The Impact of Reservation Prices on the Perceived Bias of Expert Appraisals of Fine Art

Clare Mc Andrew
Arts Economics

James L. Smith
Southern Methodist University

Rex Thompson
Southern Methodist University

February 12, 2009

(JEL: G11, G14, G12, C160, Z10)

Abstract

We examine whether the expert appraisals provided to bidders before major art auctions are unbiased indicators of market value. Despite a strong grounding in theory, this aspect of optimal auction design has been frequently challenged in previous empirical research, particularly in the market for fine art. We adopt a valuation benchmark that incorporates sellers' reservation prices as well as high bids, and recognize latent censoring due to works that fail to make reserve. Although the auction houses never divulge sellers' reserve prices, which are therefore unobservable, we exploit the fact that they can be observed indirectly via their impact on buy-in rates. Using the set of French Impressionist paintings brought to auction from 1985 to 2001, we estimate the distribution of seller reserve prices, establish their link to a proper valuation benchmark, and isolate the selection bias due to bought-in works on the perceived market value of fine art works brought to auction. After controlling for the impact of reservation prices, and considering all works brought to auction (and therefore all works appraised), we find no evidence of bias in the experts' pre-sale estimates.

We thank Kusin & Co. for access to the data used in this study, and G. Shockley, V. Ginsburgh, R. Towse, and participants at the First European Workshop on Applied Cultural Economics, San Leo, Italy, 2003 for helpful comments. All remaining errors are our own.

I. Introduction

The monetary values of most fine works of art are established through public sales organized by the major auction houses.[1] Bids are an important indicator of value, but other signals generated in the auction process also convey information: the seller's reserve price and the auction house's own expert appraisal, for example. We consider how these disparate indicators of value are related to each other, and demonstrate with an application to the market for French Impressionist paintings how they can be integrated to arrive at a consistent and useful measure of value.

Central to our inquiry is the question of whether, as often alleged, the appraised values provided by the auction houses before each sale are biased. Despite Milgrom and Weber's (1982) general argument that truthful revelation of private information is the seller's best policy, subsequent researchers have suggested that appraisals are biased (Salamon, 1992; Chanel et. al., 1996; Beggs and Grady, 1997; Bauwens and Ginsburgh, 2000), and even that sellers are motivated by perceived self-interest to misrepresent the value of auctioned goods (Mei and Moses, 2005). Despite the potential intuitive appeal of arguments couched in terms of vested interest, we find such conclusions to be somewhat peremptory and aim to subject these claims to closer scrutiny.

Testing for bias requires a clear benchmark. If we consider a particular work, say the i^{th} item in a specific auction, we might compare the appraised value (A_i) to the high bid (H_i). But, if the bid falls short of the seller's reserve price (R_i), the bid is rejected and the item is said to be "bought in." Only if the high bid makes reserve ($H_i \geq R_i$) is the item

[1] Sotheby's and Christie's handle the majority of all art works brought to auction, but more than one hundred smaller auction houses are also active. A few highly specialized niche art markets conduct the majority of their trade privately, outside the auction process. For example, results of dealer interviews and

"hammered down" and sold to the highest bidder. The "successful bid" benchmark, which attempts to detect bias by comparing appraisals to high bids over the subset of works that make reserve, has a major problem: it excludes from the sample appraisals rendered on works that are subsequently bought in, and the excluded works may not be a random draw from the population of appraisals. That problem is remedied if the scope is broadened to include all works brought to auction (and therefore all works appraised), which we call the "all-bid" benchmark, but then another problem arises: the highest bid for a work of art may be a distorted measure of value if it is not high enough to actually trigger a sale.

In light of these problems, an alternative and more plausible benchmark is given by $V_i = \max(H_i, R_i)$, the amount it would take to actually win a particular work.[2] This measure encompasses all works brought to auction, whether bought-in or sold, and represents a more comprehensive measure of value than either the successful-bid or all-bid benchmarks. V_i is the valuation that seems most clearly relevant to the interests of the bidders—who presumably are interested in purchasing art at least cost—and therefore we propose to use it as a "composite" benchmark to measure the degree of bias in pre-sale appraisals. By construction, the composite benchmark falls between the other two.

The idea of using the seller's reserve along with transaction prices to create a benchmark value is not new. Goetzmann and Peng (2006) adopted this approach in their pioneering study of real estate valuation and demonstrated that a "reserve-conditional benchmark" mitigates bias in real estate price indexes. In their context, market

---

polling conducted by Kusin & Company in 1999 indicted that around 85 percent of sales of paintings by Canaletto (1697-1768) are conducted exclusively through dealers.

transactions result from sequential search, so the seller's reserve price determines the number of "days on the market" rather than the probability that a work is bought in. But the rational for using a composite benchmark in auction settings is fundamentally the same as the rational for using a reserve-conditional benchmark in search models.

To conduct an analysis that spans many diverse works, it is useful to normalize all works brought to auction on the basis of appraised values.[3] For i = 1, …, N, we define $h_i = H_i/A_i$ (the "hammer ratio"), $r_i = R_i/A_i$ (the "reserve ratio"), and $v_i = V_i/A_i$ (the "valuation ratio"). If these ratios are regarded as realizations of three jointly distributed random variables (**h**, **r**, and **v**) with a fixed distribution, then the hypothesis that pre-sale appraisals are unbiased can be stated simply: H_o: $E[\mathbf{v}] = 1$.

Unfortunately, the sample data needed for a direct test of H_o are incomplete. High bids on works that are bought-in are not recorded. Thus, the hammer ratios $\{h_i\}$ are censored; we observe h_i only if $h_i \geq r_i$. In addition, the censoring threshold for each individual work (r_i) is never observed because, whether a work is bought in or hammered down, it is auction house policy not to reveal the seller's reserve price.

These circumstances pose a specific difficulty for assessing the bias in appraisals. To see why, consider the following decomposition:

$$E[\mathbf{v}] \;=\; E[\mathbf{h} \mid \mathbf{h}>\mathbf{r}]\times\Pr(\mathbf{h}>\mathbf{r}) + E[\mathbf{r} \mid \mathbf{r}>\mathbf{h}]\times\Pr(\mathbf{r}>\mathbf{h}). \qquad (1)$$

All terms but the third can be consistently estimated from the sample data. The third is problematic. Since reserve ratios are unobservable, it is not immediately apparent how

---

[2] In ascending auctions of the type used for art, the high bid reflects the valuation placed by the second highest bidder since the winning bidder must only outbid his or her rivals, including the seller who participates via the reserve price.
[3] As we show later, normalizing by the appraised value corrects for heteroscedasticity in the distribution of high bids.

the distribution of **r**, or its conditional mean, can be estimated. We overcome this difficulty by exploiting the fact that sellers' reserves, although never observed directly, can be observed indirectly via their impact on buy-in rates. In other words, we proceed by attempting to identify the distribution of reserve ratios that maximizes the likelihood of the observed pattern of rejected bids. Roughly speaking, this is not unlike attempting to deduce the astronomical properties of an invisible star by observing its gravitational effects on an adjacent twin.

