The comparison of the hedonic, repeat sales, and hybrid models: Evidence from the Chinese paintings market

Fang Wang and Xu Zheng

Abstract: With the objective of evaluating the accuracy of price index models, we adopt a series of techniques to compares the performances of the hedonic, repeat sales, and hybrid models based on the data from the Chinese most representative painter, Qi Baishi during the period from 2000 to 2016. When applying the mean squared error (MSE) technique, the repeat sales model outperforms alternative models. However, according to the correlations and width confidence intervals, the hybrid model provides the most reliable estimates of price indices. The study also shows that the repeat sales model obtains relatively a lower total return estimate as well as a higher volatility than other two models. Our findings have important implications in identifying the precision of index models and provide supplements to art investment studies.

Subjects: Economics; Finance; Business, Management and Accounting

Keywords: Chinese paintings market; hedonic model; repeat sales model; hybrid model; art investment

JEL codes: C52; Z1

1. Introduction

The rapid growth in the Chinese paintings market has attracted increasingly attention in recent years. Until now, studies focusing on returns of Chinese artworks generally utilize the common hedonic or repeat sales models and there has been no study comparing the results and no uniform conclusion on these issues, thus creating confusion for investors. The aim of this study is to investigate the performances of the hedonic, hybrid along with repeat sales models in the Chinese paintings market in terms of (i) the accuracy of the various models and (ii) the influence on art return estimation using the data from the Chinese most influential painter, Qi Baishi during the period from...
2000 to 2016. Notably, it is the first time to combine main index construction models to provide an overall perspective on their performances in the Chinese paintings market.

The reason we choose data from Qi Baishi as the representative of the Chinese paintings market is that he has been known as the renowned Chinese painter of the twentieth century and one of the top-selling artists. According to the 2016 market annual report from Artpiece, he ranked the third of the top world 500 artists by auction revenue and the twentieth of the top 100 auction performance in 2016. It is widely acknowledged that the price behavior of Qi Baishi’s artworks exhibits consistent patterns of the Chinese paintings market and reveals useful aspects in terms of performance.

This paper, using both 8,218 hedonic data and 2,207 repeat sales data from Qi Baishi sold at China and global auction houses over the period 2000–2016, examines the performances of three common price index models: the hedonic, hybrid and the repeat sales models. In the hedonic model, all available transaction data are pooled. The logarithm of price is regressed as the dependent variable on a set of value-determining attributes and time dummies. One advantage of the hedonic model is that it makes efficient use of available data and may therefore give more reliable estimates of price indices (Renneboog & Spaenjers, 2013). However, it is criticized for disadvantages of implicit functional form and the change of item characteristics over time (Collins, Scorcu, & Roberto, 2009; Scorcu & Zanola, 2011; Wolverton & Senteza, 2000; Worthington & Higgs, 2004).

In contrast, by confining the analysis to a series of consecutive transactions of a given item at least twice during the sample period, the repeat sales model estimates a price index by regressing the price change of each item on a set of dummy variables. Researchers, such as Goetzmann (1993), Mei and Moses (2002), Pesando (1993), and Pesando and Shum (2008), employ this method to art investments. The main strength of the repeat sales model is that it does not require the specific information on item attributes. Nevertheless, since this approach is based on only a part of total sales observations and this subset may not represent the full sample, the weaknesses of using repeat sales model include the difficulty of controlling for the market trend during the sample period, as well as suffering from selection issues (Biey & Zanola, 2005; Kraeussl & Elsland, 2008; Renneboog & Spaenjers, 2013).

The hybrid model, first advanced by Case and Quigley (1991) and then extended by Carter Hill, Knight, and Sirmans (1997), constructs the index by combining features of both the hedonic and repeat sales models. This method goes beyond the previous techniques since it allows one to correct for the effects of heterogeneity and serial correlation (Biey & Zanola, 2005). However, the shortcoming of the hybrid model is that it needs an extensive data-set together with the details on the concerned observations. In addition, the requirement of complicated econometric techniques and implementation skills also makes it difficult for some people using this method. Actually, the hybrid model are mainly used in the house market and only a few applications exist on the art market due to the difficulties to apply it to art items (Chanel, Gérard-Varet, & Ginsburgh, 1996).

Our paper makes several important contributions to the analysis of performances of the hedonic, hybrid, and repeat sales models in the Chinese paintings market. First, this is the first time to apply the hybrid model to construct the price index of the Chinese paintings market. Secondly, we present an overall perspective on the comparison of these price index models by examining the results concerning the accuracy and their effects on return estimation. Finally, we examine some fundamental characteristics of the Chinese artworks and their interactions on return performance. Consequently, our findings are expected to be helpful to the evaluation of model construction and art investments.

The remainder of the paper is organized as follows. In Section 2 we provide a brief explanation of each model. Section 3 describes details on the data-set. Empirical results are analyzed in Section 4. In the final Section 5 we summarize the key conclusions.
2. The models

2.1. The hedonic model
Let $i = 1, \ldots, N$ be the set of selling artworks, define $P_i$ as the logarithm price of the painting $i$ at time $t$; $X_i$ as a $K \times 1$ vector of characteristics for the painting $i$ and $\alpha$ as a $K \times 1$ vector of implicit prices of $X_i$; $\delta$ as a $T \times 1$ vector of index numbers and $D_i$ as a $T \times 1$ vector of time dummies taking a value of one if a transaction on the painting $i$ occurs in time $t$ and zero otherwise, and $\epsilon_i$ as an error term. The hedonic regression equation can be expressed as:

$$P_i = \alpha'X_i + \delta'D_i + \epsilon_i$$

(1)

2.2. The repeat sales model
Define $P_{jt+s}$ as the logarithm price of the painting $j$ at time $t+s$. The price difference between two sales of the painting $j$ can be written as:

$$P_{jt+s} - P_{jt} = \Delta P_{jt} = \delta'\Delta D_{jt} + \Delta \epsilon_{jt}$$

(2a)

where $\Delta \epsilon_{jt} = \epsilon_{jt+s} - \epsilon_{jt}$ and

$$\Delta D_{jt} = \begin{cases} 
-1 & \text{if } \tau = t \\
1 & \text{if } \tau = t + s \\
0 & \text{otherwise}
\end{cases}$$

(2b)

2.3. The hybrid model
Separate the whole set $N$ into the subset of artworks that were sold only once during the sample period $w = 1, \ldots, I$ and the subset of artworks that were transacted more than once $j = 1, \ldots, J$. The error term $\epsilon_i$ of Equation (1) is now decomposed into two parts: a time dependent random error term $e_i$ and a time independent error term $\eta_i$. The hybrid model can then be represented as:

$$P_{wt} = \alpha'X_{wt} + \delta'D_{wt} + \eta_w + e_{wt}$$

(3a)

$$P_i = \alpha'X_i + \delta'D_i + \eta_i + e_i$$

(3b)

$$P_{jt+s} = \alpha'X_{jt+s} + \delta'D_{jt+s} + \eta_j + e_{jt+s}$$

(3c)

$$P_{jt+s} - P_{jt} = \Delta P_{jt} = \delta'\Delta D_{jt} + \Delta \epsilon_{jt}$$

(3d)

where all other variables are as previously defined.

