2009s-41

Dimension Reduction and Model Averaging for Estimation of Artists' Age-Valuation Profiles

John W. Galbraith, Douglas J. Hodgson

Série Scientifique
Scientific Series

Montréal
Septembre 2009

© 2009 *John W. Galbraith, Douglas J. Hodgson.* Tous droits réservés. *All rights reserved*. Reproduction partielle permise avec citation du document source, incluant la notice ©.
Short sections may be quoted without explicit permission, if full credit, including © notice, is given to the source.

CIRANO
Allier savoir et décision

Centre interuniversitaire de recherche en analyse des organisations

CIRANO

Le CIRANO est un organisme sans but lucratif constitué en vertu de la Loi des compagnies du Québec. Le financement de son infrastructure et de ses activités de recherche provient des cotisations de ses organisations-membres, d'une subvention d'infrastructure du Ministère du Développement économique et régional et de la Recherche, de même que des subventions et mandats obtenus par ses équipes de recherche.

CIRANO is a private non-profit organization incorporated under the Québec Companies Act. Its infrastructure and research activities are funded through fees paid by member organizations, an infrastructure grant from the Ministère du Développement économique et régional et de la Recherche, and grants and research mandates obtained by its research teams.

Les partenaires du CIRANO

Partenaire majeur
Ministère du Développement économique, de l'Innovation et de l'Exportation

Partenaires corporatifs
Banque de développement du Canada
Banque du Canada
Banque Laurentienne du Canada
Banque Nationale du Canada
Banque Royale du Canada
Banque Scotia
Bell Canada
BMO Groupe financier
Caisse de dépôt et placement du Québec
DMR
Fédération des caisses Desjardins du Québec
Gaz de France
Gaz Métro
Hydro-Québec
Industrie Canada
Investissements PSP
Ministère des Finances du Québec
Power Corporation du Canada
Raymond Chabot Grant Thornton
Rio Tinto
State Street Global Advisors
Transat A.T.
Ville de Montréal

Partenaires universitaires
École Polytechnique de Montréal
HEC Montréal
McGill University
Université Concordia
Université de Montréal
Université de Sherbrooke
Université du Québec
Université du Québec à Montréal
Université Laval

Le CIRANO collabore avec de nombreux centres et chaires de recherche universitaires dont on peut consulter la liste sur son site web.

Les cahiers de la série scientifique (CS) visent à rendre accessibles des résultats de recherche effectuée au CIRANO afin de susciter échanges et commentaires. Ces cahiers sont écrits dans le style des publications scientifiques. Les idées et les opinions émises sont sous l'unique responsabilité des auteurs et ne représentent pas nécessairement les positions du CIRANO ou de ses partenaires.
This paper presents research carried out at CIRANO and aims at encouraging discussion and comment. The observations and viewpoints expressed are the sole responsibility of the authors. They do not necessarily represent positions of CIRANO or its partners.

ISSN 1198-8177

Partenaire financier
Développement économique, Innovation et Exportation
Québec

Dimension Reduction and Model Averaging for Estimation of Artists' Age-Valuation Profiles *

John W. Galbraith†, *Douglas J. Hodgson*‡

Résumé / *Abstract*

Dans des modèles de régression hédonique de l'évaluation des oeuvres d'art, on trouve souvent que la valeur explicative de l'âge de l'artiste lors de la création d'une œuvre (ou d'une variable indicatrice des périodes de sa carrière) est très significative. La plupart des études préalables estiment des modèles où sont mis ensemble plusieurs artistes. Bien qu'il soit intéressant d'estimer de tels modèles pour des artistes individuels, les échantillons disponibles sont d'habitude de taille insuffisante en considérant le grand nombre de toutes les autres variables explicatives qu'on voudrait aussi inclure dans la régression. Dans ce papier, on adresse ce problème de degrés de liberté insuffisants en se servant de deux méthodes statistiques (moyenne de modèles et réduction de dimension) afin d'incorporer de l'information contenue dans un nombre possiblement grand de variables explicatives nous permettant l'estimation avec un petit échantillon. On trouve que les profils « âge-prix » des peintres individuels sont souvent considérablement différents de ceux estimés pour un groupe de peintres rassemblés, suggérant que l'application de nos méthodes dans des régressions hédoniques pourrait fournir de meilleures prévisions de prix réalisés dans des ventes aux enchères.

Mots clés : réduction de dimension, modèle de facteur augmenté, moyenne de modèles

In hedonic regression models of the valuation of works of art, the age at which an artist produces a particular work, or an indicator variable for periods in his or her artistic career, is often found to have highly significant predictive value. Most existing results are based on regressions that pool large groups of painters. Although it is of interest to estimate such regressions for individual artists, the sample sizes are often inadequate for a model that would also include the large number of other relevant variables. We address this problem of inadequate degrees of freedom in individual artist regressions by using two statistical methods (model averaging and dimension reduction) to incorporate information from a potentially large number of predictor variables, allowing us to work with relatively small samples. We find that individual age-valuation profiles can differ substantially from general pooled profiles, suggesting that methods that are more responsive to the unique features of individual artists may provide better predictions of art valuations at auction.

Keywords: *Dimension reduction, factor-augmented model, model averaging*

---

* 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.
† Department of Economics, McGill University and CIRANO, john.galbraith@cirano.qc.ca.
‡ Département de sciences économiques, Université du Québec à Montréal.

