The Effect of Seller Reserves on Market Index Estimation

William N. Goetzmann
Yale School of Management

First Draft: June 9, 1994
Current Draft January 25, 1999

William N. Goetzmann
Yale School of Management
Box 208200 New Haven
Connecticut 06520-8200
email will@viking.som.yale.edu

Acknowledgements

Thanks to Matthew Spiegel and John Clapp for helpful suggestions . All errors are the sole responsibility of the author.

Abstract

This paper examines the effect of seller reserves on market index construction. It reports the results of simulations in which transactions are conditioned upon various reservation strategies. Indices constructed by averaging across observed conditional prices each period differ dramatically from unconditional indices. Not only are conditional price levels higher, but the dynamics of the price path are changed. Time-series' of conditional mean returns are not highly correlated to the time-series' of unconditional mean returns, and average return estimates are biased upwards.

Alternate estimation procedures provide clear improvements to the conditional mean estimate. Volume of sales is a significant predictor of returns in the presences of certain types of reservation behavior. Hedonic control via the repeat-sales regression provides significant improvements, generating an index that is well correlated to the unconditional mean estimate. The repeat-sales regression fails to reduce the upward bias in mean estimates. The simulation results have particular application to the art and housing markets, in which private values may be used to set reservation prices.

The Effect of Seller Reserves on Market Index Estimation

I. Introduction

Analysts of trends in asset markets typically rely upon observed transactions prices. For very liquid markets such as the New York Stock Exchange, in which most assets trade every day, the average price adequately captures market dynamics. For less liquid markets, such as real estate or art, the average observed transaction price may not be as informative. In both of these markets, sellers set reservation prices based in part upon private values. Thus, the prices observed in these markets are *conditioned* upon seller reserves. We show in this paper that the average price index can change dramatically depending upon what kind of reservation rules sellers employ. Given that a reservation rule truncates the observed price distribution each period, it is not surprising to find that the conditional average price index is biased upward. More surprising is that the times series of returns based upon the conditional index may not even be correlated to the unconditional series. In other words, observed market trends may be entirely spurious.

Sellers of all sorts of properties set reservation prices. For purely *common value* goods, the reserve indicates the seller's assessment of the economic value of the property. For a share of stock, for instance, the reserve might reflect an assessment of the net present value of future sales price, plus the discounted dividend stream over the time until sale. For private value goods, the seller reserve also reflects the personal satisfaction the owner gets from ownership and use. For instance, certain things such as a lock of hair have sentimental value to the owner, but virtually no value to a potential buyer. Other goods, like paintings, are a combination of private and public valuation. While many artworks are purchased for investment, certain owners derive extraordinary pleasure from owning particular pieces, and might set a reserve higher than the going "price" for the painting. The same can be said for single family houses. Much home improvement actually detracts from the resale value of the house, while enhancing the private value to the owner.[1]

Seller reserves are integral to the auction process.   For instance, virtually all paintings offered at major auctions have a secret "reserve" price known only to the consignor and the auction house.[2] Since houses, art, antiques, stamps coins and other collectibles have a private value component which is difficult to quantify, there are no clear economic rules about how to set such reserves. Reservation prices might be chosen based upon the seller's deep-felt conviction about how painful it would be to part with the piece, or they might be based upon what the latest example of such a work happened to go for. It is important to point out that a seller's choice to set a reserve based upon his or her private value for the good is not irrational. Indeed, the seller may be setting the reserve in such a way as to maximize utility based upon the future enjoyment of the artwork and upon revenues derived from sale. If the seller reserve were observable in the auction record it would be reasonable to use this value as an observed price, implicitly paid by the current owner. The problem for index estimation is that the seller reserve is typically unobserved. The only prices observable are those which exceed the owner's valuation.

---

[1] Homes are rarely auctioned in the typical sense, however sellers often set reserves, since bids not meeting the asking price may be refused.   Other types of real estate may be auctioned with reserve. See for instance, the Ashenfelter and Genovese (1992) study of condominium auctions.

[2] Some private value goods are auctioned without reserve.   Stamp auctions, studied by Taylor (1983), for instance were without reserve.

