant price premiums of 35.7%, 52.8%, 44.2%, and 41.2%, respectively).

Certification also enhances trust: Panel B of Table 2 shows that all physical and non-physical (testimonial) certification is positively correlated with prices (Column (2)). Certificates by experts have the biggest price impact (40.7%). In the case of a non-physical certification by the artist (e.g., when the provenance text refers to a testimonial of an oral statement by the artist), the price impact amounts to 32.5%.

3.3 Endogeneity

A first endogeneity concern may emerge when the provision of provenance information is driven by past prices of a painting (or of similar paintings). For a painting expected to attract a high price, more resources can be made available to research and document its provenance. To rule out (or at least attenuate the possibility of) the endogeneity concerns, we perform the following four analyses. (i) In a DiD setting, we examine the provenance effects on the artworks made by artists whose works were

faked or forged (Subsection 3.3.1). (ii) We exploit Christie's change in provenance policy in 2012 in a DiD setting and use the provenance policy by Sotheby's (which is regarded as having similar reputation and quality) as a control (Subsection 3.3.2). (iii) To address potential endogeneity induced by reverse causality between Past Prices (capturing higher quality) and Provenance Changes, we run a two-staged model on a repeat sales sample (Subsection 3.3.3). A second endogeneity concern relates to the possibility that sale decisions may be endogenous in that sellers may offer a painting for sale to an auction house after having observed recent prices for similar paintings (of the same artist, same school, or in general in the entire market). Therefore, we resort to (iv) a subsample analysis for auction sessions that comprise sales decisions that we expect to be made exogenously (Subsection 3.3.4). These are auction sales from estates of deceased previous owners. The heirs might be forced to sell to pay inheritance tax,[27] or might not be able to afford or be disinterested in the upkeep of an estate including art inheritance. Thus, the decision to offer art for auction in the period after the decease of the collector might not be (or be less) endogenous.

3.3.1 Discoveries of Fakes and Forged Paintings

Art markets offer a congenial environment for forgeries, which has a detrimental effect on the trustworthiness of the market. Large numbers of fake and forged paintings are rumored to circulate. According to a former director of The Museum of Modern Art, up to 40% of the high-end art market may consist of forged art (Thompson 2010).[28]

This may undermine the art world's confidence in the authenticity of paintings, and hence distort price formation and depress value. In this respect, provenance is a prominent factor underlying the proper functioning of the art market—although provenance may also be subject to forgery. The discovery of a fake painting may negatively affect the prices of the paintings by the forged artist, and could shift transactions to the most trustworthy intermediaries who could offer an insurance (a guarantee to take forged paintings back). To investigate the impact of the discoveries of fakes/forgeries on the price of artworks through the effect of the provenance dimensions, we collect the disclosure dates of fakes and forgery cases from three primary sources: the specialized art journals Artsjournal.com and Artnews.com, and the general news database Factiva, which includes the worldwide print media in 28 languages. For each article that we retrieve on this topic, we collect the event date of the discovery, title of the fake or forged painting, name of the artist, name of the forger (if available), title of the original painting, and the auction house(s) involved in the case. We take the first date that the rumor, motivated

---

[27] In some countries (e.g., the UK), it is possible that the fiscal authorities accept the donation of art to public museums as payment of inheritance tax.

[28] Bocart and Oosterlinck (2011) show that fraud discoveries shift the market toward intermediaries with higher reputation (e.g., Christie's and Sotheby's). Given that high-end auction houses provide insurance in that they offer to repurchase an art object if it has been shown after the auction to be a counterfeit, they have a strong interest to collect detailed and reliable provenance information. Thus, the provision of provenance not only protects an auction house against claims and court cases, but also helps the intermediary to build a reputation of trustworthiness.

suspicion, proof, or ruling by a judge was mentioned in the press as the date of discovery of fraud. We retain only the cases that were ultimately confirmed as fakes or forgeries. In this manner, we identified 54 cases of fraud related to paintings in our sample period.[29]

Equation (4) presents the DiD regression in which *Treated* equals one for the paintings of artists of whom one or more paintings were discovered as being faked or forged and whose paintings were offered for sale in an auction at a date after the date of the discovery. *Provenance* is a dummy variable capturing if the painting contains information on provenance (and on its dimensions). The control variables are the hedonic variables and fixed effects consistent with Equation (1):

$$Ln(Price_{i,t}) = \alpha_{i,t} + \beta_r Treated_{i,t} + \beta_k Treated_{i,t} \times \sum_{a=1}^{A} Provenance_{a,i,t}$$

$$+ \sum_{a=1}^{A} \beta_a Provenance_{a,i,t} + \sum_{m=1}^{M} \beta_m X_{m,i,t} + \sum_{t=1}^{T} \gamma_t D_{i,t} + \varepsilon_{i,t} \qquad (4)$$

The variable of interest is the interaction term *Treated* and *Provenance*. We expect a negative coefficient *Treated*, because after a fake is discovered for a particular artist, doubts about the authenticity of all of this artist's paintings may arise, which may translate into lower prices for their paintings without provenance information. The provision of provenance information may undo or reduce the negative effect of the discovery because in case of mistrust, additional information as an authenticity signal is vital to restore confidence, which is why we expect a positive coefficient of the interaction term *Treated* × *Provenance*.

The results in Column (1) of Table 3 reveal a sharp decline in the prices of paintings without provenance information by artists whose paintings have been forged (the parameter estimate of Treated is -0.2670, which represents a 23.4% price decline (*exp(-0.2670)-1*). The provision of provenance has a significantly positive impact on the hammer price subsequent to the discoveries of fakes/forgeries (in Column (1), *Treated* × *Provenance* = 0.7054), which results in a net positive price impact of 55.0% (*exp(-0.2670+0.7054)-1*) for the paintings with provenance of the affected artists. The provision of provenance thus more than compensates the negative impact of the discovery of a forged painting. The price increase may also reflect a "supply shock," as the number of authentic paintings (with strong and reliable provenance) may be scarcer.

[Insert Table 3 about here]

Column (2) of Table 3 disentangles provenance into its four dimensions and depicts a similar and consistent result: there is a negative price impact on auctioned paintings of artists whose work is forged as soon as the media reports rumors or proof of forgeries and fakes. From that moment on, provenance information on pedigree, exhibitions, and literature becomes more important and undoes and even more than compensates the negative price effect of forgeries.

Since most of the artists affected by fraud in our databases were active in the 20th century (the vast majority were born between 1880 and 1940), we restrict our sample to these affected artists as well as

---

[29] The sample of fake and forgery cases can be found in Online Appendix Table A.IV.

the unaffected ones born in the same time span and therefore productive in the same era. We show that Columns (3) and (4) of Table 3 yield very similar results.

In addition, we examine the impact of provenance information on the probability of being sold of (non-forged) paintings after the discovery of fakes and forgeries of paintings of a specific artist. We expect a positive coefficient for the interaction term *Treated × Provenance*, as a painting by an affected artist that is offered for sale with provision of provenance information may sell more easily. We confirm that provenance has a significant positive impact on the probability of being sold (Appendix C). Higher bids arise in times of mistrust for items containing provenance information.

3.3.2 Provenance Policy Change

As disclosed in the Opinion & Order of "Waren v. Christie's Inc. 2018" court case, Christie's confirmed during the Interrogatory Responses that it had amended its provenance policy in 2012 and required *all* consignors to provide detailed provenance information.[30] This change to Christie's disclosure policy on provenance was applied to all art types offered for auction (or private sale) at all of its branches. We treat this policy change as a quasi-natural experiment using a DiD setting by contrasting the "shock's" impact on Christie's sales (treated sample) with Sotheby's sales (which serves as control sample). The latter is often considered Christie's "twin" auction house, as both auction houses are similar in terms of auction history, reputation, quality of auction lots offered, branch locations, international clientele, the networks of sellers and buyers, and valuation expertise. Over the period 2007 to 2016, 198,076 paintings were sold through Christie's and Sotheby's, which represents 17.8% of our sample. In Table 4, the generic *Provenance* variable and its four main dimensions equal 1 if any information on provenance, pedigree, exhibition, literature, and certification is provided. The variables of interests are the interaction terms of *Treated* with the abovementioned provenance terms, which we expect to be positive not only because of the mere provision of provenance information (which we test) but also because of the provision of higher quality provenance (which we cannot test). We find that sales with provenance information by Christie's since 2013 earn a DiD premium (Models (1) to (5) in Table 4). This suggests that the provision of provenance creates trust and is reflected in the hammer prices. In particular, Model (6) shows that information on pedigree, exhibitions, and certification affect hammer prices, and that the certification, which comprises the physical and logged oral authenticity confirmation, has the largest price impact. An average sale with mandatory certification reporting sold through Christie's since 2013 enjoys a premium of 37.1% (*exp(0.3157)-1*), compared to the average sale through Sotheby's.

---

[30] See Opinion & Order of Waren v. Christie's Inc. 16cv1386 signed by Judge William H. Pauley, III on 5/31/2018 at Southern District of New York, U.S. Christie's amended its provenance policy in 2012 such that "…consignors were required to submit verifiable documentation to substantiate the Provenance and Country of Origin (PCOO) Form. Verifiable documentation could include receipts, invoices, inheritance documents, insurance listings, photographs, letters, or any other source that the clients can provide above and beyond their own testimony." Source: https://law.justia.com/cases/federal/district-courts/new-york/nysdce/1:2016cv01386 /453955/112/

We follow Roberts and Whited (2013) by running a series of placebo tests to confirm the validity of the above quasi-natural experiment on a subsample of sales sold via Christie's and Sotheby's during the pre-policy-change period (namely from the beginning of 2007 to the end of 2011) and set the placebo-policy change in the beginning of the years 2008, 2009, 2010, and 2011. We then repeat the regressions of Table 4 and find that *Treated* interacted with *Pedigree*, *Literature*, *Exhibition*, and *Certification* are not statistically significant (not tabulated).

[Insert Table 4 about here]

3.3.3 Addressing Reverse Causality between Past Prices and Provenance Changes

To address potential endogeneity induced by reverse causality between Past Prices (capturing higher quality) and Provenance Changes, we run a two-stage model on the repeat sales sample.[31] Repeat sales are identified by matching the exact artist's name, a painting's measurements (length and width), title, medium, and presence of a signature and date. After eliminating paintings from the same artists and of about the same size, but with indiscriminate titles ("landscape," "view of the sea," "portrait of a lady"), we obtain 6,647 repeat sales pairs, which we are certain are real repeat sale transactions.

First, we regress the changes in provenance, $\Delta Provenance_{a,i,(1-\delta,2-\delta)}$, (*a* representing four provenance dimensions: *Pedigree*, *Exhibition*, *Literature,* and *Certification*) on the past price of the painting ($Ln(Price_{i,1})$), the log of the first sale's hammer price in deflated USD. The control variables are *Time Span* (the number of days between the two sales), the initial provenance predating the first sale $Provenance_{a,i,1-\delta}$ (with provenance characteristic *a* of item *i* at time *1-δ*), and all the hedonic control variables included in Equation (1) (e.g., Artist, Year, Month, and Auction House Branch Level (AH) fixed effects). The subscripts *1*, *2*, *(1,2),* and *δ* refer to the time of the first sale, the second sale, the holding period, and a short time lag of 4–6 weeks, respectively, which are consistent with Equation (3). Hence, we estimate

$$\Delta Provenance_{a,i,(1-\delta,2-\delta)} = \alpha_{i,(1,2)} + \beta_x Ln(Price_{i,1}) + \beta_s Time\ Span_{i,(1,2)}$$
$$+ \sum_{a=1}^{A} \beta_a Provenance_{a,i,1-\delta} + \sum_{m=1}^{M} \beta_m X_{m,i,t} + \sum_{t=1}^{T} \gamma_t D_{i,t} + \epsilon_{i,(1,2)}$$
$$(5)$$

In a second model, we regress the dependent variable *Ln(Price₂)* of the second sale on the residuals from provenance dimensions *ε(ΔProvenance),* which represents *ε(ΔPedigree), ε(ΔExhibition), ε(ΔLiterature),* and *ε(ΔCertification)* of Equation (5), with the same control variables as in Equation (3):

$$Ln(Price_{i,2}) = \alpha_{i,2} + \sum_{a=1}^{A} \beta_a \epsilon(\Delta Provenance_{a,i,(1-\delta,2-\delta)}) + \beta_x Ln(Price_{i,1})$$
$$+ \beta_s Time\ Span_{i,(1,2)} + \sum_{a=1}^{A} \beta_a Provenance_{a,i,1-\delta} + \sum_{m=1}^{M} \beta_m X_{m,i,2} + \sum_{t=1}^{T} \gamma_t D_{i,2} + \varepsilon_{i,2} \quad (6)$$

---

[31] We exclude between-sales periods of less than 180 days to avoid speculative transactions. As pointed out by Pénasse et al. (2021), a short holding period usually indicates that a "flipper" (speculator) is able to purchase at a low price and quickly offers the object at a higher price. For this type of speculator, the selling decision may be more endogenous and is driven by recent prices for similar paintings.

We report the results of the estimations of Equations (5) and (6) in Panels A and B of Table 5, respectively. For the changes in the four dimensions of provenance, we use two specifications: (i) dummy variables capturing the change, and (ii) changes in the textual length of this provenance dimension. In the first specification, the indicator variable equals 1 when new provenance information is provided over the holding period, and otherwise is 0. For the changes in provenance length, we take the natural logarithms of the difference between corresponding text character length of the two sales. Panel A of Table 5 shows that the past price $Ln(Price_1)$ has no statistically significant impact on the decision to offer more *Pedigree* information (Models (1) and (5)), but a higher past price induces the provision of more information on *Exhibition*, *Literature*, and *Certification* (Models (2) to (4)) and increases the amount of information (text length) offered on *Literature* and *Certification* in between the two sales (Models (7) and (8)). However, the economic effects are minimal. When for instance, we take Δ*Literature* in Models (3) and (7), we observe that when a painting's past price has doubled, we observe an increase of merely 2% in the probability that Literature information is provided in the period between the first auction transaction and the publication of the catalog of the second auction and by 11.6% in the *Literature* text length, ceteris paribus. As the average character length of Literature amounts to about 242 characters, a doubling of the past price would only lead to a four-word increase in the literature information ($242 \times 11.59\% = 27$ characters). These results attenuate reverse causality concerns, as the possibility that past prices drive the effort to do new provenance research seems rather limited. It should be noted that we control for many fixed effects (including auction house branch) but also for the time span between the two transactions, because a longer holding period gives a greater opportunity to generate more provenance information as well as to collect additional information).

To further alleviate concerns about reverse causality, we include in the pricing model of the second transaction the residuals from Equation (5), which capture provenance information that is not predicted by past prices, as independent variables in Equation (6) (Panel B, Table 5). The unpredicted information of *Exhibition, Literature,* and *Certification* is significantly positively related to the price of the second sale in all four models. The specification with changes in text length, show positive correlations for *Pedigree*, *Exhibition,* and *Literature*. In summary, the results from Panel B imply that art prices are affected by the provenance information but not by the provenance information that is provided or augmented following high past prices in previous transactions (as the provenance information unpredicted by past prices can predict the future prices).

[Insert Table 5 about here]

3.3.4 Exogenous Sales Decisions

We explore a subsample of what we expect to be "exogenous" sales. These sales might not be affected (or at least be less affected) by past price trends as we select sales related to the four-Ds (divorce, debt, death, and disaster). To do so, we search in the auction session titles: "estate," "property of," "legacy," "bequest," "heritage," "gift," "endowment," "charity," in singular and plural forms, and

"late" plus a person's name. An example of such an auction is titled: "The property of the late M.H.D. McAlpine: Paintings, Ceramics, Silver, Works of Art and Furniture." In this way, we obtain 37,851 paintings of which 25,904 (68.4 %) were sold. Admittedly, this subsample choice cannot completely exclude endogeneity because, while death may be exogeneous, the sales decision by the heir may still be stalled. Still, we expect that endogeneity concerns are somewhat lower for this subsample.

The results in Table 6 demonstrate that the provenance dimensions *Pedigree* and *Exhibition* are significantly positively correlated with the probability of being sold, and respectively indicate a 3.7% and 7.9% higher probability of being sold than for paintings lacking this kind of documentation. Including the reserve price proxy (lowest estimate) shows that a high reserve makes a sale more difficult. Turning to the price regressions in Table 6, we observe that the provenance dimensions *Pedigree*, *Exhibition*, and *Literature* are significantly positively associated with a higher price level of paintings with a price impact of 18.1% (*exp(0.1659)-1*), 51.0%, and 66.3%, respectively. By exploiting (what we expect to be more) exogenous sales decisions that may be less affected by unobservable price-trend related motivations of the sellers, we show that the impact of provenance factors on the probability of being sold and price level is upheld. These findings are consistent with our full sample results (Table 2), and show even larger economic magnitudes.

[Insert Table 6 about here]

3.4 Provenance Effects and Returns for Repeat Sales Transactions

To investigate the relation between provenance and returns, we turn to the return model of Subsection 2.2.3 (Equation (3)) and the repeat sales sample used in Subsection 3.3.3. The dependent variable is the annualized geometric returns of the painting in the repeat sales transaction, which we regress on the changes in the provenance while also controlling for the initial provenance predating the first sale. Columns (1) to (4) of Table 7 are based on the full repeat sales sample, while Columns (5) to (8) are based on the subsample where the repeated sales take place at the same auction house branch to alleviate concerns about policy discrepancies among auction houses (or branches within the same auction house) in the provision of provenance information. For changes in provenance, we have two specifications: (i) indicator variables capturing whether a change occurred from no information to an insertion of provenance information (the reverse case does not occur), and (ii) changes in the provenance text's character-length.[32]

Table 7 shows that changes in information related to the *Exhibition* and *Literature* dimensions have a material impact on annual *Returns*. If a painting has no exhibition history prior to the first sale but is exhibited during the holding period, the annualized returns rise by 16.5 percentage points (Column (3)). Likewise, when there was no information on the painting in the (art history) literature was not present prior to the first sale but was subsequently present, prior to the second sale, the annualized return

---

[32] We present only the analysis on the four provenance dimensions in Table 7, but also provide an analysis with detailed provenance elements in Online Appendix Table A.V.

rises by 14.9% (Column (3)). An extension of the catalog text on the exhibition history of the painting also positively affects returns (Columns (2) and (4)). Columns (3) and (4) show that the initial provenance dimensions (e.g., *Pedigree₁*), disclosed prior to the first transaction of the repeat sale, have no statistically significant impact on the subsequent returns. This implies that this information was already priced at the first auction and that only additional information affects the second sale and hence returns. We also control for a price effect that runs through the "upgrade" of the auction house, which captures that the second transaction is made at a more prestigious auction house than the previous one. Higher hammer prices are reached, reflecting that more prestigious auction houses may reach a wealthier clientele and may provide a "quality stamp" for an auctioned painting. We also want to eliminate the effect of idiosyncratic provenance provision policies of auction houses by limiting the repeat sales sample to the transactions that took place in the same auction house (branch). We observe that the provision of more information highlighting the role of the painting in the literature and its exhibition history both have a stronger economic impact on returns (Columns (7) and (8) of Table 7).