1. Introduction

It is no accident that the first painting to break the $100 million barrier at auction was a work by Picasso. As one of the most influential and celebrated of all artists, his works are more highly appreciated than those of most other artists, whether by art lovers, critics, historians, or the collectors who are willing to pay more cash for a painting if it is known to have been created by Picasso. This sale is an illustration of the fact that the price that a work of art will command in the market is heavily dependent on the identity of the painter; consequently, dummy variables for painter's identity are invariably included in hedonic regression models of art prices where more than one painter is included in the analysis (see Hodgson and Vorkink (2004) for an analysis of the parameter estimates associated with identity dummies in the case of the Canadian art market).

The work in question, "Boy With a Pipe", was sold at Sotheby's New York on 5 May 2004 for $104 million. However, many works by Picasso have been sold, before and since 2004, for considerably smaller sums. This consideration will tend to lead any person interested in the art market, such as a dealer, collector, artist, museum curator, or casual observer, to enquire as to whether there are any objectively observable properties of "Boy With a Pipe", other than the fact that is a Picasso, that could have helped in predicting that it would sell for an unusually high price. Such characteristics as its subject matter, size, or medium and support, could certainly contribute to its value, and these variables are indeed often included as explanatory variables in hedonic regressions.

One factor that must have been of primary importance in distinguishing "Boy With a Pipe" from many of the other Picassos that regularly appear on the market is its date of execution, 1905, when Picasso was 24 years old and was making the transition between two of the most celebrated stylistic "periods" of his career, the Rose and Blue. In fact, a hedonic pricing model of Picasso paintings sold at auction between 1963 and 1994 by Czujack (1997) included dummy variables for several different career periods, and found the Blue and Rose to be among the most highly valued. Galenson and Weinberg (2001), in a hedonic study of the influence of age on the price of paintings for four generations of artists active in Paris in the late nineteenth and early twentieth centuries, found that artists in Picasso's cohort (birth dates between 1880 and 1900) generally "peaked" early in their careers, at about age 27.

The importance of the age of a painter at the date of execution of a work to the quality, and thus market value, of the work, is well recognized, and has been investigated in a number of econometric studies of hedonic models: much work by David Galenson and collaborators studying European and American artists, including Galenson (2000, 2001) and Galenson and Weinberg (2000,2001), and Hodgson's (2009) study of Canadian painters. In most of these papers, polynomial functions linking age to log-price are included in a larger hedonic specification in which many painters are grouped together, with a single age-valuation profile being estimated for an entire category of artists, whether defined by nationality or cohort or both.[1]

[1]We use the term 'age-valuation' profile here to refer to a relationship between the typical

Although such hedonic regressions may have hundreds of regressors, the sample sizes are generally over a thousand, so that ample degrees of freedom are available to deliver useful parameter estimates.

For various reasons, disaggregating a sample of artists by cohort, interesting as it may be in itself, may not be satisfactory. As the discussion of Picasso illustrates, it is often the career effects of an individual artist that may be of principal interest to the analyst. As is clear from the work of Galenson and Weinberg (2000, 2001), these career effects are not the same for all cohorts within a national market, and thus the estimation of a disaggregated model can provide us with more precise information regarding the influence of age on quality and price for any given painter by making reference to his or her cohort. It follows that further disaggregation would be desirable, to the level of the individual artist, if possible. Certainly, if age-valuation profiles can vary across cohorts, then given the very individual nature of artistic creation and the variety of creative paths that can conceivably be followed, it is reasonable to suppose that age-valuation profiles can vary significantly across individual artists, even if they are born in the same place at the same time. We would thus prefer to estimate individual age-valuation profiles, if possible.

It is in the estimation of such profiles that data availability and degree of freedom considerations can pose significant obstacles to the obtention of useful econometric estimates. For many artists of interest, the number of observed sales at auction for which date of execution is available may be fewer than one hundred and could thus be quite small in relation to, even less than, the number of hedonic regressors available for the artist. In the present study, we work with a sample of sales at auction for Canadian painters similar to that analyzed by Hodgson (2009). We have selected eleven Canadian painters for the estimation of age-valuation profiles, with individual sample sizes ranging from 55 to 780, with an average of 192.

Since a paucity of degrees of freedom renders the estimation of the hedonic regression models by conventional techniques such as ordinary least squares (OLS) doubtful or even impossible in many cases, we are led to enquire as to the existence of alternative approaches. Since many of the regressors in the hedonic model are dummy variables, we could in principle increase the degrees of freedom (or, in other words, reduce the number of columns in the regressor matrix) by either omitting certain variables, or by redefining certain variables so as to reduce the number of columns associated with them (effectively "collapsing" together the columns). This can be done in a more or less arbitrary manner, or it can be done in a systematic manner employing optimality criteria for the omission or combination of columns in the regressor matrix.

In this paper, we will investigate the second of these approaches through the application of recent methods in the optimal reduction of dimensionality in linear regression models with degree of freedom deficiencies. There are two general approaches that we will apply: in the first, we use an optimally weighted combination of estimates delivered by differently-dimensioned smaller models, where the weights are delivered by the Mallows

value placed on a work of art and the artist's age, controlling for other factors such as the size of the painting.

criterion as proposed by Hansen (2007); in the second, we use selection of factors or components derived from columns of the regressor matrix, the properties of which have been analyzed Galbraith and Zinde-Walsh (2009).

In the next section, we describe our data set and provide more details on the age-valuation relation, in Section 3 we describe the econometric methodology employed, and in Section 4 our empirical results are presented and discussed.

2. Art auction data and the age-valuation relation

Age-valuation profiles for individual painters' works will be estimated in the framework of a hedonic regression model in which the dependent variable is log-price of the sale of a painting at auction and the artist's age at the time of execution of the painting is included in quadratic form among a possibly large set of independent variables. The data set and modelling approach are similar to those used by Hodgson (2009), whose analysis of the Canadian market is similar to analyses of the American and European markets reported by Galenson (2000, 2001) and Galenson and Weinberg (2000, 2001) and of the Picasso market as reported by Czujack (1997).