Our work is related to a growing literature on the econometric analysis of auction data.[4] Paarsch and Hong (2006), for example, treat the case where reserve prices are fixed and known to all bidders, and apply maxmimum likelihood techniques in that context to obtain consistent estimates of the distribution of bidder's value. However, many types of goods are auctioned subject to *secret* reserves, including oil and gas leases, timberland, construction contracts, and wine—in addition to fine art. Empirical studies of auctions with secret reserves, for example Elyakime, Laffont, Loisel, and Vuong (1997) and Ji and Li (2008), have so far been limited to situations where the reserve price is ultimately revealed—before the model is empirically estimated. Our research extends the econometric analysis of auctions into the realm of an unknown and unobservable reserve.

In Section II, we outline our approach, which is based on a two-step maximum likelihood estimation procedure that overcomes the limitations of our sample data and provides a consistent and asymptotically efficient estimator of the latent distribution of reserve ratios. In Section III, we discuss previous empirical research on art auctions and

[4] An important collection of papers that explore the econometrics of auctions, published in honor of Jean-Jacques Laffont, can be found in the November 2008 issue of the *Journal of Applied Econometrics*.

explain relevant institutional aspects of the auction process. The French Impressionist data set used in our analysis is described in Section IV. Section V explains how our estimation method has been applied to the sample and gives a full account of results. In Section VI, we summarize the major conclusions of our research.

II. Estimating a Latent Distribution: the Reserve Ratio

We start with a sample of N works of art brought to auction, of which a known subset of M works made reserve and were hammered down. We observe the presale appraised value of every work ($A_1, \ldots, A_N$) and the hammer ratios for those items that were actually sold ($h_1, \ldots, h_M$).

The first step towards identifying the latent distribution of reserve ratios is to estimate the (uncensored) cumulative distribution of hammer ratios, which (for reasons discussed later) we assume belongs to the lognormal family denoted $F(\mathbf{h}|\mu, \sigma)$.[5] The empirical distribution of hammer ratios is oddly shaped, not lognormal, due to the censoring of bought-in works, which occurs predominantly when the hammer ratio is low (i.e., when the highest bid falls below the appraisal). To eliminate the impact of censored data, we take advantage of a universal auction house rule that bounds the seller's reserve price. This bound effectively caps the reserve ratio at a known level, h_0, beyond which no censoring of hammer ratios could have occurred in our sample. Thus, we can apply Cohen's (1959) method to obtain maximum likelihood estimates of the mean and standard deviation of the uncensored normal distribution $(\hat{\mu}, \hat{\sigma})$.[6]

---

[5] This is consistent with Paarsch and Hong's (2006) approach, which puts additional structure on the distribution of bidder values by assuming that it comes from a family of parametric distributions.
[6] We apply the formulae from Cohen (1959, p. 228) to the log transform of the high bid data, $z_i = \ln(h_i)$, which produces maximum likelihood estimates of the parameters of the uncensored distribution of **z**:

$$\hat{\mu}_z = \bar{h} - \lambda(\bar{h} - h_0) \quad \text{and} \quad \hat{\sigma}_z^2 = s^2 + \lambda(\bar{h} - h_0)^2,$$

Next, we approximate the lognormal density of (uncensored) hammer ratios by a discrete multinomial distribution: $f(x_i) = F(a_{i+1}|\hat{\mu},\hat{\sigma})-F(a_i|\hat{\mu},\hat{\sigma})$, for i = 1 to 30, where the a_i represent uniform bin boundaries that span the range of hammer ratios, and where $x_i = (a_i \times a_{i+1})^{1/2}$ represents the midpoint (geometric mean) of the given bin.[7] We use these estimated frequencies to compute the number of works expected to have fallen into each bin: $n_i = f(x_i) \times N$, for i = 1, …, 30. From the sample, we observe the *actual* number of works (sold) that fell into each bin: $m_1, \ldots, m_{30}$. Therefore, an estimate of the number of works in each bin that failed to make reserve is given by the difference: $n_1-m_1, \ldots, n_{30}-m_{30}$. This completes the first step of the estimation process.

The second step exploits the fact that differences between the $\{n_i\}$ and $\{m_i\}$ reflect the distribution of hidden reserves. Industry sources indicate that some works are sold without reserve, depending on the seller's preference, and that reserve prices for the rest are distributed over a wide range. We therefore assume that a fixed, but unknown, percentage (κ) of all works are offered without reserve ($r_i = 0$), and that this is determined independently of the work's hammer ratio or appraised value. For all the rest, we assume the reserve ratio follows a fixed beta distribution, $B(\mathbf{r}|\nu,\omega)$, over the interval (0, 0.9).[8] The upper bound is a reflection of auction house rules that prevent sellers from setting reserve prices that exceed approximately 90% of the appraised value.

---

where $\bar{h}$ and s^2 represent the sample mean and variance of $\mathbf{z}$ measured above the censoring threshold, and λ is a transformation of the relative number of censored observations in the sample. Corresponding maximum likelihood estimates of the parameters of the lognormal distribution $(\hat{\mu},\hat{\sigma})$ are obtained by the standard translation.

[7] The continuous density would serve our purpose just as well. However, the discrete version facilitates a multinomial approximation in the second step that simplifies the estimation procedure.

[8] Parameters κ, ν, and ω are to be estimated from the sample data. We chose the beta distribution because it is relatively flexible and encompasses U-shaped, J-shaped, unimodal, rectangular, and many other shapes, depending on the values of ν and ω, which leaves room for the data to determine the shape of the

Thus, the conditional probability that a work will make reserve if its hammer ratio is x_i can be expressed as:

$$prob(H \geq R \mid h = x_i) \quad = \quad \kappa + (1-\kappa) \times B(x_i \mid \nu, \omega). \tag{2}$$

The joint probability that a work's hammer ratio falls into the i^{th} bin and the work is sold, denoted p_i, must then be:

$$p_i(\kappa,\nu,\omega) \quad = \quad f(x_i) \times [\kappa + (1-\kappa) \times B(x_i \mid \nu,\omega)], \quad \text{for i = 1, …, 30.} \tag{3}$$

Likewise, the joint probability that a work's hammer ratio falls into the i^{th} bin and the work is *not* sold, denoted q_i, must be:

$$q_i(\kappa,\nu,\omega) \quad = \quad f(x_i) \times (1-\kappa)[1 - B(x_i \mid \nu,\omega)], \quad \text{for i = 1, …, 29,} \tag{4}$$

and where q_{30} is determined by the requirement that all p_i and q_i sum to unity.