Estimations of the hybrid model involve a series of intermediate steps. Note that specification error term $\eta_i$ in the hybrid model is directly analogous to the unobserved heterogeneity term in the random effects model outlined in Greene (2003). Under this assumption, the covariance matrix associated with the hybrid model can be written as:

$$\Omega = \begin{bmatrix} \sigma^2_e I & 0 & 0 \\
0 & \sigma^2_e I & -\sigma^2_{eT} \\
0 & -\sigma^2_{eT} & 2\sigma^2_e I \end{bmatrix}$$

(4)

Following Jones (2010), we estimate Equation (3d) using all $J$ observations to obtain the residuals as a consistent estimate of $\sigma^2_e$ and then apply a degrees of freedom correction:

$$\hat{\sigma}^2_e = \frac{1}{2} \left( \frac{1}{J - T} \right) \sum_{j=1}^{J} \hat{\epsilon}^2_{jt}$$

(5)
Analogously, a consistent estimate of $\sigma^2$ is found by estimating Equation (1) using all $W$ observations to obtain the residuals and then applying a degrees of freedom correction:

$$\hat{\sigma}^2 = \left( \frac{1}{W - K - T - 1} \right) \sum_{i=1}^{N} \hat{\varepsilon}_i^2$$

The covariance matrix $\hat{\Omega}$ could then be constructed using $\sigma^2$ and $\hat{\sigma}^2$. Theoretically, $\hat{\Omega}^{-1}$ can be found, and according to Cholesky decomposition, we could find the $P$ matrix, where $P^T P = \hat{\Omega}^{-1}$, and use this $P$ matrix to transform the data prior to estimation via least squares. Practically, finding $\hat{\Omega}^{-1}$ and $P$ directly requires intensive computations, and it is not possible for large data sets. Fortunately, the structure of $\hat{\Omega}$ is such that non-zero elements only exist on the on-diagonal of each block, and that the non-zero values within each block are identical. This means that the inverse of a proportionally much smaller version of $\hat{\Omega}$ can be found to obtain the values and the location of the non-zero elements (Fogarty & Jones, 2011).

3. Data
The data employed in this study are obtained from Artron.net and span the period 2000 to 2016. The Artron database contains a number of relevant information of artworks sold at China and global major auctions, (e.g. sales record, artist’s name, title of the work, year of production, materials, sales date, prices, dimension, signed or not). Buyers and sellers’ transaction fees paid to auction houses are included in prices and are recorded in both renminbi (RMB) and local currencies.

Specifically, the explanatory variables $X$ and $D$ included in models are:

3.1. Dimension
The variable area is a product of both width and length, which is measured in square meters. To examine whether the price increases with a diminishing effect, the variable area$^2$ is also included.

3.2. Medium
In our study, the Chinese artworks are categorized in two different mediums. The categories are paper and silk, with silk as the reference variable.

3.3. Mounting
Normally, the Chinese artworks can be mounted in several ways. We specify dummies album of sheets, folding fan, folding screen, framed, hand scroll, and hanging scroll to reflect in which type the painting is adopted, with album of sheets as the reference variable.

3.4. Type
Traditional Chinese artworks are done either in black ink or colored. In this case, two dummy variables ink (as the reference variable) and color are created to identify whether the painting is in ink or colored.

3.5. Auction houses
In this paper, the auction house dummies are specified by beijing hanhai, beijing kuangshi, christie, guardian, poly, and sotheby, which are main dominant auction houses in China and worldwide. We assign the dummy variable to take the value of one if the painting is classified into any of these six auction houses. Other auction houses are assigned to the references and take the value of zero.

3.6. Period
Both semiannual and annual period are applied in this paper. On the semiannual basis, a series of dummy variables, $D_t$ with $t = 2000s, ..., 2016a$, are introduced for each semiannual between the spring of 2000 and the autumn of 2016, regarding the spring of 2000 as the baseline variable and standardized to 100 in the price index. Similarly, $t = 2000, ..., 2016$ are introduced annually, considering 2000 as the baseline variable and standardized to 100 in the price index.
Table 1 summarizes the descriptive statistics of artworks sold. The number of samples, means, and standard deviations are presented, along with the minimum and maximum. The average artwork price of Qi Baishi from 2000 to 2016 is 1,742,045 RMB, ranging from 330 RMB to 425,500,000 RMB. According to the dimension, the average area is 0.3501m², with the standard deviation of 0.3068. There are 8,163 artworks which have been painted on paper, accounting for 99.33% of the total samples. Among all mounting types, hanging scroll is the most prevalent (5,115), followed by framed (2,320), whereas there are only 27 artworks mounted in hand scroll. Furthermore, the artworks in color (6,254) are approximately three times more than those in ink (1,964).

| Variable | Number | Mean    | S.D.   | Minimum | Maximum  |
|----------|--------|---------|--------|---------|----------|
| Price    | 8,218  | 1,742,045 | 7,009,035 | 330    | 425,500,000 |
| Area     | 8,218  | 0.3501   | 0.3068  | 0       | 13,4000  |
| Area²    | 8,218  | 0.2166   | 2.2939  | 0       | 179.5600 |
| Medium   |        |         |        |         |          |
| Paper    | 8,163  | 0.9933   | 0.0815  | 0       | 1        |
| Silk     | 55     | 0.0067   | 0.0815  | 0       | 1        |
| Mounting |        |         |        |         |          |
| Album sheets | 84  | 0.0102   | 0.1006  | 0       | 1        |
| Folding fan | 610 | 0.0742   | 0.2622  | 0       | 1        |
| Folding screen | 62  | 0.0075   | 0.0865  | 0       | 1        |
| Framed   | 2,320  | 0.2823   | 0.4501  | 0       | 1        |
| Hand scroll | 27  | 0.0033   | 0.0572  | 0       | 1        |
| Hanging scroll | 5,115 | 0.6224  | 0.4848  | 0       | 1        |
| Type     |        |         |        |         |          |
| Color    | 6,254  | 0.7610   | 0.4265  | 0       | 1        |
| Ink      | 1,964  | 0.2390   | 0.4265  | 0       | 1        |
| Auction houses | | | | | |
| Beijing Hanhai | 943 | 0.1147 | 0.3187 | 0 | 1 |
| Beijing Kuangshi | 632 | 0.0769 | 0.2665 | 0 | 1 |
| Christie | 608 | 0.0740 | 0.2618 | 0 | 1 |
| Guardian | 2,111 | 0.2569 | 0.4369 | 0 | 1 |
| Poly     | 1,108  | 0.1348   | 0.3416  | 0       | 1        |
| Sotheby  | 495    | 0.0602   | 0.2379  | 0       | 1        |
| Others   | 2,321  | 0.2824   | 0.4502  | 0       | 1        |
| Period   |        |         |        |         |          |
| 2000s    | 62     | 0.0075   | 0.0865  | 0       | 1        |
| 2000a    | 105    | 0.0128   | 0.1123  | 0       | 1        |
| 2001s    | 85     | 0.0103   | 0.1012  | 0       | 1        |
| 2001a    | 76     | 0.0092   | 0.0957  | 0       | 1        |
| 2002s    | 131    | 0.0159   | 0.1253  | 0       | 1        |
| 2002a    | 102    | 0.0124   | 0.1107  | 0       | 1        |
| 2003s    | 11     | 0.0013   | 0.0366  | 0       | 1        |

(Continued)
### Table 1. (Continued)