[Insert Table 7 about here]

4. Extensions and Robustness Tests

4.1 Subsample of Artworks with Provenance Information

A potential concern is that for the majority of the auctioned paintings no provenance information is provided in catalogs and only traditional hedonic variables are provided (artist's name, title, measurements, medium, lot number, auction house (branch), and transaction date). We therefore restrict our sample to only those observations with available provenance information and replicate the baseline results for the full sample as presented in Table 2. The findings are consistent with our baseline results; *Pedigree*, *Exhibition*, *Literature,* and *Certification* are qualitatively and statistically similar in the probability of being sold and price regressions (Online Appendix Table A.VI).

4.2 LASSO Estimations

A methodological concern may be that we estimated high dimensional fixed effects models, which could result in a decrease in predictive power for each added variable. Therefore, we verify our results by means of LASSO estimation, which enables us to examine which of the many variables (we have more than one hundred thousand fixed effects—about 10% of the total number of observations as we include for instance artists fixed effects) are the most important ones for purposes of prediction (Belloni Chernozhukov, and Hansen 2014). Effectively, this method chooses a simpler model with fewer variables, reduces over-fitting, increases out-of-sample prediction, and generates a more efficient algorithm. The LASSO results for the hedonic probability of being sold and price regressions indicate that the statistical significance of the provenance variables remains valid (Online Appendix Table A.VII). In fact, the economic significance of the provenance variables becomes even larger relative to the baseline results of Table 2. It should be noted, however, that the LASSO method is mainly designed

for prediction such that drawing inferences from model parameters could still be problematic in the sense that when some of the parameter estimates are set to 0, an omitted variable bias could arise (which is the cost paid for a reduction in variance). Therefore, we apply this method only as a robustness test to verify whether the provenance factors are of prime importance in relation to probability of being sold and auction pricing.

4.3 Non-linearities

We also study non-linear relations between prices and provenance. We use the text length of the provenance and the number of items mentioned in the provenance (e.g., the number of exhibitions in which a painting was shown; the number of art books in which the painting is depicted). This analysis is less suitable for the pedigree dimension (uninterrupted ownership chain, type of owners) and certification (based on presence or not). We find that the prices increase with the amount of provenance information provided in the auction catalog, but only regressively so (given the negative second derivative) (see Model (2) of Online Appendix Table A.VIII). Similar patterns are presented in Models (4), (6), and (8). Thus, what matters is that a painting was, for instance, exhibited in a prestigious exhibition once or a few times but a high frequency loses its effect, as reflected in the prices.

4.4 Liquidity

It is possible that the provision of provenance is related to past liquidity and hence, that the significance of provenance in the pricing models is induced by past liquidity. We estimate models with the provision of provenance and hammer prices as dependent variables, and include various measures of liquidity. These measures include past sales by artist measures of varying time windows (past 5. 4, 3, 2, and 1 years), and the sales ratio (sales/paintings offered for sale) by artist (over past 5. 4, 3, 2, and 1 years). The former measure captures liquidity, the latter measure captures liquidity including market sentiment. The liquidity measures may affect the subjective risk related to the quality of the painting. For each measure, we can consider global sales or sales by country depending on different assumptions about what auction information potential buyers would consider when contemplating a purchase of a painting at a specific auction house branch. Would they look up information on past auction transactions in the country where a painting was currently being sold, or would they collect information on worldwide sales of an artist's oeuvre? An argument for the latter option would be that it is rather easy to look up global sales when one subscribes to specialized auction sales databases, such as Artnet, Artprice, or Blouin, or when one collects basic information from the internet (which is de facto international). Our conclusion from the pricing models is that past liquidity (past sales and sales ratio) affects current hammer prices because buyers may be willing to pay a premium for the oeuvre of an artist that has been shown to be liquid, in the sense that many paintings were offered for sale and many of those paintings were successfully auctioned (Online Appendix Table A.IX). We demonstrate that in these models, the impacts of provenance on prices documented in previous subsections are not

qualitatively affected by the introduction of liquidity (Online Appendix Table A.IX). Furthermore, the interaction between liquidity and provenance does not affect prices (Online Appendix Table A.X). As the provision of provenance may be affected by liquidity and as this relation may go either way—auction houses may offer more provenance for the most liquid artists given that they sell well or they may offer more provenance for a less liquid oeuvre, as more effort would be needed to sell the paintings—we examine whether provenance is related to lagged liquidity. We find that all models directly relating current provenance provision to past liquidity exhibit insignificant relations (Online Appendix Table A.XI). Our overall conclusion is that liquidity does not affect prices through provenance.

4.5 Artistic Style

Given that provenance may be more important for specific schools of art, we investigate whether the presence of a provenance premium depends on artistic style. We distinguish between 13 styles: (1) Medieval and Renaissance, (2) Baroque, (3) Rococo, (4) Neoclassicism, (5) Romanticism, (6) Realism, (7) Impressionism and Symbolism, (8) Fauvism and Expressionism, (9) Cubism, Futurism, and Constructivism, (10) Dada and Surrealism, (11) Abstract Expressionism, (12) Pop Art, and (13) Minimalism and Contemporary. We observe that provenance and its four dimensions (pedigree, literature, exhibitions, and certification) strongly and significantly affect prices for each subsample by school of art, controlling for the extensive set of hedonic controls and fixed effects (including auction house, year, seasonality). The economic effects are large for each subsample and similar; there is no evidence that younger schools of art have a different provenance premium (Appendix D).

4.6 Auction House Types

Some auction houses have a global reach whereas others are smaller and focus on regional art buyers. While we have controlled for auction house branch fixed effects in the price and provenance provision regressions, we zoom in on the provenance effects by auction house type and relate this to paintings of specific price ranges. The purpose is to study whether a substitution effect exists between auction house types and the provenance information provided. We distinguish between (1) large auction houses (Christie's and Sotheby's); (2) medium-sized auction houses (Bonhams and Phillips, and other important auction houses in United States and Europe[33]; and (3) other small auction houses. In Panel A of Online Appendix Table A.XII, we observe that the hammer prices are strongly related to provenance (ex-ante provided in the catalog) for each type of auction house (Models (1)–(3)). Thus, regardless of the auction house type (from prestigious to small), provenance information is strongly correlated with prices. The same holds for all four dimensions of provenance (Models (4)–(6)) except certification for the largest international auction houses (Christie's and Sotheby's), where certification may matter less

---

[33] For auction house details, see Appendix A.

given that these auction houses' transactions are insured, in that they will repay the purchase of paintings that could later be exposed as forgeries. In Panel B, we repeat the relationship between provenance and auction house for different price ranges (quartiles). We observe that provenance matters for paintings of each price quartile and for each type of auction house (the only exception is the cheapest paintings offered by the largest auction houses—also because of subsample size, as Christie's and Sotheby's sell relatively few paintings within this price range).

4.7 (Non-)established Artists

One could wonder whether paintings by established artists (e.g., Pablo Picasso) would need less provenance to sell or reach high prices. Alternatively, established artists would need more provenance information because (a) artists with a more expensive oeuvre are more likely to be subject to being forged and (b) prolific artists for which there do not exist comprehensive *catalogues raisonnés* are also more likely to be forged. In addition, it is rumored (but difficult to prove) that about 40 to 50% of the contemporary art market consists of forgeries. Consequently, for an established artist, provenance information would be very important. We perform the following test based on price estimates in the auction catalogs: we take the lowest price estimate and consider artists above the 75th percentile as established artists and those below the 25th percentile as non-established. We take only the lowest estimate, which is close to the secret reserve price, and not the highest estimate, as this may be used to create anchoring effects. We document in Online Appendix Table A.XIII that for both established and non-established artists, a strong correlation exists between prices and provenance and its dimensions. The relationship between prices and provenance is even stronger is for established artists. As a further illustration of the importance of provenance within the oeuvre of established artists, we focus on Picasso, Raoul Dufy, and Andy Warhol and confirm the above conclusion (see Online Appendix Table A.XIV).

4.8 Trust versus Quality

Of the four provenance measures, pedigree and certification seem prima facie the dimensions that are expected to create most trust in the offered object of art, whereas exhibitions and literature might not only capture trust but also to some degree reflect or enhance the quality of a painting. *Pedigree* captures ownership and *Certification* captures whether there is physical/oral evidence of authenticity. In what follows, we make a few caveats about quality versus trust versus glamour.

First, with regard to pedigree, ideally, there is an uninterrupted ownership chain between the current seller and the artist; a painting was, for example, in the possession of a family for decades/centuries or in the personal collection of a private collector. However, for paintings that were created decades or centuries ago, an uninterrupted ownership chain is rare. In many cases, the ownership link is severed, which may be caused by the non-availability of intermediate ownership information or originate from the discretion that art buyers cherish about their art purchases/collections. In auctions, the current seller is hardly ever mentioned, and in the vast majority of cases, the counterparty prefers

discretion (one often only learns ex post—if at all—who has purchased a painting). As a consequence of this preference for discretion, the average ownership chain almost inevitably has some lacunas. Missing information may arise when this information is not deemed important (e.g., when nothing is known about an individual intermediate owner apart from temporary ownership, the catalog might not offer such information). Still, even in cases of interrupted ownership chains, the number of ownership indications and the detail of this information could help enhance trust. It may be sufficient for potential buyers that there is information available on the first transaction (between an artist or gallery and a first buyer), as this may then be sufficient proof that the painting is not a forgery. Our textual analysis picks this up as well as transactions through the generations (we capture inheritance/descent—lemmatized). In addition, we identify whether specific types of owners appearing in the pedigree generate trust— either because these owners are considered as knowledgeable about art (investments) or may be considered as wealthy enough to hire the necessary expertise in order to verify art quality/authenticity. As such, we identify prominent collectors (we gather information on 3,885 important collections around the world), but also of nobility/royalty, as art has often been in such families across generations.

Thus, while pedigree may be more related to trust than to quality, some types of owners in the pedigree may unavoidably be related to art quality (famous collectors) or even glamour (wealthy families, CEOs, and celebrities). The presence of such past owners can be related to (a) trust in authenticity, because these past owners can afford to buy in expertise; (b) quality of the art object, as past owners' wealth could enable them to focus on the "best" art (e.g., by art school, period, and artist) and (c) glamour in case current buyers desire to own a piece of art that belonged to a person they admire. Hence, even in the pedigree dimension, for some types of owners, trust, quality, and glamour are inextricably connected. Included in pedigree is the reputation of the (historical) auction house/dealer responsible for past transactions, because the quality of expertise by the research departments of auction houses varies. Therefore, we distinguish in detailed regression models among Christie's, Sotheby's, Bonhams, Phillips, historically important auction houses (e.g., Hôtel Drouot, Dorotheum, and many others), and important current and renowned historic dealers. The fact that past transactions have passed through these auction houses/dealers can induce trust in that art might have been screened well in the past. Still, such paintings might be related to quality in that these auction houses/dealers might have been focused on auctioning "quality" art.

Second, *Certification* is most clearly related to trust, as here, we search for certificates, photographs, witness statements (by the artist's family, pupils, descendants), and statements by foundations and experts. The caveat here is that certificates provided by experts are not foolproof. Ample examples of errors made by experts could be cited; one of the most prominent examples is the certification of Han van Meegeren's biblical Vermeers, which were authenticated by the leading expert Abraham Bredius (whose mistake enabled van Meegeren to mislead the art scene in the first half of the 20th century, as all the Biblical Vermeers were fakes). Van Meegeren also benefited from the expert Cornelis Hofstede de Groot's mistake in identifying a fake Frans Hals as authentic.

While one could regard the provenance dimensions *Literature* and *Exhibition* as capturing quality (rather than trust), as one could expect that the most important paintings would appear in the literature and or would be exhibited, this is only partially true, because both dimensions also have a strong trust aspect in the following sense. In relation to *Literature*: *catalogues raisonnés* strive by definition to list and describe an artist's oeuvre exhaustively. Moreover, art books on specific artists often cover a substantial part of their works. The *Literature* dimension also captures whether (historic) photos/images/illustrations of a particular work are included in art books and this dimension in this respect also plays a role in certification. We focus on the most authoritative press (e.g., scientific books by university presses) and the most reputable publishers, which all use refereeing committees) Thus, these books are based on art historical/art market research. Consequently, we argue that the literature dimension includes an important aspect of trust (likely even dominating the "quality" label of this dimension), but acknowledge that a complete separation between quality and trust cannot be achieved. Regarding the *Exhibition* dimension, it may indeed be the case that an artist's highest quality paintings are exhibited as part of a permanent collection of museums and included in exhibitions organized by other renowned museums/galleries. Nonetheless, exposure through the latter channel may face impediments, as museums or collectors often do not let their most important works travel. In addition, exhibitions focusing on artists' impact and significance in art history exhibit paintings from their early or late periods to offer a comprehensive picture of their careers or to show influences by or on other artists. Moreover, exhibitions often trigger new research on the exhibited work and the artist, with scientific articles published in the exhibition catalog, such that the dimensions *Exhibition* and *Literature* can be correlated (but not to the extent of causing multicollinearity in our models). Finally, over the past three decades, new exhibitions with loaned paintings often lead to technical examinations (infra-red reflectography, non-invasive spectroscopic imaging, X-rays, chemical analysis of paint, etc.) or restoration. Consequently, exhibited works are thoroughly scrutinized such that exhibitions contribute to deeper insights affecting trust. Of all the four dimensions, the exhibition dimension is most related to quality or salience.

In summary, trust and quality (and even status or glamour) are embedded in provenance. *Pedigree* and *Certification* capture trust, but may reflect to some degree quality and status. As argued above, *Literature* may capture trust to a much larger degree than it reflects the quality of a painting. *Exhibition* may indeed be more related to quality than trust, and the latter aspect originates from additional exhibition-induced research.

If trust were priced, one could expect a higher probability of being sold and higher price premiums for the provenance dimensions in this order: pedigree/certification, literature, and exhibitions. However, the statistical and economic significance also depends on the degree to which each of the dimensions (and certainly their constituting elements) (a) captures quality/ status/glamour; and (b) is present within the provenance records of the auction catalogs. The pricing models of Table 2 (Panel A) report that each of the four dimensions has a strong price impact, with *Literature* exhibiting the strongest and

Exhibition the second strongest effect, such that a high price premium may capture not merely trust but also some aspects of the quality of the painting (while it should be recognized that the models control for artist reputation, auction house reputation, and physical aspects of the painting as well as transaction characteristics). The dimensions *Pedigree* and *Certification* may capture trust to a much larger extent than quality, which may explain why the price premium is lower (but still strongly significant and economically large). When we investigate in more detail (Panel B), we find that within the *Certification* dimension, a physical certification by the artist (trust) has the strongest impact on the probability of being sold. Conditional on being sold, we find that all types of certificates affect hammer prices (with the highest impact by a physical certificate by an expert and the non-physical testimonial by the artist).

Within the *Literature* dimension, the probability of being sold is positively affected, first, by the presence of the picture on the cover page (capturing a combination of trust, quality, and salience) and second, by inclusion in the *catalogue raisonné* (which is also one of the strongest trust measures). Again, all constituting elements of the literature dimension affect prices (cover page, publication by authoritative press, etc.).

For the *Pedigree* dimension, past ownership (prominent collectors, corporate collections, nobility/royalty, wealthy families) strongly affects the probability of being sold. It is also remarkable that the category of famous sportspeople affects the probability and prices—here, the aspect of glamour (owning a piece acquired by an idol) may add to both trust and quality (as wealthy sportspeople can afford quality and pay for expertise to select art). Descent information is priced (also when a first sale can be traced directly to an artist).

Turning to the *Exhibition* dimension, we find that the most prominent exhibitions (organized by prominent museums and renown galleries) affect the probability of being sold as well as prices. These results could be explained by selection of the highest quality paintings, but the arguments outlined above explain why there may also be a strong trust aspect.

5. Conclusion

Trust is key for any type of market, but particularly for illiquid, opaque, and largely unregulated markets, such as the art market. A lack of trust undermines sales and prices, which is worsened when fakes and forgeries are rumored to circulate in the art market. Therefore, guarantees about the authenticity of an art object are pivotal in creating trust, and the provision of provenance can be a (partial) solution by emitting a signal about the art's authenticity. Provenance comprises records of ownership or pedigree, exhibition history, literature coverage, and certification, all of which relate to the artwork's authenticity and could enhance trustworthiness of the object offered for sale. For example, if there is evidence that a painting was originally purchased from the artist or artist's family, from a first buyer whose family has held the painting for generations, or from famous collectors, and there are corroborating documents, the potential buyer's caution will be attenuated.

We investigate the impact of providing detailed provenance information, measured by hundreds of variables resulting from textual analysis applied to auction catalogs, along with the set of traditional art value determinants, on the probability of being sold, hammer prices, and returns of about two million paintings and works on paper. We find that provenance information provision increases the probability of being sold by 2% to 4%, leads to a price premium of 14% to 54%, and increases the annualized returns by 5 to 16 percentage points after controlling for artwork and transactions characteristics (e.g., topic, signature, medium, and measurements), as well as artist, time, and auction house branch fixed effects.