Since m_i is the observed number of works in a given bin that made reserve and n_i-m_i represents the estimated number of works in that bin that failed to make reserve, we can write the log-likelihood of the sample in terms of the multinomial distribution:

$\mathcal{L}(n_1, \ldots, n_{30}, m_1, \ldots, m_{30} \mid \kappa,\nu,\omega)$ =

$$c + \sum_{1}^{30} (n_i - m_i) ln(q_i(\kappa,\nu,\omega)) + \sum_{1}^{30} m_i \, ln(p_i(\kappa,\nu,\omega)) \tag{5}$$

where c represents a fixed scalar. The values of κ, ν, and ω that maximize this function constitute our estimate of the latent distribution of sellers' reserves.

---

reserve distribution. Ji and Li (2008), in contrast, assume that the distribution of the unknown reserve is exponential.

III. Art Auctions and Price Formation

Significant institutional characteristics of art auction markets have been explored by Ashenfelter (1989), who notes that auction markets for wine and art are often not arranged as assumed in auction theory or observed in pure financial markets. These arrangements play an important role in the price formation and reporting process. In the major art auction houses, such as Sotheby's and Christie's, bidding starts low and rises as the auctioneer calls out higher and higher prices. Ashenfelter estimated that about one-third of the Impressionist paintings he examined did not find buyers because they failed to make reserve when the bidding stopped. Our sample of French Impressionist art exhibits a similar buy-in rate, approximately 30 percent.

The significance of reserve prices and their relation to the auction houses' presale estimates have captured the interest of a number of scholars. In auction catalogues, the houses publish a presale lower (L) and upper (U) estimate of value of each work. The sellers' reserve prices (R) are not published, but they are subject to auction house rules and must not exceed the lower presale estimate. There may be feedback from expert appraisals to the reserve prices set by consignors since the auction houses provide expert opinion to sellers as they establish their reserve. On the other hand, experts may place some weight on sellers' reserve demands when setting their final presale estimates.

Ashenfelter and Graddy (2003) examine the question of motivation in setting final upper and lower presale estimates and the associated spread (S=U/L). They entertain the notion that experts might be influenced by the desires of sellers when the presale low is established. If a seller wishes to set a high reserve, and experts accommodate by increasing their low estimate, the spread may be compressed. Ashenfelter and Graddy

(2002), and Ekelund, Ressler, and Watson (1998) interpret their findings of a negative correlation between buy-ins and high-low spread as evidence of this.[9] We hesitate to infer from this logic, however, that experts are coerced. Rather, we view the seller's reserve price as a value signal that is useful in estimating the price required to buy the work.

The accuracy and unbiasedness of art experts' opinions, as expressed in presale appraisals, have been questioned in several papers, with mixed results. From the theoretical perspective, Milgrom and Weber (1982) show that, in most auction settings, including the type of ascending bid auctions used to sell paintings, "honesty is the best policy" for sellers.[10] Empirically, Ashenfelter (1989) argues that auction houses are generally truthful since estimates are highly correlated with the prices obtained. Of course high correlation is not a rigorous test of unbiasedness. Going beyond correlation, Beggs and Graddy (1997), Chanel, et al (1996), and Bauwens and Ginsburgh (2000) all find evidence of bias—but they define bias relative to the successful-bid benchmark.[11] Those findings are ambiguous since, even if appraisals do accurately measure the amounts needed to purchase works brought to auction (the composite benchmark), they will appear low relative to the subset of successful bids (the successful-bid benchmark).

In a more recent paper, Mei and Moses (2005) also investigate presale art appraisals, building on their earlier finding that investors seem to overpay for

---

[9] We ran similar probit regressions on the French Impressionist data set and found that while the likelihood of a buy-in is negatively related to the mean of the high and low estimates, it is unrelated to the spread between the high and low.

[10] "Honesty is the best policy" is known to break down under certain circumstances in *multi-unit* auctions, where multiple copies of the same item are available. In our sample, however, each work offered at auction is unique, so Milgrom and Weber's results (which are based on single-unit auctions) should apply. See, for example, Perry and Reny (1999).

[11] Beggs and Graddy (1997) also conclude that pre-sale estimates are biased relative to the all-bids benchmark, at least in the subsample of data where this benchmark could be directly observed.

masterpieces (the so-called "masterpiece effect").[12] In a sample of prices obtained in repeated sales, which necessarily limits the scope to the censored sample of works that made reserve, they find that while estimates are highly correlated with prices paid, there is an apparent upward bias in estimates for very highly-priced paintings. They attribute this to the auctioneer's "vested interest" in high prices (which generate high commissions), and suggest that the auction houses may attempt to tilt pre-sale appraisals upward for expensive paintings (since they can benefit most from such bias if investors are credulous) while trying to remain unbiased overall.[13] Although Mei and Moses interpret their results as evidence that investors are credulous—significantly and unduly influenced by the estimates of experts—Gershkov and Toxvaerd (2008) subsequently demonstrate that the type of behavior observed and documented by Mei and Moses is not inconsistent with the hypothesis that bidders are fully rational and pre-sale estimates unbiased.

To subject the hypothesis of biased appraisals to more rigorous empirical examination, we take an approach that, unlike previous studies, considers all works brought to auction, and thus all works appraised—whether they were subsequently bought-in or hammered down for sale.

[12] Mei and Moses found in their earlier (2002) paper that expensive paintings tend to under-perform their "art market index." Pesando (1993) studied the market for lithographic prints and reached a similar conclusion; i.e., no evidence of masterpieces outperforming the market because most of the desirable characteristics of these works are apparently capitalized into their prices. If a "masterpiece" is defined to be an inspirational work of art whose owner predictably earns a positive abnormal return, then perhaps Gertrude Stein (1935) was correct to opine, albeit in a different context, that we can expect to observe few of them.

[13] Mei and Moses (2005) propose a test of unbiasedness in logged data and find, in their sample, evidence of such bias. We have replicated their regressions using our sample and obtain similar results and significance levels. But, unbiasedness in the logs is very different from (and does not imply) unbiasedness in the levels. It is not possible to be unbiased in both transformations of the data. Why experts should want to be unbiased in the logs at the expense of bias in the levels has never been motivated satisfactorily.

IV. Data Description

The data set used in this analysis was compiled by Kusin & Company using art auction data from ArtNet and ArtFact.[14] The works sold are French Impressionist paintings, which, using the Kusin & Company Classification Code, comprise the complete set of 14 artists.[15] The data consists of 4,260 attempted auction sales from the 16-year period between January 1985 and December 2001 from a cohort of 130 international auction houses. Prior to an auction, a presale catalogue is published with information on artist, title, date of sale, auction house/location of sale, size, medium, lot number, year, and a presale lower and upper price estimate.