#### Panel A: Hedonic model

| Variable | Number | Mean   | S.D.   | Minimum | Maximum |
|----------|--------|--------|--------|---------|---------|
| 2003a    | 260    | 0.0316 | 0.1750 | 0       | 1       |
| 2004s    | 338    | 0.0411 | 0.1986 | 0       | 1       |
| 2004a    | 317    | 0.0386 | 0.1926 | 0       | 1       |
| 2005s    | 269    | 0.0327 | 0.1779 | 0       | 1       |
| 2005a    | 342    | 0.0416 | 0.1997 | 0       | 1       |
| 2006s    | 230    | 0.0280 | 0.1649 | 0       | 1       |
| 2006a    | 215    | 0.0262 | 0.1596 | 0       | 1       |
| 2007s    | 224    | 0.0273 | 0.1628 | 0       | 1       |
| 2007a    | 265    | 0.0322 | 0.1767 | 0       | 1       |
| 2008s    | 246    | 0.0299 | 0.1704 | 0       | 1       |
| 2008a    | 151    | 0.0184 | 0.1343 | 0       | 1       |
| 2009s    | 235    | 0.0286 | 0.1667 | 0       | 1       |
| 2009a    | 494    | 0.0601 | 0.2377 | 0       | 1       |
| 2010s    | 457    | 0.0556 | 0.2292 | 0       | 1       |
| 2010a    | 432    | 0.0526 | 0.2232 | 0       | 1       |
| 2011s    | 528    | 0.0642 | 0.2452 | 0       | 1       |
| 2011a    | 430    | 0.0523 | 0.2227 | 0       | 1       |
| 2012s    | 240    | 0.0292 | 0.1684 | 0       | 1       |
| 2012a    | 250    | 0.0304 | 0.1718 | 0       | 1       |
| 2013s    | 293    | 0.0357 | 0.1854 | 0       | 1       |
| 2013a    | 226    | 0.0275 | 0.1635 | 0       | 1       |
| 2014s    | 268    | 0.0326 | 0.1776 | 0       | 1       |
| 2014a    | 231    | 0.0281 | 0.1653 | 0       | 1       |
| 2015s    | 162    | 0.0197 | 0.1390 | 0       | 1       |
| 2015a    | 176    | 0.0214 | 0.1448 | 0       | 1       |
| 2016s    | 173    | 0.0211 | 0.1436 | 0       | 1       |
| 2016a    | 194    | 0.0236 | 0.1518 | 0       | 1       |

#### Panel B: Repeat sales model

| Variable | Number | Mean   | S.D.   | Minimum | Maximum |
|----------|--------|--------|--------|---------|---------|
| $\text{Price}_{t-1}$ | 1,609 | 977,891 | 3,044,355 | 6,600 | 92,000,000 |
| $\text{Price}_{t-1}$ | 2,207 | 2,375,188 | 7,464,974 | 1,150 | 195,500,000 |
| 2000s    | 23     | −0.0177 | 0.1318 | −1      | 0       |
| 2000a    | 37     | −0.0276 | 0.1640 | −1      | 0       |
| 2001s    | 27     | −0.0168 | 0.1319 | −1      | 1       |
| 2001a    | 36     | −0.0217 | 0.1520 | −1      | 1       |
| 2002s    | 49     | −0.0267 | 0.1774 | −1      | 1       |
| 2002a    | 32     | −0.0190 | 0.1463 | −1      | 1       |
| 2003s    | 7      | −0.0032 | 0.0562 | −1      | 0       |
| 2003a    | 100    | −0.0476 | 0.2537 | −1      | 1       |
| 2004s    | 158    | −0.0648 | 0.3069 | −1      | 1       |
| 2004a    | 153    | −0.0503 | 0.2992 | −1      | 1       |

(Continued)
Table 1 includes the statistics of artworks according to different auction houses as well. Of the six dominant auction houses, guardian is the most popular, with the number of 2,111. Relative to christie (608) and sotheby (495), poly (1,108), beijing hanhai (943) and beijing kuangshi (632) achieve higher proportions, suggesting that the majority of Chinese artworks of Qi Baishi has been sold in China.

4. Empirical results

4.1. Comparison of three price index models

The performances of three models are now evaluated. Table 2 reports the coefficient estimates and corresponding standard errors for semiannual and annual periods using the White (1980) heteroscedasticity-robust procedure. Columns (1) and (2) reveal the hedonic regression results. Results of the repeat sales method are shown on Columns (3) and (4). The hybrid model is estimated using the unrelated regression methodology and displayed in columns (5) and (6).

To evaluate and compare price index models in depth, we then use three key measures to assess the precisions of three price index models:

(1) The mean squared error (MSE) based on an out-of-sample forecast (Goh, Costello, & Schwann, 2012).

(2) The squared correlation between the actual and predicted values (Case & Szymanoski, 1995).

(3) The width of a 95% confidence interval (Case & Szymanoski, 1995).
Table 2. Empirical results
Panel A. Semiannual bases

| Variables | Hedonic model | Repeat sales model | Hybrid model |
|-----------|---------------|--------------------|-------------|
|           | Coefficient   | S.E.               | Coefficient | S.E. | Coefficient | S.E. |
|           | (1)           | (2)                | (3)         | (4)  | (5)         | (6)  |
| Constant  | 11.1982***    | 0.2069             | 0.2503***   | 0.0248 | 3.2991***  | 0.1682 |
| Dimension |               |                    |             |       |             |      |
| Area      | 2.5463***     | 0.0559             |             |       | 3.3623***   | 0.0630 |
| Area^2    | -0.1952***    | 0.0070             |             |       | -0.2650***  | 0.0080 |
| Medium    |               |                    |             |       |             |      |
| Paper     | -0.0357       | 0.1302             |             |       | 7.9030***   | 0.1110 |
| Silk      | [Left out]     |                    |             |       | [Left out]  |      |
| Mounting  |               |                    |             |       |             |      |
| Album sheets | [Left out]   |                    | [Left out]  |       |             |      |
| Folding fan | -1.0341***   | 0.1173             |             |       | 3.0767***   | 0.1152 |
| Folding screen | -0.8137***  | 0.1600             |             |       | 2.5819***   | 0.1704 |
| Framed    | -0.6175***    | 0.1098             |             |       | 3.4814***   | 0.1061 |
| Hand scroll | -0.7271***   | 0.2119             |             |       | 3.6666***   | 0.2105 |
| Hanging scroll | -0.6397***  | 0.1076             |             |       | 3.3154***   | 0.1043 |
| Type      |               |                    |             |       |             |      |
| Color     | 0.2341***     | 0.0250             |             |       | 0.5400***   | 0.0328 |
| Ink       | [Left out]     |                    |             |       | [Left out]  |      |
| Auction houses |          |                    |             |       |             |      |
| Beijing Hanhai | -0.1169***  | 0.0383             |             |       | -0.1971***  | 0.0435 |
| Beijing Kuangshi | 0.0508      | 0.0440             |             |       | 0.2081***   | 0.0462 |
| Christie  | 0.1623***     | 0.0439             |             |       | 0.2321***   | 0.0525 |
| Guardian  | -0.0096      | 0.0290             |             |       | -0.0058     | 0.0333 |
| Poly      | -0.0996***    | 0.0361             |             |       | 0.1665***   | 0.0375 |
| Sotheby   | 0.3218***     | 0.0481             |             |       | 0.2298***   | 0.0553 |
| Others    | [Left out]     |                    |             |       | [Left out]  |      |
| Period    |               |                    |             |       |             |      |
| 2000s     |               |                    |             |       |             |      |
| 2000a     | -0.0264       | 0.1534             | -0.8872***  | 0.1344 | -0.6201***  | 0.0955 |
| 2001s     | -0.0767       | 0.1590             | -0.8274***  | 0.1513 | -0.8157***  | 0.1108 |
| 2002a     | -0.0751       | 0.1634             | -0.5556***  | 0.1389 | -0.5787***  | 0.1069 |
| 2002s     | 0.0605        | 0.1469             | -0.5348***  | 0.1298 | -0.6576***  | 0.0877 |
| 2003s     | -0.1999       | 0.1536             | -1.0432***  | 0.1409 | -0.8220***  | 0.1019 |
| 2003a     | 0.7585**      | 0.3136             | -0.0483     | 0.2894 | -0.6024***  | 0.2834 |
| 2004s     | 0.4682***     | 0.1347             | -0.1850*    | 0.1106 | -0.4837***  | 0.0638 |
| 2004a     | 0.7113***     | 0.1317             | 0.2401**    | 0.1064 | -0.2264***  | 0.0556 |
| 2005s     | 0.9443***     | 0.1323             | 0.4113***   | 0.1080 | -0.2174***  | 0.0575 |
| 2005a     | 1.1220***     | 0.1342             | 0.7463***   | 0.1088 | -0.0725     | 0.0614 |
| 2006s     | 1.4316***     | 0.1316             | 0.8193***   | 0.1064 | 0.1089**    | 0.0552 |
| 2006a     | 1.1819***     | 0.1365             | 0.5380***   | 0.1120 | -0.1077     | 0.0668 |
| 2006a     | 1.0318***     | 0.1375             | 0.5180***   | 0.1187 | -0.1395*    | 0.0727 |

