Econometric fine art valuation by combining hedonic and repeat-sales information

John W. Galbraith
Department of Economics, McGill University

Douglas J. Hodgson *
Dépt. de sciences économiques, Université du Québec à Montréal

December 11, 2017

ABSTRACT

Statistical methods are widely used for valuation (prediction of the value at sale or auction) of a unique object such as a work of art. The usual approach is estimation of a hedonic model for objects of a given class, such as paintings from a particular school or period, or in the context of real estate, houses in a neighborhood. Where the object itself has previously sold, an alternative is to base an estimate on the previous sale price. The combination of these approaches has been employed in real estate price index construction (e.g. Jiang, Phillips and Yu 2015); in the present context we treat the use of these different sources of information as a forecast combination problem. We first optimize the hedonic model by considering the level of aggregation that is appropriate for pooling observations into a sample, then apply model-averaging methods to estimate predictive models at the individual-artist level, in contexts where sample sizes would otherwise be insufficient to do so. Next we consider an additional stage in which we incorporate repeat-sale information, in a subset of cases for which this information is available. The results are applied to a data set of auction prices for Canadian paintings. We compare the out-of-sample predictive accuracy of different methods and find that those that allow us to use single-artist samples produce superior results, that data-driven averaging across predictive models tends to produce clear gains, and that where available, repeat-sale information appears to yield further improvements in predictive accuracy.

Key words: art market, auction prices, hedonic model, model averaging, repeat sales
JEL Classification number: Z11

*We thank the Fonds québécois de la recherche sur la société et la culture (FQRSC) and the Social Sciences and Humanities Research Council of Canada (SSHRC) for financial support of this research, and the Centre Interuniversitaire de recherche en analyse des organisations (CIRANO) for research facilities.
1 Introduction

In the secondary art market, the only way to know with certainty the value placed on an art work at a particular point in time is to see the work be offered for sale and purchased. Nonetheless there are many art market participants who would like to have an approximate idea of the worth of a work at a point in time, for whom this datum is unavailable, because the work in question has been sold privately, has not been sold recently, or will not be sold. An owner who is considering offering a work for sale will often want to know what it might fetch before consigning it to an auction house; the valuation of an individual work or indeed an entire collection can be necessary for insurance or taxation.

Although expert appraisers can be consulted, and are employed by auction houses to set the pre-sale estimates reported in auction catalogues, in many cases expert appraisal may be more costly than necessary given the degree of precision required or the likely value of the work. In such cases, the application of a statistical model that averages over past sale prices of art works that are in some way similar to the one under consideration can provide a quick and relatively inexpensive estimate of the likely sale price. Indeed, there are agencies that provide such estimates on a commercial basis (Artprice.com, Blouin Art Index, Beautiful Asset Advisors), and statistical models may also be useful in providing a baseline estimate for a professional appraiser. It is, of course, also conceivable that sophisticated econometric models could match or even exceed the accuracy of professional appraisers, or could be a valuable complement to an expert’s opinion.

The usual statistical model of this type is the hedonic regression, in which various observable characteristics of an item are used to estimate its value. Well established in the literature on real estate pricing (see for example Meese and Wallace (1997), and Hodgson, Slade, and Vorkink (2006)), there is also a large body of research that applies hedonic regression to the analysis of the pricing of collectible objects, including art works (see for example Hodgson and Vorkink 2004). In recent real estate literature, hedonic models have also been combined with information from repeat sales of the same property; see in particular Jiang, Phillips and Yu (2015). Hedonic regression is applied when we have a number of observations on sales of objects that, although individually unique, have some degree of commonality, so that the likely value of an item can be inferred from sales of other items. In Hodgson and Vorkink (2004), these objects are paintings created by important (as judged by major art-historical survey books) Canadian painters; Arvin and Scigliano (2004) consider the Canadian Group of Seven painters; Atukeren and Seckin (2009) study Turkish paintings, and Helmanzik (2009, 2010), studies prominent modern artists. A typical hedonic data set contains sale-specific data such as the price and date of sale (as well as the auction house when the observations come from auction sales), as well as object-specific characteristics such as the size, materials, date of production, and artist and subject matter (in the case of paintings).