After dropping observations with incomplete data, there remain 4,174 fully attributed paintings.[16] If the work was actually sold, the hammer price and/or the premium price are listed.[17] Where only a premium price is recorded, it is converted to a hammer price using the buyers' premiums from Christie's, Sotheby's or an average of the other major auction houses. Summary information about the paintings in our sample appears in Table 1. In total, 2,925 works were sold for the total amount $3.02 billion. An additional 1,249 works failed to make reserve and were bought in. The two largest auction houses, Christie's and Sotheby's, offered 40 percent and 45 percent of the paintings respectively.

---

[14] Some data from both of these sources were incorrect and/or incomplete, for example missing a price or presale estimates or outside the realms of marketplace reality, for example a Renoir with a $50 hammer price. Where possible, Kusin & Company cleaned and corrected the data by reference to the original auction catalogues or correspondence with the relevant auction house. Auction data without an attributable hammer price or without presale estimates were omitted from the study.

[15] Works by these 14 artists represent the complete accepted scholarly canon of French Impressionist painters: Frederic Bazille, Gustave Caillebotte, Mary Cassatt, Paul Cezanne, Edgar Degas, Eva Gonzales, Paul Gauguin, Armand Guillaumin, Edouard Manet, Claude Monet, Berthe Morisot, Camille Pissarro, Auguste Renoir, and Alfred Sisley.

[16] A fully attributed painting is one for which there exists no controversy over the artist that created it. All works with attributions that include statements such as "in the style of", "in the school or circle of", or "attributed to" are not included in the data.

Since expert opinion regarding fine art is, by convention, expressed in terms of a *range* of values, we must consider how lower and upper estimates should be aggregated to obtain the expected monetary value of a given work. If estimation errors were distributed symmetrically, the natural approach would be to take a simple arithmetic average of U_i and L_i.[18] If, for a given work, experts choose U_i and L_i to create equal tail probabilities, and if the distribution is symmetric, then the arithmetic average would correspond exactly to the expected value. However, the distribution of high bids is heavily skewed to the right, at least for works that are actually sold. As we show later, after accounting for works that are bought in, the distribution of hammer ratios conforms closely to the lognormal, and for a lognormal distribution, the arithmetic mean of lower and upper cut points chosen to equate tail probabilities would understate the population mean. The geometric mean would do likewise, since it must fall below the arithmetic mean.

When the distribution of estimation errors is skewed, the criterion by which experts would choose upper and lower estimates is open to question. Announcing a range with equal tail probabilities makes it impossible to extract an unbiased estimate of the population expected value unless the tail probabilities are reported—which they are not. On the other hand, if experts choose high and low estimates that equate the deviations of mean values in upper and lower tails from the population mean (rather than equating tail probabilities), then the *geometric* mean of lower and upper estimates would correspond exactly with the mean of the lognormal.[19] For this reason, we hypothesize

[17] The premium price includes a fee paid by buyers to compensate the auction house for services rendered.
[18] This is the approach taken in all of the papers cited in the introduction.
[19] These deviations are measured as the distance between the population mean and probability-weighted values in the right and left tails. Proof of the proposition stated in the text is available from the authors.

that the presale estimation process is one where an unbiased expert first establishes the expected value of an item, and then selects upper and lower bounds whose geometric mean equals this expected value.[20] In terms of the model presented in Section II, the appraised value of a given work is represented by the geometric mean of upper and lower estimates: $A_i = (L_i \times U_i)^{½}$.

Because some art works fail to make reserve, the observed distribution of hammer ratios in our sample is censored. Figure 1 depicts the censored distribution of hammer ratios, and a fitted lognormal distribution is superimposed. The parameters of the lognormal were chosen to minimize the Kolmogorov-Smirnov D-statistic. Although the fitted distribution resembles the observed frequencies, the KS test easily rejects the hypothesis that observed frequencies were drawn from a lognormal distribution.

Of course, fitting any distribution to the sample of observed hammer ratios ignores the fact that bought-in works have been censored. However, auction house rules cap reserve prices at about 90% of the geometric mean appraisal, which means that the upper portion of the distribution is unaffected.[21] Figure 2 shows a fit of the lognormal distribution based on just the right tail of the empirical distribution (i.e., hammer ratios above 0.90), which avoids the impact of censoring. Above this cutoff, the empirical and lognormal distributions show a close correspondence. Below the cutoff, the fitted lognormal distribution reflects the uncensored frequency of hammer ratios. Notwithstanding a few spikes around unity and a longer right tail in the empirical

---

[20] Given this approach, the spread between U and L affords an additional degree of freedom that could be used by the expert to communicate an assessment of price uncertainty independent of the signal that the geometric mean gives about the expected value of the distribution.

[21] The bound actually limits the seller's reserve to not exceed the lower pre-sale estimate of value, but the lower value averages about 86% of the geometric mean. By setting the threshold at 90% we have allowed for a small margin of error. Later, when discussing estimation results, we show that the estimated distribution of reserve ratios has virtually no mass above 0.85.

distribution, the Kolmogorov-Smirnov test fails to reject, even at a 50% significance level, the hypothesis that the uncensored distribution of hammer ratios is lognormal. Although the test statistic is based only on the upper part of the distribution, the results are highly suggestive of a close fit between the distributions. A truncated normal distribution, by comparison, does not fit the data nearly as well, with a KS D-statistic three times as large. Thus, the lognormal form is a defensible characterization of the uncensored distribution of hammer ratios.

Table 2 provides additional summary statistics regarding hammer prices and presale estimates for the 2,925 works *actually sold* at auction. The table reveals that, within this subset, the mean hammer ratio significantly exceeds unity, 1.14 with a t-ratio (relative to unity) above 10. Thus, it appears that auction experts are systematically low in their estimate of French Impressionist works relative to the successful-bid benchmark. This result contradicts Mei and Moses (2005) but is consistent with the rest of the previous literature summarized in Section III that concluded that appraised values tend to understate the hammer prices of works *actually sold*. However, nothing about this result indicates that appraisals tend to understate the composite benchmark, i.e., the amount needed to actually purchase a given work.

Table 2 also reports two measures of heteroscedasticity that justify normalizing bids by appraised values. The sample correlation between A_i and $|H_i\text{-}A_i|$ is 0.65 and highly significant. This confirms, not surprisingly, that hammer price variation is closely tied to prior measures of value. However, the correlation between A_i and $|h_i\text{-}\bar{h}|$ is only -0.03 and not significantly different than zero, which suggests that taking ratios corrects the natural heteroscedasticity in hammer prices.