Regardless of whether seller behavior is rational or irrational, seller reservation rules "censor" the observed market transactions. For some simple rules such as a known price threshold, it may be possible to estimate the unconditional distribution. However, in most settings it is not. In this paper, we simulate a few reservation rules and observe their effects. These simulated rules reflect a range of seller attitudes about the auction price, and condition upon past purchases of their asset, or past prices of similar assets. For instance, one rule sets the reserve at the seller's purchase price. Another uses the maximum past observed transaction price for the asset. Another bases the reserve upon the quality-adjusted maximum of the most recent period's sales. Another uses the asset value in the first period. We find that, when an index is based upon the average price of works *that sell*, then the fluctuation in that index might be completely unrelated to the index based upon an average across all works in the market. In fact, under certain conditions, we find no correlation between uncensored and the estimated indexes. Part of this variation is due to the changing average quality of the work. This variation in quality from one period to another is a major motivation for the use of hedonic index methods and repeat-sales index methods in real estate research (see Bailey, Muth and Nourse, 1963 and Case and Shiller, 1986, for example). In the art market, Taylor (1983, 1992) uses signal extraction regression to control for the fluctuating quality of the average asset at stamp auctions.

The results of these simulations have a direct bearing on performance measures used in a number of asset markets. When trend estimates are based upon mean transaction prices, they may be incorrect. In addition, estimates of asset price levels are almost surely too high. Techniques used to control for quality variation help to control the first problem but not the second.

II. Methodology

II.1 Stochastic Specification

In order to examine the effects of different reservation rules upon the average price index, we repeatedly simulate an asset market over a period of time in the following manner. For each simulation, prices for 100 assets in the initial period are distributed uniformly over the interval [0,1]. Returns in the subsequent 40 periods are generated by a multiplicative market model, in which all assets have the same sensitivity to the market[3]. For simplicity, returns are lognormal, and simulations are performed in logs. Log market returns are distributed normal [$N(\mu_m,\sigma_m)$] with positive drift of .05 and standard deviation of .10. Residuals are distributed normal [$N(\mu_i,\sigma_i)$] with zero drift and standard deviation of .4. Thus, the log asset value for asset i in period t may be expressed as:

$$
\begin{aligned}
P_{i,t} &= P_{i,0} + \sum_{\tau=1}^{t} r_{i,\tau} \\
&= P_{i,0} + \sum_{\tau=1}^{t} r_{m,\tau} + \sum_{\tau=1}^{t} \varepsilon_{i,\tau}
\end{aligned}
\qquad \textbf{(1)}
$$

II.2 Reservation Rules

Seller reserves are expressed as a conditioning rule at time t. That is $P_{i,t}$ is observed [denoted as $\boldsymbol{P}_{i,t}$] conditional upon it exceeding a seller threshold. Rule one conditions upon price at time 0.

[3] i.e. $r_{i,t} = r_{m,t}\ (\varepsilon_{i,t})$

That is, $\boldsymbol{P}_{i,t} = P_{i,t} \mid P_{i,t} > P_{i,0}$. This assumes that all prices are observed in the initial period, and that a sale is only made if the price increases. The rationale for rule one is that a sale occurs only when the price increases beyond its beginning value. Rule two conditions upon a previous purchase price, chosen randomly from earlier observed transactions for the asset, assuming that all assets trade in the first period: $\boldsymbol{P}_{i,t} = P_{i,t} \mid P_{i,t} > \mathit{random} \in \{\boldsymbol{P}_{i,0} \ldots \boldsymbol{P}_{i,\tau<t}\}$. The rationale for rule two is that a sale occurs only when the price exceeds the buyer's purchase price, where the buyer may have acquired the asset at a transaction chosen at random from the asset's transaction history. Rule three conditions upon the maximum past observed value: $\boldsymbol{P}_{i,t} = P_{i,t} \mid P_{i,t} > \mathit{max} \{\boldsymbol{P}_{i,0} \ldots \boldsymbol{P}_{i,\tau<t}\}$. The rationale for rule three is that the seller will wait until the asset price exceeds its historical high. Rule four conditions upon the N transaction prices for assets observed last period, and scales them according to the known quality differences observed in the first period: $\boldsymbol{P}_{i,t} = P_{i,t} \mid P_{i,t} > (\boldsymbol{P}_{i,0})(\mathit{max} \{\boldsymbol{P}_{1,t-1} \ldots \boldsymbol{P}_{N,t-1}\})$. The rationale for rule four is that the seller observes the best asset sold for the highest price last period, and then scales the asking price for quality variation.[4] When the condition cannot be satisfied, the price is not observed. We also estimate two additional indexes, based upon rule three transactions. Instead of averaging across observed transactions, we estimate two return series via a maximum likelihood procedure called the *repeat-sales regression*. This procedure is discussed in the following section.