Hedonic regressions in the context of a particular art market often group all artists together in one estimation sample, in which most of the hedonic covariates are presumed to have the same proportionate effect on the value of a painting regardless of artist. Differences among artists are captured by individual-artist dummies. Such a specification has clear shortcomings, because the effect on valuation of a given characteristic, for example the age
of the artist at the date of execution of the work, may differ across artists; in fact there is a growing literature that explores the effects of various forms of disaggregation of a large group of artists on the estimated age-valuation profile (e.g. Galenson (2000,2001), Galenson and Weinberg (2000,2001), Hellmanzik (2009, 2010), Accominotti (2009), and Hodgson (2011)). The effects of disaggregation by, for example, birth cohort, can be striking. Researchers have generally avoided the estimation of individual-artist level hedonic models, however, because the number of degrees of freedom, although ample in a model where many artists are pooled together, can be prohibitively small at the level of individual artists.

When a work is known to have sold previously, this prior-sale information can also be used to predict the value at the next sale. However, market conditions may have evolved since the previous sale, particularly if it was relatively long ago; combining both types of information may lead to better estimates, as noted earlier with respect to construction of price indices for real estate. Using the hedonic characteristics together with the previous-sale information may be seen as a forecast combination problem.

The present paper investigates the problem of improving upon art valuations from simple hedonic regression methods, using two potential efficiency improvements. First, as in Galbraith and Hodgson (2012), we consider the estimation of hedonic regressions for individual artists with relatively few degrees of freedom by applying model-averaging methods. We can then examine whether an artist-specific hedonic model average may be able to deliver hedonic predictions of the sale price of a painting that are superior to those provided by a single model. Second, we suggest methods for incorporating repeat-sale information, where available for a particular work of art, into a hybrid prediction. These hybrid hedonic model average/repeat-sale methods are evaluated with respect to predictive performance.

We examine the out-of-sample predictive ability of these methods using auction data for a set of sixty-four prominent Canadian painters over the period 1968-2015. A sequence of rolling estimation samples is used, with all data up to a particular year used to obtain parameter estimates which are then used to predict the prices of all art works for the following year. Data from the years 2001 through 2015 are used for pseudo-out-of-sample prediction. Measures of prediction loss (such as root mean squared error) are then computed for each method and artist, and comparisons are drawn between the predictive power of alternative methods. Model-average results are compared with simple least-squares regressions, and the usefulness of repeat-sale information is evaluated.

The next section of the paper describes the econometric methods, and the data and empirical results are presented in Section 3.

2 Econometric methods

The econometric valuation of art objects has traditionally relied on hedonic methods, by which the value of an object is assumed to derive from characteristics that are at least in part observable, one of which is the identity of the artist. While these methods have

\[1\text{Galbraith and Hodgson (2012) focus on the problem of estimating individual-artist age-valuation profiles and find that the methods adopted—both model averaging and dimension reduction—do yield useful results at the level of individual artists.}\]
also long been used in real estate valuation, recent literature (particularly with respect to price index construction) has increasingly moved toward estimates based on past sales (the repeat-sales method, ‘RSM’), or the combination of hedonic methods with repeat sales.

There are many similarities between the problems of valuing individual art objects, and real estate. In each case there is a market for individual items which trade infrequently, and which are individually unique but nonetheless have some degree of commonality with other items. In each case hedonic models have had some success in capturing characteristics which clearly have important effects on sale prices, but the processes determining the price may differ markedly for different classes of object. In each case as well the number of observations which are repeat sales tends to be a small proportion of a sample of data.\(^2\)

2.1 Hedonic models

We will adopt the notation of Jiang, Phillips and Yu (2015), making only such modifications as are necessary for the problem at hand. Let, therefore, \(y_{i,j,z}\) be the logarithm of the price of the \(i\)th work of art, sold for the \(j\)th time, by artist \(z\). In a large majority of cases, a work is observed to sell only once, and so \(j = 1\). Nevertheless the data set does contain many examples with \(j \geq 2\). The date of sale corresponding to this observation is \(t(i,j,z)\), and is observed here only at the level of the year; the last year of the sample is \(T = 2015\). A general hedonic formulation (Jiang, Phillips and Yu 2015, eq. (1)) is:

\[
y_{i,j,z} = c + \beta t(i,j,z) + \gamma'X_{i,,z} + h_i + \epsilon_{i,j,z},
\]

