EVALUATING REAL ESTATE VALUATION SYSTEMS

BY

ROBERT J. SHILLER
AND
ALLAN N. WEISS

COWLES FOUNDATION PAPER NO. 983

COWLES FOUNDATION FOR RESEARCH IN ECONOMICS
AT YALE UNIVERSITY
Box 208281
New Haven, Connecticut 06520-8281
1999

Journal of Real Estate Finance and Economics, 18:2, 147–161 (1999)

Evaluating Real Estate Valuation Systems

ROBERT J. SHILLER
Cowles Foundation, Yale University, 30 Hillhouse Avenue, New Haven, CT 06520

ALLAN N. WEISS
Case Shiller Weiss, Inc., 1698 Massachusetts Avenue, Cambridge, MA 02138

Abstract

A framework for comparing real estate valuation systems (including automated valuation models (AVMs) and current appraisal methods) is proposed. The density estimation and profit simulation (DEPS) method measures quality of a valuation system by simulating benefits to the mortgage lender who uses this method in mortgage underwriting to limit mortgage portfolio losses due to default. Related simple measures relevant to the selection of a valuation system are also discussed: skewness of the distribution of errors, correlation of valuation errors with current selling price errors, correlation of errors of the valuation system with errors of valuation systems used by competing mortgage lenders, and other measures.

Key Words: appraisal, automated valuation models (AVMs), accuracy, mortgage, default, foreclosure, density estimation and profit simulation (DEPS) method

For a real estate valuation system (or appraisal method) to perform well, it should help the mortgage originator, who uses the system to judge whether to approve mortgage applications, prevent default-related expenses and loss of principal without incurring too many other costs, and put the lender at an advantage, appropriately defined, in terms of other firms. Unfortunately, while this objective is simple to state, there are many difficulties in translating it into a criterion for judging real estate valuation systems. There are many complicated links from a real estate valuation system to its ultimate costs and benefits to the mortgage lender. Because of these difficulties, faulty decisions are being made in making comparisons between competing statistical valuation systems or between a purely statistical valuation system and an appraisal.

At this time of rapid development of automated underwriting systems and automated valuation models (AVMs), the importance of evaluating real estate valuation systems is greater than ever, and at the same time, because the competitive situation for mortgage lenders is complex and automated underwriting systems are themselves in flux, the costs and benefits of competing valuation systems are difficult to define.

We assume here that data are available for testing valuation systems. The data are assumed to be in the form of (1) current estimated values rendered by each of the competing valuation systems (including appraisal methods) for individual properties, (2) actual unforced (not defaulted) selling prices, current as well as subsequent, for the same properties, (3) loan amount, as well as (4) other information that may be available about the properties or the valuations. Data sets with some of this information are often constructed by mortgage originators when they are considering changing their valuation systems.

Many tests evaluating valuation systems using such data do nothing more than compare measures of dispersion of the errors—of the percentage difference of an actual sales price relative to the predicted price. Dispersion may be measured by variance or mean absolute error (see, for example, Cole, Guilkey, and Miles, 1986, or Webb, 1994). With these tests, the valuation system with the lowest dispersion is judged the best. While these tests are useful, there are many reasons why they may be misleading, and other measures of valuation system success are also important.

We provide here a framework that will enable us to compare different valuation systems (including comparing appraisals with statistical valuation systems) in a manner that is directly revealing of the results for the mortgage underwriting process. The framework that we propose has two aspects. First, we use the framework to develop a density estimation and profit simulation (DEPS) method that, given an assumed relation of mortgage lender profits to the true loan-to-value ratio, provides a direct estimate of the expected value of the valuation system to the mortgage lender. The DEPS method also allows us to define an optimal competitive strategy for the mortgage originator. Second, the framework suggests various simple measures of the quality of valuation systems that are generally overlooked today. Some of these simple measures are (1) skewness of the distribution of valuation errors, (2) correlation of valuation error made by the valuation system with actual current selling price error for the same property, and (3) correlation of valuation error on a property by the method with valuation errors by methods used by competing mortgage lenders. We also discuss other simple measures and estimation issues.

Our analysis shows the extreme complexity of the issues involved in judging valuation systems and shows how the complexity can be dealt with using the DEPS method. At the same time, our analysis reveals that the DEPS method is not the only way to understand the differences among valuation systems; the simple measures of valuation system performance should also be useful.