A first type of endogeneity is embedded in the decision to offer provenance, which may be affected by recent price increases of the painting (in case of a repeat sales) or of similar paintings (e.g., by the same artists or school), as well as by expected prices (proxied by price estimates). To address the concern of reverse causality in the relationship between past prices and changes in provenance, we study the provenance effects for artists affected by the discoveries of fakes and forgeries, and exploit Christie's policy change in the provision of provenance information following a litigation case against the firm as a quasi-natural experiment in a DiD setting. We also run two-staged regressions for repeat sales transactions to control for changes in provenance induced by past high prices. To address the potential endogeneity concern in the sales decision, we examine subsamples of "exogenous" sales, which we expect to be less affected by past price trends (e.g., sales following the death of collectors). These attempts to address endogeneities yield results similar to those of our baseline models. In conclusion, provenance information is an important factor corroborating an artwork's authenticity and creating trust in art markets as reflected in sales, prices, and returns.

REFERENCES

Adams, RB, Kräussl, R, Navone, M, Verwijmeren, P (2021) Gendered prices. *Rev. Financ. Stud.* 34(8):3789–3839.

Beggs A, Graddy K (2009) Anchoring effects: Evidence from art auctions. *Am. Econ. Rev.* 99(3):1027–1039.

Belloni A, Chernozhukov V, Hansen C (2014) High-dimensional methods and inference on structural and treatment effects. *J. Econ. Perspect.* 28(2):29–50.

Bocart F, Oosterlinck K (2011) Discoveries of fakes: Their impact on the art market. *Econ. Lett.* 113(2):124–126.

Cameron L, Goetzmann WN, Nozari M (2019) Art and gender: market bias or selection bias? *J. Cult. Econ.* 1–29.

Chohan UW (2021). Non-fungible tokens: Blockchains, scarcity, and value. *Critical Blockchain Research Initiative (CBRI) Working Papers*.

Deloitte Luxembourg and ArtTactic (2014) Art & Finance Report 2014. Accessed February 7, 2022. https://www2.deloitte.com/content/dam/Deloitte/es/Documents/acerca-de-deloitte/Deloitte-ES-Opera_Europa_Deloitte_Art_Finance_Report2014.pdf

Deloitte Luxembourg and ArtTactic (2016) Art & Finance Report 2016. Accessed February 7, 2022. https://www2.deloitte.com/content/dam/Deloitte/lu/Documents/financial-services/artandfinance/ lu-en-artandfinancereport-21042016.pdf

Fairfield J (2022). Tokenized: The law of non-fungible tokens and unique digital property. *Indiana Law J.* Forthcoming.

Goetzmann WN, Renneboog L, Spaenjers C (2011) Art and money. *Am. Econ. Rev.* 101(3):222–226.

Graddy K, Loewenstein L, Mei J, Moses M, Pownall R (2015) Empirical evidence of anchoring and loss aversion from art auctions. Working paper No. 73, Brandeis University.

Guiso L, Sapienza P, Zingales L (2008) Trusting the stock market. *J. Financ.* 63(6):2557–2600.

Korteweg A, Kräussl R, Verwijmeren P (2016) Does it pay to invest in art? A selection-corrected returns perspective. *Rev. Financ. Stud.* 29(4):1007–1038.

Li Y, Ma M, Renneboog L (2022). Pricing art and the art of pricing: On returns and risk in the art auction markets. *Eur. Financ. Manag.* (No. 12348). Forthcoming.

Lovo S, Spaenjers C (2018) A model of trading in the art market. *Am. Econ. Rev.* 108(3):744–774.

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

Pénasse J, Renneboog L (2022) Speculative trading and bubbles: Evidence from the art market. *Manag. Sci.* Forthcoming.

Pénasse J, Renneboog L, Spaenjers C (2014) Sentiment and art prices. *Econ. Lett.* 122(3):432–434.

Pénasse J, Renneboog L, Scheinkman J (2021) When a master dies: Speculation and asset float. *Rev. Financ. Stud.* 34(8):3840–3879.

Pownall R (2017) TEFAF Art Market Annual Report 2017. The European Fine Art Foundation. Accessed February 7, 2022. http://made2measure.org/tefaf/amr2017/

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

Roberts MR, Whited TM (2013) Endogeneity in empirical corporate finance. Constantinides G, Harris M, Stulz R, eds. *Handbook of the Economics of Finance*, vol. 2 (Elsevier, Amsterdam), 493–572.

Thompson D (2010) *The $12 Million Stuffed Shark: The Curious Economics of Contemporary Art* (Aurum Press, London).

Table 1. Descriptive Statistics on Provenance

This table presents the descriptive statistics of the provenance variables. *Pedigree (Text Length)* stands for the number of characters. *Prominent Collectors*, *Royalty / Nobility*, *Wealthy Families*, *CEOs*, *Influential People* (*Time 100)*, *Celebrities*, and *Famous Sportspeople* are indicator variables equal one if the artwork has been in the collections of those types of collectors, respectively. *Corporate Collection* and *Private Collection (Anonymous)* equal one if the artwork was at one point part of corporate and private collections, respectively. *Directly from Artist*, *From Artist Family*, and *From Sitter* equal one if the paintings were acquired directly from the above respective categories. *Descent* equals one if the artwork's pedigree information contains any descendance information (not included in the above categories). *Sold by Sotheby's or Christie's*, *Sold by Bonhams or Phillips*, *Sold by Historic Auction Houses*, *Sold by Other Important Auction Houses*, and *Sold by Prominent Dealers* equal one if the artwork was once sold via the respective channels. *Other Pedigree Information* equals one if the artworks have other unclassified pedigree information. *Exhibition (Text Length)* is the number of characters of exhibition information, and *Exhibition (Number Count)* is the number of past exhibitions by the painting. *Prominent Exhibition, Prominent Art Fair*, *Prominent Museum, Other Museum, Cultural City*, and *Gallery Exhibition* equal one if the painting was at one point exhibited in the above types of exhibitions/museums/fairs/cities, respectively. *Literature (Text Length)* is the number of characters on literature information, and *Literature (number count)* is the number of times that the painting is referred to in the art literature. *Catalogue Raisonné, Cover Page* (of an art history book), *Illustration* (in an art history book), and *Authoritative Press* equal one if the artwork was illustrated in the above ways, respectively. *Other Literature* equals one if the catalog information refers to other literature information not included in the above categories. *Certification (Text Length)* is the number of characters related to certification. *Artist (Physical)*, *Artist Family (Physical)*, (artists') *Association (Physical)*, *Expert (Physical)*, and *Other People (Physical)* equal one if the painting has a physical certification (e.g., "photo certificate of authenticity by artist") issued by the above sources, respectively. *Artist (Non-Physical)*, *Artist Family (Non-Physical)*, *Association (Non-Physical)*, *Expert (Non-Physical)*, and *Other People (Non-Physical)* equal one if the painting has non-physical certification (e.g., "the authenticity was orally confirmed by Paul Vogt, Essen") issued by the above sources, respectively. Variables with "Number Count" are count variables. For each variable, we report the number of observations (N), the conditional mean, the standard deviation (S.D.), the minimum, and the maximum.

	N	Mean	S.D.	Min	Max
Pedigree					
Pedigree (Text Length)	256,560	103.99	152.79	0	9,034
Owned by:					
Prominent Collectors	256,560	2.07%	14.20%	0	1
Royalty / Nobility	256,560	2.39%	15.30%	0	1
Wealthy Families	256,560	0.68%	8.22%	0	1
CEOs	256,560	0.06%	2.45%	0	1
Influential People (Time 100)	256,560	0.08%	2.89%	0	1
Celebrities	256,560	0.20%	4.46%	0	1
Famous Sportspeople	256,560	0.12%	3.52%	0	1
Corporate Collection	256,560	0.16%	4.01%	0	1
Private Collection (Anonymous)	256,560	26.00%	43.80%	0	1
Descent:					
Directly from Artist	256,560	10.51%	30.67%	0	1
From Artist's Family	256,560	5.29%	22.39%	0	1
From Sitter	256,560	0.33%	5.76%	0	1
Other Descent Information	256,560	9.81%	29.75%	0	1
Past Sales Channel:					
Sold by Sotheby's or Christie's	256,560	14.96%	35.67%	0	1
Sold by Bonhams or Phillips	256,560	0.90%	9.42%	0	1
Sold by Historic Auction Houses	256,560	1.33%	11.46%	0	1
Sold by Other Important Auction Houses	256,560	1.30%	11.30%	0	1
Sold by Prominent Dealers	256,560	5.75%	23.29%	0	1
Other Collections:					
Other Pedigree Information	256,560	42.48%	49.43%	0	1
Number Count by Painting:					
Prominent Collectors (Number Count)	256,560	0.0213	0.1690	0	8
Descent (Number Count)	256,560	0.1090	0.3610	0	11

	N	Mean	S.D.	Min	Max
Sold by Sotheby's or Christie's (Number Count)	256,560	0.1780	0.4660	0	12
Sold by Bonhams or Phillips (Number Count)	256,560	0.0092	0.0983	0	5
Sold by Other Important Auction Houses (Number Count)	256,560	0.0137	0.1230	0	3
Sold by Historic Auction Houses (Number Count)	256,560	0.0150	0.1370	0	5
Sold by Prominent Dealers (Number Count)	256,560	0.0624	0.2640	0	6
Exhibition					
Exhibition (Text Length)	67,713	208.87	283.31	1	6,828
Exhibition (Number Count)	67,713	1.9760	2.6480	0	46
Prominent Exhibition	67,713	6.20%	24.10%	0	1
Prominent Art Fair	67,713	0.39%	6.20%	0	1
Prominent Museum	67,713	17.20%	37.80%	0	1
Other Museum	67,713	29.90%	45.80%	0	1
Cultural City	67,713	74.10%	43.80%	0	1
Gallery Exhibition	67,713	14.90%	35.60%	0	1
Number Count by Painting:					
Prominent Exhibition (Number Count)	67,713	0.0708	0.3050	0	18
Prominent Art Fair (Number Count)	67,713	0.0041	0.0674	0	4
Prominent Museum (Number Count)	67,713	0.2730	0.7860	0	20
Other Museum (Number Count)	67,713	0.2992	0.4579	0	1
Cultural City (Number Count)	67,713	1.5740	2.1210	0	41
Literature					
Literature (Text Length)	72,906	241.79	388.37	0	22,413
Literature (Number Count)	72,906	1.5300	2.3970	0	150
Catalogue Raisonné	72,906	15.70%	36.40%	0	1
Cover Page	72,906	1.66%	12.80%	0	1
Illustration	72,906	45.90%	49.80%	0	1
Authoritative Press	72,906	1.15%	10.60%	0	1
Other Literature	72,906	48.10%	50.00%	0	1
Number Count by Painting:					
Catalogue Raisonné (Number Count)	72,906	0.1690	0.4100	0	6
Cover Page (Number Count)	72,906	0.0181	0.1490	0	6
Illustration (Number Count)	72,906	0.8450	1.6170	0	89
Authoritative Press (Number Count)	72,906	0.0122	0.1180	0	4
Certification					
Certification (Text Length)	70,556	66.90	63.35	6	4,101
Certification by:					
Artist (Physical)	70,556	31.70%	46.50%	0	1
Artist Family (Physical)	70,556	6.16%	24.00%	0	1
Association (Physical)	70,556	15.10%	35.80%	0	1
Expert (Physical)	70,556	2.72%	16.30%	0	1
Other People (Physical)	70,556	27.60%	44.70%	0	1
Artist (Non-Physical)	70,556	5.31%	22.40%	0	1
Artist's Family (Non-Physical)	70,556	2.20%	14.70%	0	1
Association (Non-Physical)	70,556	4.45%	20.60%	0	1
Expert (Non-Physical)	70,556	2.88%	16.70%	0	1
Other People (Non-Physical)	70,556	6.15%	24.00%	0	1

Table 2. Provenance Effects on Probability of Being Sold and Hammer Price

This table presents the relation between Provenance, and Probability of Being Sold and Hammer Price. The dependent variables are: (1) *Sold[0,1]*, which takes one if the painting is successfully sold (and zero when bought in) , and (2) *Ln(Price)*, the natural logarithm of the real hammer price in real (2007) USD. In Panel A, Columns (1) and (3), *Pedigree*, *Exhibition*, *Literature*, and *Certification* are dummy variables capturing if the artwork's catalog comprises any information on these respective dimensions. In Panel A, Columns (2) and (5), these four provenance variables stand for the natural log of the number of characters used in the provenance text for each of these respective dimensions. Column (4) in Panel A reports the corresponding price impact (relative price change) of Column (3) calculated as $exp(estimated\ coefficient) - 1$. Panel B has a similar setup but the independent variables are the detailed elements that constitute *Pedigree*, *Exhibition*, *Literature*, and *Certification* are included (for definitions, see Appendix A). In both panels, *Ln(Low Price Estimate)*, the log of the low price estimate (in real USD), is a proxy for the reserve price. All regressions in both panels include Hedonic Controls (detailed in Online Appendix Table A.I), and Artist, Year, Month, and Auction House Branch Level (AH) Fixed Effects. *, **, and *** indicate statistical significance at the 10%, 5%, and 1% levels, respectively. Standard errors (S.E.) are reported in parentheses and clustered at the auction branch level.

Panel A: Provenance Effects on Probability of Being Sold and Hammer Price

Dept. Var.:	Sold[0,1]		Ln(Price)		
	(1)	(2)	(3)	(4)	(5)
Provenance as:	Indicator	Text Length	Indicator	Price Impact	Text Length
Pedigree	0.0173***	0.0044***	0.1885***	20.74%	0.0531***
	(0.0062)	(0.0014)	(0.0170)		(0.0045)
Exhibition	0.0379***	0.0079***	0.3499***	41.89%	0.0734***
	(0.0041)	(0.0008)	(0.0208)		(0.0041)
Literature	0.0250***	0.0056***	0.4288***	53.54%	0.0869***
	(0.0063)	(0.0011)	(0.0336)		(0.0063)
Certification	0.0102	0.0115***	0.1297***	13.85%	0.0614***
	(0.0133)	(0.0033)	(0.0226)		(0.0100)
Ln(Low Price Estimate)	-0.0661***	-0.0668***			
	(0.0036)	(0.0036)			
Artist FE	Yes	Yes	Yes		Yes
Year FE	Yes	Yes	Yes		Yes
Month FE	Yes	Yes	Yes		Yes
Auct. House Branch FE	Yes	Yes	Yes		Yes
Hedonic Controls	Yes	Yes	Yes		Yes
# of Obs.	1,707,136	1,707,136	1,111,220		1,111,220
R-squared	0.1750	0.1752	0.7805		0.7817

Panel B: Detailed Provenance Effects on Probability of Being Sold and Hammer Price

Dept. Var.:	Sold [0,1]	Ln(Price)	
	(1)	(2)	(3)
Provenance as:	Indicator	Indicator	Price Impact
Ln(Low Price Estimate)	-0.0665***		
	(0.0036)		
Pedigree			
Past Ownership			
Prominent Collectors	0.0524***	0.2183***	24.40%
	(0.0113)	(0.0405)	
Royalty / Nobility	0.0656***	0.2724***	31.31%
	(0.0148)	(0.0262)	
Wealthy Families	0.0855***	0.3538***	42.45%
	(0.0232)	(0.0479)	
CEOs	-0.0101	0.1479	15.94%
	(0.0343)	(0.1021)	
Influential People (Time 100 list)	-0.0167	0.0841	8.77%
	(0.0372)	(0.0842)	
Celebrities	0.0372**	0.1288	13.75%
	(0.0170)	(0.1308)	
Famous Sportspeople	0.0948***	0.4062***	50.11%
	(0.0289)	(0.1358)	
Corporate Collection	0.0976***	0.1038**	10.94%
	(0.0376)	(0.0482)	
Private Collection (Anonymous)	-0.0051	0.1793***	19.64%
	(0.0084)	(0.0222)	
Descent			
Directly from Artist	0.0145***	0.1296***	13.84%
	(0.0056)	(0.0142)	
From Artist's Family	-0.0027	0.0070	0.70%
	(0.0049)	(0.0239)	
From Sitter	-0.0892***	0.1068**	11.27%
	(0.0185)	(0.0515)	
Other Descent Information	0.0288***	0.2121***	23.63%
	(0.0045)	(0.0190)	
Past Sales Channel			
Sold by Sotheby's or Christie's	-0.0060	0.2121***	23.63%
	(0.0042)	(0.0243)	
Sold by Bonhams or Phillips	-0.0277*	0.0687*	7.11%
	(0.0147)	(0.0387)	
Sold by Historic Auction Houses	-0.0164**	0.0957***	10.04%
	(0.0072)	(0.0363)	
Sold by Other Important Auction Houses	-0.0133	-0.0463	-4.52%
	(0.0083)	(0.0393)	
Sold by Prominent Dealers	0.0377***	0.2842***	32.87%
	(0.0064)	(0.0378)	
Other Collections			
Other Pedigree Information	0.0274***	0.1323***	14.15%
	(0.0058)	(0.0133)	
Exhibition			
Prominent Exhibition	0.0263***	0.2327***	26.20%
	(0.0081)	(0.0247)	
Prominent Art Fair	0.0015	-0.0055	-0.55%
	(0.0298)	(0.0648)	
Prominent Museum	0.0543***	0.4631***	58.90%
	(0.0083)	(0.0331)	