(Continued)
### Panel A. Semiannual bases

| Variables | Hedonic model | Repeat sales model | Hybrid model |
|-----------|---------------|-------------------|--------------|
|           | Coefficient   | S.E.   | Coefficient   | S.E.   | Coefficient   | S.E.   |
| (1)       | (2)           |        | (3)           | (4)    | (5)           | (6)    |
| 2007s     | 1.0493***     | 0.1370 | 0.5025***     | 0.1152 | −0.1851***    | 0.0696 |
| 2007a     | 1.2939***     | 0.1346 | 0.7060***     | 0.1125 | −0.0843       | 0.0664 |
| 2008s     | 1.2630***     | 0.1354 | 0.6854***     | 0.1131 | −0.0735       | 0.0669 |
| 2008a     | 1.2180***     | 0.1438 | 0.6828***     | 0.1203 | 0.0905        | 0.0803 |
| 2009s     | 1.3615***     | 0.1362 | 0.6772***     | 0.1153 | 0.0070        | 0.0688 |
| 2009a     | 1.7093***     | 0.1286 | 0.9874***     | 0.1061 | 0.3114***     | 0.0518 |
| 2010s     | 2.0639***     | 0.1291 | 1.2625***     | 0.1076 | 0.6708***     | 0.0547 |
| 2010a     | 2.5414***     | 0.1297 | 1.6880***     | 0.1094 | 1.1266***     | 0.0580 |
| 2011s     | 2.8310***     | 0.1282 | 2.0089***     | 0.1075 | 1.3616***     | 0.0543 |
| 2011a     | 2.7568***     | 0.1297 | 1.8981***     | 0.1098 | 1.4460***     | 0.0586 |
| 2012s     | 2.5319***     | 0.1360 | 1.8316***     | 0.1221 | 1.1901***     | 0.0706 |
| 2012a     | 2.3981***     | 0.1356 | 1.7820***     | 0.1167 | 1.5522***     | 0.0844 |
| 2013s     | 2.6978***     | 0.1335 | 1.9308***     | 0.1181 | 1.4586***     | 0.0662 |
| 2013a     | 2.6555***     | 0.1370 | 1.7434***     | 0.1237 | 1.6879***     | 0.0717 |
| 2014s     | 2.6223***     | 0.1346 | 1.9391***     | 0.1206 | 1.6128***     | 0.0686 |
| 2014a     | 2.5841***     | 0.1366 | 1.7879***     | 0.1228 | 1.4346***     | 0.0721 |
| 2015s     | 2.6356***     | 0.1427 | 2.0247***     | 0.1401 | 1.5967***     | 0.0850 |
| 2015a     | 2.0664***     | 0.1411 | 1.6134***     | 0.1406 | 1.0829***     | 0.0835 |
| 2016s     | 2.3942***     | 0.1413 | 1.7709***     | 0.1419 | 1.3095***     | 0.0844 |
| 2016a     | 2.4238***     | 0.1395 | 1.6387***     | 0.1361 | 1.3108***     | 0.0807 |
| Adjusted $R^2$ | 0.5628 | 0.5042 | 0.9343 |
| $F$-statistics | 221.4017 | 68.9894 | 3197.5440 |
| Included observations | 8,218 | 2207 | 10,422 |

### Panel B. Annual bases

| Variables | Hedonic model | Repeat sales model | Hybrid model |
|-----------|---------------|-------------------|--------------|
|           | Coefficient   | S.E.   | Coefficient   | S.E.   | Coefficient   | S.E.   |
| (1)       | (2)           |        | (3)           | (4)    | (5)           | (6)    |
| Constant  | 11.1804***    | 0.1855 | 0.2946***     | 0.0274 | 0.8303***     | 0.1044 |
| Dimension |               |        |               |        |               |        |
| Area      | 2.5603***     | 0.0563 | 3.3257***     | 0.0630 |
|          | −0.1957***    | 0.0070 | −0.2601***    | 0.0079 |
| Medium    |               |        |               |        |               |        |
| Paper     | −0.0183       | 0.1310 | 7.3831***     | 0.1153 |
| Silk      | [Left out]     |        | [Left out]     |        |
| Mounting  |               |        |               |        |               |        |
| Album sheets | [Left out]  |        | [Left out]     |        |
| Folding fan | −1.0383***   | 0.1182 | 2.9726***     | 0.1162 |
| Folding screen | −0.7899*** | 0.1613 | 2.5463***     | 0.1719 |
| Framed    | −0.6233***    | 0.1106 | 3.3672***     | 0.1071 |

(Continued)
4.1.1. The mean squared error (MSE) based on an out-of-sample forecast

The first measure we adopt is the MSE based on an out-of-sample forecast. This procedure, proposed by Goh et al. (2012), is simple to deal with and effective in examining the precisions of the price index models.