Hodgson (2009) estimates an aggregate age-valuation profile for a set of 248 Candian painters grouped together, and also divides the set of painters into three chronological "cohorts" based on birth date (pre-1880, 1880-1920, post-1920), with a separate profile being estimated for each one. Finally, a model is estimated in which the age-valuation profile is allowed to shift continuously as a function of the birth date of the painter. In the two latter specifications, Hodgson (2009) finds that the age-valuation profile tends, on average, to shift to the left the later the birth year of the painter. The data set used by Hodgson (2009) uses auction results for the period from 1968 to 2007 and consists of 9051 individual auction records. There are fewer than 400 regressors in the hedonic regression, with dummies for painter, medium and support, auction house, genre, date of sale, and measures of the height, width, and surface area of the painting, with the regression thus having over 8600 degress of freedom.

Age-valuation profiles may differ substantially for different painters, and it is therefore interesting to try to obtain estimates specific to the individual. To do so, a separate hedonic regression can in principle be estimated for each painter, as is done by Galenson (2000) for modern American painters. The problem that arises in this case is that degrees of freedom problems arise, since the number of observations for a given painter may not be much more than – in fact, may be less than – the number of regressors. One can in practice reduce the number of regressors by arbitrarily excluding variables (such as, for example, genre or auction house) or by arbitrarily redefining certain dummy variables to reduce their numbers (individual time-period dummies for five- or ten-year periods rather than for each year, for example). Both expedients are used by Galenson (2000). However, it is possible that important information contained in the regressors is being lost when such arbitrary dimension-reduction measures are imposed. In this paper, we will investigate the possibility of reducing the dimension of the regressor matrix in a data-dependent fashion that adjusts optimally to all the information contained in the regressors.

For each painter, we recorded every sale of an oil or acrylic work sold at auction during the years 1968-2008 for which date of execution was provided in the auction record, with data also collected on the variables mentioned above. 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 (1981-2009). 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 being included.

In the literature on the age-valuation relationship, age generally enters as a polynomial function with between two and four terms, reflecting a similar modelling of age-wage or experience-wage profiles in the labour economics literature (e.g. Mincer (1974)). Hodgson (2009) estimates a fourth-order polynomial for the overall and cohort regressions, but uses a quadratic for the specification with continuously shifting profile. Galenson (2000) allows for up to three polynomial terms in each individual-artist regression, with the exact number in each case being selected based on the best R^2 fit. In this paper, we will work with a quadratic specification for each artist, with the dimension of the regressor matrix being selected according to the algorithms described in the next section.

One interesting aspect of estimating age-valuation profiles for individual artists is to see to what degree they conform to typical profiles obtained from aggregated models such as that of Hodgson (2009). As mentioned, the third model specification estimated by Hodgson (2009) is a quadratic function where the parameters are allowed to drift as a function of the artist's date of birth. It follows that for any particular artist, if we know the year of birth, we can use the parameter estimates of Hodgson (2009) to plot a "typical" or "expected" age-valuation profile for any artist born in that particular year. Such plots for the artists considered in the present paper are given by the solid lines in Figure 1. The artists we consider, in order of year of birth (which is given in parentheses along with the appropriate Figure) are as follows: Cornelius Krieghoff (1815), George Reid (1860), Emily Carr (1871), Clarence Gagnon (1881), A.Y. Jackson (1882), Paul-Emile Borduas (1905), Jack Shadbolt (1909), Stanley Cosgrove (1911), Jean-Paul Riopelle (1923), Jean McEwen (1923), Harold Town (1924). The tendency for such aggregated age-valuation profiles to shift to the left over time is evident from a perusal of these graphs. However, art historical considerations (see Reid (1988) for a survey of Canadian art history with some detail on the careers of the better known of the artists considered here), along with the usual variations in individual character, suggest that individual profiles may differ considerably from those estimated from an aggregated model. Carr and Borduas, for example, are two artists who developed their best-known styles later in life, so that one would expect the individual profiles in these cases to peak later than those derived from aggregate data. Krieghoff, Jackson, and Riopelle, on the other hand, seem to have followed career paths that are fairly typical of their respective generations.

3. Statistical methods

Selection of appropriate regressors is a traditional difficulty in econometric modelling, and it is often observed that results for a variable of interest are sensitive to inclusion or

exclusion of other variables in a model.

Two recent classes of development offer potentially useful ways of mitigating this problem. Dimension reduction methods, exemplified by factor-augmented forecasting models, attempts to extract information from a large set of potential explanatory variables rather than to choose a subset of the variables for inclusion. Another class of method is model averaging, where results from models with different sets of included regressors are averaged, as opposed to choosing only one of the models. Although model averaging is a traditional idea, substantial recent progress has been made in criteria for choosing weights on different models. We will use methods of each type.

The data sets described above have a further feature that makes such methods especially useful. While the number of observations on all artists taken together is large, sample sizes for individual artists are much more limited. While we can estimate an 'aggregate' age-valuation profile using data on all artists, because the sample size is large enough that we could in principle include data on all explanatory variables, sample sizes for most individual artists are approximately the same size as, or smaller than, the number of explanatory variables. It is therefore impossible to include all potential effects in regressions applying to individual artists. However, while regression on all explanatory effects is not possible, both dimension reduction and model averaging allow us to use information from a large number of potential explanatory effects even where sample sizes preclude including each one in a single regression.

This study will use methods of each type; we will now describe each briefly.