(1)

The \(X_{i,,z}\) are observable characteristics of the work (unchanging with \(j\)), which are presumed in a hedonic specification to give utility and therefore affect value. In this specification there is an artist-specific effect and a work-specific effect as well. This specification is appropriate to a sample that pools the works of all artists; a corresponding specification for an individual artist is

\[
y_{i,j} = c + \beta t(i,j) + \gamma'X_i + h_i + \epsilon_{i,j},
\]

(2)

where we now suppress the subscript \(z\) since it is a constant within each artist-specific sample. Unless parameters are identical for each \(z\), in which case specification (1) is unambiguously more efficient, there is a bias–variance trade-off between (1) and (2) which we examine empirically below. The small number of repeat sales \((j > 1)\) in our sample means that we have little information with which to estimate \(h_i\) separately from the constant; it is therefore suppressed in our individual-artist specifications.

Estimation by least squares (LS) is generally straightforward: endogeneity of regressors is not normally a problem in modeling secondary market prices at auctions which take place many years after the determination of characteristics of the artwork (a possible exception to this rule lies in auction house dummy variables, if these are used as characteristics of the work).

\(^2\)Repeat sales occur when the same item is sold at least twice a given sample data, so that an earlier sale may be used to predict the price of a later one.
2.2 Model averaging

The common problem of uncertainty concerning model specification particularly afflicts hedonic models of art sales where, as here, we have a large set of potential explanatory variables. In the pooled sample of nearly 10,000 observations, there are sufficient degrees of freedom to include all regressors; by contrast if we wish to allow all parameters to differ by artist, the number of degrees of freedom for each individual is too small to allow inclusion of all observable indicators. In this case we need to consider subsets of the possible regressors, but with many available indicators it may be unclear which to include for optimal prediction.

One approach to conserving degrees of freedom that has been successful in a number of forecasting contexts is model averaging, where results from models with different sets of included regressors are averaged, as opposed to choosing only one of the models; model averaging allows us to use information from a large number of potential explanatory effects even where sample sizes preclude including each one in a single regression.\(^3\)

Model averaging is usually accomplished using sample-based criteria, such as information criteria, to select weights. Hansen (2007) proposed an alternative weighting based on the Mallows (1973) \(C_p\) criterion. Within a class of discrete averaging estimators, Hansen showed that this criterion has a minimum mean squared error property, and we adopt this criterion here. For use as a forecasting procedure, we retain the parameter estimates from each of the estimated models with its corresponding weight, produce forecasts from each of the sets of parameters, and produce an overall forecast by weighting these forecasts by the set of model weights.

Galbraith and Hodgson (2012) use the model-average method in the context of artists’ age–valuation profiles, and summarize the steps as follows, but see Hansen (2007) for a detailed exposition. Consider, in the notation of Hansen, the model

\[
y = c + W\Theta + \epsilon, \tag{3}
\]

where \(W\) represents a \(k\)-dimensional set of potential explanatory variables; the columns of \(W\) are ordered in accordance with an \(a\ priori\) view of their importance. Select a maximum number \(M\) of regressors to include, such that \(M < N\) where \(N\) is the sample size, and estimate all models with constant and \(m\) regressors, \(m = 1, 2, \ldots, M\). The model average estimator of \(\Theta\) is \(\hat{\Theta} = \sum_{m=1}^{M} \omega_m \hat{\theta}_m\), where \(\omega_m\) is the weight given to model \(m\), \(\sum_{m=1}^{M} \omega_m = 1\), and \(\hat{\theta}_m\) is the parameter vector estimated using a model with \(m\) regressors, padded with zeroes for excluded variables in the cases for which \(m < M\). The weights \(\omega_m\) are chosen to minimize the Mallows criterion \(SSR + 2\sigma^2 k(\omega)\), where \(\sigma^2\) is replaced with an estimate and where \(k(\omega) = \sum_{m=1}^{M} \omega_m k_m\) is the effective number of parameters, \(k_m\) being the number of parameters estimated in model \(m\).