| Variables | Hedonic model | Repeat sales model | Hybrid model |
|-----------|---------------|--------------------|--------------|
|           | Coefficient   | S.E.               | Coefficient  | S.E.      | Coefficient | S.E. |
|           | (1)           | (2)                | (3)          | (4)       | (5)         | (6)  |
| Hand scroll | -0.7345***    | 0.2134             | 3.5191***    | 0.2140    |
| Hanging scroll | -0.6452***    | 0.1085             | 3.2086***    | 0.1053    |
| Type Color | 0.2293         | 0.0252             | 0.5766***    | 0.0321    |
| Ink        | [Left out]     | [Left out]         | [Left out]   |           |
| Beijing Hanhai | -0.1308***    | 0.0382             | -0.2294***   | 0.0434    |
| Beijing Kuangshi | 0.0383        | 0.0442             | 0.2101***    | 0.0464    |
| Christie  | 0.1638***      | 0.0441             | 0.2011***    | 0.0520    |
| Guardian | -0.0248        | 0.0291             | -0.0318      | 0.0333    |
| Poly      | -0.0942***     | 0.0362             | 0.1427***    | 0.0377    |
| Sotheby  | 0.3179***      | 0.0479             | 0.2057***    | 0.0549    |
| Others    | [Left out]     | [Left out]         | [Left out]   |           |
| Period    |               |                    |              |           |
| 2000      | -             | -                  | -            | -         |
| 2001      | -0.0602       | 0.1060             | -0.8032***   | 0.0995    |
| 2002      | -0.0417       | 0.0982             | -0.8789***   | 0.0915    |
| 2003      | 0.4928***     | 0.0953             | -0.3059***   | 0.0834    |
| 2004      | 0.8368***     | 0.0838             | 0.1884***    | 0.0656    |
| 2005      | 1.3074***     | 0.0847             | 0.6344***    | 0.0683    |
| 2006      | 1.1205***     | 0.0884             | 0.3623***    | 0.0755    |
| 2007      | 1.1914***     | 0.0874             | 0.4632***    | 0.0752    |
| 2008      | 1.2572***     | 0.0899             | 0.5248***    | 0.0779    |
| 2009      | 1.6070***     | 0.0839             | 0.7106***    | 0.0729    |
| 2010      | 2.3057***     | 0.0826             | 1.2671***    | 0.0730    |
| 2011      | 2.8075***     | 0.0821             | 1.7245***    | 0.0748    |
| 2012      | 2.4723***     | 0.0878             | 1.5805***    | 0.0859    |
| 2013      | 2.6892***     | 0.0873             | 1.6269***    | 0.0882    |
| 2014      | 2.6144***     | 0.0876             | 1.6307***    | 0.0905    |
| 2015      | 2.3473***     | 0.0927             | 1.5074***    | 0.1072    |
| 2016      | 2.4183***     | 0.0916             | 1.4931***    | 0.1056    |

Adjusted $R^2$ 0.5556 0.4098 0.9220
F-statistics 332.3532 96.7448 3977.9050
Included observations 8,218 2,207 10,425

* indicates significance at the 10%, respectively.
** indicates significance at the 5%, respectively.
*** indicates significance at the 1%, respectively.
Specifically, for each price index model, we split its data-set into two sub-samples. The sub-sample 1 consists of 70% of the full sample and is used to compute the indices. The parameter estimates obtained in the first sub-sample are then applied to the sub-sample 2, which involves the remaining 30% of the full sample and is used to forecast prices for each observation. In this way, there are two sets of prices in the sub-sample 2: one is the forecast price, $\hat{y}_i$, and the other is the actual price $y_i$. To evaluate their relative precisions, we then use the standard forecast analysis statistics in terms of the MSE for each price index model, which is represented by $\frac{1}{n} \sum_{i=1}^{n} (\hat{y}_i - y_i)^2$. In this equation, $n$ denotes the number of observations used to forecast the price, $\hat{y}_i$ is the forecast price for observation $i$, and $y_i$ is the actual price of observation $i$. Since the MSE is combined with both the bias and the variance, a lower MSE value for a particular price index generally indicates a better accuracy.

Table 3 presents the estimated semiannual and annual MSEs for the three price index models, respectively. It reveals that at both levels of time frequency, the repeat sales model generates the lowest MSEs. However, Case and Szymanoski (1995) note that the fact that standard deviation of the disturbance term in the repeat sales model smaller than those in the hedonic and hybrid models could be explained for two reasons. One reason is that the pure random noise component of the disturbance term is assumed to be positively related to the length of time elapsed between transactions, which is likely to be smaller for the repeat sales model than for other two model. The other reason is that the repeat sales model normally incorporates additional information on repeat transactions that are ignored in the hedonic model. Therefore, it is probably that the smaller estimated MSEs of the repeat sales model could be interpreted merely as a reflection of the smaller standard deviation of the disturbance term rather than an evidence that the repeat sales model is more precise than other two models.

Additionally, although the hybrid model stands out as a relative accurate price index model in most housing literatures (Jones, 2010; Goh et al., 2012; Quigley, 1995), that is not the case in our study. Even if the hybrid model has the advantage of combing features of both the hedonic and repeat sales model, it yields the highest estimated MSEs at both levels of time frequency (1.5265 semiannual and 1.5554 annual), underperforming the other price index models. Biey and Zanola (2005) also find that the standard deviation of the disturbance term of the hybrid model are higher than those of the hedonic and repeat sales models when investigating the performance of Picasso prints and attribute to the limited number of repeat sales in the sample. Similarly, our finding provides a certain support to Meese and Wallace (1991), who argue that the hybrid model appears not to produce reliable estimates of the price indices. Figure 1 depicts the MSEs of the three models based on an out-of-sample forecast at both semiannual and annual time frequencies, where the red line and the blue line represent the forecast value and the standard error, respectively.

### 4.1.2. Correlation between actual and predicted values

The second measure commonly used as means of assessing price index models is the correlation between actual and predicted values from econometric models, which provides a direct measure of the reliability. Normally, a high correlation can be interpreted as an indication of little error in the model, while a low correlation would indicate a large amount of error.

| Index models | Index frequency | Semiannual | Annual |
|--------------|----------------|------------|--------|
| Hedonic      |                | 1.1536     | 1.1619 |
| Hybrid       |                | 1.5265     | 1.5554 |
| Repeat sales |                | 0.9497     | 1.0994 |
Table 4 presents the squares of the estimated correlation coefficients between actual and predicted values based on the hedonic, hybrid, and repeat sales models ($r^2_{y\hat{y}}$). As shown, the hybrid model generates overwhelming higher correlations than the corresponding hedonic and repeat sales models, demonstrating its more accurate prediction ability. Moreover, correlations of the three price index models on the basis of semiannual frequency are slightly higher than those on the basis of annual frequency, suggesting relative better predictions.

### 4.1.3. Width of a 95% confidence interval

The third measure used is the width of a 95% confidence interval, which reflects both the estimated standard deviation of the disturbance term and the precision with which the individual parameters of the model. In view of this, we calculate the estimated semiannual and annual price indexes along with their corresponding 95% confidence intervals, which are reported in Table 5. Overall, the
Table 5. Annual and semiannual prices indexes