II.3 Repeat-Sales Regression

[4] Notice that rule four implicitly assumes that the seller takes the maximum price observed last period as the price for the highest quality good. This price is then scaled down by the quality differential observed in the first period between the good held by the buyer and the highest quality good. This assumes that the buyer knows exactly the relative quality of his or her good. Taylor (1983) shows how uncertainty about quality can lead to the actual choice to bring a good to market.

The repeat-sales regression is used in both real estate and art market research to address the problem of infrequent transactions in asset markets, and to control for quality variation in transacting assets from auction to auction (see Bailey, Muth and Nourse (1963), Case and Shiller (1986) for real estate market examples, and Anderson (1974) and Goetzmann (1993) for art market examples). The regression uses only matched buy prices and sale prices to infer changes in returns from period to period. When prices for all assets are observed each period the regression is equivalent to a simple average across log returns each period, however in relatively illiquid markets such as the art market or the real estate market, this is clearly not the case. For such markets, when the time-series of returns of a given asset are log-normally distributed with i.i.d. errors around a market index, the repeat-sales regression represents the maximum-likelihood estimate of the equal-weighted index of all assets in the sample, calculated each period, regardless of whether the prices are observed or not. The model has also been adapted to conditions where returns contain a non-temporal component. Goetzmann and Spiegel (1994), for instance, include an intercept term in the regression to control for the components of asset returns unrelated to the stochastic errors or the trend in the index. There are reasons to expect that the selection bias due to seller reserves is non-temporal in nature. For instance, the magnitude of the "jump" from one transaction to the next caused by the selection bias is unrelated to the interval between sales. Thus, it is possible that the Goetzmann and Spiegel (1994) methodology may mitigate the selection bias. We report the results based upon the repeat-sales regression with and without the intercept term.

For purposes of study, we apply the repeat-sales regression to the conditional transactions generated under the third rule, that is, when owners only sell if the bid exceeds their own purchase price. The regression is performed in the following manner. Note in equation one that the log return over any period, say, from time b to time s, for s>b, is specified as the a log price difference $\boldsymbol{P}_{i,s}$ - $\boldsymbol{P}_{i,b}$, (which we denote as $\mathbf{R}_{i,s,b}$). The log market returns for each time period are estimated via a regression of the form: $\mathbf{R} = \mathbf{X}\boldsymbol{\mu} + \epsilon$, where $\mathbf{R}$ is a vector of all available returns formed from observed repeated asset sales, $\mathbf{X}$ is a dummy matrix with the number of columns equal to the number of time periods over which market data is available. Indicator variables in columns of $\mathbf{X}$ identify the time periods from the purchase date to the sale date. $\boldsymbol{\mu}$ is a vector of market log returns to be estimated by the regression. ϵ is the regression error, which is proportional to the number of time periods between the purchase and sale dates.[5] The advantage of the repeat-sales regression is that it does not limit the estimation of the mean to only those assets transacting in a given period t. Sales in periods later than t may contain information about the period t return, and the regression makes use of this information.

On additional consideration is what to do about first period prices. The simulation implicitly assumes that prices in the first period are unconditional. The return from an unconditional price to a price conditional upon exceeding a lower bound is certainly upwardly biased. Because of the effect that first period prices will have on the repeat-sales regression, we omit all observations for which the purchase price occurs in the first period.

---

[5] The GLS form of this regression scales observations by the square root of the number of holding periods.