1. Type I and Type II Errors

Our analysis of the usefulness of real estate valuation systems must begin with the recognition that there are two kinds of collateral valuation errors that are costly to mortgage lenders. Borrowing terminology from theoretical statistics, we use the term *type I error* to refer to the error of rejecting a mortgage applicant who would not have defaulted. The costs to this error are the cost of the loss of profit from writing a good mortgage, the loss of good will and further business and referrals from the rejected applicant, and the cost of the underwriting process (including the cost of purchasing the real estate valuation) when no loan is written. We use the term *type II error* to refer to the error of accepting a mortgage applicant who defaults. The costs to this error are the costs of problem-loan servicing, including collection efforts, cash-flow advances, foreclosure administration, legal costs, and principal losses related to a lender's sales of the foreclosed property (as well as the cost of the original underwriting process). Our description of the method here presumes that the mortgage originator is a portfolio lender that is directly impacted by loss of principal in default, though much the same analysis might be used in other cases.

Presumably the potential cost for each property from type I error is less than that of type II error: potential default loss is greater than potential profit from writing a good mortgage. Thus, in some sense one might think that type I error is less important than type II error.[1] However, from the standpoint of comparing valuation systems, the type I error is no less important since the approval criteria are set so that type I error will occur more often, at the margin, than does type II error. We must consider both kinds of error in evaluating approval methods. If we considered only type I error, we could reduce this to zero by selecting a valuation system that always gives a very high value for each property, so that no applications are ever rejected. If we considered only type II error, we could reduce this error to zero by selecting a valuation system that always gives a very low value for each property, so that all applications are rejected. Neither of these extremes makes any sense for mortgage lenders.[2]

Simple variances of real estate valuation errors are not sufficient statistics for the type I and type II errors that are relevant to judging the valuation systems. They do not take account of the shape of the distribution of errors. Taking full account of the distribution of errors of the valuation system is not something that can be done formally until we specify the precise use of the valuation in the underwriting method and the nature of the costs associated with underwriting errors.

2. The Simplest Model for the DEPS Method

Let us consider first a very simple profit-simulation model that allows us to take account of type I error and type II error with the DEPS method and that suggests some of the complexity of the problem of comparing valuation systems. The model assumes that the mortgage lender operates in isolation, so that there is no issue of selection bias that would arise if the lender's underwriting guidelines permit acceptance of borrower loan applications that other lenders reject. This simple profit-simulation model also makes no use of the selling price of the property itself: we assume that we are either applying our analysis to refinance mortgages, for which no current purchase price is available, or our mortgage applications have already been prescreened on the basis of selling price relative to loan applied for. Also, the context here is evaluating valuation systems only, and so we take credit-underwriting considerations as a constant.

The value ν is the natural log of the true value of the property (the price it would obtain in a normal sale and not a foreclosure sale) scaled by subtracting the natural log of the loan balance applied for. We use the symbol $\hat{\nu}$ to refer to the natural log estimated value of the property also scaled by subtracting the natural log of the loan balance applied for. Let $f(\nu, \hat{\nu})$ be the joint probability density function for ν and $\hat{\nu}$ among the loan applications that the mortgage lender wishes to consider with a valuation, including both applications that will be accepted and applications that will be rejected on the basis of the valuation. A simple measure of the quality of the valuation system, similar to measures that are widely used to evaluate valuation systems today, is σ_1^2, the second moment (or if $\hat{\nu}$ is unbiased, the variance) of $\nu - \hat{\nu}$. In terms of $f(\nu, \hat{\nu})$ this is

$$\sigma_1^2 = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} (\nu - \hat{\nu})^2 f(\nu, \hat{\nu}) d\nu d\hat{\nu}. \tag{1}$$

The valuation system with the lowest σ_1^2 is judged the best.[3] This simple measure will be contrasted with other measures.