Panel A. Semiannual price indexes

| Period | Hedonic model | Repeat sales model | Hybrid model |
|--------|---------------|--------------------|--------------|
|        | Index | 95% C.I. | Index | 95% C.I. | Index | 95% C.I. |
|        | Low   | High    | Low   | High    | Low   | High    |
| 2000s  | 100.00| 100.00  | 100.00| 100.00  | 100.00| 100.00  |
| 2000a  | 97.39 | 97.07   | 97.72 | 41.18   | 40.95 | 41.41   |
| 2001s  | 92.62 | 92.30   | 92.94 | 43.72   | 43.44 | 44.00   |
| 2001a  | 92.76 | 92.44   | 93.09 | 57.38   | 57.04 | 57.71   |
| 2002s  | 106.24| 105.90  | 106.58| 58.58   | 58.26 | 58.90   |
| 2002a  | 81.88 | 81.61   | 82.15 | 35.23   | 35.03 | 35.44   |
| 2003s  | 213.51| 212.07  | 214.97| 95.29   | 94.15 | 96.45   |
| 2003a  | 159.72| 159.25  | 160.18| 83.11   | 82.73 | 83.50   |
| 2004s  | 203.66| 203.08  | 204.24| 127.14  | 126.58| 127.71  |
| 2004a  | 257.11| 256.38  | 257.85| 150.88  | 150.20| 151.56  |
| 2005s  | 307.10| 306.21  | 307.99| 210.91  | 209.95| 211.87  |
| 2005a  | 418.55| 417.37  | 419.75| 226.90  | 225.89| 227.91  |
| 2006s  | 326.05| 325.09  | 327.01| 171.25  | 170.45| 172.05  |
| 2006a  | 280.62| 279.78  | 281.45| 167.87  | 167.04| 168.70  |
| 2007s  | 285.57| 284.73  | 286.42| 165.29  | 164.49| 166.08  |
| 2007a  | 159.72| 159.25  | 160.18| 83.11   | 82.73 | 83.50   |
| 2008s  | 213.51| 212.07  | 214.97| 95.29   | 94.15 | 96.45   |
| 2008a  | 159.72| 159.25  | 160.18| 83.11   | 82.73 | 83.50   |
| 2009s  | 390.21| 389.06  | 391.36| 196.84  | 195.89| 197.79  |
| 2009a  | 552.51| 550.98  | 554.05| 268.44  | 267.25| 269.63  |
| 2010s  | 787.65| 785.45  | 789.85| 353.41  | 351.82| 355.00  |
| 2010a  | 1269.78| 1266.23| 1273.35| 540.86 | 538.39| 543.33 |
| 2011s  | 1696.16| 1691.47| 1700.87| 745.53 | 742.19| 748.88 |
| 2011a  | 1574.89| 1570.48| 1579.31| 667.31 | 664.26| 670.37 |
| 2012s  | 1257.74| 1254.05| 1261.44| 624.60 | 621.23| 627.59 |
| 2012a  | 1100.23| 1097.01| 1103.46| 594.17 | 591.28| 597.07 |
| 2013s  | 1484.70| 1480.43| 1489.01| 689.50 | 686.11| 692.91 |
| 2013a  | 1423.20| 1419.00| 1427.43| 571.67 | 568.72| 574.62 |
| 2014s  | 1376.75| 1372.75| 1380.77| 695.23 | 691.74| 698.74 |
| 2014a  | 1325.07| 1321.17| 1329.00| 597.72 | 594.66| 600.79 |
| 2015s  | 1395.15| 1390.87| 1399.48| 757.37 | 752.96| 761.81 |
| 2015a  | 789.63| 787.23| 792.05| 410.98 | 408.58| 413.40 |
| 2016s  | 1095.91| 1092.57| 1099.27| 587.58 | 584.12| 591.07 |
| 2016a  | 1128.89| 1125.49| 1132.30| 514.86 | 511.94| 517.79 |

Panel B. Annual price indexes

| Period | Hedonic model | Repeat sales model | Hybrid model |
|--------|---------------|--------------------|--------------|
|        | Index | 95% C.I. | Index | 95% C.I. | Index | 95% C.I. |
|        | Low   | High    | Low   | High    | Low   | High    |
| 2000   | 100.00| 100.00  | 100.00| 100.00  | 100.00| 100.00  |
| 2001   | 94.16 | 93.94   | 94.38 | 44.79   | 44.60 | 44.97   |

(Continued)
Wang & Zheng, Cogent Economics & Finance (2018), 6: 1443372
https://doi.org/10.1080/23322039.2018.1443372

hedonic, hybrid, and repeat sales models show a similar trend in Table 5: there has been a noticeable increase from the year 2000 to 2011, and then followed by a strongly fluctuation till 2016. Such situation could also be demonstrated in Figure 2, which presents the indices constructed with three models at the annual and semiannual time levels, respectively. Figure 2 reveals that in general, the patterns of three indices seem almost alike, although the repeat sales index always tracks below the hedonic and the hybrid indices at different time levels.

To make the estimate efficiency comparison across models much clearer, Figure 3 plots the 95% confidence interval for each model around a constant index value of unity. The dashed lines, dotted lines, and continuous lines represent the 95% confidence intervals (the level of price volatility) estimated from the hedonic model, the repeat sales model and the hybrid model, respectively. In line with our expectation, for both time levels, the hybrid model provides the narrowest confidence interval among three models. This result, combined with the result of the correlation between actual and predicted values, suggests that the hybrid model to some extent provide the most efficient estimates, which confirms the findings by Biey and Zanola (2005) and Fogarty and Jones (2011).

4.2. Evaluations of three price index models
Besides the estimate efficiency, we also compare different returns evaluated by different models. In general, the estimated return on art varies with data, methodology, and the time period under consideration (Ashenfelter & Graddy, 2003). Specifically, an important difference between the repeat

| Panel B. Annual price indexes |
|-----------------------------|
|                             | Hedonic model | Repeat sales model | Hybrid model |
|                             | Index | 95% C.I. | Index | 95% C.I. | Index | 95% C.I. |
|                             | Low   | High | Low | High | Low | High | Low | High |
| 2002                        | 95.92 | 95.71 | 96.12 | 41.53 | 41.37 | 41.68 | 95.36 | 95.23 | 95.49 |
| 2003                        | 163.69 | 163.35 | 164.03 | 73.65 | 73.39 | 73.90 | 125.87 | 125.72 | 126.03 |
| 2004                        | 230.90 | 230.48 | 231.32 | 120.73 | 120.40 | 121.06 | 166.60 | 166.45 | 166.74 |
| 2005                        | 369.65 | 368.98 | 370.33 | 188.59 | 188.05 | 189.13 | 227.70 | 227.47 | 227.93 |
| 2006                        | 306.64 | 306.05 | 307.23 | 143.66 | 143.21 | 144.11 | 187.31 | 187.11 | 187.50 |
| 2007                        | 329.17 | 328.55 | 329.80 | 158.92 | 158.42 | 159.42 | 188.55 | 188.35 | 188.75 |
| 2008                        | 351.54 | 350.86 | 352.22 | 169.02 | 168.47 | 169.57 | 211.48 | 211.25 | 211.72 |
| 2009                        | 498.78 | 497.88 | 499.69 | 203.51 | 202.89 | 204.13 | 268.93 | 268.68 | 269.18 |
| 2010                        | 1003.15 | 1001.36 | 1004.94 | 355.06 | 353.98 | 356.14 | 531.19 | 530.70 | 531.69 |
| 2011                        | 1656.89 | 1653.96 | 1659.84 | 560.98 | 559.23 | 562.73 | 896.93 | 896.08 | 897.79 |
| 2012                        | 1184.93 | 1182.68 | 1187.18 | 485.75 | 484.01 | 487.50 | 781.30 | 780.42 | 782.19 |
| 2013                        | 1472.02 | 1469.25 | 1474.81 | 508.79 | 506.92 | 510.66 | 1088.62 | 1087.38 | 1089.86 |
| 2014                        | 1365.92 | 1363.34 | 1368.51 | 510.75 | 508.83 | 512.68 | 1047.36 | 1046.14 | 1048.59 |
| 2015                        | 1045.71 | 1043.62 | 1047.81 | 451.49 | 449.48 | 453.51 | 837.53 | 836.41 | 838.64 |
| 2016                        | 1122.69 | 1120.47 | 1124.92 | 445.11 | 443.15 | 447.07 | 833.86 | 832.77 | 834.95 |