III. Results

Deviations of the conditional series from the unconditional series are calculated by subtracting the unconditional return from the conditional return each period, for each simulation. Table one reports the summary statistics of these deviations for each behavioral rule. Mean residuals represent the annual average deviation from the unconditional index. Note that this number is positive for each rule, *including* the repeat-sales regression. The Goetzmann and Spiegel specification of the repeat-sales regression reduces the bias only slightly. This positive deviation results in an upward bias in price indices that is of the same order of magnitude as the mean annual return of the index itself, about 5%. In general, the magnitude of the bias in will depend upon the cross-sectional variation in the asset market. If all assets moved closely with the market with little residual variation, then the conditioning rules would have little effect. Thus, the bias in table one is a function of the residual variation of 40%. The bias in returns, as will as the standard deviation of the series' also appear to be related to the average percentage of the market that "transacts" each period. Certain rules appear to censor the market more than others. For instance, the "Max" and "Random" rules reduce the percentage of asset prices observed each period to around 30%. Table one also reports the autoregression coefficients for the conditional estimates. Note that they are negative and significant for the "Random" and "Max" reserve rules. In other words, average prices will appear to be mean-reverting from year to year. This is not the case, however, for the repeat-sale regression estimate. After adjusting for quality variation, the reversion disappears.[6]

---

[6] The results of quality-variation adjustment are consistent with Taylor's (1983) observation that heterogeneity in quality in the context of auctions without reserve can induce spurious negative autocorrelation. In our simulated setting we show how reseration rules may induce similar effects.

How well do the conditional indices capture the dynamics of the unconditional index? To answer this question, we regressed the unconditional return series' on the conditional return series.' Table two reports the results of one hundred regressions of the form

$$r_{m,t} = \alpha_j + \beta_j \bar{r}_{j,t} + \epsilon_{j,t} \qquad \textbf{(2)}$$

where j indexes the one hundred simulations. Because the conditional indices are more volatile than the unconditional series, the regression coefficients are typically below one. The t-statistics indicate that most coefficients are significantly different from zero, indicating that the estimates provide some information about the actual market. The exception is the "Max" rule which appears to result in a completely uninformative index. The R^2 's reported in the table suggest that less that half of the variation in the unconditional index can be explained by variation in the conditional indices. This suggests that information about the actual index behavior is lost as a result of the censoring process. Changes in the conditional indices due to quality variation mask fluctuations in the unconditional indices due to the stochastic process of returns. Note that the repeat-sales regression improves dramatically upon all of the conditional series.' Not only is the regression coefficient closest to one, but the R^2 is about three times greater than that of the "Random" rule. Clearly, the repeat-sales regression filters out a considerable amount of the noise due to quality variation. Given that the repeat-sales regression controls for quality variation by exactly matching purchase and sales prices of the assets, Why doesn't it do better? The answer is that the procedure only uses prices of 30% of the assets in the market each period. Thus, the remaining error in the index is due to small sample variation and truncation.

Little can be done to address the small sample variation since this is dependent upon the reserve rule, however it may be possible to address the truncation. Heckman (1979), for instance, suggests an estimation procedure when the probability of observing the truncated dependent variable can be estimated. Gatzlaff and Haurin (1992) apply the Heckman procedure to estimating a repeat-sales index of housing prices. In the current framework, the results in table one suggest that market volume or percent of properties transacting in a given period might contain some information about the probability of observing an increase in the index. In table three, we use the transaction percentage as an instrument for the probability of observing an increase. As before, the unconditional return series is the dependent variable. The independent variables are the conditional series and the percentage transactions measure. The coefficient on transactions in three of the four reservation strategies is significant and positive. In other words, high volume indicates an up market and low volume indicates a down market. Only the "Last" rule, and the repeat-sales regression show no improvement when percentage volume is included.