V. Estimation Results

A. The Uncensored Distribution of Hammer Ratios

Because auction house rules prevent sellers from setting reserve ratios that exceed approximately 90% of the geometric mean appraisal, no hammer ratios falling above that threshold could have been censored from our sample. Therefore, we apply Cohen's method, setting $h_0 = 0.90$, to obtain maximum likelihood estimates of the parameters of the uncensored lognormal distribution of hammer ratios. The results are $\hat{\mu} = \mathit{0.89}$ and $\hat{\sigma} = \mathit{0.59}$. The fact that $\hat{\mu} < 1$ indicates that, when the full sample is considered, including works that are either hammered down or bought in, expert appraisals tend to exceed the high bid (i.e., the all-bids benchmark). This contrasts, but is not inconsistent, with our earlier finding that expert appraisals tend to fall short of the successful-bid benchmark (cf. Table 2). Nor is it inconsistent with the possibility that appraisals are unbiased relative to the composite benchmark, which lays somewhere in the middle. By estimating the distribution of seller reserves, as in the next section, we are able to account for the influence of hidden reserves, and therefore to quantify the composite benchmark and test the null hypothesis that appraisals are unbiased against that standard.

B. The Latent Distribution of Reserve Ratios

As described in Section II, we use the estimated hammer ratio distribution, $F(\mathbf{h}|0.89,0.59)$, to estimate the number of works within each subinterval, $n_i = f(x_i)\times 4{,}174$. Paired with the observed sales, this provides an estimate of the number of bought-in works within each subinterval, n_i-m_i. These estimates, in turn, are treated as if they had been observed and used to identify the distribution of reserve ratios that maximizes the likelihood function, Eq. 5. The results are summarized in Table 3.

The first estimate $(\hat{\kappa})$ implies that 5.1% of all works are offered without reserve. Mei and Moses (2005, p. 2412) speculate that about 5% of all works are offered without reserve; our empirical result provides apparent confirmation. The next two parameters $(\hat{\nu},\hat{\omega})$ define the moments of the beta distribution, and determine the unconditional mean reserve ratio according to:[22]

$$E(r) = 0.9\times\left(1-\hat{k}\right)\times\left(\frac{\hat{v}}{\hat{v}+\hat{w}}\right) = 0.598. \tag{6}$$

Alternatively, the unconditional mean can be calculated directly from the estimated multinomial distribution:

$$E(r) = \sum_{i=1}^{30} x_i \times b(x_i), \tag{7}$$

where $$b(x_i) = \left(1-\hat{k}\right)\times\left[B(x_{i+1} \mid \hat{v},\hat{w}) - B(x_i \mid \hat{v},\hat{w})\right] = 0.598. \tag{8}$$

The fact that Equations 7 and 8 produce the same mean reserve ratio demonstrates that our discrete approximation of the continuous distribution has not distorted the results.

Another useful benchmark is the expected reserve ratio for the subset of works that are offered with *non-zero* reserve:

$$E(r \mid r > 0) = E(r)/(1-\hat{\kappa}) = 0.631. \tag{9}$$

Of utmost importance, given its role in Equation 1, is the expected reserve ratio for works that are bought-in, which may be calculated as follows:

---

[22] Moments of the beta distribution are discussed by Evan, Hastings, and Peacock (1993, pp. 31-33).

$$E(r \mid r > h) \;=\; \sum_{j=1}^{30} \sum_{i>j} x_i \, b(x_i) \, f(x_j) / \Pr ob(r > h) = 0.637 , \qquad (10)$$

We obtain some corroborating evidence for these results by comparing our estimated reserve prices to the lower pre-sale estimates. As shown in Table 3, the mean reserve for works carrying non-zero reserve averages 74% of the lower presale estimate (L), which is consistent with information provided by the major auction houses, who report that reserve prices are typically set at roughly 75 percent of the presale low estimate.[23] The estimated distribution of reserve ratios is shown in its entirety in Figure 3. Apart from the works that are offered without reserve, most carry reserve prices that range between 50% to 80% of the pre-sale geometric mean estimate. The fact that the estimated distribution places virtually no reserves beyond 85% of the appraised value is consistent with our prior assumption (based on auction house rules) that 90% is an effective cap on the reserve ratio.

The joint variation in hammer ratios and reserve ratios determines the distribution of bought-in works in our sample, as shown in Figure 4. The relative frequency of low versus high hammer ratios is represented by the overall height of the stacked bar chart—the lower segment reflects the portion of all works with a given hammer ratio that are bought in, and the upper segment the portion that are sold. The conditional probability of making reserve, given a particular hammer ratio, is represented by the ratio of the height of the upper segment to the total. This conditional probability quickly falls as the

[23] This figure is based on the reported standard guideline of the major auction houses of Sotheby's and Christie's with respect to sellers' reserves. Christie's stated policy is that the reserve price must be set below the low estimate and advises that it is usually between 70 and 80 percent for paintings. The only stipulation in Sotheby's policy, on the other hand, is that the reserve be below the low estimate, and while the standard is 75 percent of the low for paintings, it typically ranges between 50 to 100 percent of the low estimate depending on the value of the item, and sellers' preferences.

hammer ratio falls below 0.50, but is bounded at 5.1%, since that portion of all works are offered without reserve.

C. The Bias in Pre-Sale Estimates

We have shown that experts, in setting presale estimates, respectively over- or under-estimate *hammer prices* depending upon whether or not we account for the subset of works that were subsequently bought-in. As explained earlier, however, our null hypothesis is that experts provide an unbiased estimate of **v**, the minimum bid sufficient to purchase a given work. Using the results reported in Tables 1 and 3, we can now evaluate Equation 1 to examine the hypothesis that E(**v**) = 1:

$$E[\mathbf{v}] = 1.135\times0.701 + 0.637\times0.299 = 0.986. \quad (12)$$

Considering all 4,174 works that were brought to auction (and therefore all works that were appraised), the valuation ratio (**v**) averages 98.6%. Thus, expert appraisals are slightly high relative to the composite benchmark, but is the discrepancy statistically significant? To investigate this, we assume that for each work that was bought-in, the seller’s reserve ratio corresponds to E[**r**|**r**>**h**], or 63.7% based on Table 3. In this way we create for each work an “adjusted hammer ratio” that reflects the *true hammer* prices for works sold and the *conditional reserve estimate* for works bought-in. The variation in the adjusted hammer ratio across paintings in our sample is large relative to its deviation from 100%, yielding a t-ratio of –1.24. Moreover, not every bought-in work would (if we could observe the actual data) carry a reserve ratio equal to the conditional mean, which means there exists additional variation within our sample that we cannot observe or measure. Taking that additional independent variation into account would only reduce the reported t-statistic and reinforce our conclusion that, on average, the expert appraisal is statistically indistinguishable from the minimum amount a successful bidder would have had to pay in order to purchase the given work. Therefore, if we assume that experts take into account the valuations of both buyers and sellers, and consider the set of all works that are appraised, it is not possible to reject the null hypothesis that experts are unbiased.