Let us use the symbol π to represent the expected profit as a fraction of the loan amount applied for with a single mortgage application. (π is not in logs, as are ν and $\hat{\nu}$.) To allow us to account for the relevant costs of the two types of error, let us suppose that π is a function of ν. We suppose that the function $\pi(\nu)$ is concave down, is very negative for small values of ν (reflecting such factors as foreclosure costs and the loss of property value in a foreclosure sale), is only slightly positive for large values of ν, and becomes flat (π becomes unaffected by ν) for high values of ν. This shape for the function embodies our assumption that costs of type II error (of accepting an application that was bad because of low ν) can be much larger than the benefits of accepting a good application. Figure 1 shows a hypothetical π function. We show this function over a range of ν from -0.5 to $+0.5$, a wider range than mortgage lenders hope to see very often in their loans. They do not intentionally make loans with negative ν; preventing mortgage loan approval on the negative values of ν is what the valuation system is supposed to do. But we must hypothesize a π function that is defined over wide range, if we are to use it properly to construct a valuation system that prevents approving mortgage loans when ν is very low. The function shown in figure 1 represents that being able to write a good mortgage loan is worth 1.5% of the loan balance to the mortgage lender, while, at the other extreme, making the catastrophic error of writing a mortgage loan where ν is -0.5 (meaning granting credit when the collateral is worth $e^{-0.5} = 0.61$ of the principal balance that is, writing a mortgage loan with a loan-to-value ratio of $e^{0.5} = 1.65$) would mean that the mortgage lender expects to lose 50% of the loan made.[4]

The mortgage lender decides on a threshold value $\hat{\nu}^*$ so that only mortgage applications for which $\hat{\nu}$ exceeds $\hat{\nu}^*$ will be accepted. The expected profit Π that the mortgage lender obtains—accounting for the profit function, the joint distribution $f(\nu, \hat{\nu})$ using this valuation, as well as the cost C of obtaining the valuation $\hat{\nu}$ (including associated application-processing costs)—is then

$$\Pi = \int_{\hat{\nu}^*}^{\infty} \int_{-\infty}^{\infty} \pi(\nu) f(\nu, \hat{\nu}) d\nu d\hat{\nu} - C. \tag{2}$$

The valuation system with the highest Π; is judged the best. Note that the expected profit Φ is not the expectation of $\pi(\nu)$ conditional on $\hat{\nu}$ being greater than $\hat{\nu}^*$, since such a conditional expectation would not take account of the probability of the type I error of rejecting a good mortgage. The expression embodies the assumption that the cost C of purchasing the valuation is borne even if the mortgage application is rejected. The expected profit Π is a sort of weighted integral of the probability density function, $f(\nu, \hat{\nu})$,

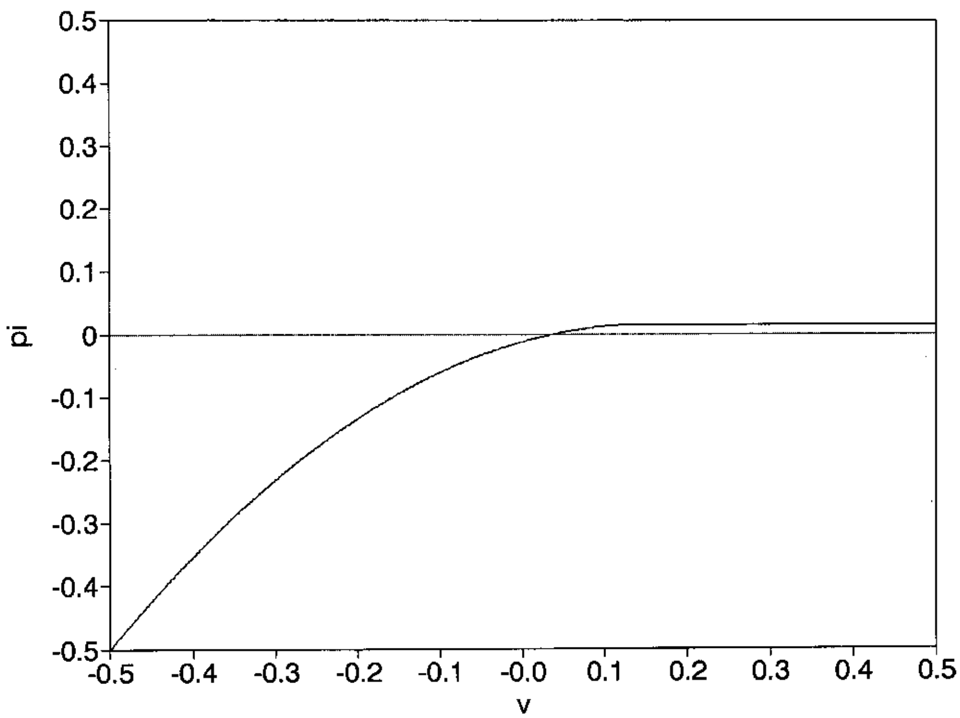

Figure 1. Hypothetical $\pi(\nu)$ function, π (expected profit divided by loan balance) against ν (natural log of value of property minus natural log of loan balance).