Let the vector of forecasts from model \(m\) be \(\hat{y}_m = \hat{c}_m + W\hat{\theta}_m\). Then the model-average forecast is \(\hat{y}_{MMA} = \sum_{m=1}^{M} \omega_m (\hat{c}_m + W\hat{\theta}_m)\), again with \(\sum_{m=1}^{M} \omega_m = 1\).

\(^3\)Jiang, Phillips and Yu (2015) quote Shiller (2008) as follows: “there are too many possible hedonic variables that might be included, and if there are \(n\) possible hedonic variables, then there are \(n!\) possible sets of independent variables in a hedonic regression, often a very large number. One could strategically vary the list of included variables until one found the results one wanted.” Model averaging mitigates this problem, providing results based on objective criteria rather than the investigator’s selections.
2.3 Combining hedonic and repeat-sale information

A traditional statistical principle states that, when one has information from multiple different sources, it is in general optimal to combine information from the different sources rather than to use one of them alone. In the special case where independent unbiased sources have known finite variances, and one is interested in minimizing the overall variance of a prediction, it is optimal to combine the sources in inverse proportion to their variances (Bates and Granger, 1969).

Here, albeit in a relatively small subset of cases, we identify a previous sale of a given artwork in the data set which can serve as the basis of an alternative forecast. We make this identification by searching for a previously-sold work by a given artist that has identical measured characteristics. This means that we could in principle incorrectly identify as a match two different works with identical dimensions, year of production, medium, genre and so on; the converse error, failing to identify a genuine repeat sale, should never arise unless there are undiscovered coding errors in the characteristics.

The previous sale price, updated with any information we have concerning changes in average prices over the relevant time period, provides an alternative forecast of the current value at auction. We treat the incorporation of this information, where available, in the classical forecast combination framework.

For regressions on data from a single artist, let $\hat{y}_{i,j}$ be the hedonic regression forecast of the log price, and $\tilde{y}_{i,j}$ be the repeat sale forecast, that is $\tilde{y}_{i,j} = \tilde{y}_{i,j-1} + \beta_{t(i,j)} - \beta_{t(i,j-1)}$. A combined forecast is then

$$\hat{y}_{i,j} = \alpha \hat{y}_{i,j} + (1 - \alpha) \tilde{y}_{i,j}. \quad (4)$$

An optimal weight $\alpha$ may be estimated by training based on cases observed at times preceding $t_{i,j}$. However, it has often been observed in the literature on forecast combination (see for example Clemen 1989 for a review) that it is difficult to improve upon equal weighting ($\alpha = 0.5$) for two forecasts that are close to being of equal variance. The small individual-artist sample sizes here constrain us to use this equal weighting in this problem.

3 Empirical predictions of value at auction

The data set that we use for empirical comparison of methods is an extended version of that used by Hodgson (2011) in a study of age-valuation profiles for Canadian painters. In the pooled model, all of the artists included in the sample are grouped together, whereas in the individual-artist models, a separate regression is estimated for each artist. The artists included, along with numbers of observations, are listed in Appendix 1. For each of the sixty-three painters analyzed, we record every sale of an oil or acrylic work sold at auction during the years 1968-2015 for which date of execution was provided in the auction record, with data also collected on the variables mentioned below. The sources for the data were summaries of sales of Canadian paintings at auction as reported in Campbell (1970-75, 1980), Sotheby’s (1975, 1980), and Westbridge (1981-2015). Painters from various periods of Canadian art history were considered, with artists of different gender and working in a variety of styles, and of different geographical and linguistic backgrounds.
The full pooled sample size comprises 9891 works, individual sample sizes range from 50 to 997 with a median of 107, and we have as many as 80 regressors, many of which are indicators for qualitative features of the work of art.

Densities of values of paintings are generally heavily right-skewed, and models are usually defined on the logarithms of prices. Figure 1 shows boxplots of these log prices. We see that, as is typical in data on sale prices of works of art, there is wide variation in the locations and dispersions of the 64 distributions, and that there is relatively modest asymmetry remaining in the logarithmically transformed prices.
Figure 1: logarithms of sale prices at auction, 9891 paintings, 64 artists

Each of the individual plots corresponds to one of the 64 artists named in Appendix 1, in the order given there.