C. Additional Orthogonality Conditions

Table 4 shows regression results from alternative models of hammer prices and ratios fitted to our sample of Impressionist paintings. These equations reveal the extent to which appraisals, although unbiased overall, might nevertheless be skewed within certain sub-samples. The first column contains results based on the sub-sample of works that made reserve and were sold, using actual hammer prices (H) as the dependent variable and the arithmetic mean estimate, AM =(L+U)/2, and simple spread, U-L, as the independent variables.

$$H_i \quad = \quad \beta_0 + \beta_1 \times AM_i + \beta_2 \times \left(U_i - L_i\right) + \xi_i \tag{13}$$

The second column also contains results based on the sub-sample of works that made reserve and were sold, using as dependent variable the actual hammer ratio (h), and as independent variables the geometric mean estimate (A) and pre-sale spread (S=U/L) measured as percentage deviations from their mean levels in the overall sample.[24]

$$h_i \quad = \quad \beta_0 + \beta_1 \times \left(\frac{G_i - \overline{G}}{\overline{G}}\right) + \beta_2 \times \left(\frac{S_i - \overline{S}}{\overline{S}}\right) + \xi_i \,. \tag{14}$$

[24] Linear transformations of G and S have no impact on the significance of the estimated coefficients but do affect the magnitude of the intercept. We employ deviations from means in these variables so that the intercept measures the mean of h.

The third column is the same regression as column two except that it encompasses the whole sample, including buy-ins, for which the dependent variable is measured as the *adjusted* hammer ratio.

Before discussing the results, it is important to appreciate that these regressions present a very high bar for testing the hypothesis that estimates are free of bias. In columns two and three, the independent variables are components of the dependent variable; thus measurement error in the reporting of auction estimates may show up as spurious dependence in the regression. In the first column, measurement errors in pre-sale estimates would shift the slope downwards toward zero, which in turn confounds the intercept, biasing it upwards. In addition, if reserve ratios vary across the sample as a function of the independent variables, then the slope coefficients could be reflecting that variation rather than a bias in expert appraisals. Moreover, since the reserve ratios used to form the adjusted hammer ratios that enter the third regression are presumably measured with error (but without bias), the significance levels of estimated coefficients are very likely overstated.

We have not excluded any observations based on trimming or outlier considerations. This is because the prospect of a legitimate high auction bid can have a profound effect on presale estimates if the goal of the expert is to incorporate all possible auction outcomes. While we have used care in checking the data, there may still be a few mis-measured observations because of errors in the original auction catalogues. These would also tend to skew results towards a finding of significant departures from unbiasedness.

In the first regression, ignoring buy-ins and using raw data, everything in the prediction model is significant and highly so. The negative intercept gives the impression that experts overestimate value holding everything else constant. The significant slope coefficients in this regression could be interpreted as evidence that certain types of work (e.g. high-value & low-uncertainty) are subject to different (less) bias than others.

In the second regression, where the actual hammer ratio serves as dependent variable, the results simplify considerably. The intercept (1.168) is significantly greater than unity, with a t-ratio of 5.19. In combination with the negative coefficient attached to the appraisal (albeit not quite significant), this pattern could be given a similar interpretation to the results in Mei and Moses (2005); i.e., that experts intentionally overestimate the value of high priced works, where the slope dominates but underestimate the value of low priced works where the intercept dominates.

In the third regression, where the adjusted hammer ratio is the dependent variable, the sample includes all works that were brought to auction and appraised, and incorporates the estimated reserve price for works bought in-house. Here, the intercept is not significantly different from unity (signifying unbiased overall appraisal) and neither the relative value of a given work nor its spread seems to skew the appraisal. Experts appear to be unbiased overall, unbiased on both the most and least valuable works, and unbiased on works with wide or narrow pre-sale spreads. From this equation we still cannot reject the hypothesis that experts are unbiased when all of the data are accounted for, including the seller's reserve valuation. From the first two regressions, however, it is clear that model mis-specification and selection bias may create the appearance of bias from a data set in which none exists.

VI. Conclusions

Without re-examining every previous data set used to investigate bias in expert appraisals of fine works of art, it is hard to conclusively argue that no biases exist or have ever existed. We know, after all, that traces of bias have been reported in the previous literature. Based on our own inquiry, however, we also recognize that in all such investigations subtle issues can taint the experiment and, unlike many predictive models, an incomplete experimental design may give the appearance of both average bias and systematic variations in degree. Moreover, the apparent significance levels in experiments concerning estimation bias will also be overstated if there are cyclical patterns in the market demand for art works. If time trends in demand are present but unpredictable, the reported precision of estimated coefficients will be overstated if the data are treated as independent observations. In effect, there are fewer independent observations than any given data set might suggest.

While the specification tests we employ fail to find evidence of bias, there are many additional variables that may be worth investigating in future research. For example, differences in appraisal methods and accuracy across auction houses and time periods merits attention. Also, the benchmark for judging value (and bias) itself may deserve further consideration. The spirit of our composite benchmark is hardly new; it follows from the standard economic definition of *market value*: the value that one could expect to emerge from negotiations between a willing buyer and willing seller. The closest comparable concept in auction research requires explicit account of the unwillingness of sellers to accept less than their reserve.[25] Although we strongly believe

[25] In an auction setting, the highest bid presumably leaves some rents to the winner and, as such, understates their maximum willingness to pay. On the other hand, the seller's reserve also tends to

that a composite appraisal—one that reflects anticipated bids and the seller's reserve—is more useful to potential bidders than any other, we acknowledge that the definition of the relevant price target has a profound impact on the question of bias—as it should.

As a final comment, the influence of buy-ins should also be recognized when analyzing the monetary returns to investments in art, or other real assets where transactions are subject to a seller's reserve. As Goetzmann and Peng (2006) have demonstrated in the context of search models, if buy-ins are ignored, estimated holding period returns and the degree of financial risk will both be misrepresented. Despite this, buy-in data have been almost exclusively ignored in the growing volume of studies on art investment and pricing. This means that the implicit consumption yield of art ownership required to offset the difference between rates of return earned on financial instruments and those on investment in fine art has been, so far, understated. Interestingly, the bias from ignoring buy-ins is not avoided by focusing on repeat sales since these constitute an incomplete and biased sample. Treating the buy-in dimension within a time series model of art valuation should prove a challenging avenue of future research.

---

overstate the seller's valuation if the seller uses the reserve to entice higher bids (see Krishna [2002] and references). Our conclusions abstract from these more subtle, and somewhat offsetting, valuation issues which impinge on the link between valuation and all transaction data.

References

Ashenfelter, O. (1989) "How Auctions Work for Wine and Art." *Journal of Economic Perspectives*. 3(3), 23-36.

Ashenfelter, O. and Graddy, K. (2002) "Art Auctions: A Survey of Empirical Studies." NBER Working Paper No. 8997. National Bureau of Economic Research: Cambridge, Massachusetts.

Ashenfelter, O. and Graddy, K. (2003) "Auctions and the Price of Art." *Journal of Economic Literature.* 41(3), 763-787.