with weights $\pi(\nu)$ that are very different from the $(\hat{\nu} - \nu)^2$ that appeared in equation (1). The isoquants of $(\nu - \hat{\nu})^2$ are (plotting ν on the horizontal axis and $\hat{\nu}$ on the vertical axis) parallel straight lines with slopes of one, while the isoquants of $\pi(\nu, \hat{\nu})$ are parallel straight lines with slopes of infinity; only the region above of $\hat{\nu}^*$ is used. This shows immediately that the measure of the quality of the valuation system may be very different between (1) and (2).

It must be recognized that $\hat{\nu}^*$ is a choice variable for the mortgage lender. Assuming that the mortgage lender operates so as to maximize expected profits, then the lender will choose $\hat{\nu}^*$ to maximize (2)—so that $\partial\Pi/\partial\hat{\nu}^* = 0$, equation (3) is

$$\int_{-\infty}^{\infty} \pi(\nu)\mathrm{f}(\nu, \hat{\nu}^*)d\nu = 0, \tag{3}$$

and the mortgage lender should proceed to obtain the valuation only if the optimized Π is positive.

Loan applications where $\nu = \hat{\nu}^*$ are the marginal loans, for which the expected profits earned when the decision to lend proves successful are exactly matched by expected losses when the decision to lend proves detrimental. The lender is indifferent between making these loans or not making them, but for mortgage applications where $\nu > \hat{\nu}^*$, the mortgage lender can expect to make a profit. Note that the choice of $\hat{\nu}^*$—made as it is with expression (3), which involves $f(\nu, \hat{\nu})$—depends on characteristics of the valuation system, and so $\hat{\nu}^*$ will differ across valuation systems. This variation across valuation systems in the optimal $\hat{\nu}^*$ is very important to account for: one valuation system may have advantages over another only because it allows the mortgage lender to lower $\hat{\nu}^*$ and thereby lower type I error.

Given this model, we propose that a good method for comparing valuation systems, the DEPS method, for each valuation system (1) estimates the joint distribution $f(\nu, \hat{\nu})$ using data on actual values of properties and their estimated values using the valuation system and loan balances applied for, (2) hypothesizes a profit function $\pi(\nu)$, (3) uses equation (3) to derive the $\hat{\nu}^*$ for that valuation system, and then (4) uses equation (2) to derive Π. The valuation system with the highest Π would be, in terms of this simple model, the best method. We suppose that a nonparametric method or a very general parametric method of estimating the probability density function should be used if there are sufficient data, so that the shape of the function can be captured in some detail (see Tapia and Thompson, 1978; Révész, 1984; or Silverman, 1986).

Consideration of this model also suggests simple measures that can be obtained without going through the DEPS method. These measures go beyond just variance of valuation errors and would be considered as reflecting on the quality of a real estate valuation system. One such measure, which might be looked at in addition to the variance shown in equation (1), is the skewness of the distribution of $\nu - \hat{\nu}$.[5] Suppose we are comparing two real estate valuation systems that both are unbiased ($E(\nu - \hat{\nu})$ equals zero) and that both have the same variance given by equation (1). We would then generally prefer the valuation system whose error $\nu - \hat{\nu}$ is skewed to the right—whose large errors tend to undervalue a property. Really large undervaluation errors tend not to hurt the mortgage lender, since if the errors were somewhat smaller, then the mortgage application would be rejected anyway, and so there is no difference in costs to the mortgage lender. On the other hand, really large negative errors can be catastrophic to mortgage lenders if they cause lenders to issue mortgages on properties whose values are much less than their respective loan amounts. That is why positive outliers matter less than negative outliers. Thus, valuation systems with positively skewed distributions of $\nu - \hat{\nu}$ tend, other things equal, to perform better in the DEPS method with this model. Of course, saying that *other things equal* we would prefer the distribution to be skewed does not mean that designers of valuation systems should try to design their valuations so that the errors are skewed. They too should look at the whole set of factors, as summed up by the DEPS method analysis. There may, however, be a tendency for positively skewed errors in well-designed valuation systems: if there is a cost to reducing the right tail of the error, then perhaps the designer of the valuation system may choose not to bear this cost.