3.1 Dimension reduction

Consider the linear model

$$y = c + X\beta + Z\gamma + \epsilon, \tag{3.1}$$

with the $\ell-$dimensional parameter of interest β, and where Z has elements which are also relevant to y and may be correlated with elements of X. We need to include information from Z in order to minimize omitted variables bias for β, but the number of columns K of Z may be too large to allow inclusion of each.

Rather than estimate (3.1), we orthogonalize the regressors in Z. Define a $K \times K$ matrix C such that the columns of C contain the K eigenvectors of $Z'Z$; that is, $C'Z'ZC = \Lambda$, where Λ is the $K \times K$ matrix with the K eigenvalues of $Z'Z$ on the main diagonal, zeroes elsewhere. Next define the matrix $C(\kappa)$ as any $K \times \kappa$ matrix which contains κ of the K eigenvectors: κ will be the number of control regressors included in the model. Finally define the auxiliary model regressors $S(\kappa, K) = ZC(\kappa)$; note that where $\kappa = K$, $S(K, K)$ spans the same space as Z.

We then estimate β using the auxiliary model

$$y = c + X\beta + S(\kappa, K)\delta + e, \tag{3.2}$$

which uses the subset $S(\kappa, K)$ of the available orthogonalized regressors contained in $S(K, K)$. Galbraith and Zinde-Walsh (2009) give general conditions for consistency of the

LS estimator of β in this model as κ grows with sample size. Criteria for selection of κ elements in $S(\kappa, K)$ are also given; note that if components are selected according to largest eigenvalue, then $S(\kappa, K)$ contains the first κ principal components of Z. In the present study we will use the Akaike Information Criterion to select the number κ of these principal components; Galbraith and Zinde-Walsh recommend this criterion, which tends to be generous in the number of terms selected, because of the asymmetry observed between selections greater than or less than the unknown optimal order: an order greater than optimal typically yields lower root mean squared error of estimates than does an order equally far below the optimum.[2]

Note that estimation using (3.2) does not require *a priori* knowledge of which elements of Z are important, but incorporates information from all columns of Z. The estimates of the parameters β can be interpreted in the usual way; here these will correspond with age effects in the valuation of an artists' works.

3.2 Model averaging

An alternative to combining the explanatory power of different data series is to combine the explanatory power of different models. Models are usually averaged using information criteria to select weights. Hansen (2007) has recently proposed an alternative weighting based on the Mallows C_p criterion. Within a class of discrete averaging estimators, Hansen shows that this criterion has a minimum squared error property, and we adopt this criterion here.

Consider again the model (3.1),

$$y = c + X\beta + Z\gamma + \epsilon = c + W\Theta + \epsilon, \tag{3.1}$$

where again X represents an $\ell-$ dimensional effect of interest to be included in all models and $W = X|Z$. Order the $\ell + K$ elements of W in accordance with of an *a priori* view of their importance.

Select a maximum number of regressors to include, $M < N$ where N is the sample size, and estimate all models with m regressors, $m = \ell, \ell + 1, \ldots, M$. The model average estimator is $\hat{\Theta} = \sum_{m=\ell}^{M} \omega_m \hat{\theta}_m$, where ω_m is the weight given to model m and $\hat{\theta}_m$ is the parameter vector estimated using a model with m regressors, padded with zeroes where $m < M$. The estimates of the parameters of interest β are then the weighted averages of the estimates of these parameters across all models. The weights ω_m are chosen to minimize the Mallows criterion $SSR + 2\sigma^2 k(\omega)$, where σ^2 is replaced with an estimate and where $k(\omega) = \sum_{m=\ell}^{M} \omega_m k_m$ is the effective number of parameters, k_m being the number of parameters estimated in model m.

Either of these methods allow us in principle to obtain information from a large set of explanatory variables despite a restricted sample size. With the model averaging estimator

[2]Although they are not reported here, results with orders selected by the more parsimonious SIC are very similar.

the number must nonetheless be less than the sample size, whereas for the dimension reduction estimator this constraint need not apply.

We now use these estimators to obtain age-valuation profiles for a number of individual artists.

4. Empirical results

In considering results of application of these methods to particular painters we are interested both in characterizing the careers of these individuals an a useful and interpretable way, and more generally in comparing results for individuals with the results that would be expected from aggregate models, to discover whether aggregate patterns are good predictors for individuals. We are also interested in comparing the methods for handling limited numbers of degrees of freedom, and in whether they tend to provide similar characterizations.

The results below address each of these aims. We present results graphically in the form of fitted age-valuation profiles, and numerically in the form of estimates of the parameters of the quadratic relation used for the effect of artist's age. In each case we do so for each of the methods that allow us to handle small samples (model averaging, dimension reduction), and compare with the aggregate results in Hodgson (2009) that use parameter estimates on the sample of all artists, with the year of birth for a particular artist used to obtain the predicted age-valuation profile for that artist. Model averaging is performed over the largest feasible set of models and dimension reduction methods optimize over all feasible values of the number of included components. The results are presented in Figure 1 and Table 1. In the graphical results presented in Figure 1, the quadratic fits are normalized to zero at age 20 for ease of comparison of slope and curvature.

The number of sample points (documented auction sales) available for each artist is recorded in Table 1. There are several hundred conditioning variables available in the data set; however, only a subset of these is available for each artist, primarily because many are dummy variables which may show no variation for a particular artist. For example, a dummy may record a '1' if a painting used acrylic paints, '0' otherwise. For an artist who never painting in acrylic, all observations will be zeroes, and although the series would be usable in an aggregate regression over all artists it will not be usable for this individual artist. The number of available regressors for these artists ranges from 43 (Borduas) to 104 (Jackson).