IV. Conclusion

Simulations of a private values asset market with prices conditional upon seller reserves reveal a disturbing phenomenon. An index of returns estimated from observed prices may bear little or no relationship to the unconditional series. The conditional series' have several unwanted properties. They have an upward bias which is conditional upon the variance of asset returns. They are typically more volatile that the unconditional series, and they often have a significant, but spurious negative autocorrelation from year to year. Measures of market volume may improve predictions of the unconditional index somewhat, but the percentage of variation in the unconditional index explained by variation in the conditional index remains small. The greatest improvement in estimating the unconditional index is afforded by using hedonic methods to control for fluctuations due to variations in average quality. One commonly used procedure, the repeat-sales regression, estimates an index that has higher explanatory power. The repeat-sales procedure eliminates the spurious negative autocorrelation in the simulations and it reduces the "excess volatility" of the index. Unfortunately, due to the one-sided censorship of observed prices each period, even repeat-sales estimates suffer from positive biases which cannot be controlled by including transactions volume. While we had some hope that the inclusion of an intercept term in the repeat-sale regression would mitigate the effect of seller reserves, it did not. We can only conclude that, in the presence of one or more reservation strategies used by asset owners, the repeat-sales index is biased upwards.

This bias may have serious consequences. The repeat-sales regression is used as a method of estimating capital appreciation returns in the housing markets. Previous estimates of the risk and return of investing in the single family home (see Goetzmann, 1993, for instance) have been based upon summary statistics about repeat-sales returns. Repeat-sales indices are currently used by mortgage insurers to estimate loan to asset values in residential housing markets. Case, Shiller and Weiss (1994) have suggested using the repeat-sales regression index as the basis for settlement of real estate futures contracts. All of these applications are potentially affected by biases induced by investor behavior.

References

Anderson, R.C., 1974, "Paintings as Investment," *Economic Inquiry* March, 12, 13-26.

Ashenfelter, O. and D. Genovese, "Testing for Price Anomolies in Real-Estate Auctions," *American Economic Review*, 80:2, 501-505.

Bailey, M.J. , R. F. Muth and H. O. Nourse, 1963, "A Regression Method For real Estate Index Construction," *Journal of the American Statistical Association*, 58, 933-942.

Case, K., and R. Shiller, 1987, "Prices of Single Family Homes Since 1970: New Indexes for Four Cities," *New England Economic Review* (September/October) 45-56.

Case, K., R. Shiller, A. Weiss, 1993, "Index-Based Futures and Options Markets in Real Estate," *Journal of Portfolio Management* Winter, 1993.

Gatzlaff, D., and R. Haurin, 1993, "Selection Bias and Real Estate Index Construction (?)," Working Paper, Florida State University, Tallahassee.

Goetzmann, W., 1993, "The Single Family Home in the Investment Portfolio," *The Journal of Real Estate Economics and Finance* 6, 201-222.

Goetzmann, W. and M. Spiegel, 1994, "Non-temporal Components of Housing Price Appreciation," *forthcoming, Review of Economics and Statstics*.

Taylor, W., 1992, "The Estimation of Quality-Adjusted Auction REturns With Varying Transaction Intervals," *Journal of Financial and Quantitative Analysis*, 27 (1), 131-142.

Taylor, W., 1983, "The Estimation of Quality-Adjusted Rates of Returns in Stamp Auctions," *Journal of Finance*, 38, 1095-1110.

Table 1
Summary Statistics of Deviations From Unconditional Average Series'

Summary statistics in this table are calculated over 100 simulations of market histories of forty years, in a market comprised of 100 assets of quality varying continuously from zero to one. Log price relatives for each year are generated according to the model:

$$r_{i,t} = r_{m,t} + e_{i,t}$$

where $r_{m,t}$ is distributed normally with mean of .05 and standard deviation of .10, and $e_{i,t}$ is distributed normally with mean of zero and standard deviation of .40. The unconditional series' mean for each simulation is calculated as the equal-weighted index of returns for all 100 assets. The conditioning rules are described in the text. The repeat-sales regression is applied to the transactions observed conditional upon selling only when the price exceeds the owner's purchase price. Transactions refers to the average percentage of the market transacting each period. Averages of autoregression statistics are calculated across statistics from 100 regressions of the form:

$$\bar{r}_{j,t} = \alpha_j + \beta_j \bar{r}_{j,t-1} + \epsilon_{j,t}$$

	"Start"	"Last"	"Random"	"Max"	"RSR"	"GS-RSR"
	$P_{i,t} > P_{i,0}$	$P_{i,t} >$ *max* $\{\mathbf{P}_{i,0} \ldots \mathbf{P}_{i,\tau<t}\}$	$P_{i,t} >$ *random* $\in \{\mathbf{P}_{i,0} \ldots \mathbf{P}_{i,\tau<t}\}$	$P_{i,t} > (\mathbf{P}_{i,0})$(*max* $\{\mathbf{P}_{1,t-1} \ldots \mathbf{P}_{N,t-1}\}$)	Repeat-Sale Regression	RSR w/ intercept
mean	.0243	.0115	.0484	.0501	.0472	.0424
standard deviation	.1061	.1582	.1678	.3456	.177	.177
Autoregression Coefficient	-.079	-.192	-.342	-.353	.0423	.044
Autoregression t-statistic	-.505	-1.22	-2.26	-2.27	.268	.279
Transactions %	.713	.746	.316	.270	.316	.316

Table 2
Regressions of Unconditional Return Series on Conditional Return Series'

Averages are calculated across statistics from 100 regressions of the form:

$$r_{m,t} = \alpha_j + \beta_j \overline{r}_{j,t} + \epsilon_{j,t}$$

Where j indexes the 100 simulations, "Mean" represents the average beta coefficient, "t-stat" represents the average t-statistic, "Std" represents the standard deviation of the regression coefficient distribution and "Med R^2" represents the median R^2.

	"Start"	"Last"	"Random"	"Max"	"RSR"	"GS-RSR"
	$P_{i,t} > P_{i,0}$	$P_{i,t} > \mathit{max}\ \{\mathbf{P}_{i,0} \ldots \mathbf{P}_{i,\tau<t}\}$	$P_{i,t} > \mathit{random} \in \{\mathbf{P}_{i,0} \ldots \mathbf{P}_{i,\tau<t}\}$	$P_{i,t} > (\mathbf{P}_{i,0})(\mathit{max}\ \{\mathbf{P}_{1,t-1} \ldots \mathbf{P}_{N,t-1}\})$	Repeat-Sale Regression	RSR w/ intercept
Mean β	0.524	0.302	0.25	0.019	0.577	0.577
t-stat β	4.05	3.5	2.59	0.36	5.93	5.93
Standard Deviation of β	0.178	0.0925	0.121	0.059	0.115	0.114
Median R^2	0.281	0.235	0.162	0.015	0.483	0.484

Table 3
Volume as a Predictor of Returns

Averages are calculated across statistics from 100 regressions of the form:

$$r_{m,t} = \alpha_j + \beta_j \bar{r}_{j,t} + \gamma_j \bar{L}_{j,t} + \epsilon_{j,t}$$

Where $L_{i,t}$ represents the average fraction of the jth market for which transactions are observed in period t. "Mean" represents the average beta coefficient, "Std" represents the standard deviation of the regression coefficient distribution, Liq and "Med R^2" represents the median R^2.

	"Start"	"Last"	"Random"	"Max"	"RSR"	"GS-RSR"
	$P_{i,t} > P_{i,0}$	$P_{i,t} > \mathit{max}\ \{\mathbf{P}_{i,0} \ldots \mathbf{P}_{i,\tau<t}\}$	$P_{i,t} > \mathit{random} \in \{\mathbf{P}_{i,0} \ldots \mathbf{P}_{i,\tau<t}\}$	$P_{i,t} > (\mathbf{P}_{i,0})(\mathit{max}\ \{\mathbf{P}_{1,t-1} \ldots \mathbf{P}_{N,t-1}\})$	Repeat-Sale Regression	RSR w/ intercept
mean β	0.579	0.241	0.309	0.008	0.568	0.57
standard deviation of β	0.169	0.125	0.084	0.058	0.12	0.118
Transactions γ	0.004	0.001	0.005	0.003	0.004	0.004
standard deviation of γ	0.001	0.001	0.001	0.001	0.005	0.005
t-statistic of γ	2.21	0.768	2.88	2.35	1.35	1.25
Median R^2	0.364	0.176	0.385	0.144	0.538	0.537