Measuring the skewness of the distribution of $\nu - \hat{\nu}$ is not a substitute for our DEPS method. For example, skewness might be related to the level of $\hat{\nu}$, and only when $\hat{\nu}$ is above $\hat{\nu}^*$ does skewness matter. One could derive other revealing simple measures, more closely related to the DEPS method, of the quality of a valuation system by simplifying the π function so that it is, for example, a step function. We might suppose that for ν below some threshold $\nu^- \pi$ is the negative constant π^- and that above some threshold ν^+ (where

ν^+ is greater than ν^-) π is the positive constant π^+ and π is zero between these two values. Then $\Pi = \pi^- \text{Prob}(\nu < \nu^- \text{ and } \hat{\nu} > \hat{\nu}^*) + \pi^+ \text{Prob}(\nu > \nu^+ \text{ and } \hat{\nu} > \hat{\nu}^*) - C$. Then useful summary statistics would be the two estimated probabilities. Users could use their own values for π^-, π^+, and C to evaluate Π given these probabilities.[6]

Reporting such measures, the skewness or the probabilities defined just above, is especially important in comparing statistical valuation systems with appraisals. Using measures of dispersion of valuation errors alone to make comparisons may be unfair to the appraisers, since appraisers operate with a sense of the costs of type I and type II error and tend to be (correctly) mindful of these costs in allocating their time and effort in the appraisal process. In appraising properties where the value is obviously sufficiently above the loan amount, the appraiser may see less benefit in expending extra effort to refine the valuation. The result could be a misleadingly high standard deviation of errors for appraisers, which should be considered offset by the high skewness.

3. Correlation of Valuation Error with Selling Price Error

In the preceding section, we disregarded the fact that for purchase mortgages, in contrast to refinance mortgages, the mortgage lender actually has two pieces of information: the valuation placed on the property, which provides $\hat{\nu}$, and the selling price today. We will use the symbol p to represent the natural log of the selling price today scaled by subtracting the natural log of the loan balance applied for. Note that the true log value used to compute ν is not the current selling price but is the price at next sale, on a date a few years down the road when mortgagor default propensity is greatest.

Let us define $f(\nu, \hat{\nu}, p)$ as the joint probability density function of the three quantities among all purchase mortgage applications that the mortgage lender is considering, even including those applications that might be ultimately rejected because the loan amount applied for is too high relative to the current selling price. The most common measure of the accuracy of valuation systems used in practice is actually σ_2^2—the second moment or, if $\hat{\nu}$ is unbiased, the variance of $p - \hat{\nu}$. This is not the variance σ_1^2 of $\nu - \hat{\nu}$. It is instructive, for comparison, to write σ_2^2 in terms of $f(\nu, \hat{\nu}, p)$:

$$\sigma_2^2 = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} (p - \hat{\nu})^2 f(\nu, \hat{\nu}, p)\, d\nu\, d\hat{\nu}\, dp. \tag{4}$$

As before, by the conventional criterion, the valuation system with the lowest σ_2^2 is judged the best.[7]

In contrast, the expected profit to a mortgage lender that accepts a loan application for which p is above some threshold value p^* and for which $\hat{\nu}$ is above $\hat{\nu}^*$ is

$$\Pi = \int_{p^*}^{\infty}\int_{\hat{\nu}^*}^{\infty}\int_{-\infty}^{\infty} \pi(\nu) f(\nu, \hat{\nu}, p)\, d\nu\, d\hat{\nu}\, dp - C, \tag{5}$$

and, as before, the valuation system with the highest Π is judged the best. Note that, much as in the preceding section, the expected profit Π is a sort of weighted integral of the probability density function, here $f(\nu, \hat{\nu}, p)$, with weights, determined by $\pi(v)$ and the thresholds $\hat{\nu}^*$ and p^*, that are totally different from the $(p - \hat{\nu})^2$ that appeared in equation (4).