Consider first the numerical results of Table 1. These values are estimates of the age-related subset of the parameters of β in (3.1); other variables included in β are the dimensions (height and width) of a painting, which are very strong predictors of value.[3] The first four columns of numbers are from regressions (3.1) specific to the painter; the final two columns are from the aggregate data, as described earlier.

[3]Since area is a non-linear function of height and width, it can be included separately in the regression; however we in general find substantial predictive value in only two of the three variables and include only two in the regressions.

Note first from the aggregate data (final two columns) the gradual and monotonic change in coefficient values as birth year increases from Krieghoff through Town. The coefficient on age begins with positive values and falls as we move down through the table, ending with negative values; for age squared there is very little variation, although the results again change monotonically. Correspondingly, the quadratic fit of valuation to age from aggregate data begins with mid-career peaks for the earlier-born artists, but this pattern does not appear for those born later; the aggregate pattern is that later-born artists are valued more highly in their early career phases. These patterns are reflected in Figure 1. Earlier-born artists (Krieghoff, Reid, to a lesser extent Carr) show valuation peaks for works completed in the neighbourhood of age 55; Borduas and Shadbolt show approximately linear valuation profiles (sloping downward); the later-born artists show some upward curvature in a downward-sloping relation.

Next, we observe that there is considerable variation in individual artists from the aggregate pattern. Clearly, while the aggregate pattern has predictive value in cases where we cannot obtain individual-specific estimates, information specific to the artist gives us much more information.

The methods allowing individual-specific estimates (model averaging and dimension reduction) give qualitatively similar results in each case; that is, whether the parabolic fit opens downward or upward (mid-career peak or trough) is the same in each case, although of course specific parameter values differ and so the precise form is not in general the same; it is noteworthy however that for the individual with largest sample size here, A.Y. Jackson, the two methods give virtually identical estimates. In a majority of cases the age and age^2 parameters are significant at least at the 10% level despite the small sample sizes of most of these samples; for Stanley Cosgrove, by contrast, there is no substantial evidence of any age effect at all, even a linear effect, on valuation of the art.

The results from individual-specific methods do differ from those derived from aggregate estimation in numerous cases. Borduas, Carr and Reid show a different sign on the age effect in individual-specific methods than arises from the aggregate estimation, although these are each cases in which the individual-specific methods give coefficients which are poorly determined. The individual-specific results show, of course, much greater variability than the aggregate results, and it is of course a goal of individual-level analysis to be able to estimate these effects, obscured by averaging of artists with each other.

Overall, we observe that methods that allow us to obtain results on the small sample of data from individual artists, while retaining information from a large set of possible control variables for valuation of works of art, do allow us to uncover individual-specific information about deviations from the common pattern. Given the uniqueness of each artist's body of work, it is not surprising that uncovering this information tends to reveal very different patterns for different individuals. The numerical values corresponding to those patterns offer the potential to use statistical methods to suggest reserve prices at auction or to predict final sale values; while such statistical projections will not replace expert judgment, they may have value in pricing large numbers of works or as supplements to expert judgments in important cases.

5. Concluding remarks

The valuation of an art object must be one of the more subjective of the activities that fall within the scope of economic analysis. It is a natural challenge for econometric analysis, therefore, to attempt to provide an objective characterization of that subjective activity, to the extent possible, through observable and measurable factors. Hedonic pricing models are a traditional avenue for doing so; they are limited in their effectiveness for individual artists, however, because of small sample sizes of auction sales for any but the best known artists; in many cases the number of observable characteristics that we might condition on exceeds the number of sample points. Nonetheless the uniqueness of the bodies of work of different artists suggests strongly that analysis using only average results for a large number of artists will not characterize well the patterns observable for many individuals.

Appendix
Standard errors of function estimates

The age-valuation profiles have the quadratic form

$$q(t) = \alpha t + \beta t^2,$$

where t is age and $q(t)$ is the effect on log valuation of age. In practice, α and β are unknown and must be replaced by estimates $\widehat{\alpha}$ and $\widehat{\beta}$. This yields the estimated age-valuation function:

$$\widehat{q}(t) = \widehat{\alpha} t + \widehat{\beta} t^2.$$

As the parameter estimates are random, so is the functional estimate $\widehat{q}(t)$, whose standard error at any point t will be the square root of the estimated variance:

$$V\left(\widehat{q}(t)\right) = t^2 V\left(\widehat{\alpha}\right) + t^4 V\left(\widehat{\beta}\right) + 2t^3 C\left(\widehat{\alpha}, \widehat{\beta}\right)$$

where V$(\cdot)$ and C$(\cdot)$ denote variance and covariance, respectively.

For the model of Hodgson (2009) in which age-valuation profiles depend on date of birth, the computations are a bit more involved. In this model, the parameters of the quadratic are allowed to drift as a linear function of date of birth s, so that $\alpha = \alpha_0 + \alpha_1 s$ and $\beta = \beta_0 + \beta_1 s$, and therefore the age-valuation profile and its estimate are given by:

$$q_s(t) = \alpha_0 t + \alpha_1 st + \beta_0 t^2 + \beta_1 st^2$$

and

$$\widehat{q}_s(t) = \widehat{\alpha}_0 t + \widehat{\alpha}_1 st + \widehat{\beta}_0 t^2 + \widehat{\beta}_1 st^2.$$