Dept. Var.:	Sold [0,1]	Ln(Price)	
	(1)	(2)	(3)
Provenance as:	Indicator	Indicator	Price Impact
Other Museum	0.0176***	0.1878***	20.66%
	(0.0042)	(0.0147)	
Cultural City	0.0192***	0.2165***	24.17%
	(0.0034)	(0.0147)	
Gallery Exhibition	0.0430***	0.2535***	28.85%
	(0.0077)	(0.0220)	
Literature			
Catalogue Raisonné	0.0236**	0.3056***	35.74%
	(0.0095)	(0.0448)	
Cover Page	0.0509***	0.4237***	52.76%
	(0.0143)	(0.0513)	
Illustration	0.0207***	0.3660***	44.20%
	(0.0079)	(0.0366)	
Authoritative Press	0.0130	0.3450***	41.20%
	(0.0183)	(0.0876)	
Other Literature	0.0149*	0.3186***	37.52%
	(0.0081)	(0.0252)	
Certification by			
Artist (Physical)	0.0599***	0.0828***	8.63%
	(0.0172)	(0.0262)	
Artist's Family (Physical)	-0.0166	0.0674*	6.97%
	(0.0152)	(0.0385)	
Association (Physical)	0.0348**	0.1199***	12.74%
	(0.0148)	(0.0410)	
Expert (Physical)	-0.0169	0.3416***	40.72%
	(0.0179)	(0.0517)	
Other People (Physical)	0.0011	0.1148***	12.16%
	(0.0129)	(0.0302)	
Artist (Non-Physical)	0.0348	0.2810***	32.45%
	(0.0224)	(0.0287)	
Artist's Family (Non-Physical)	-0.0312*	0.0039	0.39%
	(0.0171)	(0.0597)	
Association (Non-Physical)	-0.0381	0.1889***	20.79%
	(0.0240)	(0.0352)	
Expert (Non-Physical)	-0.0292	0.1654***	17.99%
	(0.0211)	(0.0482)	
Other People (Non-Physical)	-0.0206	0.1340***	14.34%
	(0.0136)	(0.0393)	
Artist FE	Yes	Yes	
Year FE	Yes	Yes	
Month FE	Yes	Yes	
Auct. House Branch FE	Yes	Yes	
Hedonic Variables	Yes	Yes	
# of Obs.	1,707,136	1,111,220	
R-squared	0.1755	0.7819	

Table 3. Provenance Effects on Hammer Price after the Discovery of Fakes and Forgeries

This table presents the difference-in-differences (DiD) estimators related to the discoveries of faked or forged paintings. The dependent variable is the natural logarithm of deflated hammer price in USD. *Treated* equals one if the auction date of a specific artist's painting falls after the date of the discovery of fakes/forgeries for this artist. *Provenance, Pedigree, Exhibition, Literature,* and *Certification* are dummy variables equal to one if the painting contains any such information from the auction catalog. The control variables are discussed in Section 2. Full sample results are presented in Columns (1) and (2) and results for the subsample of artists active in 20th century are in Columns (3) and (4). *, **, and *** indicate statistical significance at the 10%, 5%, and 1% levels, respectively. Standard errors (S.E.) are reported in parentheses and clustered at the auction branch level.

| | Full sample | | Artists active in 20th century | |
| | (1) | (2) | (3) | (4) |
Dept. Var.:	Ln(Price)	Ln(Price)	Ln(Price)	Ln(Price)
Treated	-0.2670**	-0.1977*	-0.2747**	-0.2023
	(0.1168)	(0.1142)	(0.1267)	(0.1262)
Treated × Provenance	0.7054***		0.7115***	
	(0.0895)		(0.1017)	
Treated × Pedigree		0.3520***		0.3459***
		(0.0926)		(0.0988)
Treated × Exhibition		0.2215**		0.2652***
		(0.0915)		(0.0729)
Treated × Literature		0.3529***		0.3253***
		(0.1024)		(0.1245)
Treated × Certification		0.0656		0.0163
		(0.2314)		(0.2626)
Provenance	0.2715***		0.2526***	
	(0.0201)		(0.0228)	
Pedigree		0.1863***		0.1995***
		(0.0171)		(0.0210)
Exhibition		0.3450***		0.3300***
		(0.0193)		(0.0253)
Literature		0.4178***		0.4131***
		(0.0318)		(0.0351)
Certification		0.1302***		0.1117***
		(0.0227)		(0.0247)
Artist FE	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes
Month FE	Yes	Yes	Yes	Yes
Hedonic Controls	Yes	Yes	Yes	Yes
Auct. House Branch FE	Yes	Yes	Yes	Yes
# of Obs.	1,111,220	1,111,220	548,684	548,684
R-squared	0.7773	0.7807	0.7963	0.7996

Table 4. Christie's Provenance Policy Change

This table presents the difference-in-differences (DiD) results exploiting the Christie's Provenance Policy Shock in 2012. The dependent variable is *Ln(Price)*, the natural logarithm of deflated hammer price in USD. The sample consists of all sales by Christie's and Sotheby's. *Treated* is a dummy that equals one when a sale takes place through Christie's since 2013. *Provenance*, *Pedigree*, *Exhibition*, *Literature*, and *Certification* equal one if the catalog provides information on provenance and its dimensions. All regressions include Hedonic Controls (see Online Appendix Table A.I), and Artist, Year, Month, and Auction House Branch Level (AH) Fixed Effects. *, **, and *** indicate statistical significance at the 10%, 5%, and 1% levels, respectively.

	(1)	(2)	(3)	(4)	(5)	(6)
Dept. Var:	Ln(Price)	Ln(Price)	Ln(Price)	Ln(Price)	Ln(Price)	Ln(Price)
Treated	-0.0165	-0.0026	0.0547***	0.0260***	0.0478***	-0.0019
	(0.0132)	(0.0128)	(0.0102)	(0.0100)	(0.0098)	(0.0126)
Treated × Provenance	0.1179***					
	(0.0125)					
Treated × Pedigree		0.1017***				0.0805***
		(0.0122)				(0.0125)
Treated × Exhibition			0.0904***			0.0481***
			(0.0136)			(0.0151)
Treated × Literature				0.0753***		0.0145
				(0.0133)		(0.0150)
Treated × Certification					0.3581***	0.3157***
					(0.0805)	(0.0781)
Provenance	0.2897***					
	(0.0061)					
Pedigree		0.2574***				0.1729***
		(0.0061)				(0.0061)
Exhibition			0.4454***			0.3061***
			(0.0070)			(0.0073)
Literature				0.5407***		0.4185***
				(0.0072)		(0.0075)
Certification					-0.1225***	-0.0743***
					(0.0253)	(0.0246)
Artist FE	Yes	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes	Yes
Month FE	Yes	Yes	Yes	Yes	Yes	Yes
Hedonic Controls	Yes	Yes	Yes	Yes	Yes	Yes
Auct. House Branch FE	Yes	Yes	Yes	Yes	Yes	Yes
# of Obs.	198,076	198,076	198,076	198,076	198,076	198,076
R-squared	0.7715	0.7708	0.7745	0.7765	0.7677	0.7708

Table 5. Addressing Endogeneity in the Relationship between Past Prices and Provenance Changes

This table shows models addressing endogeneity concerns in the relation *Past Price and Provenance Changes*. The dependent variables *ΔPedigree, ΔExhibition, ΔLiterature* and *ΔCertification* in Panel A are changes in the information related to these provenance dimensions between the first and second sales of each repeat sales pair. The changes are captured by (i) dummy variables equal one when new provenance information arises between the first and second sale, and are zero otherwise or (ii) changes in the textual length of each dimension. *Ln(Price$_1$)* is the natural log of the first sale hammer prices in real USD. *Time Span* is the number of days between the two sales. All regressions include initial pedigree, exhibition, literature, and certification information provided just prior to the first transaction of the repeat sales pair. In Panel B, *ε(ΔPedigree), ε(ΔExhibition), ε(ΔLiterature)* and *ε(ΔCertification)* are the corresponding residuals from the regressions in Panel A. All regressions include hedonic controls (see Online Appendix Table A.I), and Artist, Year, Month, and Auction House Branch Level (AH) Fixed Effects. *, **, and *** indicate statistical significance at the 10%, 5%, and 1% levels, respectively. Standard errors (S.E.) are reported in parentheses and clustered at the auction branch level.

Panel A: The Impact of First Sale Price on the Provenance Changes

Dept. Var.:	Changes in provenance provision (Dummy variables)				Changes in the text length of provenance information			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	ΔPedigree	ΔExhibition	ΔLiterature	ΔCertification	ΔPedigree	ΔExhibition	ΔLiterature	ΔCertification
Ln(Price$_1$)	-0.0001	0.0109*	0.0226***	0.0098**	0.0407	0.0479	0.1159***	0.0407*
	(0.0084)	(0.0056)	(0.0071)	(0.0050)	(0.0443)	(0.0349)	(0.0392)	(0.0220)
Time Span	0.0001**	-0.0001	0.0001	0.0001**	0.0002***	0.0001*	0.0001***	0.0002**
	(0.0001)	(0.0001)	(0.0001)	(0.0001)	(0.0001)	(0.0001)	(0.0001)	(0.0001)
Pedigree$_1$	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Exhibition$_1$	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Literature$_1$	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Certification$_1$	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Artist FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Month FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Hedonic Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Auct. House Branch FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
# of Obs.	6,612	6,612	6,612	6,612	6,612	6,612	6,612	6,612
R-squared	0.6578	0.4311	0.3773	0.4291	0.6327	0.3286	0.3267	0.3764

Panel B: Second Sale Price and Unpredicted Provenance Changes

Dept. Var.:	Changes in Provenance (Dummies)		Changes in Text Length	
	(1)	(2)	(3)	(4)
	Ln(Price$_2$)	Ln(Price$_2$)	Ln(Price$_2$)	Ln(Price$_2$)
$\varepsilon(\Delta\text{Pedigree})$	0.0329	0.0322	0.0213**	0.0213**
	(0.0410)	(0.0413)	(0.0083)	(0.0084)
$\varepsilon(\Delta\text{Exhibition})$	0.1050**	0.1050**	0.0206*	0.0203**
	(0.0512)	(0.0508)	(0.0110)	(0.0101)
$\varepsilon(\Delta\text{Literature})$	0.1479***	0.1462***	0.0270***	0.0266***
	(0.0273)	(0.0270)	(0.0058)	(0.0056)
$\varepsilon(\Delta\text{Certification})$	0.0504**	0.0502**	0.0011	0.0012
	(0.0235)	(0.0242)	(0.0054)	(0.0054)
Ln(Price$_1$)	0.6233***	0.6183***	0.6229***	0.6164***
	(0.0556)	(0.0559)	(0.0557)	(0.0561)
Time Span		-0.0001***		-0.0001***
		(0.0001)		(0.0001)
Pedigree$_1$	Yes	Yes	Yes	Yes
Exhibition$_1$	Yes	Yes	Yes	Yes
Literature$_1$	Yes	Yes	Yes	Yes
Certification$_1$	Yes	Yes	Yes	Yes
Artist FE	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes
Month FE	Yes	Yes	Yes	Yes
Hedonic Controls	Yes	Yes	Yes	Yes
Auct. House Branch FE	Yes	Yes	Yes	Yes
# of Obs.	6,346	6,346	6,346	6,346
R-squared	0.9504	0.9506	0.9505	0.9506

Table 6. Exogenous Sales Decisions

This table presents the models relating provenance effects to the probability of being sold or hammer price for a subsample of transactions of which the sales decision may be taken more exogenously (less dependent on past prices). The sample employed includes sales retained when their auction title refers to "Estate Sale," "Property of," "Legacy," "Bequest," "Heritage," "Gift," "Endowment," or "Charity" (in singular or plural) or contains "Late" plus a person's name. The dependent variables are: (i) *Sold[0,1]* that equals one if the auction lot is successfully sold, and (ii) *Ln(Price),* which is the natural log of deflated hammer price in real USD. *Pedigree*, *Exhibition*, *Literature*, and *Certification* are defined in Appendix A. In Columns (1) and (3), the provenance variables equal one if a painting has corresponding provenance information (by provenance dimension) and zero otherwise. In Columns (2) and (4), the provenance variables are the natural log of text character length. *Ln(Low Price Estimate)*, the log of the low price estimate, is a proxy for the reserve price. All regressions include Hedonic Controls (see Online Appendix Table A.I), and Artist, Year, Month, and Auction House Branch Level (AH) Fixed Effects. *, **, and *** indicate statistical significance at the 10%, 5%, and 1% levels, respectively.

Dept. Var.:	Sold[0,1]		Ln(Price)	
	(1)	(2)	(3)	(4)
Provenance as:	Indicator	Text Length	Indicator	Text Length
Pedigree	0.0368**	0.0104**	0.1659***	0.0439***
	(0.0180)	(0.0044)	(0.0513)	(0.0125)
Exhibition	0.0791***	0.0158***	0.4118***	0.0866***
	(0.0277)	(0.0055)	(0.0910)	(0.0180)
Literature	0.0347	0.0073	0.5084***	0.1082***
	(0.0232)	(0.0049)	(0.1180)	(0.0274)
Certification	0.0470	0.0091	0.1366	0.0211
	(0.0302)	(0.0116)	(0.1016)	(0.0164)
Ln(Low Price Estimate)	-0.1221***	-0.1231***		
	(0.0146)	(0.0145)		
Artist FE	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes
Month FE	Yes	Yes	Yes	Yes
Hedonic Controls	Yes	Yes	Yes	Yes
Auct. House Branch FE	Yes	Yes	Yes	Yes
# of Obs.	26,733	26,733	16,527	16,527
R-squared	0.4369	0.4371	0.8499	0.8509

Table 7. Provenance Effects and Returns for Repeat Sales Transactions

The regressions are based on the repeat sales (RS) sample. The dependent variable *Return* is the geometrically annualized return; results are for the full RS sample (Models (1) to (4)) and for the RS transactions that took place in the same auction house (branch) (Model (5) to (8)). *ΔPedigree, ΔExhibition, ΔLiterature* and *ΔCertification* are the changes in provenance information in between the two sales. The change in provenance precedes the return as provenance information in the auction catalogue usually precedes the auction by 4 to 6 weeks. They are defined as either (i) a change from no provenance information (pedigree, exhibition, literature, and certification) at the first sale to available information at the second sale or (ii) a change in textual length of provenance information (by dimension) from the first sale to the second sale. *Auction House Upgrade* equals one when a second sale moves up from a small auction house to a medium or big auction house or when a second sale moves from a medium auction house to a large one. All regressions include the hedonic controls that are presented in Online Appendix Table A.I, and Artist, Year, Month, and Auction House Branch Level (AH) Fixed Effects. *, **, and *** indicate statistical significance at the 10%, 5%, and 1% levels, respectively. Standard errors (S.E.) are reported in parentheses and clustered at the auction branch level.

Dept. Var.: Return	Full RS Sample				RS Sample with Transactions in the Same Auction House			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Provenance as:	Indicator	Text Length	Indicator	Text Length	Indicator	Text Length	Indicator	Text Length
ΔPedigree	-0.1123	0.0032	-0.0489	0.0065	-0.4264	0.0014	-0.1382	0.0160
	(0.1064)	(0.0199)	(0.1474)	(0.0218)	(0.3403)	(0.0634)	(0.4136)	(0.0691)
ΔExhibition	0.1401**	0.0246**	0.1646**	0.0250**	0.1865*	0.0341	0.2761*	0.0548**
	(0.0621)	(0.0125)	(0.0715)	(0.0123)	(0.0960)	(0.0253)	(0.1468)	(0.0273)
ΔLiterature	0.1206*	0.0173	0.1487**	0.0176	0.2332***	0.0437**	0.2706***	0.0413*
	(0.0692)	(0.0138)	(0.0627)	(0.0134)	(0.0748)	(0.0214)	(0.0713)	(0.0223)
ΔCertification	0.0853	0.0062	0.0010	-0.0006	0.1299*	-0.0013	0.0227	-0.0054
	(0.0850)	(0.0200)	(0.0607)	(0.0185)	(0.0673)	(0.0271)	(0.0612)	(0.0235)
$Pedigree_1$			0.0985	0.0374			0.6753	0.1977
			(0.2082)	(0.0407)			(0.5931)	(0.1282)
$Exhibition_1$			0.0555	-0.0070			0.1232	-0.0019
			(0.0711)	(0.0133)			(0.2382)	(0.0408)
$Literature_1$			0.0847	-0.0010			0.1619	0.0062
			(0.0700)	(0.0166)			(0.1221)	(0.0232)
$Certification_1$			-0.1658	-0.0655*			-0.1784*	-0.1153**
			(0.1152)	(0.0390)			(0.0943)	(0.0554)
Auction House Upgrade	0.3462***	0.2931**	0.3888***	0.3886***				
	(0.1075)	(0.1220)	(0.1218)	(0.1181)				
Artist FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Month FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Auct. House Branch FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Hedonic Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
# of Obs.	6,647	6,647	6,647	6,647	4,236	4,236	4,236	4,236
R-squared	0.3361	0.3357	0.3366	0.3364	0.3919	0.3907	0.3940	0.3947

Auction European	*Auction European:* The category includes important auction houses in Europe: Lyon & Turnbull (Scotland), Francis Briest / Artcurial Briest (France), Ader, Picard & Tajan / Ader & Tajan / Tajan (France), Bruun Rasmussen (Denmark), Dorotheum (Austria), Koller (Switzerland), Lempertz (Germany), Neumeister (Germany), Finarte (Italy), Bukowskis (Sweden), and Stockholms Auktionsverk (Sweden).
Auction American	*Auction American:* The category includes important auction houses in the US: Butterfields (until 2002), Swann Auction Galleries, Skinner, Doyle New York, Freeman's, and Leslie Hindman.
Auction House Upgrade	*Auction House Upgrade* is a dummy variable equal to one when a second sale (in a repeat sale) "moves up" from a small auction house to a sub-top or large prominent auction house, or when it "moves up" from a medium-sized one to a large prominent one. Large prominent auction houses include all the branches of Christie's and Sotheby's; the sub-top auction houses include all branches of Bonhams and Phillips, or other important European and US auction hoses which we grouped as Other Important European Auction Houses, Other Important American Auction Houses (see the classification above). Small auction houses include all the unclassified auction houses.
Pedigree (Past ownership)	*Prominent Collectors*, *Royalty/Nobility*, *Wealthy Families*, *CEOs, Influential People (Time 100)*, *Celebrities*, and *Famous Sportspeople* are dummy variables equal to one if the painting has been in the collections of those respective types of collectors.
Pedigree (Descent)	*Directly from Artist*, *From Artist Family*, and *From Sitter* are dummy variables and equal one if the artworks are acquired directly from the above categories, respectively. *Other Descent Information* equals one if the artwork's pedigree information contains any descent information.
Pedigree (Past sales)	*Sold by Sotheby's or Christie's*, *Sold by Bonhams or Phillips*, *Sold by Historic Auction Houses*, *Sold by Other Important Auction Houses*, and *Sold by Prominent Dealers* are dummy variables equal to one if the painting was sold in the past at the auction house (types)/dealer, respectively.
Pedigree (Other)	*Corporate Collection* and *Private Collection (Anonymous)* are dummy variables equal to one if the artwork was at one point in corporate and private collections, respectively. *Other Pedigree Information* is a dummy variable equals to one if the painting includes pedigree information not in any of the above pedigree categories.
Exhibition	*Prominent Exhibition*, *Prominent Art Fair*, *Prominent Museum*, *Other Museum*, *Cultural City*, and *Gallery Exhibition* are dummy variables equal to one if the artwork was at one point exhibited in the above types of exhibitions / fairs / museums / cities / galleries, respectively.
Literature	*Catalogue Raisonné*, *Cover Page* (of an art history book), *Illustration* (in an art history book), and *Authoritative Press* are dummy variables equal to one if the artworks are illustrated in the above ways, respectively. *Other Literature* is a dummy variable and equals one if the artwork's catalog information refers to information in other types of publication.
Certification	*Artist (Physical)*, *Artist Family (Physical)*, (artists') *Association (Physical)*, *Expert (Physical)*, and *Other People (Physical)* are dummy variables equal to one if the artwork is auctions with physical certification issued by the above sources, respectively. *Artist (Non-Physical)*, *Artist Family (Non-Physical)*, *Association (Non-Physical)*, *Expert (Non-Physical)*, and *Other People (Non-Physical)* are dummy variables and equal one if the artworks are auctioned with non-physical certification issued by the above sources, respectively.