As with the simpler model of the preceding method, the thresholds $\hat{\nu}^*$ and p^* are both decision variables for the mortgage lender. Taking the partial derivatives of the expected profit, equation (5), with respect to $\hat{\nu}^*$ and p^* and setting these to zero, we have the pair of equations:

$$\int_{\hat{\nu}^*}^{\infty} \int_{-\infty}^{\infty} \pi(\nu) f(\nu, \hat{\nu}, p^*) d\nu d\hat{\nu} = 0 \tag{6}$$

$$\int_{p^*}^{\infty} \int_{-\infty}^{\infty} \pi(\nu) f(\nu, \hat{\nu}^* p) d\nu dp = 0. \tag{7}$$

These two equations, solved together, will yield the optimal thresholds $\hat{\nu}^*$ and p^* for purchase mortgages. Note that changing the valuation system will in general cause changes in both $\hat{\nu}^*$ and p^*.[8]

Proceeding as in the preceding section, the DEPS method of comparing valuation systems is, for each valuation system, first to estimate the probability density function $f(\nu, \hat{\nu}, p)$, then to use equations (6) and (7) to derive the optimal $\hat{\nu}^*$ and p^*, and finally to evaluate (5). The valuation system with the highest Π will be judged best. The model we have defined is actually rather complex: in comparing just two valuation systems there are four thresholds to be derived, and getting these wrong in the comparison process could reverse the conclusions.

A simpler statistic that bears on the results that we will obtain using the DEPS method with this model is the correlation of the valuation error $\nu - \hat{\nu}$ with the current selling-price error $\nu - p$. It is important to realize that errors in valuations usually do not result in default losses among purchase mortgages: the losses tend to occur only when the selling price of the house is too high relative to true value *and* the valuation is also too high relative to true value; only then will the mortgage lender be influenced to make a costly error. Selling-price error is the error made by the buyer of the house who pays too much for the house relative to its true market value—that is, too much relative to what the typical buyer would pay. If this error is independent of the valuation system error, then the probability of a loss is related essentially to the product of the probabilities of the two large errors; if both probabilities are small, this product will be much smaller. It is important to note this, since a tendency to make occasional large errors is common among statistical valuation systems, and yet this tendency is really not very costly if the errors are uncorrelated with current selling-price errors.

There are a number of reasons why there might be a correlation between the valuation error and the current selling price error. Notably, since with purchase mortgages FHLMC and FNMA underwriting guidelines stipulate that the actual selling price is information available to the appraiser, the price might influence the valuation. An unethical appraiser or other valuer might always just support the contract selling price, in which case the valuation is worthless, even though the variance of its error with respect to this selling price is zero.

However, even when judging statistical valuation systems that do not make use of the actual selling price, there is a chance for a correlation of errors. Even when the actual selling price of the property is not used by the appraiser or other valuer, there is still a possible correlation between the selling error and the valuation error if the prices of comparables used by the appraiser are correlated with the selling-price error on the house. Such correlation is possible, since the reasons that cause buyers to pay too much for a house may apply to neighboring houses as well. Moreover, buyers of houses might be looking at the same comparables in trying to judge how much the subject house is worth. Because of buyer's tendency to do this, there may be temporary local mispricings of real estate, mispricings that may not remain later should such mortgagors default on the loans secured by the defaulted properties. There is the possibility of significant correlation between valuation errors and selling-price errors if there is too great reliance in the valuation system on the recent selling prices of nearby properties. Appraisers face a difficult decision that can be made only with careful judgment—how nearby or similar comparables should be weighed (Isakson, 1986). This correlation between sales price errors and valuation errors occurs because the nearby properties are not only similar but are viewed as similar by others and because, since there is likely to be some inefficiency of the market, the errors may be expected to decline to zero later.

Reporting the correlation between sales-price errors and valuation errors is not the same as reporting the outcome of the DEPS method, and one should be mindful of some important problems that may emerge when this correlation is relied on in judging valuation systems. It was noted above that appraisers tend (correctly) to allocate less time and effort to refining valuations in which the value is obviously sufficiently above the loan amount. Doing this could tend to generate a correlation between appraisal errors and sales-price errors that is perfectly harmless and that is of no consequence to the mortgage lender.