For given date of birth s, the standard error of the estimated function $\widehat{q}_s(t)$ at the point t will be the square root of the estimated variance:

$$\begin{aligned} V\left(\widehat{q}_s(t)\right) = {} & t^2 V\left(\widehat{\alpha}_0\right) + s^2 t^2 V\left(\widehat{\alpha}_1\right) + t^4 V\left(\widehat{\beta}_0\right) + s^2 t^4 V\left(\widehat{\beta}_1\right) + 2st^2 C\left(\widehat{\alpha}_0, \widehat{\alpha}_1\right) \\ & + 2t^3 C\left(\widehat{\alpha}_0, \widehat{\beta}_0\right) + 2st^3 C\left(\widehat{\alpha}_0, \widehat{\beta}_1\right) + 2st^3 C\left(\widehat{\alpha}_1, \widehat{\beta}_0\right) \\ & + 2s^2 t^3 C\left(\widehat{\alpha}_1, \widehat{\beta}_1\right) + 2st^4 C\left(\widehat{\beta}_0, \widehat{\beta}_1\right). \end{aligned}$$

References

Bai, J. and S. Ng (2002) Determining the number of factors in approximate factor models. *Econometrica* 70, 191-221.

Bai, J. and S. Ng (2006) Confidence intervals for diffusion index forecasts and inference for factor-augmented regressions. *Econometrica* 74, 1133-1150.

Campbell, H. 1970-75. *Canadian Art Auction Record 1969-1974,* vols. 1-6. Toronto: Canadian Antiques and Fine Arts; and Montreal, Bernard Amtmann.

Campbell, H. 1980. *Canadian Art Auctions, Sales, and Prices, 1976-1978.* Don Mills, Ont., General.

Czujack, C. 1997. Picasso paintings at auction, 1963-1994. *Journal of Cultural Economics* 21, 229-247.

Galbraith, J.W. and V. Zinde-Walsh (2009) Some asymptotics for partially dimension-reduced regressions. Working paper.

Galenson, D.W. 2000. The careers of modern artists. *Journal of Cultural Economics* 24, 87-112.

Galenson, D.W. 2001. *Painting Outside the Lines.* Harvard.

Galenson, D.W and Weinberg, B.A. 2000. Age and quality of work: The case of modern American painters. *Journal of Political Economy* 108, 761-777.

Galenson, D.W and Weinberg, B.A. 2001. Creating modern art: The changing careers of painters in France from impressionism to cubism. *American Economic Review* 91, 1063-1071.

Hansen, B. (2007) Least squares model averaging. *Econometrica* 75, 1175-1189.

Hodgson, D.J. and K. Vorkink (2004) Asset pricing theory and the valuation of Canadian paintings. *Canadian Journal of Economics* 37, 629-655.

Hodgson, D.J. (2009) Age-price profiles for Canadian painters at auction. Working paper, UQAM.

Mincer, J. 1974. *Schooling, Experience, and Earnings.* Columbia University Press.

Reid, D. 1988. *A Concise History of Canadian Painting (2nd ed.).* Toronto: Oxford.

Sotheby's. 1975. *Canadian Art at Auction, 1968-1975.* Toronto: Sotheby's.

Sotheby's. 1980. *Canadian Art at Auction, 1975-1980.* Toronto: Sotheby's.

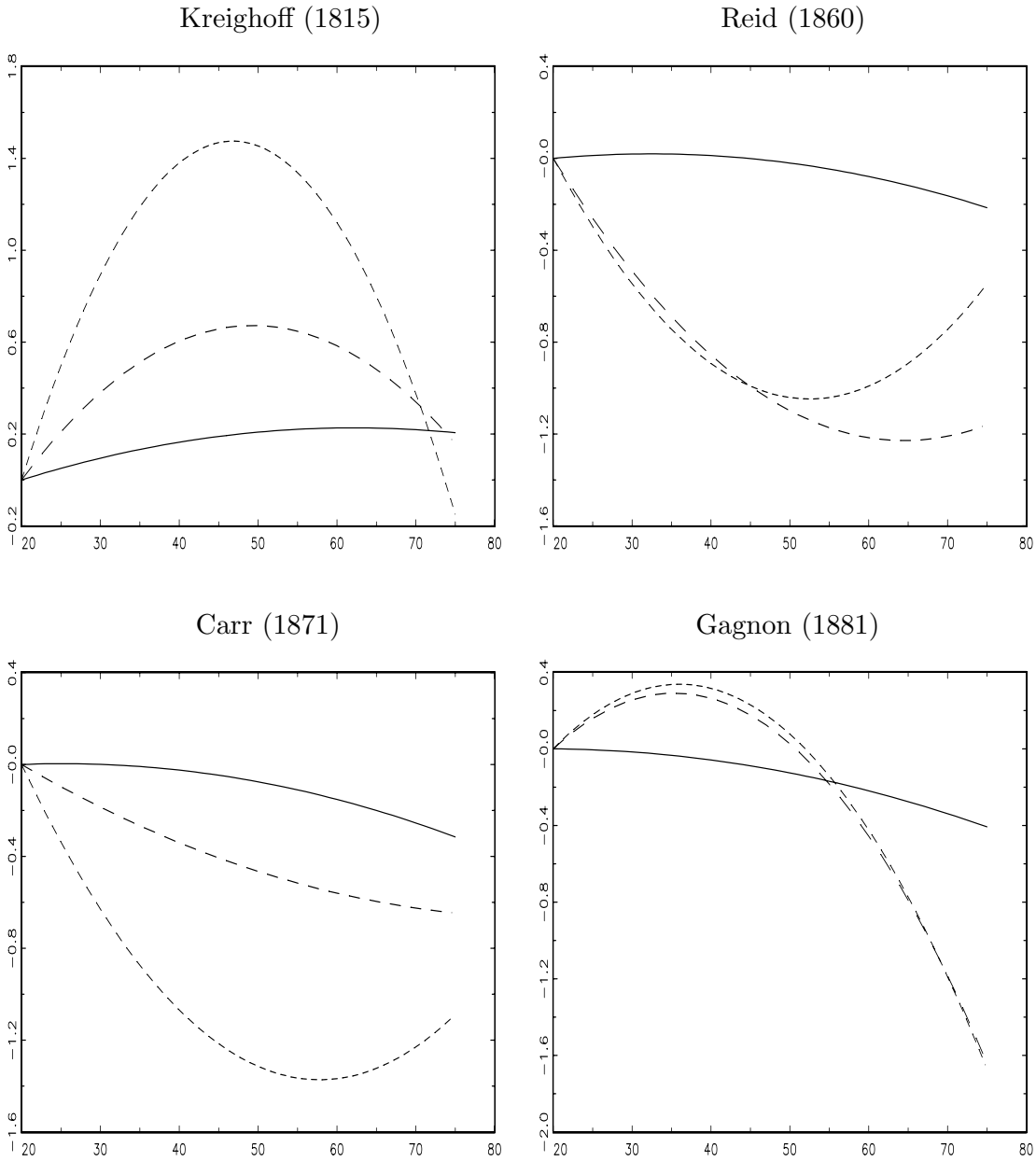
Westbridge, A.R. 1981-2009. *Canadian Art Sales Index, 1977-2008.* Vancouver, Westbridge.