---

4Each of the individual plots corresponds to one of the 64 artists named in Appendix 1, in the order given there.
3.1 Model elements

Hedonic models for the prices of art works are usually estimated with the dependent variable as log-price of the sale of a painting at auction, and the (possibly large) set of independent variables includes the artist’s age at the time of execution of the painting, included in polynomial form, the height, width, and surface area of the painting, dummy variables for the date of the sale, the medium/support combination, the genre of the painting’s subject matter, and, in specifications where different artists are pooled together, dummies for artist identity.

In the literature on the age-valuation relationship, age generally enters as a polynomial function with two to four terms, reflecting the modeling of age-wage or experience-wage profiles in the labor economics literature (e.g. Mincer 1974). Hodgson (2011) estimates a fourth-order polynomial for the overall and cohort regressions, as well as for the specification with continuously shifting profile. Galenson (2000) allows for up to three polynomial terms in each individual-artist regression, selecting order by $R^2$. Here we include a quadratic specification for each artist, with the parameters in the pooled version of the model allowed to shift continuously in year of birth in the fashion described by Hodgson (2011).

3.2 Forecast evaluation

We have data on auctions that take place over $T$ periods of time (years): we choose a break point, $T^*$, and initially estimate the model for all data observed for $t = 1, \ldots, T^*$. The resulting parameter estimate, $\hat{\beta}(T^*)$, is then used to calculate an out-of-sample prediction for each painting $i$ sold in the period $T^* + 1$, using the fitted value $\hat{p}_i = \hat{\beta}(T^*)'x_i$. The quality of the prediction will be judged by a comparison of $\hat{p}_i$ with the realized auction price for this painting, $p_i$. The only unusual aspect arising here is that, for any painting observed in period $T^* + 1$, we do not have a parameter estimate for the appropriate time dummy, as of course this dummy did not appear as a regressor in the estimation sample that extended only to period $T^*$. There are various ways to deal with this problem where time dummies are used; in principle, a forecasting model could be applied to use the time dummies for periods $t = 1, \ldots, T^*$ to compute a time series forecast of the $T^* + 1$ dummy. We will suppose however that market values can be approximated by a simple random-walk process and use the estimated time dummy for $T^*$ as our forecast of that for $T^* + 1$.\footnote{In the context of housing markets, forecasting models of the time dimension have been considered by, for example, Clapp and Giacotto (2002). In many cases of both art and real estate, however, a valuation is made conditional on sales already observed in the current year, so that the time lag between price index update and sale is very small, and any deviation of the price process from a random walk should have correspondingly small effect.}

The above procedure is then repeated with the estimation sample extended to $T^* + 1$ and predictions of sale prices in period $T^* + 2$, and so on. Our initial estimation sample comprises data up to and including the year 2000 ($T^*$), and predicted values are computed recursively for the years 2001 through 2015. The data set does not contain information on the month or date of sale, so no timing distinction is made among sales taking place in a given year. This method is applied to the data for all artists pooled together, where at each stage the model is estimated by LS. We then estimate the regressions at the individual-
artist and clustered levels, as described above. For learning algorithms the model structure differs, but the essentials of the exercise are the same.

After computing forecasts \( \hat{y} \) for each observation in the successive out-of-sample periods 2001...2015, we compute overall measures of forecast loss. In particular, we report the root mean squared error (RMSE), \( \ell_{\text{rmse}} = \left[ n_f^{-1} \sum_{i=1}^{n_f} (\hat{p}_i - p_i)^2 \right]^{1/2} \), where there are \( n_f \) forecasts to evaluate. Because the loss associated with an inaccurate forecast may be asymmetrical—in the case where a forecast is made for a potential seller, for example, it may be more costly to underestimate the potential sale price at auction than to overestimate it—we also report a widely used asymmetric loss function, the linex (linear-exponential) loss, in the form \( \ell_{\text{lex}} = n_f^{-1} \sum_{i=1}^{n_f} [\exp(-a(\hat{p}_i - p_i)) + a(\hat{p}_i - p_i) - 1] \), where \( a \) is a shape parameter which we set to 0.5.