Appendix B. String Searches and Sources

Panel A: Pedigree

Past ownership

- **Prominent Collectors:**

  Sources: various lists from *Artnet* World's Top Art Collectors 1990–2017; *Artnet* 20 of the World's Most Innovative Art Collectors; *Forbes* Top Billionaire Art Collectors; *Grove Art Online*; art collectors from 18th century to 21st century in *Wikipedia*; 3885 names[34] in total.

- **Royalty / Nobility:**

  Sources: textual analysis by searching the royal and noble ranks in seven languages (English, Latin, Dutch, French, German, Italian, and Spanish); Including Imperial titles, High royal titles, Royal titles, Princely, ducal, and other sovereign titles, Tribal titles, Religious titles, Other sovereigns, royalty, peers, and major nobility, Minor nobility, gentry, and other aristocracy from various areas, cultures, and countries in history from *Wikipedia*[35]; 364 ranks and titles in total.

- **Wealthy Families:**

  Sources: *Forbes* World's Billionaires 1987–2017; Contemporary Wealthiest Family List from *Wikipedia*[36]; 8479 names in total.

- **CEOs:**

  Sources: various sources including *Chief Executive* CEOs of the Year 1986–2017; *Chief Executive* CEO1000 Tracker Full List; *Forbes Most Powerful People 2007–2016 (CEO, Founder, Cofounder, Chairman, Executive Vice President, Co-Chief Investment Officer, Chief Investment Officer, Director-General, etc.)*; *Forbes* America's Top 20 Favorite Bosses; *Forbes* World's 10 Most Powerful CEOs 2016; *Industry Week 10 Most Popular Manufacturing CEOs; Industry Week CEO Of the Year 2004; Industry Week CEO of the Century; Time Magazine* Person of the Year (1991, 1997, 1999, 2005, 2010); *Cable News Network (CNN)* Top 25 Influential business leaders 2005; *Ernst & Young* Entrepreneur of the Year Award; 2001–2017; *Atlantic Business* CEO of the Year 2005–2017; *The Finance Monthly* CEO Awards 2016–2017; *Harvard Business Review* Best-Performing CEOs in the World 2010–2017; *The New York Times* Equilar *200 2016*; *Barron's* World's Best CEOs 2016; *Time* Person of the Year (1991, 1997, 1999, 2005, 2010); *Fortune* 25 most powerful people in business; *Wikipedia* CEOs of notable companies[37], etc.; 2703 names in total.

- **Influential People (Time 100)**:

  Sources: *Time* 100 lists of Titans, Pioneers, Artists, Leaders, and Icons 2004–2017 and *Time* 100 Persons of the 20 Century; 3519 names in total.

- **Celebrities**:

  Sources: *IMDb* Top 1000 Actors/actresses/directors/producers; *Forbes* World's Highest-Paid Celebrities 1999–2017 (including actors, actresses, comedians, models, musicians, vocalists, directors, producers, filmmaker, TV Personalities); *Forbes* World's Highest-Paid TV Show Hosts 2016; British television personalities in *Wikipedia* [38]; American television talk show hosts in *Wikipedia*[39]; 6255 names in total.

- **Famous Sportspeople**:

  Sources: *Forbes* World's Highest-Paid Sportspeople 2012–2017 (including Boxing, Golf, Basketball, Tennis, Soccer, Football, Baseball, Racing, Motorcycle, Cricket, Track, Auto Racing, Mixed Martial Arts, etc.); World champions and superstars of sports including Golf, Basketball, Tennis, Soccer/Football, Baseball, Motorsport, Cricket, and Hockey; 4872 names in total.

  Golf: Official World Golf Ranking Top 100 2003–2017; ESPN Golf World Rankings Top 100 2017; U.S. Open champions 1895–2017; The Open Championship 1860–2017; Masters Tournament champions 1934–2017; PGA Championship 1916–2017.

    Basketball: NBA All-Stars; Hall of Fame.

    *Tenni*s: Association of Tennis Professionals Rankings Top 100; Women's Tennis Association) Rankings Top 100; Australian Open champions 1969–2017; French Open champions

---

1891–2017; Wimbledon champions 1877–2017; US Open champions 1881–2017; Grand Slam related tennis records.

Soccer: FIFA 100; *English International Football Magazine* the Greatest Players of the 20th century; *World Soccer Magazine* World Player of the Year 1982–2016; *the Guardian* Top 100 Footballers 2013–2016; *the Guardian* World Cup's Top 100 Footballers of All Time.

Baseball: Baseball Hall of Fame 1936–2017.

Motorsport: Formula One World Drivers' Champions 1950–2017; Formula Two Champions 1967–2012; 500cc/MotoGP Motorcycle World Champions 1949–2017; Motorsport Drivers Current Standings Top 32.

Cricket: ICC ODI Championship Batsmen Top 100; Historical Test Cricket Rankings 1877–2016; Current Test Rankings Top 10; Current ODI Rankings Top 10; Current T20I Rankings Top 10.

Hockey: The Hockey News Top 100 National Hockey League Players of All-Time; NHL Top 200 Rankings 2017–2018; ESPN Top 300 Fantasy Hockey Rankings 2017–2018.

- **Corporate Collection:**
  String search: corporate collection.
- **Private Collection (Anonymous):**
  String search examples: private collection.

Descent

- **Directly from Artist:**
  String search examples: from artist; from the artist; directly from artist; directly from the artist; by artist; by the artist; gift(s) (courtesy / donation(s) / goodwill(s) / bequest(s) / endowment(s) / present(s)) of (the) artist.
- **From Artist's Family:**
  String search examples: by descent (by inheritance / estate / legacy / inherited / descended / collection) from artist (the artist / by artist / by the artist / of artists / of the artist / from painter / from the painters); artist's (artist`s / artist's / artist's) + family (son / daughter / wife / husband / partner / spouse / girlfriend / boyfriend / widow / brother / sister / sibling / cousin / grandson / granddaughter / uncle / aunt / nephew / niece / heirs / heir / grandnephew / grandniece).
- **From Sitter:**
  String search examples: sitters; sitter; from sitter; from the sitter; from sitters; from the sitters.
- **Other Descent Information:**
  String search examples: descent; descended; inheritance; inherited.

Past sales

- **Sold by Sotheby's or Christie's:**
  String search examples: Christie; Sotheby.
- **Sold by Bonhams or Phillips:**
  String search examples: Bonhams; Phillips.
- **Sold by Historic Auction Houses:**
  Sources: Getty Provenance Index
  String search examples: Achenbach; Anderson & Garland; Thomas Dodd; F. Dörling; Dorotheum; Dowell's; Hôtel Drouot; Galerie Fischer; Edward Foster & Son; Messrs Foster; Frederik Muller & Co.; John Gerard; Gerard-Tasset-Juge; Gilhofer & Ranschburg; Goesin-Verhaeghe; Pierre François; Paul Graupe; Heinrich Hahn; Hugo Helbing; Galerie Helbing; Internationales Kunst Auktionshaus; George Jones; Albert Kende; S. Kende; Thomas King; August Klipstein; Galerie Kornfeld; Knight Frank & Rutley; W. S. Kündig; Hans W. Lange; Langford; Mathias Lempertz; Heinrich Lempertz; Gallery Lempertz Contempora; Venator & Hanstein; Kunsthaus Lempertz; Leo Spik; Rudolph Lepke; Bignell Marle; P. L. Mastraeten; Franz A. Menna; Corneille Moor; Morrison Mcchlery; Max Perl; Thomas Philipe; Harry Phillips; Mr. Prestage; Puttick & Simpson; William Richardson; George Henry Robins; Henry J. Robins; Robinson & Foster; Robert Saunders; Hodgson & Co; Saunders & Hodgson; Philippus Van Der Schley; James Webber Southgate; George Squibb; Squibb & Son; Rushworth, Abbott & Co; George Stanley; J. A. Stargardt; William Stewart; E. J. Terlinck; ; De Vries; Adolf Weinmüller; Munich Auction House; Benjamin Wheatley; Willis's Rooms; Winstanley & Sons; Puttick & Simpson; Stewart, Wheatley & Adlard; Wheatley & Adlard.
- **Sold by Other Important Auction Houses:**

String search examples: Butterfields; Lyon & Turnbull; Francis Briest; Artcurial Briest; Tajan; Bruun Rasmussen; Dorotheum; Koller; Lempertz; Neumeister; Finarte; Bukowskis; Stockholms Auktionsverk; Swann Auction Galleries; Swann Galleries; Skinner; Doyle New York; Freeman's; Freeman`s; Freeman's; Leslie Hindman.

- **Sold by Prominent Dealers:**
Sources: *Artnet* 100 best galleries, *Forbes* contemporary dealers, and *Grove Art Online* famous historic dealers; 233 names in total.

Panel B: Exhibition

- **Prominent Exhibition:**
String search examples: retrospective; rétrospective; anniversary; anniversaire; biennale; triennale; biannual; biennial; triannual; triennial.
- **Prominent Art Fair:**
Sources: *Artnet*; *Artprice*
String search examples: ARCO Madrid; Armory Show New York; Art Basel; Art Basel HK; Art Basel Miami Beach; Art Cologne; Art Miami; Art Santa Fe; ARTISSIMA; Documenta Kassel; FIAC Paris; Frieze London; Frieze New York; India Art Fair; PAN Amsterdam; TEFAF Maastricht; TEFAF New York; Venice Biennale; BRAFA Brussels.
- **Prominent Museum:**
Sources: most important museums of paintings in important art cities from *National Geographic*, *Wikipedia, Reuters, and The Telegraph*; 517 museums in total.
- **Other Museum:**
String search examples: museum; musée; museo; museu; museums; musea; museen; musées; museos; museus; musei.
- **Cultural City:**
Sources: European Capital of Culture[40]; UN City 2016; City Mayors EU 500; City Mayors World 300 and other cultural cities (defined by locations with considerable amount of museums, galleries and auction houses) around the world; City names in English, French and original languages; 236 cities in total.

Panel C: Literature

- **Catalogue Raisonné:**
String search examples: catalog(ue)/catalog(ue) raisonne; catalogue/catalogue raisonné
- **Cover Page:**
String search examples: cover.
- **Illustration:**
String search examples: illustration; illustrated; cover; images; image; photos; photo.
- **Authoritative Press:**
Sources: 280 notable university presses from *Wikipedia*[41] and World's 57 largest book publishers from *Publishers Weekly Magazine*.

Panel D: Certification

- **Certification:**
String search examples: echtheitsbestätigung; gutachten; essay(s); assessment(s); opinion(s); appraisal(s); expert(s); expertise(s); report(s); mail(s); photo-certificate(s); photocopy; photocopies; issued; verified; witnessed; authenticity; authentication.
- **Forms-physical:**
String search examples: photocertificate(s); report; written; handwritten; photocopy; photocopies; photo(s); photograaphy; photographic; photograph; foto(s); foto's; photography; fotografische; fotografie; fotografie; fotografie; fotografien; photography; photographique; photographie(s).
- **Issuers-artist:**
String search examples: issued (verified / witnessed / certificates / certificate / certificate + signed / certified / authenticity / authenticity signed / authentication / authentication signed / authenticated / identified / identification / confirmed / confirmation / confirmatory information / registered /

---

[40] See https://en.wikipedia.org/wiki/European_Capital_of_Culture
[41] See https://en.wikipedia.org/wiki/List_of_university_presses

registration / registration card / registered / recorded / documentation / letter(s) / photo(s) / photo(s) signed / photograph(s) / photograph(s) signed) + by artist (by the artist / from artist / from the artist / of artist / of the artist).

- **Issuers-artist's family:**
  String search examples: son; daughter; wife; husband; partner; spouse; girlfriend; boyfriend; widow; brother; sister; sibling; cousin; grandson; granddaughter; uncle; aunt; nephew; niece; family; descendants; descendant; biographer; pupils; pupil; students; student.

- **Issuers-association:**
  String search examples: authentication; board; estate; foundation(s); fundament; stiftung; fondation; fundación; fundação; fondazione; association; vereniging; verband; asociación; associação; associazione; committee; commissie; ausschuss.

- **Issuers-expert:**
  String search examples: Dr; Prof; curator(s); custodian(s); professor(s); doctor(s); director(s); expert(s); expertise(s); professoren; professore; professoressa; professeur(s); professore(s); professori; profesor.

Appendix C. Impact on Probability of Being Sold after Discoveries of Fakes

This table presents the DiD estimators for the discoveries of fakes cases. The dependent variable is the outcome of the sale (sold or unsold). *Treated* equals one if the auction date falls after the date of discovery of fakes for artists whose paintings have been forged. *Provenance* (*Pedigree*, *Exhibition*, *Literature* and *Certification*) is the dummy variable equaling 1 if the paintings contain any of the above types of provenance information. As controls, all hedonic variables discussed in Section 1 are included. In Columns (1)–(5), the interaction terms of *Treated* with each provenance dimension (*Provenance*, *Pedigree*, *Exhibition*, *Literature* and *Certification*) are presented separately, while in Column (6), all these interactions are combined. *, **, and *** indicate statistical significance at the 10%, 5%, and 1% levels, respectively. Standard errors (S.E.) are reported in parentheses and clustered at the auction branch level.

	(1)	(2)	(3)	(4)	(5)	(6)
Dept. Var.:	Sold[0,1]	Sold[0,1]	Sold[0,1]	Sold[0,1]	Sold[0,1]	Sold[0,1]
Treated	-0.0319	-0.0247	-0.0094	-0.0204	0.0004	-0.0233
	(0.0262)	(0.0248)	(0.0205)	(0.0208)	(0.0191)	(0.0249)
Treated × Provenance	0.0514**					
	(0.0234)					
Treated × Pedigree		0.0466**				0.0089
		(0.0212)				(0.0219)
Treated × Exibition			0.0364**			0.0017
			(0.0145)			(0.0120)
Treated × Literature				0.0666***		0.0617***
				(0.0182)		(0.0187)
Treated ×Certification					-0.0214	-0.0021
					(0.0331)	(0.0358)
Provenance	0.0235***					
	(0.0055)					
Pedigree		0.0242***				0.0172***
		(0.0063)				(0.0062)
Exihibition			0.0496***			0.0379***
			(0.0053)			(0.0041)
Literature				0.0402***		0.0236***
				(0.0065)		(0.0062)
Certification					0.0085	0.0104
					(0.0135)	(0.0134)
Ln(Low Price Estimate)	-0.0650***	-0.0647***	-0.0652***	-0.0652***	-0.0641***	-0.0662***
	(0.0036)	(0.0036)	(0.0036)	(0.0036)	(0.0035)	(0.0036)
Artist FE	Yes	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes	Yes
Month FE	Yes	Yes	Yes	Yes	Yes	Yes
Hedonic Controls	Yes	Yes	Yes	Yes	Yes	Yes
Auct. House Branch FE	Yes	Yes	Yes	Yes	Yes	Yes
# of Obs.	1,707,136	1,707,136	1,707,136	1,707,136	1,707,136	1,707,136
R-squared	0.1747	0.1747	0.1748	0.1747	0.1745	0.1750

Appendix D. Hedonic Pricing Regressions for Subsamples of Schools and Movements

We match and classify artists into 13 major movements in art history based on the biography information from Oxford Art Online: (1) *Medieval and Renaissance*, (2) *Baroque*, (3) *Rococo*, (4) *Neoclassicism*, (5) *Romanticism*, (6) *Realism*, (7) *Impressionism and Symbolism*, (8) *Fauvism and Expressionism*, (9) *Cubism, Futurism, and Constructivism*, (10) *Dada and Surrealism*, (11) *Abstract Expressionism*, (12) *Pop Art*, and (13) *Minimalism and Contemporary.* The dependent variable is *Ln(Price)*, the natural logarithm of the real hammer price in real (2007) USD. In Panel A, *Provenance* is a dummy variable capturing if the artwork's catalogue comprises any provenance information. In Panel B, *Pedigree*, *Exhibition*, *Literature*, and *Certification* are dummy variables if the artworks have any information on this provenance dimension, respectively. All regressions include Hedonic Controls (detailed in Online Appendix Table A.I), and Artist, Year, Month, and Auction House Branch Level (AH) Fixed Effects. *, **, and *** indicate statistical significance at the 10%, 5%, and 1% levels, respectively. Standard errors (S.E.) are reported in parentheses.