TABLE 1
Estimates of the quadratic parameters in the age-valuation relation [4] , [5]

		MMA		DR-AIC		Aggregate	
	N	Age	Age^2	Age	Age^2	Age	Age^2
Krieghoff (1815)	155	.077	-.00078	.192	-.00205	.016	-1.28e-4
		0.82	-0.69	2.30	-2.03	4.18	-4.52
Reid (1860)	73	-.080	.00062	-.104	.00099	.008	-1.29e-4
		-1.24	0.97	-1.77	1.72	2.54	-4.45
Carr (1871)	79	-.026	.00015	-.111	.00096	.007	-1.29e-4
		-0.21	0.13	-0.88	0.82	2.03	-4.43
Gagnon (1881)	118	.087	-.00123	.095	-.00132	.005	-1.29e-4
		1.74	-2.13	1.90	-2.25	1.54	-4.40
Jackson (1882)	780	.054	-.00057	.054	-.00056	.005	-1.29e-4
		6.54	-8.24	6.73	-8.47	1.48	-4.39
Borduas (1905)	55	.301	-.00282	.195	-.00173	.0008	-1.30e-4
		1.98	-1.62	1.04	-0.80	0.27	-4.38
Shadbolt (1909)	120	-.149	.00119	-.122	.00097	.0002	-1.30e-4
		-2.82	2.65	-1.87	1.76	0.06	-4.38
Cosgrove (1911)	243	-.011	.00003	-.006	-.00001	-.0002	-1.30e-4
		-0.56	0.19	-0.31	-0.04	-0.05	-4.35
Riopelle (1923)	302	-.263	.00252	-.285	.00279	-.0022	-1.30e-4
		-6.09	4.91	-6.67	5.46	-0.69	-4.35
McEwen (1923)	99	-.033	-.00002	-.112	.00080	-.0022	-1.30e-4
		-0.49	-0.03	-1.63	1.20	-0.69	-4.35
Town (1924)	84	-.081	.00068	-.228	.00233	-.0023	-1.30e-4
		-0.79	0.58	-2.34	2.06	-0.74	-4.35

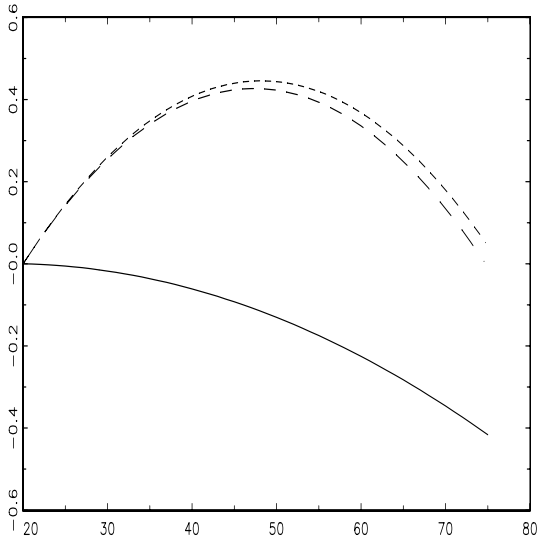
[4]MMA: Model averaging by Mallows criterion to choose weights; DR-AIC: dimension reduction with Akaike information criterion to choose number of parameters; Aggregate: estimates based on parameters estimated in aggregate model by Hodgson 2009.
[5]$t-$type statistics are reported below the estimates.

FIGURE 1
Estimated age-valuation profiles[6]

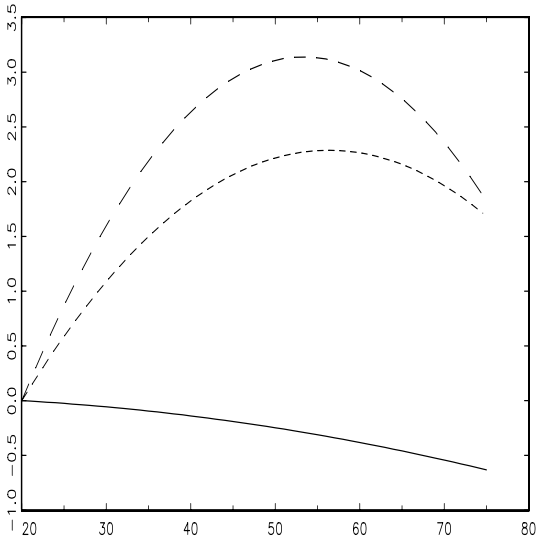


[6]long dashes: model averaging with weights selected by Mallows criterion (Hansen 2007); short dashes: dimension reduction with order selection by AIC; solid line: profile predicted for an artist born in the particular artists's actual birth year, from model estimated on data for these and other artists; see Hodgson 2009.