3.3 Predictive results

We now consider whether refinements of the standard hedonic technique allow us to improve forecasts, on the sample of 9891 works sold at auction, with 5157 observations in the period 2001-2015 used for pseudo-out-of-sample evaluation, as described above. We compare the standard pooled OLS regression on all available regressors (‘OLS-broad’, ‘pooled’) and individual-artist regressions using one of five specifications: the largest set of regressors which can be included (‘OLS-broad’), a minimal OLS model involving only the core characteristics of painting height and width, a quadratic polynomial in artist’s age at completion of the painting and a linear specification of time (‘OLS-core’), the Hansen-Mallows model average (‘model average’), and finally the equally weighted combination of repeat-sales information with each of the two previous methods. In the latter two cases we again report in Table 1 the global results on the set of 5157 test cases, although repeat-sale information is available only in about one-fifth of these cases. Results specific to the cases where repeat sales are identified are reported in Table 3.

The empirical results address several questions:

- What level of aggregation of artists into groups produces the best bias-variance trade-off (and therefore lowest forecast loss)? Table 1 and Figures 1 and 2 summarize the differences in forecast loss.

- Does model averaging produce statistically superior forecasts in individual-level regressions? Table 1 reports these forecast losses, and Table 2 reports tests (Diebold-Mariano) of the null of equal forecast variance across various methods.

- Does incorporation of repeat-sale information (via previous sale hedonic regression residual, in this case) offer potential further improvement? Table 3 treats the subset of cases in which repeat-sale information is available, compares estimated forecast losses, and gives tests of the null of equal forecast variance with and without repeat-sale information.

- What are the key observable factors that predict sale prices at auction?
For symmetric loss, Table 1 clearly suggests that results at the individual-artist level tend to be superior to pooled results. The preferred model is an individual-level model-average forecast that incorporates repeat-sale information where available, whichever of the loss functions one uses. Statistical inference on these differences is provided in Table 2 below.

Table 1
Forecast loss (RMSE, linex) by method and aggregation:
full sample of 5157 predicted values

<table>
<thead>
<tr>
<th>Method</th>
<th>RMSE:</th>
<th>Linex:</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>pooled</td>
<td>indiv.</td>
</tr>
<tr>
<td>OLS-broad</td>
<td>0.824</td>
<td>0.783</td>
</tr>
<tr>
<td>OLS-core</td>
<td>—</td>
<td>0.769</td>
</tr>
<tr>
<td>model avg</td>
<td>—</td>
<td>0.750</td>
</tr>
<tr>
<td>OLS-core+RSM</td>
<td>—</td>
<td>0.754</td>
</tr>
<tr>
<td>model avg+RSM</td>
<td>—</td>
<td><strong>0.736</strong></td>
</tr>
</tbody>
</table>

Figure 2 shows a sequence of forecast RMSE loss reductions, \(\frac{1-[\text{loss(model average+repeat sales)}]}{\text{loss(single hedonic model)}}\), across the sample of sixty-four artists (i.e. positive values indicate an improvement in forecast performance measured by RMSE). Linex results are qualitatively similar and are not depicted. Hansen-Mallows model averaging plus repeat-sales information (‘MMA+RSM’) is compared with two alternative hedonic regression models with explanatory variables chosen a priori; the first includes as broad a set of indicator variables as is feasible on the limited sample for the particular artist, and the second is OLS-core.\(^7\)

---

\(^6\)Table 1 records loss function measures by method; the best (lowest) is in **bold** for each measure. One large extreme outlier is removed for the individual broad OLS regression and equally-weighted model average, in Table 1 and 2 results involving either of these two methods; these methods are not competitive and so rankings are not affected, but removal of the outlier gives a more accurate impression of their typical performance.

\(^7\)The size of dot for each artist is proportional to individual sample size. Results from equally weighted model averages are relatively erratic and so are not depicted.
Despite the fact that in overall results it is clear (Table 1) that the refinements (model averaging, repeat sales) tend to produce some gains, there is considerable variation in results across individual artists visible in the figure, and numerous artists for whom simple hedonic models provided slightly better results in this sample.