Panel A: Hedonic Pricing Regressions with Provenance Dummy for Subsamples of Schools and Movements

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)
Sample:	School 1	School 2	School 3	School 4	School 5	School 6	School 7	School 8	School 9	School 10	School 11	School 12	School 13
Dept. Var.:	Ln(Price)	Ln(Price)	Ln(Price)	Ln(Price)	Ln(Price)	Ln(Price)	Ln(Price)	Ln(Price)	Ln(Price)	Ln(Price)	Ln(Price)	Ln(Price)	Ln(Price)
Provenance	0.2751***	0.3505***	0.3733***	0.1849***	0.3923***	0.3569***	0.3401***	0.2460***	0.2212***	0.2123***	0.2055***	0.6051***	0.2417***
	(0.0487)	(0.0319)	(0.0532)	(0.0573)	(0.0439)	(0.0514)	(0.0304)	(0.0341)	(0.0389)	(0.0537)	(0.0345)	(0.1296)	(0.0596)
Artist FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Month FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Auct. House Branch FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Hedonic Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
# of Obs.	4,297	17,290	4,463	2,502	8,762	9,626	22,181	17,180	10,959	13,761	15,538	8,691	12,739
R-squared	0.6733	0.6426	0.6548	0.6453	0.6935	0.7149	0.7756	0.7765	0.7772	0.7663	0.7575	0.7628	0.7668

Panel B: Hedonic Pricing Regressions with Four Provenance Dimensions for Subsamples of Schools and Movements

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)
Sample:	School 1	School 2	School 3	School 4	School 5	School 6	School 7	School 8	School 9	School 10	School 11	School 12	School 13
Dept. Var.:	Ln(Price)	Ln(Price)	Ln(Price)	Ln(Price)	Ln(Price)	Ln(Price)	Ln(Price)	Ln(Price)	Ln(Price)	Ln(Price)	Ln(Price)	Ln(Price)	Ln(Price)
Pedigree	0.1439***	0.1672***	0.1243**	0.0291	0.2457***	0.1639***	0.1801***	0.1004**	0.1039***	0.1188*	0.1534***	0.5391***	0.1180
	(0.0434)	(0.0281)	(0.0523)	(0.0598)	(0.0535)	(0.0475)	(0.0322)	(0.0393)	(0.0349)	(0.0648)	(0.0343)	(0.1171)	(0.0726)
Exhibition	0.3504***	0.4265***	0.4145***	0.1884*	0.3433***	0.3707***	0.4456***	0.3088***	0.3040***	0.3771***	0.1927***	0.4790***	0.3483***
	(0.0811)	(0.0804)	(0.0874)	(0.1017)	(0.0851)	(0.0554)	(0.0304)	(0.0381)	(0.0420)	(0.0397)	(0.0299)	(0.0480)	(0.0205)
Literature	0.4245***	0.3719***	0.4613***	0.4719***	0.3782***	0.4247***	0.2947***	0.3124***	0.4161***	0.4683***	0.5163***	0.4366***	0.4414***
	(0.0777)	(0.0380)	(0.0683)	(0.1126)	(0.0884)	(0.0438)	(0.0404)	(0.0432)	(0.0623)	(0.0530)	(0.0401)	(0.0729)	(0.0842)
Certification	-0.2583*	0.2022***	0.2767**	-0.0652	0.4481***	0.2122***	0.1423***	0.1311***	0.1882***	0.0991**	-0.0065	0.1042*	0.2566***
	(0.1363)	(0.0494)	(0.1332)	(0.1986)	(0.0898)	(0.0719)	(0.0417)	(0.0411)	(0.0505)	(0.0455)	(0.0446)	(0.0629)	(0.0647)
Artist FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Month FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Auct. House Branch FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Hedonic Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
# of Obs.	4,297	17,290	4,463	2,502	8,762	9,626	22,181	17,180	10,959	13,761	15,538	8,691	12,739
R-squared	0.6834	0.6513	0.6651	0.6508	0.6998	0.7223	0.7820	0.7807	0.7836	0.7767	0.7663	0.7757	0.7754

In Art We Trust

Online Appendix

Table A.I Descriptive Statistics of Hedonic Variables

This table presents the descriptive statistics of the hedonic variables. *Deceased* equals one in case the artist is dead at the time of the sale. The attribution dummies *Attributed*, *Studio*, *Circle*, *School*, *After*, and *Style* equal one if the auction catalog identifies the work as "attributed to" the artist, from his "studio," from his "circle," from the artist's "school," "after" the artist, or "in the style of" the artist, respectively. The dummies *Signed*, *Dated*, and *Inscribed* take the value one if the work carries a signature, is dated, or is inscribed, respectively. The medium dummies *Oil*, *Watercolor*, and *Drawing* indicate whether the work is an oil painting, a watercolor, or a drawing. The variables *Height* and *Width* measure the height and the width of the work in centimeters. The month dummies indicate the sales month (January is the benchmark). The auction house branch dummies equal one in case of a sale at, respectively, *Sotheby's London*, *Sotheby's New York*, *Sotheby's Other Branches*, *Christie's London*, *Christie's New York, Christie's Other Branches*, *Bonhams London*, *Bonhams Other Branches*, *Phillips London*, and *Phillips Other Branches*. *Auction European* and *Auction American* are dummy variables equal to one if the sale takes place at a large Continental European or a large American auction house, respectively (see Appendix A.). *Pedigree*, *Exhibition*, *Literature*, and *Certification* equal one if the artwork has any information on these dimensions, respectively. *Pedigree (Text Length)*, *Exhibition (Text Length)*, *Literature (Text Length)* and *Certification (Text Length)* measure the number of characters of the text for these dimensions. *Exhibition (Number Count)* and *Literature (Number Count)* count the number of exhibitions and the number of the cases for which information on a painting is given in the literature. For each variable, we report the number of observations (N), the mean, the standard deviation (S.D.), the minimum, and the maximum.

	N	Mean	S.D.	Min	Max
Artist Characteristics					
Deceased	1,812,807	76.68%	42.28%	0	1
Artwork Characteristics					
Attribution Dummies					
Attributed	1,812,807	2.75%	16.35%	0	1
Studio	1,812,807	0.12%	3.44%	0	1
Circle	1,812,807	0.25%	4.95%	0	1
School	1,812,807	0.02%	1.42%	0	1
After	1,812,807	0.20%	4.52%	0	1
Style	1,812,807	0.04%	2.10%	0	1
Signature Dummies					
Signed	1,812,807	80.41%	39.69%	0	1
Dated	1,812,807	36.28%	48.08%	0	1
Inscribed	1,812,807	11.38%	31.76%	0	1
Medium Dummies					
Oil	1,812,807	68.13%	46.60%	0	1
Watercolor	1,812,807	20.48%	40.35%	0	1
Drawing	1,812,807	11.39%	31.77%	0	1
Measurement Variables					
Height (cm)	1,812,807	59.46	51.14	11.00	200.00
Width (cm)	1,806,082	61.33	73.86	11.40	215.00
Topic Dummies					
Abstract	1,812,807	2.33%	15.08%	0	1
Animals	1,812,807	4.74%	21.25%	0	1
Landscape	1,812,807	14.92%	35.63%	0	1
Seascape	1,812,807	4.26%	20.18%	0	1
Cityscape	1,812,807	8.77%	28.28%	0	1
Nude	1,812,807	1.37%	11.61%	0	1
People	1,812,807	8.41%	27.76%	0	1
Self Portrait	1,812,807	0.33%	5.72%	0	1
Portrait	1,812,807	3.56%	18.53%	0	1
Religion	1,812,807	1.99%	13.96%	0	1
Still Life	1,812,807	5.84%	23.45%	0	1
Study	1,812,807	1.28%	11.26%	0	1

	N	Mean	S.D.	Min	Max
Untitled	1,812,807	6.44%	24.54%	0	1

Transaction Characteristics

Month

	N	Mean	S.D.	Min	Max
January	1,812,807	3.77%	19.05%	0	1
February	1,812,807	5.16%	22.12%	0	1
March	1,812,807	8.69%	28.17%	0	1
April	1,812,807	7.74%	26.73%	0	1
May	1,812,807	11.83%	32.30%	0	1
June	1,812,807	13.01%	33.64%	0	1
July	1,812,807	4.24%	20.15%	0	1
August	1,812,807	2.21%	14.70%	0	1
September	1,812,807	6.76%	25.11%	0	1
October	1,812,807	9.77%	29.68%	0	1
November	1,812,807	14.56%	35.27%	0	1
December	1,812,807	12.26%	32.80%	0	1

Auction House

	N	Mean	S.D.	Min	Max
Sotheby's London	1,812,807	2.24%	14.81%	0	1
Sotheby's New York	1,812,807	2.36%	15.18%	0	1
Sotheby's Other Branches	1,812,807	2.03%	14.10%	0	1
Christie's London	1,812,807	1.53%	12.27%	0	1
Christie's New York	1,812,807	2.36%	15.19%	0	1
Christie's Other Branches	1,812,807	5.01%	21.82%	0	1
Bonhams London	1,812,807	1.03%	10.11%	0	1
Bonhams Other Branches	1,812,807	4.38%	20.47%	0	1
Phillips London	1,812,807	0.26%	5.14%	0	1
Phillips Other Branches	1,812,807	0.36%	6.00%	0	1
Auction American	1,812,807	3.16%	17.49%	0	1
Auction European	1,812,807	11.82%	32.28%	0	1

Provenance Information

	N	Mean	S.D.	Min	Max
Pedigree	1,812,807	14.20%	34.90%	0	1
Exhibition	1,812,807	3.74%	19.00%	0	1
Literature	1,812,807	4.02%	19.60%	0	1
Certification	1,812,807	3.89%	19.30%	0	1

Hammer Price

	N	Mean	S.D.	Median
Nominal in USD	1,165,829	53142	638180	3400
Real in USD	1,165,829	48470	575599	3105

Table A.II Provenance Effects on Probability of Being Sold

This table presents the baseline hedonic linear probability regression results (Models (1) and (2)). Models (3) and (4) is estimated by means of OLS. The dependent variable is the sale results (sold / unsold), the independent variables are given in Equation (1). In Column (1), *Provenance* equals one if an artwork has any provenance information (pedigree, exhibition, literature, or certification). In Column (2), *Pedigree*, *Exhibition*, *Literature*, and *Certification* are dummy variables capturing whether the artwork is accompanied by any such information. Column (3) uses the natural log of text length related to the information on pedigree, exhibition, literature, and certification information. Column (4) uses the count variables for exhibition and literature, and dummy variables for pedigree and certification. *, **, and *** indicate statistical significance at the 10%, 5%, and 1% levels, respectively. Standard errors (S.E.) are reported in parentheses and clustered at the auction branch level.

Dept. Var.:	Sold[0,1]			
	(1)	(2)	(3)	(4)
Provenance as:	Indicator	Indicator	Text Length	Number Count
Provenance Information				
Provenance	0.0239***			
	(0.0055)			
Pedigree		0.0173***	0.0044***	0.0211***
		(0.0062)	(0.0014)	(0.0062)
Exhibition		0.0379***	0.0079***	0.0098***
		(0.0041)	(0.0008)	(0.0009)
Literature		0.0250***	0.0056***	0.0092***
		(0.0063)	(0.0011)	(0.0013)
Certification		0.0102	0.0115***	0.0104
		(0.0133)	(0.0033)	(0.0133)
Artist Characteristics				
Deceased	0.0262***	0.0265***	0.0267***	0.0266***
	(0.0059)	(0.0059)	(0.0059)	(0.0059)
Artwork Characteristics				
Attribution				
Attributed	-0.0991***	-0.0991***	-0.0988***	-0.0995***
	(0.0073)	(0.0074)	(0.0074)	(0.0073)
Studio	-0.1044***	-0.1020***	-0.1015***	-0.1019***
	(0.0117)	(0.0116)	(0.0115)	(0.0116)
Circle	-0.1683***	-0.1646***	-0.1633***	-0.1649***
	(0.0145)	(0.0149)	(0.0150)	(0.0148)
School	-0.1631***	-0.1603***	-0.1588***	-0.1586***
	(0.0369)	(0.0372)	(0.0373)	(0.0373)
After	-0.2075***	-0.2033***	-0.2028***	-0.2028***
	(0.0169)	(0.0167)	(0.0165)	(0.0167)
Style	-0.1443***	-0.1411***	-0.1394***	-0.1408***
	(0.0226)	(0.0229)	(0.0230)	(0.0228)
Signature				
Signed	0.0311***	0.0314***	0.0313***	0.0313***
	(0.0067)	(0.0067)	(0.0066)	(0.0067)
Dated	0.0285***	0.0286***	0.0284***	0.0286***
	(0.0029)	(0.0029)	(0.0028)	(0.0029)
Inscribed	0.0125***	0.0127***	0.0121***	0.0128***
	(0.0047)	(0.0046)	(0.0046)	(0.0046)
Medium				
Oil	0.1406***	0.1404***	0.1411***	0.1404***
	(0.0079)	(0.0080)	(0.0079)	(0.0079)
Watercolor	0.0567***	0.0565***	0.0569***	0.0564***
	(0.0058)	(0.0058)	(0.0057)	(0.0057)
Measurements				
Height	0.0003***	0.0003***	0.0003***	0.0003***
	(0.0001)	(0.0001)	(0.0001)	(0.0001)
Width	0.0003***	0.0002***	0.0003***	0.0003***
	(0.0001)	(0.0001)	(0.0001)	(0.0001)

Dept. Var.:	Sold[0,1]			
	(1)	(2)	(3)	(4)
Provenance as:	Indicator	Indicator	Text Length	Number Count
Height_2	-0.0001***	-0.0001***	-0.0001***	-0.0001***
	(0.0001)	(0.0001)	(0.0001)	(0.0001)
Width_2	-0.0001***	-0.0001***	-0.0001***	-0.0001***
	(0.0001)	(0.0001)	(0.0001)	(0.0001)
Topic				
Abstract	0.0014	0.0013	0.0014	0.0010
	(0.0080)	(0.0081)	(0.0081)	(0.0081)
Animals	0.0098*	0.0096	0.0095	0.0097
	(0.0059)	(0.0059)	(0.0060)	(0.0059)
Landscape	0.0098	0.0096	0.0096	0.0098
	(0.0079)	(0.0080)	(0.0081)	(0.0080)
Seascape	0.0313***	0.0311***	0.0310***	0.0313***
	(0.0057)	(0.0058)	(0.0058)	(0.0057)
Urbanscape	0.0295***	0.0292***	0.0291***	0.0294***
	(0.0051)	(0.0051)	(0.0052)	(0.0051)
Nude	0.0133*	0.0133*	0.0132*	0.0133*
	(0.0068)	(0.0069)	(0.0070)	(0.0069)
People	0.0038	0.0035	0.0035	0.0037
	(0.0056)	(0.0056)	(0.0057)	(0.0056)
Self Portrait	0.0397***	0.0384***	0.0380***	0.0389***
	(0.0124)	(0.0125)	(0.0125)	(0.0125)
Portrait	-0.0318***	-0.0320***	-0.0321***	-0.0320***
	(0.0054)	(0.0055)	(0.0055)	(0.0055)
Religion	0.0072	0.0069	0.0065	0.0071
	(0.0059)	(0.0060)	(0.0060)	(0.0060)
Still Life	0.0092	0.0090	0.0090	0.0092
	(0.0073)	(0.0074)	(0.0075)	(0.0074)
Study	-0.0005	-0.0002	-0.0006	-0.0001
	(0.0058)	(0.0058)	(0.0059)	(0.0058)
Other Topic	0.0214***	0.0209***	0.0208***	0.0212***
	(0.0073)	(0.0074)	(0.0074)	(0.0073)
Transaction Characteristics				
Auction House				
Sotheby's London	0.1139***	0.1087***	0.1076***	0.1105***
	(0.0119)	(0.0112)	(0.0111)	(0.0113)
Sotheby's New York	0.1215***	0.1175***	0.1162***	0.1184***
	(0.0119)	(0.0114)	(0.0114)	(0.0114)
Sotheby's Other Branches	0.1099***	0.1069***	0.1075***	0.1087***
	(0.0202)	(0.0197)	(0.0196)	(0.0197)
Christie's London	0.1486***	0.1384***	0.1366***	0.1408***
	(0.0127)	(0.0119)	(0.0118)	(0.0120)
Christie's New York	0.1735***	0.1665***	0.1655***	0.1677***
	(0.0112)	(0.0107)	(0.0106)	(0.0107)
Christie's Other Branches	0.1214***	0.1182***	0.1178***	0.1201***
	(0.0177)	(0.0168)	(0.0167)	(0.0171)
Bonhams London	-0.0129	-0.0149	-0.0155	-0.0139
	(0.0108)	(0.0106)	(0.0106)	(0.0106)
Bonhams Other Branches	-0.0567***	-0.0579***	-0.0585***	-0.0576***
	(0.0139)	(0.0137)	(0.0137)	(0.0138)
Phillips London	0.0660***	0.0599***	0.0605***	0.0618***
	(0.0172)	(0.0170)	(0.0169)	(0.0170)
Phillips Other Branches	0.0799***	0.0732***	0.0736***	0.0754***
	(0.0153)	(0.0145)	(0.0143)	(0.0146)
Auction American	0.0861***	0.0858***	0.0859***	0.0856***
	(0.0112)	(0.0111)	(0.0111)	(0.0111)

Dept. Var.:	Sold[0,1]			
	(1)	(2)	(3)	(4)
Provenance as:	Indicator	Indicator	Text Length	Number Count
Auction European	-0.0130	-0.0127	-0.0126	-0.0126
	(0.0185)	(0.0186)	(0.0186)	(0.0186)
Ln(Low Price Estimate)	-0.0649***	-0.0661***	-0.0668***	-0.0662***
	(0.0036)	(0.0036)	(0.0036)	(0.0035)
Artist FE	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes
Month FE	Yes	Yes	Yes	Yes
# of Obs.	1,707,136	1,707,136	1,707,136	1,707,136
R-squared	0.1747	0.1750	0.1752	0.1750

Table A.III Provenance Effects on Hammer Price

This table presents the baseline hedonic price regression results (Equation (2)). The dependent variable is the natural log of deflated hammer price in USD. In Column (1), provenance variables are dummy variables capturing whether or not the artworks include any information on *Pedigree*, *Exhibition*, *Literature*, and *Certification*, respectively. Column (2) includes the natural log of text length of information on *Pedigree*, *Exhibition*, *Literature*, and *Certification*. Column (3) includes count variables for exhibition and literature, and dummy variables for pedigree and certification. *, **, and *** indicate statistical significance at the 10%, 5%, and 1% levels, respectively. Standard errors (S.E.) are reported in parentheses and clustered at the auction branch level.