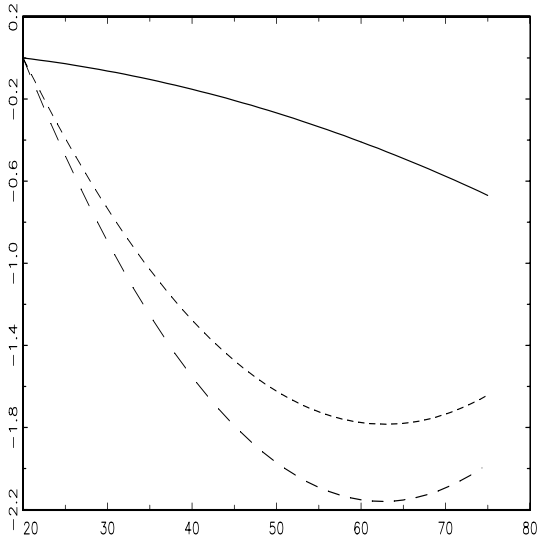
Jackson (1882)



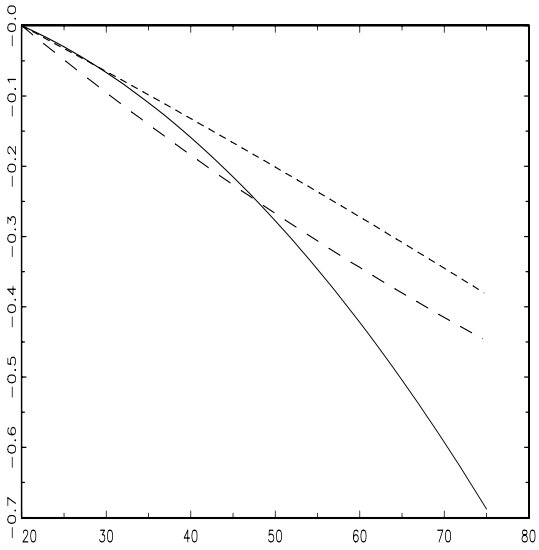
Borduas (1905)



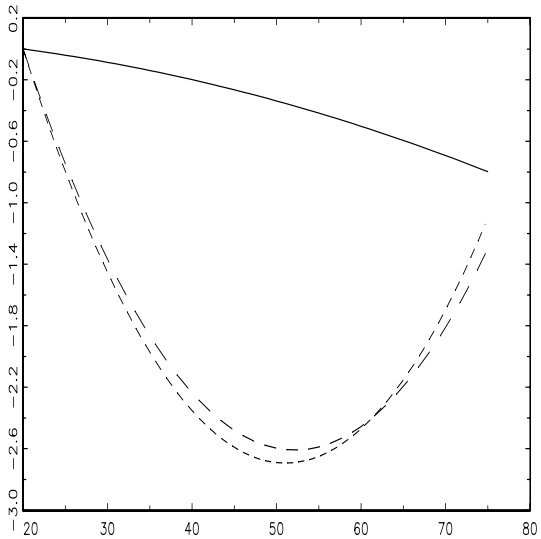
Shadbolt (1909)



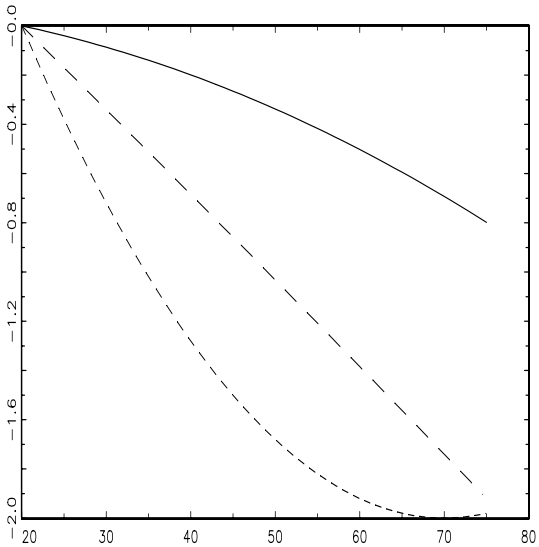
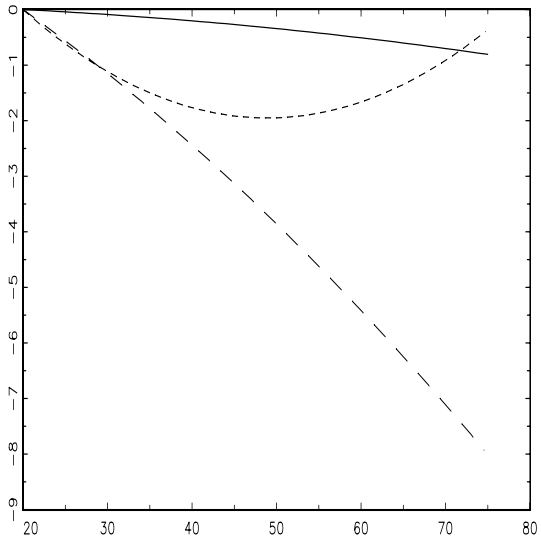
Cosgrove (1911)



Riopelle (1923)



McEwen (1923)



Town (1924)