Table 2 provides Diebold-Mariano (1995) test results. These test statistics do not account for parameter uncertainty, and so can be expected to be liberal, but are nonetheless indicative of clear differences.

---

8The figure displays \[1 - \text{loss(combined model average + RSM)}/\text{loss(hedonic model)}\] for the 64 artists named in Appendix 1. Positive values therefore indicate reduced loss (superior performance) of the model average, with repeat sales where available, relative to the single OLS hedonic model for a particular artist. Model 1: hedonic regression with selected dummies + painting dimensions, surface area, age of artist at time of execution, and age squared. Model 2, the same without dummies. The sizes of dots are proportional to total sample sizes for the artist.
Table 2

Inference on $H_0$: equal predictive accuracy of methods:
full sample of 5157 predicted values, D-M statistic$^9$

<table>
<thead>
<tr>
<th>Method:</th>
<th>vs. 1</th>
<th>vs. 2</th>
<th>vs. 3</th>
<th>vs. 4</th>
<th>vs. 5</th>
<th>vs. 6</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.pooled</td>
<td>2.64</td>
<td>3.77</td>
<td>5.15</td>
<td>4.79</td>
<td>6.13</td>
<td></td>
</tr>
<tr>
<td>2.OLS-broad</td>
<td>1.88</td>
<td>6.04</td>
<td>3.68</td>
<td>7.90</td>
<td></td>
<td></td>
</tr>
<tr>
<td>3.OLS-core</td>
<td>3.39</td>
<td>6.01</td>
<td>5.36</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4.model avg.</td>
<td></td>
<td></td>
<td>-0.68</td>
<td>5.73</td>
<td></td>
<td></td>
</tr>
<tr>
<td>5.OLS-core+RSM</td>
<td></td>
<td></td>
<td></td>
<td>3.21</td>
<td></td>
<td></td>
</tr>
<tr>
<td>6.Model avg+RSM</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

The superior results produced by model averaging, use of repeat sale information, and their combination, show strong statistical significance. The difference between model averaging without RSM, and OLS-core using RSM but not model averaging, is insignificant, reflecting the small difference between the forecast loss measures in Table 1; recall however that RSM can only be employed on approximately one-fifth of the cases, limiting the power of the test. Nonetheless OLS-core + RSM is significantly superior to all methods without RSM, and model averaging + RSM is significantly superior to all other methods.

The impact of incorporation of repeat-sale information may be obscured by the fact that only in a minority of cases can a previous sale can be identified. Table 3 focuses on the subset of 1043 cases for which a repeat sale was identified, to judge the ‘undiluted’ impact of incorporation of this information.

Table 3

Prediction loss (RMSE) and inference on equal predictive accuracy:
subset of 1043 cases with a previous identified sale

<table>
<thead>
<tr>
<th>Method:</th>
<th>RMSE</th>
<th>vs. 1</th>
<th>vs. 2</th>
<th>vs. 3</th>
<th>vs. 4</th>
<th>vs. 5</th>
<th>vs. 6</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.RSM only</td>
<td>0.700</td>
<td>1.69</td>
<td>1.08</td>
<td>1.46</td>
<td>7.67</td>
<td>7.96</td>
<td></td>
</tr>
<tr>
<td>2.OLS-broad</td>
<td>0.659</td>
<td>-2.96</td>
<td>-2.36</td>
<td>5.35</td>
<td>5.64</td>
<td></td>
<td></td>
</tr>
<tr>
<td>3.OLS-core</td>
<td>0.673</td>
<td>2.56</td>
<td>6.09</td>
<td>6.18</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4.model avg.</td>
<td>0.664</td>
<td></td>
<td>5.56</td>
<td>5.80</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5.OLS-core+RSM</td>
<td>0.585</td>
<td></td>
<td></td>
<td>1.86</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>6.Model avg+RSM</td>
<td>0.582</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 3 reinforces the earlier results. Hedonic model averaging combined with repeat sale information produces the lowest forecast loss, although the difference relative to the

$^9$Table 2 gives Diebold-Mariano statistics for the null of no difference in forecast loss between various pairs of methods; the statistic is asymptotically ($N(0, 1)$ under $H_0$). A positive value corresponds with lower mean-square forecast loss for the higher-numbered method.
core OLS model with repeat-sales information is not significant at conventional levels. The use of repeat-sales information alone, or a hedonic model alone, produces results which are dominated by the forecast combinations.