Dept. Var:	Ln(Price)		
	(1)	(2)	(3)
Provenance as:	Indicator	Text Length	Number Count
Provenance Information			
Pedigree	0.1885***	0.0531***	0.2379***
	(0.0170)	(0.0045)	(0.0194)
Exhibition	0.3499***	0.0734***	0.1043***
	(0.0208)	(0.0041)	(0.0066)
Literature	0.4288***	0.0869***	0.1266***
	(0.0336)	(0.0063)	(0.0054)
Certification	0.1297***	0.0614***	0.1305***
	(0.0226)	(0.0100)	(0.0220)
Artist Characteristics			
Deceased	0.1900***	0.1916***	0.1912***
	(0.0211)	(0.0211)	(0.0210)
Artwork Characteristics			
Attribution			
Attributed	-0.8278***	-0.8207***	-0.8339***
	(0.0360)	(0.0357)	(0.0362)
Studio	-0.7771***	-0.7638***	-0.7775***
	(0.0930)	(0.0893)	(0.0960)
Circle	-0.9265***	-0.9042***	-0.9337***
	(0.0980)	(0.0928)	(0.0976)
School	-0.9362***	-0.9096***	-0.9087***
	(0.1010)	(0.0964)	(0.0977)
After	-1.5346***	-1.5103***	-1.5307***
	(0.1212)	(0.1136)	(0.1217)
Style	-1.3193***	-1.2851***	-1.3213***
	(0.1053)	(0.0994)	(0.1044)
Signature			
Signed	0.0746**	0.0746**	0.0737**
	(0.0309)	(0.0302)	(0.0311)
Dated	0.1602***	0.1584***	0.1588***
	(0.0133)	(0.0131)	(0.0134)
Inscribed	0.0337***	0.0307***	0.0351***
	(0.0111)	(0.0112)	(0.0112)
Medium			
Oil	1.3427***	1.3368***	1.3422***
	(0.0296)	(0.0288)	(0.0292)
Watercolor	0.5227***	0.5205***	0.5207***
	(0.0310)	(0.0305)	(0.0307)
Measurements			
Height	0.0056***	0.0056***	0.0056***
	(0.0003)	(0.0003)	(0.0003)
Width	0.0049***	0.0049***	0.0049***
	(0.0003)	(0.0003)	(0.0003)
Height_2	-0.0001***	-0.0001***	-0.0001***
	(0.0001)	(0.0001)	(0.0001)
Width_2	-0.0001***	-0.0001***	-0.0001***
	(0.0001)	(0.0001)	(0.0001)

Dept. Var:	Ln(Price)		
	(1)	(2)	(3)
Provenance as:	Indicator	Text Length	Number Count
Topic			
Abstract	0.0341	0.0334	0.0302
	(0.0334)	(0.0338)	(0.0334)
Animals	0.0303	0.0298	0.0325
	(0.0350)	(0.0354)	(0.0348)
Landscape	0.0805	0.0800	0.0825
	(0.0534)	(0.0538)	(0.0532)
Seascape	0.1165***	0.1159***	0.1194***
	(0.0326)	(0.0328)	(0.0324)
Urbanscape	0.1679***	0.1667***	0.1701***
	(0.0288)	(0.0291)	(0.0287)
Nude	0.0086	0.0094	0.0102
	(0.0340)	(0.0343)	(0.0338)
People	0.0489*	0.0482	0.0510*
	(0.0292)	(0.0295)	(0.0291)
Self Portrait	0.2407***	0.2353***	0.2494***
	(0.0433)	(0.0436)	(0.0438)
Portrait	-0.1370***	-0.1366***	-0.1349***
	(0.0278)	(0.0280)	(0.0276)
Religion	0.0825***	0.0801***	0.0856***
	(0.0306)	(0.0309)	(0.0306)
Still Life	0.0847**	0.0838*	0.0875**
	(0.0431)	(0.0434)	(0.0429)
Study	-0.1375***	-0.1374***	-0.1354***
	(0.0262)	(0.0264)	(0.0262)
Other Topic	0.1497***	0.1486***	0.1535***
	(0.0453)	(0.0458)	(0.0451)
Transaction Characteristics			
Auction House			
Sotheby's London	0.9789***	0.9438***	1.0046***
	(0.0389)	(0.0398)	(0.0367)
Sotheby's New York	0.7814***	0.7450***	0.7952***
	(0.0459)	(0.0466)	(0.0440)
Sotheby's Other Branches	0.5620***	0.5517***	0.5885***
	(0.0486)	(0.0468)	(0.0488)
Christie's London	0.9361***	0.8892***	0.9625***
	(0.0421)	(0.0431)	(0.0379)
Christie's New York	0.4944***	0.4579***	0.5081***
	(0.0358)	(0.0372)	(0.0337)
Christie's Other Branches	0.3139***	0.2973***	0.3392***
	(0.0383)	(0.0365)	(0.0420)
Bonhams London	0.6413***	0.6224***	0.6517***
	(0.0276)	(0.0277)	(0.0271)
Bonhams Other Branches	0.0447	0.0338	0.0489
	(0.0627)	(0.0625)	(0.0634)
Phillips London	0.4901***	0.4865***	0.5115***
	(0.0757)	(0.0757)	(0.0767)
Phillips Other Branches	0.4337***	0.4230***	0.4611***
	(0.1366)	(0.1323)	(0.1368)
Auction American	-0.0966**	-0.0968**	-0.0987**
	(0.0388)	(0.0391)	(0.0387)
Auction European	0.2168***	0.2161***	0.2170***
	(0.0484)	(0.0483)	(0.0484)
Artist FE	Yes	Yes	Yes

Dept. Var:	Ln(Price)		
	(1)	(2)	(3)
Provenance as:	Indicator	Text Length	Number Count
Year FE	Yes	Yes	Yes
Month FE	Yes	Yes	Yes
# of Obs.	1,111,220	1,111,220	1,111,220
R-squared	0.7805	0.7817	0.7807

Table A.IV Fake and Forgery Cases

This table presents the cases of discoveries of fakes cases. We collect the disclosure dates of fakes and forgery cases from these data sources: the specialized art journals Artsjournal.com and Artnews.com, and the general news database Factiva. In each article on the topic, we collect the event date of the discovery, title of the fake or forged painting, name of the artist, name of the forger (if available), etc. We take fraud discovery date as the first date that the rumor, motivated suspicion, proof, or ruling by a judge was mentioned in the press. We only retain the cases that were ultimately confirmed as fakes or forgeries. We identify 54 fraud cases related to paintings in our sample period of 2007 to 2016.

Artist attributed	Forger	Discovery date	Fake*/Forged** painting
Gottfried Lindauer	Anonymous	10/2015	Hamiora Maioha**
Francisco de Goya	Anonymous	02/2015	Retrato de don Antonio Mará Esquivel (Portrait of Antonio Mará Esquivel)**
Emile-Othon Friesz	Wolfgang Beltracchi	12/2014	La Ciotat**
Albert Tucker	Anonymous	07/2014	Faun and Parrot**
Marc Chagall	Anonymous	02/2014	Nude**
Huang Yongyu	Anonymous	10/2013	Flower-and-Bird*
Clyfford Still	Pei-Shen Qian	12/2011	Untitled (1949)**
Mark Rothko	Pei-Shen Qian	12/2011	-**
Willem de Kooning	Pei-Shen Qian	12/2011	-**
Barnett Newman	Pei-Shen Qian	12/2011	-**
Franz Kline	Pei-Shen Qian	12/2011	-**
Jackson Pollock	Pei-Shen Qian	08/2011	Untitled (1950)**
Richard Diebenkorn	Pei-Shen Qian	08/2011	-**
Robert Motherwell8,	Pei-Shen Qian	08/2011	Spanish Elegy**
Louis Marcoussis	Wolfgang Beltracchi	11/2010	Kleines kubistiches Stilleben (Small Cubiste Still-life), Paris**
Max Pechstein	Wolfgang Beltracchi	11/2010	Seine Paris**
Pablo Picasso	Wolfgang Beltracchi	11/2010	-**
Heinrich Campendonk	Wolfgang Beltracchi	10/2010	Rotes Bild mit Pferden (Red painting with Horses)*
Heinrich Campendonk	Wolfgang Beltracchi	10/2010	Madchen mit Schwan (Girl with a Swan)**
André Derain	Wolfgang Beltracchi	10/2010	Collioure**
Heinrich Campendonk	Wolfgang Beltracchi	10/2010	Kleines Bild mit Pferden (Landscape with Horses)**
Max Ernst	Wolfgang Beltracchi	10/2010	Kleine Weiße Landschaft (Small White Landscape)**
Fernand Léger	Wolfgang Beltracchi	10/2010	Kubistisches Stilleben**
Max Pechstein	Wolfgang Beltracchi	10/2010	Frauenakt**

Artist attributed	Forger	Discovery date	Fake*/Forged** painting
Max Ernst	Wolfgang Beltracchi	10/2010	La Horde*
Louis Marcoussis	Wolfgang Beltracchi	10/2010	Portrait d'Alfred Fletchtheim**
André Derain	Wolfgang Beltracchi	10/2010	Boote in Collioure (Boats in Collioure)*
Max Ernst	Wolfgang Beltracchi	10/2010	La Mer**
Heinrich Nauen	Wolfgang Beltracchi	10/2010	Herbstwald (Autumn Forest)**
Heinrich Campendonk	Wolfgang Beltracchi	10/2010	Zwei figuren in Landschaft (Two Figures in a Landscape)**
Georges Braque	Wolfgang Beltracchi	10/2010	Guitare le compotier (Le Journal)**
Max Ernst	Wolfgang Beltracchi	10/2010	Waldbild (Forest Painting)**
Carlo Mense	Wolfgang Beltracchi	10/2010	Kleine Ansicht Waidmarkt (Small view of the Waidmarkt)**
Hans Purrmann	Wolfgang Beltracchi	10/2010	Sudliche Landschafte (Southern Landscape)**
Raoul Dufy	Wolfgang Beltracchi	10/2010	Le Havre**
Max Ernst	Wolfgang Beltracchi	10/2010	Tremblement de Terre**
Heinrich Campendonk	Wolfgang Beltracchi	10/2010	Else Laske-Schule Gewidment (Dedicated to Else Laske-Schule)**
Max Ernst	Wolfgang Beltracchi	10/2010	Oiseaux**
Max Ernst	Wolfgang Beltracchi	10/2010	La Foret**
Max Ernst	Wolfgang Beltracchi	10/2010	Vogel in Winterlandschaft (Bird in a Winter Landscape)**
Heinrich Campendonk	Wolfgang Beltracchi	10/2010	Liegender Akt mit Forsch (Reclining Nude with Frog)**
André Derain	Wolfgang Beltracchi	10/2010	Matisse Peignant**
Raoul Dufy	Wolfgang Beltracchi	10/2010	Gebaud im Wald (Building in a Forest)**
Georges Braque	Wolfgang Beltracchi	10/2010	Guitare le Compotier**
Georges Braque	Wolfgang Beltracchi	10/2010	Souvenirs d'Anvers**
Heinrich Campendonk	Wolfgang Beltracchi	10/2010	Katze in Berglandschaft (Cat in a Mountain Landscape)**
Louis Marcoussis	Wolfgang Beltracchi	10/2010	Stilleben Composition Cubiste (14th July) (Still-life Cubist Composition)**
Brett Whiteley	Peter Gant and Mohamed Aman Siddique	07/2010	Big Blue Lavender Bay**
Brett Whiteley	Peter Gant and Mohamed Aman Siddique	07/2010	Orange Lavender Bay**
Brett Whiteley	Peter Gant and Mohamed Aman Siddique	07/2010	Through the Window**
Boris Kustodiev	Anonymous	05/2009	Odalisque*
Charles Blackman,	Peter Gant	11/2008	-*

Artist attributed	Forger	Discovery date	Fake*/Forged** painting
Robert Dickerson	Peter Gant	11/2008	-**
Rover Thomas	Pamela and Ivan Liberto	10/2007	-**

Table A.V Provenance Effects for the Repeat Sales Sample

This table presents the baseline OLS return regression results of repeat sales (Equation (3)). The dependent variable is the annualized return of the repeat sales. Auction House Upgrade equals one if the artwork is sold at a larger/more reputable auction house at the second sale relative to the auction house at the first sale. The definitions of provenance variables can be found in the Appendix A. *, **, and *** indicate statistical significance at the 10%, 5%, and 1% levels, respectively. Standard errors (S.E.) are reported in parentheses and clustered at the auction branch level.

	Full sample (1) Return	Same Auction houses (2) Return
Dept. Var.:		
Changes Prominent Collectors	0.3156**	0.7381*
	(0.1514)	(0.3886)
Changes Prominent Dealers	0.4603	1.2009***
	(0.4585)	(0.4357)
Changes Museum	0.2085***	0.6230**
	(0.0722)	(0.2921)
Changes Gallery Exhibition	0.2866*	0.2700
	(0.1693)	(0.3226)
Changes Other Literature	0.1425	0.2916***
	(0.0883)	(0.0956)
Auction House Upgrade	0.3914***	
	(0.1202)	
Artist FE	Yes	Yes
Year FE	Yes	Yes
Month FE	Yes	Yes
Auct. House Branch FE	Yes	Yes
Hedonic Controls	Yes	Yes
Initial Provenance	Yes	Yes
# of Obs.	6,647	4,236
R-squared	0.3397	0.3997

Table A.VI Subsample of Paintings Including Provenance Information

This table presents OLS regressions of provenance on the probability of being sold and hammer prices for the subsample of paintings that contain provenance information. The dependent variables are: (i) *Sold[0,1],* which equals one if the auction lot is successfully sold, and (ii) *Ln(Price),* which is the natural logarithm of deflated hammer price in USD. In Columns (1) and (3), *Pedigree*, *Exhibition*, *Literature*, and *Certification* are dummy variables if the artworks have any information on these provenance dimensions, respectively. In Columns (2) and (4), *Pedigree*, *Exhibition*, *Literature*, and *Certification* are the corresponding natural logarithm of text character length. *Ln(Low Price Estimate)*, the natural logarithm of the low estimate in USD, is a proxy for reserve price. All regressions include Hedonic Controls (see Online Appendix Table A.I), and Artist, Year, Month, and Auction House Branch Level (AH) Fixed Effects. *, **, and *** indicate statistical significance at the 10%, 5%, and 1% levels, respectively. Standard errors (S.E.) are reported in parentheses and clustered at the auction branch level.

Dept. Var.:	Sold [0,1]		Ln(Price)	
	(1)	(2)	(3)	(4)
Provenance as:	Indicator	Text Length	Indicator	Text Length
Pedigree	0.0210**	0.0078***	0.2018***	0.0655***
	(0.0085)	(0.0021)	(0.0218)	(0.0065)
Exhibition	0.0393***	0.0086***	0.3365***	0.0730***
	(0.0043)	(0.0009)	(0.0143)	(0.0029)
Literature	0.0298***	0.0071***	0.4207***	0.0876***
	(0.0050)	(0.0010)	(0.0260)	(0.0053)
Certification	-0.0237*	-0.0002	0.1745***	0.0522***
	(0.0126)	(0.0039)	(0.0336)	(0.0096)
Ln(Low Price Estimate)	-0.0548***	-0.0571***		
	(0.0039)	(0.0038)		
Artist FE	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes
Month FE	Yes	Yes	Yes	Yes
Auct. House Branch FE	Yes	Yes	Yes	Yes
Hedonic Controls	Yes	Yes	Yes	Yes
# of Obs.	316,138	316,138	227,970	227,970
R-squared	0.1792	0.1795	0.8250	0.8282

Table A.VII LASSO Results of Provenance Effects on Probability of Being Sold and Hammer Price

This table presents LASSO results of the provenance effects on probability of being sold and hammer prices. The dependent variables are: (i) *Sold[0,1],* which equals one if the auction lot was successfully sold, and (ii) *Ln(Price),* which is the natural logarithm of deflated hammer price in USD. *Pedigree*, *Exhibition*, *Literature*, and *Certification* are dummy variables if the artworks have any information on this provenance dimension, respectively. Columns (1) and (4), Columns (2) and (5), and Columns (3) and (6) use subsamples of paintings by the artists whose number of sales in our sample period are above 100, 200 and 500, respectively. All regressions include Hedonic Controls (see Online Appendix Table A.I), as well as Artist, Year, Month, and Auction House Branch Level (AH) Fixed Effects. *, **, and *** indicate statistical significance at the 10%, 5%, and 1% levels, respectively. Standard errors (S.E.) are reported in parentheses and clustered at the auction branch level.