Bauwens, L., and Ginsburgh, V. (2000) "Art Experts and Art Auctions. Are Presale Estimates Fully Informative?" *Louvain Economic Review*. 66(1), 131-144.

Beggs, A., and Graddy, K. (1997) "Declining Values and the Afternoon Effect: Evidence from Art Auctions." *RAND Journal of Economics*. 28(3), 544-565.

Chanel, Olivier, Gerard-Varet, Louis-Andre, and Vincent, Stephanie (1996) "Auction Theory and Practice: Evidence from the Market for Jewellery." *Economics of the Arts—Selected Essays*, V. A. Ginsburgh and P. M. Menger (editors), pp. 135-149, Elsevier Science B.V.

Cohen, A. (1959) "Simplified Estimators for the Normal Distribution When Samples are Singly Censored or Truncated." *Technometrics*. 1, 217-237.

Ekelund, R., Ressler, R., and Watson, J. (1998) "Estimates, Bias, and 'No Sales' in Latin-American Art Auctions, 1977-1996." *Journal of Cultural Economics*. 22, 33-42.

Elyakime, Bernard, Laffont, Jean-Jacques, Loisel, Patrice, and Vuong, Quang (1997) "Auctioning and Bargaining: An Econometric Study of Timber Auctions with Secret Reservation Prices." *Journal of Business & Economic Statistics*. 15(2): 209-220.

Evans, M., Hastings, N., and Peacock, B. (1993) *Statistical Distributions*, 2nd ed. New York: John Wiley and Sons.

Gershkov, Alex, and Toxvaerd, Flavio (2008) "On Seller Estimates and Buyer Returns." Discussion Paper No. 143, GESY, University of Mannheim.

Goetzmann, William N., and Peng, Liang (2006) "Estimating House Price Indices in the Presence of Seller Reservation Prices." *Review of Economics and Statistics*. 88(1), 100-112.

Ji, Lu and Li, Tong (2008) "Multi-round Procurement Auctions with Secret Reserve Prices: Theory and Evidence." *Journal of Applied Econometrics*. 23(7), 897-923.

Krishna, Vijay (2002) *Auction Theory*. San Diego: Academic Press.

Mei, J. and Moses, M. (2005) "Vested Interest and Biased Price Estimates: Evidence from an Auction Market." *Journal of Finance*. 60(5), 2409-2435.

Mei, J. and Moses, M. (2002) "Art as an Investment and the Underperformance of Masterpieces." *American Economic Review*. 92(5), 1656-1668.

Milgrom, Paul R. and Weber, Robert J. (1982) "A Theory of Auctions and Competitive Bidding." *Econometrica*. 50(5), 1089-1122.

Paarsch, Harry J. and Hong, Han (2006) *Introduction to the Structural Econometrics of Auction Data*, Cambridge, MA. MIT Press.

Perry, Motty, and Reny, Philip J. (1999) "On the Failure of the Linkage Principle in Multi-Unit Auctions." *Econometrica*. 67(4), 895-900.

Pesando, J. (1993) "Art as Investment: The Market for Modern Prints." *American Economic Review*. 83(5), 25-32.

Salamon, Martin (1992) "Auction of Works of Art," in Ruth Towse and Abdul Khakee (eds), *Cultural Economics*, Chapter 9, Berlin/Heidelberg: Springer-Verlag, 85-89.

Stein, Gertrude (1935) "What are Master-pieces and Why are There So Few of Them," delivered at lecture at Oxford and Cambridge in 1936, and reprinted in *The Best American Essays of the Century*, Joyce Carol Oates, ed., Boston: Houghton Mifflin Co., 2000.

Table 1: Auction Statistics for French Impressionists, 1985-2001

	Number	Average Buyer's Premium[1]
Works with complete data	4174	12.54%
Works with incomplete data	86	

Summary of works with complete data:	By Number	By Value[2]
Works sold	2,925	$3,015,452,272
Buy-ins	1,249	$ 890,153,902
Works at auction	4,174	$3,905,606,174
Percentage of Buy-ins	29.9%	22.8%

[1] Fee paid by buyer (as percent of purchase price) to compensate auction house for services rendered.
[2] "Works sold" are valued as the sum of actual hammer prices. "Buy-ins" are valued (for purposes of this table) as the geometric mean of the upper and lower presale estimates.

Table 2. Auction Price Statistics for French Impressionists, 1985-2001

Total value of works sold	$3,015,452,272
Number of works sold	2,925
Average price (average value of high bid, H_i)	$1,030,924
Average expert appraisal, (average value of $A_i = (L_i \times U_i)^{1/2}$)	$960,534
Mean hammer ratio (average value of H_i/A_i)	1.135
Standard deviation of hammer ratio	0.673
t-ratio of mean hammer ratio*	10.85
Sample correlation between A_i and $\lvert H_i - A_i \rvert$	0.65
Sample correlation between A_i and $\lvert h_i - \bar{h} \rvert$ |	-0.03

*The t-ratio equals the mean hammer ratio minus unity, divided by the standard deviation. Based on the *arithmetic* average of upper and lower estimates, the mean hammer ratio is 1.11, with a t-ratio of 9.63.

Table 3: Estimated Distribution of Sellers' Reserve Ratio

Population Parameters	Estimate	
κ	0.051	
ν	51.8	
ω	22.1	
	Seller's Reserve Relative to	
Implied Reserve Ratios:	Geometric Mean	Low Estimate
a. Unconditional Mean		
- inferred from population parameters	0.598	
- calculated from discrete probabilities	0.598	0.692
b. Conditional Mean, given r > 0	0.631	0.729
c. Conditional Mean, given r > h	0.637	0.737