By assuming here that true value ν is represented in empirical work by the unforced sale price a few years after the mortgage is issued,[9] this version of the DEPS method builds into the evaluation method an incentive for those who provide valuations to provide not estimates of current value of properties but instead forecasts of future values. Many appraisers do not describe what they do as forecasting, but there is a recognition in the appraisal industry that forecasting of some sort is inevitably involved (see Lovell, 1992, or Shlaes, 1992). Real estate prices are in fact substantially forecastable, in contrast to prices of liquid financial assets (Case and Shiller, 1989; Poterba, 1991; Ito and Hirono, 1993).[10] These incentives are implied by our DEPS method because, indeed, such forecasts and not estimates of current value are what mortgage lenders really need. If we wish separately to evaluate forecasts from current valuations, then the method must somehow be amended to allow this, such as by adjusting all valuations to be compared by the amount of forecasted change according to the forecaster used by the mortgage lender.

4. Models Involving Games Against Competing Firms

In applying a valuation system, one must recognize that one's method is not used in isolation: other mortgage lenders are also using valuation systems, and the benefit to a given valuation system will tend to depend on the nature of the valuation system used by competitors. The risk that a mortgage lender faces is that it will fail to see some information that caused other mortgage lenders correctly to deny mortgage credit. It might become a haven for all such rejected applicants. Alternatively, if the mortgage lender can find a way to use the information in the valuation system to find good applications among those rejected, there may be a very large profit opportunity.

There are many different possible ways to model games between the mortgage lender and its competitors; we will consider one extreme case as an illustration. Suppose that a mortgage lender has no ability to draw prospective borrowers that are deemed as creditworthy applicants by the competitors; the mortgage lender is left solely to scavenge among the substandard applications.[11] Calling $\hat{\nu}_1$ the natural log value of the subject property (scaled by subtracting the natural log of the loan balance) produced by the valuation system used by the mortgage lender, and $\hat{\nu}_2$ the corresponding scaled natural log value used by competitors, let us define $f(\nu, \hat{\nu}_1, p, \hat{\nu}_2)$ as the joint probability-density function of these variables, among loan applicants that are assumed to apply both at the mortgage lender and the competitors. And supposing that the competitors reject properties for which $\hat{\nu}_2$ is less than $\hat{\nu}_2^*$ or p is less than p_2^*, then the expected profit Π_1 to the mortgage lender (lender 1) is given by expression (8),

$$\Pi_1 = \int_{-\infty}^{\hat{\nu}_2^*} \int_{p_1^*}^{\infty} \int_{\hat{\nu}_1^*}^{\infty} \int_{-\infty}^{\infty} \pi(\nu) f(\nu, \hat{\nu}_1, p, \hat{\nu}_2) d\nu d\hat{\nu}_1 dp d\hat{\nu}_2$$
$$+ I(p_2^* - p_1^*) \int_{\nu_2^*}^{\infty} \int_{p_1}^{p_2^*} \int_{\hat{\nu}_1^*}^{\infty} \int_{-\infty}^{\infty} \pi(\nu) f(\nu, \hat{\nu}_1, p, \hat{\nu}_2) d\nu d\hat{\nu}_1 dp d\hat{\nu}_2 - C_1, \qquad (8)$$

where $I(p_2^* - p_1^*)$ is an indicator function that is 1 if $p_2^* - p_1^*$ is positive and is zero otherwise, and C_1 is the cost of the valuation $\hat{\nu}_1$ to the mortgage lender. As in the preceding sections, the mortgage lender must partially differentiate Π_1 with respect to $\hat{\nu}_1^*$ and p_1^* and set these partial derivatives to zero to define the optimal thresholds; the mortgage lender should proceed to obtain a valuation at cost C_1 only if the optimized Π_1 is positive. The DEPS method would, with this model, as in preceding sections, go through these steps with an estimated probability density function to find the optimized Π_1 for each competing valuation system, as a way of comparing them.

A valuation system that underperforms the method used by competitors in terms of variance of errors might, by the DEPS method, be better than a valuation system that outperforms the competitors in terms of the variance of errors. This can occur if the valuation system is in effect using different information than that of the competitors. Obviously, a valuation system that gives the same valuations as the competitors' will produce zero profits for the mortgage lender by this model. At another extreme, if the information used by the valuation system is completely different from that used by competitors but no better in terms of variances of forecast errors, then there could be large expected profits for the mortgage lender. This suggests that a simple measure of the value of the valuation system would be the correlation of the error made by the valuation system with the error made by competitors' valuation system—the smaller the better.