Dept. Var.:	Sold[0,1]			Ln[Price]		
	(1)	(2)	(3)	(4)	(5)	(6)
Subsample:	#100	#200	#500	#100	#200	#500
Pedigree	0.0249***	0.0364***	0.0391***	0.7164***	0.7152***	0.6718***
	(0.0041)	(0.0049)	(0.0070)	(0.0149)	(0.0175)	(0.0254)
Exhibition	0.0196***	0.0184***	0.0122**	0.7389***	0.7681***	0.8063***
	(0.0028)	(0.0034)	(0.0049)	(0.0105)	(0.0130)	(0.0186)
Literature	0.0098***	0.0108***	0.0146***	0.9519***	1.0061***	0.9819***
	(0.0029)	(0.0035)	(0.0050)	(0.0110)	(0.0135)	(0.0187)
Certification	-0.0948***	-0.0923***	-0.1029***	0.1202***	0.1350***	0.0011
	(0.0047)	(0.0055)	(0.0075)	(0.0167)	(0.0190)	(0.0264)
Artist FE	Yes	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes	Yes
Month FE	Yes	Yes	Yes	Yes	Yes	Yes
Auct. House Branch FE	Yes	Yes	Yes	Yes	Yes	Yes
Hedonic Controls	Yes	Yes	Yes	Yes	Yes	Yes
Low Price Estimate	Yes	Yes	Yes	No	No	No
# of Obs.	185,492	127,175	58,984	135,303	93,129	43,545

Table A.VIII Hedonic Pricing Regressions with Non-Linear Provenance Effects

The dependent variable is *Ln(Price)*, the natural logarithm of the real hammer price in real (2007) USD. *Pedigree (Text Length) is the* natural log of one plus the number of characters of the provenance text. *Provenance Squared (Text Length)* is the natural log of one plus the squared term of the number of characters of the provenance text. *Exhibition (Number Count)* and *Literature (Number Count)* count the number of exhibitions and the number of the cases for which information on a painting is given in the literature. *Exhibition Squared (Number Count)* and *Literature Squared (Number Count)* are the squared terms of the *Exhibition (Number Count) and Literature (Number Count)*. All regressions include Hedonic Controls (detailed in Online Appendix Table A.I), and Artist, Year, Month, and Auction House Branch Level (AH) Fixed Effects. *, **, and *** indicate statistical significance at the 10%, 5%, and 1% levels, respectively. Standard errors (S.E.) are reported in parentheses and clustered at the auction branch level.

Provenance as:	Text Length		Number Count					
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Dept. Var.:	Ln(Price)	Ln(Price)	Ln(Price)	Ln(Price)	Ln(Price)	Ln(Price)	Ln(Price)	Ln(Price)
Provenance	0.1539***	4.5727***						
	(0.0131)	(0.7772)						
Provenance Squared		-2.1273***						
		(0.3719)						
Exhibition			0.1561***	0.2302***			0.1105***	0.1647***
			(0.0101)	(0.0153)			(0.0073)	(0.0115)
Exhibition Squared				-0.0057***				-0.0045***
				(0.0006)				(0.0005)
Literature					0.1869***	0.2577***	0.1338***	0.1849***
					(0.0090)	(0.0112)	(0.0059)	(0.0085)
Literature Squared						-0.0043***		-0.0031***
						(0.0005)		(0.0005)
Artist FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Month FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Auct. House Branch FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Hedonic Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
# of Obs.	1,111,220	1,111,220	1,111,220	1,111,220	1,111,220	1,111,220	1,111,220	1,111,220
R-squared	0.7800	0.7812	0.7782	0.7786	0.7782	0.7787	0.7795	0.7801

Table A.IX Liquidity and Provenance Effects on Hammer Price

The dependent variable is *Ln(Price)*, the natural logarithm of the real hammer price in real (2007) USD. *Provenance* is a dummy variable capturing if the artwork's catalogue comprises any provenance information. In Panel A, *Sales* is the liquidity measure defined as the natural log of one plus the number of auction sales of paintings by an artist from auctions around the world over the past 1–5 years. In Panel B, *Sales Ratio* is the percentage of an artist' paintings offered for auction that were sold from auctions around the world over the past 1–5 years. All regressions include Hedonic Controls (detailed in Online Appendix Table A.I), and Artist, Year, Month, and Auction House Branch Level (AH) Fixed Effects. *, **, and *** indicate statistical significance at the 10%, 5%, and 1% levels, respectively. Standard errors (S.E.) are reported in parentheses and clustered at the auction branch level.

Panel A: Accounting for liquidity: number of auction sales of paintings by an artist

	(1)	(2)	(3)	(4)	(5)	(6)
Dept. Var.:	Ln(Price)	Ln(Price)	Ln(Price)	Ln(Price)	Ln(Price)	Ln(Price)
Provenance	0.2762***	0.2760***	0.2760***	0.2760***	0.2759***	0.2758***
	(0.0204)	(0.0202)	(0.0202)	(0.0202)	(0.0202)	(0.0203)
Sales (past 5 years)		0.0379**				
		(0.0170)				
Sales (past 4 years)			0.0384**			
			(0.0168)			
Sales (past 3 years)				0.0384**		
				(0.0163)		
Sales (past 2 years)					0.0370**	
					(0.0152)	
Sales (past 1 year)						0.0350***
						(0.0130)
Artist FE	Yes	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes	Yes
Month FE	Yes	Yes	Yes	Yes	Yes	Yes
Auct. House Branch FE	Yes	Yes	Yes	Yes	Yes	Yes
Hedonic Controls	Yes	Yes	Yes	Yes	Yes	Yes
# of Obs.	1,111,220	1,111,220	1,111,220	1,111,220	1,111,220	1,111,220
R-squared	0.7771	0.7772	0.7772	0.7772	0.7772	0.7772

Panel B: Accounting for liquidity: sales ratio of an artist

	(1)	(2)	(3)	(4)	(5)	(6)
Dept. Var.:	Ln(Price)	Ln(Price)	Ln(Price)	Ln(Price)	Ln(Price)	Ln(Price)
Provenance	0.2762***	0.2709***	0.2710***	0.2710***	0.2710***	0.2701***
	(0.0204)	(0.0222)	(0.0223)	(0.0223)	(0.0223)	(0.0226)
Sales Ratio (past 5 years)		0.2517***				
		(0.0338)				
Sales Ratio (past 4 years)			0.2316***			
			(0.0296)			
Sales Ratio (past 3 years)				0.2079***		
				(0.0248)		
Sales Ratio (past 2 years)					0.1631***	
					(0.0187)	
Sales Ratio (past 1 year)						0.1142***
						(0.0110)
Artist FE	Yes	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes	Yes
Month FE	Yes	Yes	Yes	Yes	Yes	Yes
Auct. House Branch FE	Yes	Yes	Yes	Yes	Yes	Yes
Hedonic controls	Yes	Yes	Yes	Yes	Yes	Yes
# of Obs.	1,111,220	934,325	930,915	923,996	908,544	866,385
R-squared	0.7771	0.7750	0.7747	0.7742	0.7733	0.7725

Table A.X Accounting for Liquidity and Provenance Effects on Hammer Price: Adding Interaction Terms

The dependent variable is *Ln(Price)*, the natural logarithm of the real hammer price in real (2007) USD. *Provenance* is a dummy variable capturing if the artwork's catalogue comprises any provenance information. *Sales* is the liquidity measure defined as the natural log of one plus the number of auction sales of paintings by an artist from auctions around the world over the past 1–5 years. *Provenance × Sales* is an interaction term between *Provenance* and *Sales*. All regressions include Hedonic Controls (detailed in Online Appendix Table A.I), and Artist, Year, Month, and Auction House Branch Level (AH) Fixed Effects. *, **, and *** indicate statistical significance at the 10%, 5%, and 1% levels, respectively. Standard errors (S.E.) are reported in parentheses and clustered at the auction branch level.

Dept. Var.:	(1) Ln(Price)	(2) Ln(Price)	(3) Ln(Price)	(4) Ln(Price)	(5) Ln(Price)
Provenance	0.2686***	0.2727***	0.2775***	0.2831***	0.2863***
	(0.0186)	(0.0186)	(0.0182)	(0.0172)	(0.0157)
Sales (past 5 years)	0.0656***				
	(0.0122)				
Provenance × Sales (past 5 years)	0.0029				
	(0.0109)				
Sales (past 4 years)		0.0680***			
		(0.0125)			
Provenance × Sales (past 4 years)		0.0014			
		(0.0111)			
Sales (past 3 years)			0.0696***		
			(0.0128)		
Provenance × Sales (past 3 years)			-0.0006		
			(0.0115)		
Sales (past 2 years)				0.0715***	
				(0.0129)	
Provenance × Sales (past 2 years)				-0.0035	
				(0.0119)	
Sales (past 1 years)					0.0711***
					(0.0124)
Provenance × Sales (past 1 years)					-0.0063
					(0.0126)
Artist FE	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes
Month FE	Yes	Yes	Yes	Yes	Yes
Auct. House Branch FE	Yes	Yes	Yes	Yes	Yes
Hedonic controls	Yes	Yes	Yes	Yes	Yes
# of Obs.	1,111,220	1,111,220	1,111,220	1,111,220	1,111,220
R-squared	0.7779	0.7779	0.7779	0.7779	0.7778

Table A.XI Provenance and Liquidity (Past Sales)

This table presents the relation between liquidity and provenance. The dependent variable is *Provenance*, a dummy variable capturing if the artwork's catalogue comprises any provenance information (is then equal to 1). *Sales* is the liquidity measure defined as the natural log of one plus the number of auction sales of paintings by an artist from auctions around the world in Panel A (auctions within the same country in Panel B) over the past 1–5 years. All regressions include Hedonic Controls (detailed in Online Appendix Table A.I), and Artist, Year, Month, and Auction House Branch Level (AH) Fixed Effects. *, **, and *** indicate statistical significance at the 10%, 5%, and 1% levels, respectively. Standard errors (S.E.) are reported in parentheses and clustered at the auction branch level.

Panel A: Provenance and liquidity (worldwide auction sales)

Liquidity: Global	(1)	(2)	(3)	(4)	(5)
Dept. Var.:	Provenance	Provenance	Provenance	Provenance	Provenance
Sales (past 5 years)	0.0013 (0.0035)				
Sales (past 4 years)		0.0008 (0.0033)			
Sales (past 3 years)			0.0006 (0.0031)		
Sales (past 2 years)				0.0011 (0.0026)	
Sales (past 1 year)					0.0021 (0.0022)
Artist FE	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes
Month FE	Yes	Yes	Yes	Yes	Yes
Auct. House Branch FE	Yes	Yes	Yes	Yes	Yes
Hedonic controls	Yes	Yes	Yes	Yes	Yes
# of Obs.	1,746,514	1,746,514	1,746,514	1,746,514	1,746,514
R-squared	0.4620	0.4620	0.4620	0.4620	0.4620

Panel B: Provenance and liquidity (sales based on auctions by country)

Liquidity: Countrywide	(6)	(7)	(8)	(9)	(10)
Dept. Var.:	Provenance	Provenance	Provenance	Provenance	Provenance
Sales (past 5 years)	0.0015 (0.0026)				
Sales (past 4 years)		0.0013 (0.0026)			
Sales (past 3 years)			0.0012 (0.0026)		
Sales (past 2 years)				0.0012 (0.0026)	
Sales (past 1 year)					0.0013 (0.0026)
Artist FE	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes
Month FE	Yes	Yes	Yes	Yes	Yes
Auct. House Branch FE	Yes	Yes	Yes	Yes	Yes
Hedonic controls	Yes	Yes	Yes	Yes	Yes
# of Obs.	1,746,514	1,746,514	1,746,514	1,746,514	1,746,514
R-squared	0.4620	0.4620	0.4620	0.4620	0.4620

Table A.XII Hedonic Pricing Regressions for Sales by Big and Small Auction Houses

We category the auction houses by size as (1) large auction houses (Christie's and Sotheby's), (2) medium-sized auction houses (Bonhams and Phillips, and other important auction houses in United States and Europe listed in Appendix A.), and (3) other small auction houses. We cluster the price range by percentiles, 25th, 50th, and 75th and run regressions for each price range subsample for each auction house category. The dependent variable is *Ln(Price)*, the natural logarithm of the real hammer price in real (2007) USD. *Provenance* is a dummy variable capturing if the artwork's catalogue comprises any provenance information. *Pedigree*, *Exhibition*, *Literature*, and *Certification* are dummy variables if the artworks have any information on this provenance dimension, respectively. All regressions include Hedonic Controls (detailed in Online Appendix Table A.I), and Artist, Year, Month, and Auction House Branch Level (AH) Fixed Effects. *, **, and *** indicate statistical significance at the 10%, 5%, and 1% levels, respectively. Standard errors (S.E.) are reported in parentheses.

Panel A: Hedonic Pricing Regressions with Provenance Dummy for Subsamples of Big and Small Auction Houses

Sample	(1) Big	(2) Medium	(3) Small	(4) Big	(5) Medium	(6) Small
Dept. Var.:	Ln(Price)	Ln(Price)	Ln(Price)	Ln(Price)	Ln(Price)	Ln(Price)
Provenance	0.3570***	0.3264***	0.2242***			
	(0.0059)	(0.0069)	(0.0045)			
Pedigree				0.2340***	0.2767***	0.1604***
				(0.0058)	(0.0076)	(0.0063)
Exhibition				0.3128***	0.2851***	0.4134***
				(0.0073)	(0.0144)	(0.0140)
Literature				0.4330***	0.3178***	0.3360***
				(0.0079)	(0.0177)	(0.0119)
Certification				0.0270	0.1503***	0.1429***
				(0.0256)	(0.0136)	(0.0064)
Artist FE	Yes	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes	Yes
Month FE	Yes	Yes	Yes	Yes	Yes	Yes
Hedonic Controls	Yes	Yes	Yes	Yes	Yes	Yes
# of Obs.	198,076	214,526	670,363	198,076	214,526	670,363
R-squared	0.7611	0.7263	0.7398	0.7709	0.7283	0.7405

Panel B: Hedonic Pricing Regressions with Provenance Dummy for Subsamples of Big and Small Auction Houses across Price Ranges

Price Range:	p0-p25			p25-p50			p50-p75			p75-p100		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Auction House:	Big	Medium	Small	Big	Medium	Small	Big	Medium	Small	Big	Medium	Small
Dept. Var.:	Ln(Price)	Ln(Price)	Ln(Price)	Ln(Price)	Ln(Price)	Ln(Price)	Ln(Price)	Ln(Price)	Ln(Price)	Ln(Price)	Ln(Price)	Ln(Price)
Provenance	0.0198	0.0542***	0.0439***	0.0299***	0.0372***	0.0416***	0.0570***	0.0593***	0.0396***	0.2384***	0.1226***	0.0312***
	(0.0227)	(0.0130)	(0.0057)	(0.0072)	(0.0055)	(0.0032)	(0.0045)	(0.0047)	(0.0035)	(0.0069)	(0.0098)	(0.0076)
Artist FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Month FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Hedonic Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
# of Obs.	4,576	37,836	198,160	15,104	55,483	176,401	44,739	60,010	151,151	121,473	41,631	108,366
R-squared	0.5596	0.4292	0.4551	0.3641	0.3101	0.2710	0.3458	0.3337	0.2963	0.6271	0.5434	0.5310

Table A.XIII Hedonic Pricing Regressions for (Non-)Established Artists

The *Established* artists are defined when the low estimate is above the 75th percentile and the *Non-Established* artists are defined when the low estimate is below the 25th percentile. The dependent variable is *Ln(Price)*, the natural logarithm of the real hammer price in real (2007) USD. *Provenance* is a dummy variable capturing if the artwork's catalogue comprises any provenance information. *Pedigree*, *Exhibition*, *Literature*, and *Certification* are dummy variables if the artworks have any information on this provenance dimension, respectively. All regressions include Hedonic Controls (detailed in Online Appendix Table A.I), and Artist, Year, Month, and Auction House Branch Level (AH) Fixed Effects. *, **, and *** indicate statistical significance at the 10%, 5%, and 1% levels, respectively. Standard errors (S.E.) are reported in parentheses and clustered at the auction branch level.

	(1)	(2)	(3)	(4)
Sample:	Non-Established		Established	
Dept. Var.:	Ln(Price)	Ln(Price)	Ln(Price)	Ln(Price)
Provenance	0.0822***		0.2086***	
	(0.0073)		(0.0048)	
Pedigree		0.0826***		0.1017***
		(0.0091)		(0.0052)
Exhibition		0.1976***		0.3039***
		(0.0252)		(0.0072)
Literature		0.1447***		0.3253***
		(0.0332)		(0.0071)
Certification		0.0099		0.0316***
		(0.0125)		(0.0087)
Artist FE	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes
Month FE	Yes	Yes	Yes	Yes
Auct. House Branch FE	Yes	Yes	Yes	Yes
Hedonic Controls	Yes	Yes	Yes	Yes
# of Obs.	227,455	227,455	370,334	370,334
R-squared	0.4596	0.4599	0.7931	0.7971

Table A.XIV Hedonic Pricing Regressions of a Selection of the Most Liquid Artists

The dependent variable is *Ln(Price)*, the natural logarithm of the real hammer price in real (2007) USD. *Provenance* is a dummy variable capturing if the artwork's catalogue comprises any provenance information. *Pedigree*, *Exhibition*, *Literature*, and *Certification* are dummy variables if the artworks have any information on this provenance dimension, respectively. All regressions include Hedonic Controls (detailed in Online Appendix Table A.I), and Artist, Year, Month, and Auction House Branch Level (AH) Fixed Effects. *, **, and *** indicate statistical significance at the 10%, 5%, and 1% levels, respectively. Standard errors (S.E.) are reported in parentheses.

	(1)	(2)	(3)	(4)	(5)	(6)
Sample:	Pablo Picasso		Andy Warhol		Raoul Dufy	
Dept. Var.:	Ln(Price)	Ln(Price)	Ln(Price)	Ln(Price)	Ln(Price)	Ln(Price)
Provenance	0.3105***		0.7626***		0.3788***	
	(0.1127)		(0.0820)		(0.0813)	
Pedigree		0.1765		0.6025***		0.0061
		(0.1269)		(0.0856)		(0.1064)
Exhibition		0.4394***		0.4430***		0.1050
		(0.0677)		(0.1067)		(0.0909)
Literature		0.3644***		0.8833***		0.2501***
		(0.0774)		(0.1199)		(0.0886)
Certification		0.0936		0.0902		0.3798***
		(0.1102)		(0.1353)		(0.1054)
Artist FE	Yes	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes	Yes
Month FE	Yes	Yes	Yes	Yes	Yes	Yes
Auct. House Branch FE	Yes	Yes	Yes	Yes	Yes	Yes
Hedonic Controls	Yes	Yes	Yes	Yes	Yes	Yes
# of Obs.	1,650	1,650	1,737	1,737	1,175	1,175
R-squared	0.7045	0.7172	0.7726	0.7927	0.7739	0.7744