Note: C.I. refers to the confidence interval.
The repeat sales model and others is that the repeat sales model relies on only a subset of transactions and the utilized sample has been relatively small and selective, which may have a significant impact on the return estimates. Table 6 presents the estimated returns for the three price index models during the period 2000 to 2016. The highest return has been achieved by the hedonic model, which yields 0.0734 semiannually and 0.1511 annually. The return differences between the hybrid model and the hedonic model are not distinct, which the former is on average 13% lower than the latter in both time levels. The repeat sales model performs the worst, obtaining the lowest estimated returns and the highest risk at both time levels. By contrast, the hybrid model provides the lowest estimated volatility even at different time levels, notably surpasses all other models. This result somewhat contradicts the finding of Renneboog and Spaenjers (2013), which argues that, with respect to

| Year | Hedonic Model | Repeat Sales Model | Hybrid Model |
|------|--------------|--------------------|--------------|
| 2000 | 1000         | 900                | 800          |
| 2001 | 1100         | 1000               | 900          |
| 2002 | 1200         | 1100               | 1000         |
| 2003 | 1300         | 1200               | 1100         |
| 2004 | 1400         | 1300               | 1200         |
| 2005 | 1500         | 1400               | 1300         |
| 2006 | 1600         | 1500               | 1400         |
| 2007 | 1700         | 1600               | 1500         |
| 2008 | 1800         | 1700               | 1600         |
| 2009 | 1900         | 1800               | 1700         |
| 2010 | 2000         | 1900               | 1800         |
| 2011 | 2100         | 2000               | 1900         |
| 2012 | 2200         | 2100               | 2000         |
| 2013 | 2300         | 2200               | 2100         |
| 2014 | 2400         | 2300               | 2200         |
| 2015 | 2500         | 2400               | 2300         |
| 2016 | 2600         | 2500               | 2400         |

Figure 2. Annual and semiannual hedonic, repeat sales and hybrid price indexes of Qi Baishi during the period from 2000 to 2016.
artworks, the hedonic model has found somewhat lower returns than the repeat sales model. Considering that only part of Qi baishi’s artworks appear in the art market more often, those artworks with high appreciation and esthetic value by the collectors may be transacted infrequently.
| Period  | Hedonic model |  | Repeat sales model |  | Hybrid model |  |
|--------|---------------|---|-------------------|---|--------------|---|
|        | Index Returns |  | Index Returns     |  | Index Returns |  |
| 2000s  | 100.00 –  –   | 100.00 –  – | 100.00 –  – |  |
| 2000a  | 97.39 –0.0264 | 41.18 –0.8872 | 117.79 0.1637 |  |
| 2001s  | 92.62 –0.0503 | 43.72 0.0598  | 96.86 –0.1957 |  |
| 2001a  | 92.76 0.0016  | 57.38 0.2718  | 122.76 0.2370 |  |
| 2002s  | 106.24 0.1356 | 58.58 0.0208  | 113.45 –0.0789 |  |
| 2002a  | 81.88 –0.2604 | 35.23 –0.5084 | 96.26 –0.1643 |  |
| 2003s  | 213.51 0.9585 | 95.29 0.9950  | 119.89 0.2195 |  |
| 2003a  | 159.72 –0.2903 | 83.11 –0.1367 | 134.99 0.1187 |  |
| 2004s  | 203.66 0.2430 | 127.14 0.4251 | 174.61 0.2573 |  |
| 2004a  | 257.11 0.2331 | 150.88 0.1712 | 176.19 0.0090 |  |
| 2005s  | 307.10 0.1777 | 210.91 0.3349 | 203.67 0.1449 |  |
| 2005a  | 418.55 0.3096 | 226.90 0.0731 | 244.16 0.1813 |  |
| 2006s  | 326.05 –0.2498 | 171.25 –0.2814 | 196.62 –0.2165 |  |
| 2006a  | 280.62 –0.1501 | 167.87 –0.0200 | 190.47 –0.0318 |  |
| 2007s  | 285.57 0.0175  | 165.29 –0.0155 | 181.97 –0.0456 |  |
| 2007a  | 364.69 0.2446  | 202.58 0.2034 | 201.27 0.1008 |  |
| 2008s  | 353.59 –0.0309 | 198.45 –0.0206 | 203.46 0.0108 |  |
| 2008a  | 338.05 –0.0450 | 197.94 –0.0026 | 239.73 0.1641 |  |
| 2009s  | 390.21 0.1435  | 196.84 –0.0056 | 220.51 –0.0836 |  |
| 2009a  | 552.51 0.3478  | 268.44 0.3102 | 298.98 0.3044 |  |
| 2010s  | 787.65 0.3564  | 353.41 0.2750 | 428.26 0.3593 |  |
| 2010a  | 1269.79 0.4776 | 540.86 0.4255 | 675.61 0.6559 |  |
| 2011s  | 1696.16 0.2895 | 743.53 0.3209 | 854.52 0.2349 |  |
| 2011a  | 1574.89 –0.0742 | 667.31 –0.1108 | 929.80 0.0844 |  |
| 2012s  | 1257.74 –0.2249 | 624.40 –0.0665 | 719.86 –0.2559 |  |
| 2012a  | 1100.23 –0.1338 | 594.17 –0.0496 | 1034.02 0.3622 |  |
| 2013s  | 1484.71 0.2997  | 689.50 0.1488 | 941.57 –0.0937 |  |
| 2013a  | 1423.21 –0.0423 | 571.67 –0.1874 | 1184.27 0.2293 |  |
| 2014s  | 1376.75 –0.0332 | 695.23 0.1957 | 1098.54 –0.0752 |  |
| 2014a  | 1325.08 –0.0383 | 597.72 –0.1511 | 919.27 –0.1782 |  |
| 2015s  | 1395.17 0.0515  | 757.37 0.2367 | 1081.06 0.1621 |  |
| 2015a  | 789.63 –0.5692  | 410.98 –0.6113 | 646.71 –0.5138 |  |
| 2016s  | 1095.91 0.3278  | 587.58 0.3575 | 811.19 0.2266 |  |
| 2016a  | 1128.89 0.0297  | 514.86 –0.1321 | 812.22 0.0013 |  |
| Mean   | 0.0734 0.0497  | 0.0635 |  |
| S.E.   | 0.2787 0.3386  | 0.2112 |  |

Panel B. Annual returns

| Period  | Hedonic model |  | Repeat sales model |  | Hybrid model |  |
|--------|---------------|---|-------------------|---|--------------|---|
|        | Index Returns |  | Index Returns     |  | Index Returns |  |
| 2000   | 100.00 –  –   | 100.00 –  – | 100.00 –  – |  |
| 2001   | 94.16 –0.0602 | 44.79 –0.8032 | 99.52 –0.0048 |  |

(Continued)
For this reason, it is possible that the return estimates derived from the repeat sales model might be lower than those obtained using other models.

5. Conclusions
In this paper, we review both the advantages and disadvantages of the hedonic, hybrid and repeat sales models and apply a series of techniques to examine their performances in the Chinese paintings market by analyzing the artworks of Qi Baishi over the period 2000–2016. Applying the mean squared error (MSE) technique, the result shows that the repeat sales model outperforms alternative models. However, according to the correlations and width confidence intervals, the hybrid model provides the most efficient estimates, making it a more attractive approach for measuring price changes in the Chinese paintings market. Additionally, the empirical result reveals that, the repeat sales model gives the lowest estimated returns as well as the highest volatility among three models. Our findings, not only have important implications in identifying the accuracy of index models, but also provide a certain supplement to art investment studies.