Table 4. Regression Results and Diagnostic Tests

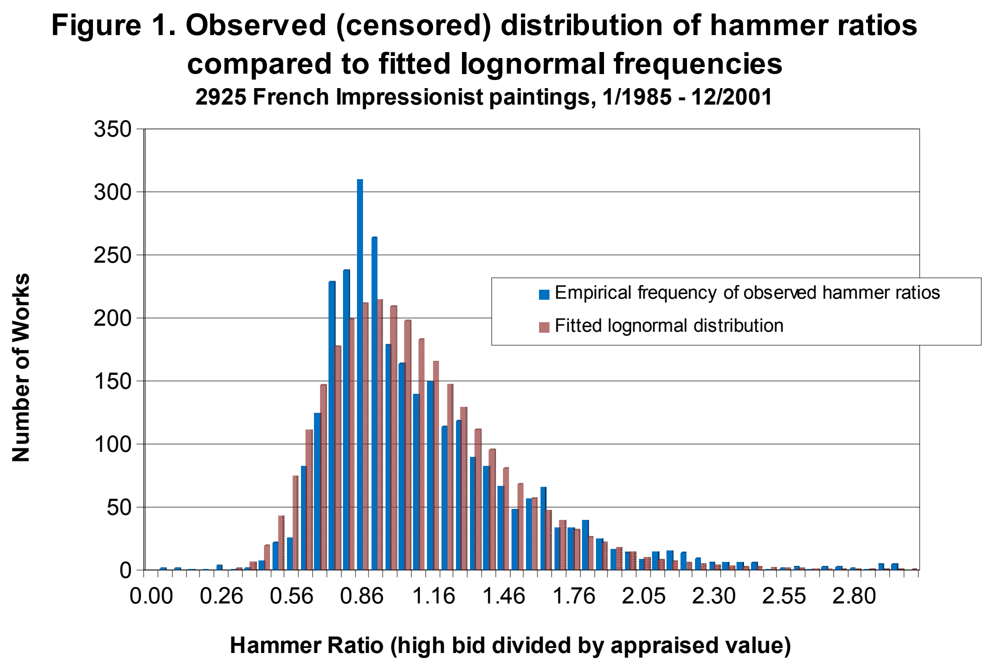
Variables (t-values)	Censored Sample 1. Hammer Price[1]	Censored Sample 2. Hammer Ratio[2]	Full Sample 3. Adjusted Hammer Ratio[3]
Intercept	-152[**] (-3.75)	1.168[**] (5.19)	0.986 (-1.51)
Mean Estimate	1.22[**] (6.11)	-0.055 (-1.48)	-0.003 (-0.80)
Spread	-115[**] (-3.31)	-0.020 (-0.54)	0.220 (1.48)

[1] Column 1 is in $1000 and uses the raw hammer prices as dependent variable. "Mean Estimate" is the arithmetic mean of upper and lower estimates, "Spread" is the upper minus lower estimate. The t-ratio for Mean Estimate is relative to unity.

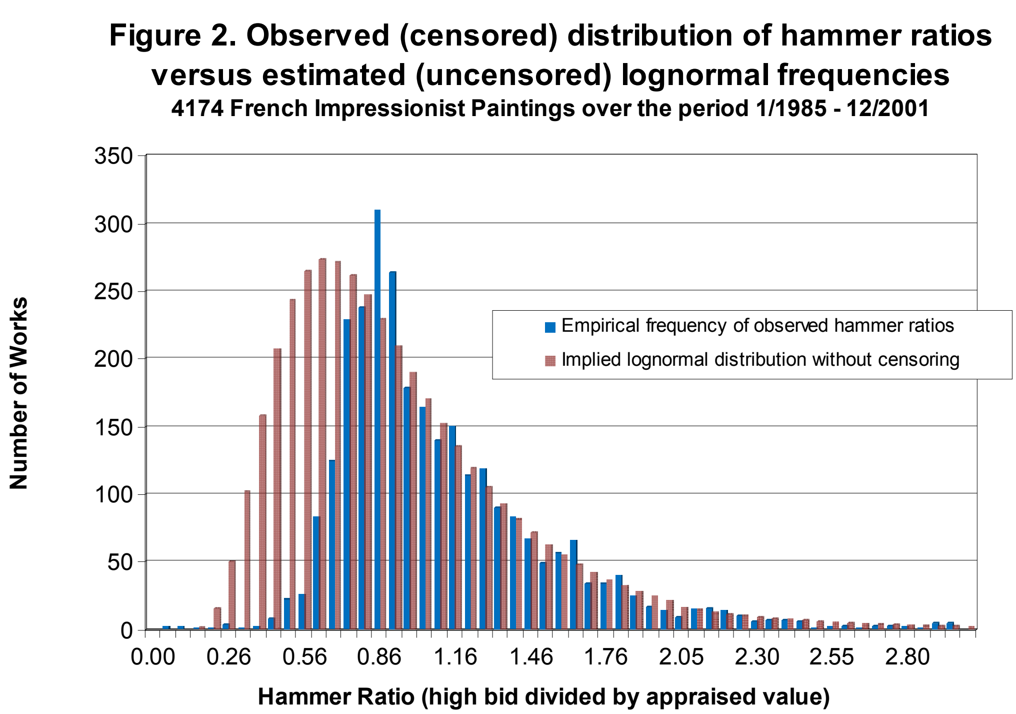
[2] Column 2 uses the hammer ratio for sold works and excludes works bought-in. "Mean Estimate" is the percentage deviation of the geometric mean estimate for a given work relative to the sample geometric mean. "Spread" is the percentage deviation of the square root of U/L for a given work relative to the sample average. The t-ratio for Intercept is relative to unity.

[3] Column 3 uses the adjusted hammer ratio and includes all works, whether sold or bought in. "Mean Estimate" and "Spread" are as in column 2. The t-ratio for Intercept is relative to unity.

[**] Coefficient is significant at 1% level.



Lognormal parameters are based on best fit to observed distribution of hammer ratios. Best fit defined as the lowest KS D-statistic.



Lognormal parameters are best fit for data above 0.9 in recognition of possible censoring

Figure 3. Inferred Distribution of Seller's Reserve

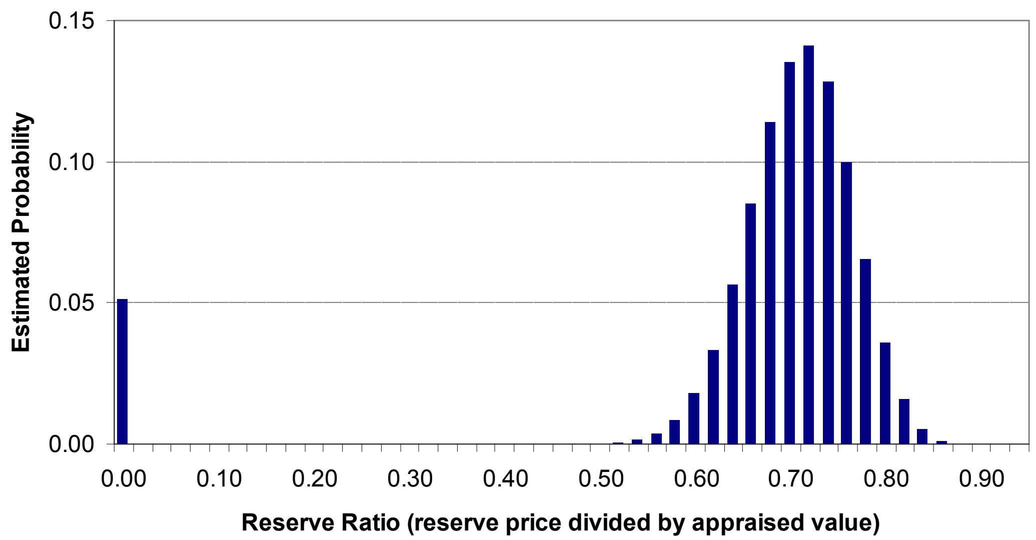


Figure 4. Joint probability of hammer ratio & sale: p(h∩h>r) and q(h∩r>h)