Globally, there is clear evidence of moderate but statistically genuine improvements in predictive efficiency resulting from the refinement of the hedonic method, and the combination with repeat-sale information, where available.

Finally, we draw attention to the best-performing of the simple hedonic regressions, the core model containing painting dimensions, simple time trend, and a polynomial in the age of the artist at time of execution of the painting; the age variables capture the period in an artist’s career at which the work was executed, commonly an important factor in the price of the work. Returning to the loss measures of Table 1, we see that simple hedonic models describing the physical dimensions of the painting, time of sale and period in the artist’s career are not far from the best-performing valuation methods that we can identify; that is, these factors evidently capture a large proportion of the explanatory power in the best hedonic models.

4 Concluding remarks

Simple least-squares hedonic regressions are commonly used for quick statistical estimates of the value of a unique object such as a work of art. However, refinements of the simple method, some of which have been employed in the literature on real estate valuation or price index construction, are feasible and appear to be potentially valuable.

The main refinements suggested here—weighted model averaging (especially to allow estimates at lower levels of aggregation) and incorporation of repeat-sale information in some form along with hedonic information—are straightforward to implement as well. While it remains to be seen whether the efficiency gains that we observe will also be observed in other data sets, the fact that these methods allow the incorporation of additional information suggests that we should expect such efficiency improvements to be a general phenomenon. We have noted also that the incorporation of repeat-sale information here is subject to a number of constraints in the data, and that further reductions in forecast loss are likely to be available. In particular, larger sample sizes would permit optimization over the weights on hedonic models and prior-sale forecasts.

It is noteworthy that hedonic regressions using physical dimensions of the painting, time of sale and period of an artist’s career appear to capture a very large proportion of available predictive power. Nonetheless, both in the hedonic regressions and in combination with repeat-sales information, data-driven averaging across a number of predictive models produces significantly better results than does the use of any single model or predictor.
Appendix 1

The following are the 64 artists included in our sample, together with the number of sample points (works sold at auction) for each artist.

William Atkinson 83; Maxwell Bates 163; William Beatty 107; Leon Bellefleur 152; Frederic Martlett Bell-Smith 53; André Bieler 61; Paul-Emile Borduas 75; Franklin Brownell 61; Jack Bush 94; Frank Carmichael 61; Emily Carr 135; A.J. Casson 573; F.S. Coburn 293; Charles Comfort 83; Stanley Cosgrove 282; Maurice Galbraith Cullen 92; Jean Dallaire 63; Jacques de Tonnancour 58; Marcelle Ferron 109; Marc-Aurèle Fortin 97; Louise Gadbois 65; Clarence Gagnon 145; John Hammond 146; Lawren Harris 216; Robert Harris 52; Edwin Holgate 80; Yvonne McKague Housser 50; E.J. Hughes 126; A.Y. Jackson 997; Otto Jacoby 88; Frank Johnston 160; Illingworth Kerr 294; Dorothy Knowles 90; Cornelius Krieghoff 185; William Kurelek 68; Jean-Paul Lemieux 184; Rita Letendre 111; Arthur Lismer 382; J.E.H. MacDonald 333; J.W.G. MacDonald 62; Henri Masson 289; Jean McEwen 140; David Milne 112; Norval Morrisseau 163; Kazuo Nakamura 51; Toni Onley 74; Paul Peel 73; Robert Pilot 330; George Reid 77; Jean-Paul Riopelle 468; Goodridge Roberts 177; Albert Robinson 122; William Ronald 105; Allen Sapp 176; Jack Shadbolt 171; Gordon Smith 98; Jori Smith 62; Marc-Aurèle Suzor-Coté 148; Takao Tanabe 70; Fernand Toupin 55; Harold Town 125; Fred Varley 59; Frederick Verner 96; W.P. Weston 120.

References