It is worth noting that there exist some limitations in our paper. Although the hybrid model has the best explanatory power in terms of correlations and width confidence intervals, it ranks the worst in the MSE measure. In other words, none of the price index models seems to be consistently superior accurate to the others. Apart from the data used in our paper, absence of a convincing proxy for measuring index accuracy and pooling data together across time may also result in such inconsistency (Ansah, 2013). We fully expect that, in further studies, there would be a methodology demonstrated to provide more accurate and robust estimations in the price index construction and be applied in more general settings.

| Panel B. Annual returns |
|-------------------------|
| Hedonic model | Repeat sales model | Hybrid model |
| Period | Index | Returns | Index | Returns | Index | Returns |
| 2002 | 95.92 | 0.0185 | 41.53 | −0.0756 | 95.36 | −0.0427 |
| 2003 | 163.69 | 0.5345 | 73.65 | 0.5730 | 125.87 | 0.2776 |
| 2004 | 230.90 | 0.3440 | 120.73 | 0.4943 | 166.60 | 0.2803 |
| 2005 | 369.65 | 0.4706 | 188.59 | 0.4460 | 227.70 | 0.3125 |
| 2006 | 306.64 | −0.1869 | 143.66 | −0.2721 | 187.31 | −0.1953 |
| 2007 | 329.17 | 0.0709 | 158.92 | 0.1010 | 188.55 | 0.0666 |
| 2008 | 351.54 | 0.0657 | 169.02 | 0.0616 | 211.48 | 0.1148 |
| 2009 | 498.78 | 0.3499 | 203.51 | 0.1857 | 268.93 | 0.2403 |
| 2010 | 1003.15 | 0.6987 | 355.06 | 0.5566 | 531.19 | 0.6807 |
| 2011 | 1656.89 | 0.5018 | 560.98 | 0.4574 | 896.93 | 0.5239 |
| 2012 | 1184.93 | −0.3353 | 485.75 | −0.1440 | 781.30 | −0.1380 |
| 2013 | 1472.02 | 0.2170 | 508.79 | 0.0463 | 1088.62 | 0.3317 |
| 2014 | 1365.92 | −0.0748 | 510.75 | 0.0039 | 1047.36 | −0.0386 |
| 2015 | 1045.71 | −0.2671 | 451.49 | −0.1233 | 837.53 | −0.2236 |
| 2016 | 1122.69 | 0.0710 | 445.11 | −0.0142 | 833.86 | −0.0044 |
| Average | 0.1511 | 0.0933 | 0.1326 |
| S.E. | 0.3068 | 0.3617 | 0.2584 |
Funding
The authors received no direct funding for this research.

Author details
Fang Wang1
E-mail: fwan6213@hotmail.com
Xu Zheng1
E-mail: xzheng@sjtu.edu.cn
1 Department of Finance, Antai College of Economics and Management, Shanghai Jiao Tong University, Shanghai, 200030, China.

References
Ansah, A. O. (2013). Construction of property price indices. Property Management, 31(2), 115–131. https://doi.org/10.1108/02637471311309427
Ashenfelter, O., & Graddy, K. (2003). Auctions and the price of art. Journal of Economic Literature, 41(3), 783–787. https://doi.org/10.1257/jel.41.3.763
Bley, M. L., & Zanola, R. (2005). The market for picasso prints: A hybrid model approach. Journal of Cultural Economics, 29(2), 127–136. https://doi.org/10.1007/s10824-005-8133-x
Carter Hill, R., Knight, J. R., & Sirmans, C. F. (1997). Estimating capital asset price indexes. The Review of Economics and Statistics, 79(2), 226–233. https://doi.org/10.1111/0034-6535.00026
Case, B., & Quigley, J. M. (1991). The dynamic of real estate prices. The Review of Economic and Statistics, 73(1), 50–58. https://doi.org/10.2307/2109686
Case, B., & Szymanski, E. J. (1995). Precision in house price indices: Findings of a comparative study of house price index methods. Journal of Housing Research, 6(3), 483–496.
Chanel, O., Gérard-Varet, L., & Ginsburgh, V. (1996). The relevance of hedonic price indices: The case of artworks. Journal of Cultural Economics, 20(1), 1–24. https://doi.org/10.1007/s10824-005-1024-3
Collins, A., Scorcu, A., & Roberto, Z. (2002). Reconsidering hedonic art price indexes. Economic Letters, 80(2), 57–60. https://doi.org/10.1016/S0160-2822(02)00226-1
Fogarty, J. J., & Jones, C. (2011). Return to wine: A comparison of the hedonic, repeat sales and hybrid approaches. Australian Economic Papers, 50(4), 147–156. https://doi.org/10.1111/j.1469-7277.2011.01586.x
Goetzmann, W. (1993). Accounting for taste: Art and the financial markets over three centuries. American Economic Review, 83(5), 1370–1376.

Goh, Y. M., Costello, G., & Schwann, G. (2012). Accuracy and robustness of house price index methods. Housing Studies, 27(5), 643–666. https://doi.org/10.1080/02673037.2012.697551
Greene, W. H. (2003). Econometric analysis (International 5th ed.). New Jersey: Prentice Hall.
Jones, C. (2010). House price measurement: The hybrid hedonic repeat-sales method. The Economic Record, 86(272), 95–97. https://doi.org/10.1111/1475-4932.2010.00861.x
Kraeussl, R., & Elsland, N. (2008). Constructing the true art market index-A novel 2-step hedonic approach and its application to the German art market. CFS Working Paper, No. 2008/11. Retrieved from http://nbn-resolving.de/urn:nbn:de:hebis:30-5439
Meese, R. A., & Wallace, N. E. (1991). Nonparametric estimation of dynamic hedonic price models and the construction of residential housing price indices. Journal of the American Real Estate and Urban Economics Association, 19, 308–332. https://doi.org/10.1111/jore.1991.19.issue-3
Mei, J., & Moses, M. (2002). Art as an investment and the underperformance of masterpieces. American Economic Review, 92(5), 1656–1668. https://doi.org/10.1257/000282802762024719
Pesando, J. E. (1993). Art as an investment: The market for modern prints. American Economic Review, 83(5), 1075–1089.
Pesando, J. E., & Shurm, P. M. (2008). The auction market for modern prints: Confirmations, contradictions, and new puzzles. Economic Inquiry, 46(2), 149–159. https://doi.org/10.1111/ecin.2008.46.issue-2
Quigley, J. M. (1995). A simple hybrid model for estimating property price indices. Journal of Housing Economics, 4(1), 1–12.
Renneboog, L., & Spoenjers, C. (2013). Buying beauty: On prices and returns in the art market. Management Science, 59(1), 36–53. https://doi.org/10.1287/mnsc.1120.1580
Scorcu, A., & Zanola, R. (2011). The right price for collectibles: A quantile hedonic regression investigation of Picasso artworks. Journal of Alternative Investments, 14(2), 89–99. https://doi.org/10.19030/joi.2011.14.2.089
White, H. (1980). A heteroscedasticity-consistent covariance matrix estimator and a direct test for heteroscedasticity. Econometrica, 48(4), 817–838. https://doi.org/10.2307/1912934
Wolverton, M. L., & Senta, J. (2000). Hedonic estimates of regional constant quality house prices. Journal of Real Estate Finance, 19(3), 235–253.
Worthington, A. C., & Higgs, H. (2004). Art as an investment: Risk, return and portfolio diversification in major painting markets. Accounting and Finance, 44(2), 257–271. https://doi.org/10.1111/aafi.2004.44.issue-2