This simple model of competition with other mortgage lenders could be extended in various ways. We need not assume that the mortgage lender is confined to consideration of mortgage loan applicants that are rejected by competitors. One could assume instead that the mortgage lender can expect to originate a certain market share of mortgages that would otherwise be approved by competitors as well. We need not assume that the thresholds $\hat{\nu}_2^*$ and p_2^* used by competitors are exogenous. We could compute for the competitors the optimal response to the subject mortgage lender's choice of thresholds by differentiating their profit function with respect to these, taking $\hat{\nu}_1^*$ and p_1^* as given, and could compute a Nash equilibrium when both the subject mortgage lenders and competitors are both optimizing.

5. Other Considerations

5.1. Risk Considerations

For a risk-neutral mortgage loan underwriter, the correlation between valuation errors of different properties represented in its portfolio of mortgages is of no account because the expected losses are the same whatever the correlation. However, if errors are correlated across properties, then the variance of the underwriting losses to the portfolio can be magnified, and this magnification can matter to risk-averse underwriters. The DEPS method could be amended to account for the variability of profit on the entire portfolio of mortgages, as well as the expected profit.[12]

The more a valuation system uses outdated information, the more, for a given variance of errors, the valuation errors are likely to be correlated across properties. The reason for this is that the more the data are out of date, the more time there is for regionwide factors to change and create errors common to all properties. By analogy, the more the valuation system relies on a same small subset of existing properties for comparables, the higher the chance that the same error will infect many different properties' valuations. The more a valuation system relies on a cost approach, the higher the chance that errors in cost measurements will affect all properties (see Marchitelli, 1992).

5.2. Costs of Failure to Make a Valuation

Many valuation systems are often unable to make a valuation at all, since the information on which the method is based may not be available. Alternatively, the producer of the valuation may be computing a standard error for the valuation and reporting only those valuations that have a standard error below some threshold decided on by the valuer, so that the errors on reported properties will tend to be small. The DEPS method should be modified to take into account the costs to the mortgage lender when no valuation is provided.

5.3. *Making Use of the Standard Error of the Valuation*

Some valuation systems may provide estimated standard errors of the valuations, which inform the mortgage lender how good the information is about the true value. These standard errors, if they vary from property to property as a reflection of the differing quality of information about them, are potentially very useful information to the mortgage lender, since the values $\hat{v}^*$ and p^* can be made a function of them. Rather than see the provider of the valuation often fail to report a valuation when the standard error is high, the mortgage lender would prefer to have the information and make use of it optimally to decide whether to accept the mortgage, even though by offering these valuations the valuation provider would be increasing the variance of the valuation error among reported properties. The DEPS method can be modified to include the reported standard errors in the probability density function and to derive the optimal Π that is possible using the information in the standard errors.

5.4. *Using Multiple Valuation Systems*

In the case where a valuation system is found to contain information that is not contained in other valuation systems, it should be explored whether there are profits to be gained by using both methods. The DEPS method could be modified to consider whether making the mortgage-approval decision depend on two valuations is profitable, taking account of their combined costs.

5.5. *Estimation Issues*

Of course, sample size for estimation of the probability density function, or for production of the simpler measures we have discussed, must be made sufficiently large that the valuation systems can be successfully compared. Many evaluations that we have seen have confined their attention to a dozen or so properties and are clearly inadequate to judge the relative merits of the competing valuation systems unless the difference between the accuracies of the methods is enormous.

Even with sufficient sample size, however, there are problems if the sample is drawn primarily from a small geographical area, a small historical interval, or a narrow price range. Errors may be serially correlated (through time) or spatially correlated (across contiguous regions), and errors may have different characters in different price ranges. The impression that we have of a large number of observations may be very misleading. One may wind up choosing as best a statistical valuation service that has only a fortuitous advantage. The likelihood that a particular service will do well in tests based on a small interval of time or a small geographical area or price range is larger the more that service tends to rely on general statistical methods applied to that area or time interval or price range.

6. Conclusion

We have seen that conventional measures of the success of valuation systems, such as measures of dispersion of valuation errors, often are unfair tests of the usefulness of the valuations to the mortgage lenders. Using such measures for comparisons of statistical valuation systems with conventional appraisals are especially likely to be unfair, since appraisers, in providing a service to mortgage lenders, have in mind the relative costs and likelihood of the different kinds of errors that are not taken into account in the conventional measures.

The DEPS method proposed here will correct these deficiencies of the conventional measures of success, but the DEPS method in practice may be somewhat difficult to implement. To be implemented properly, all of the considerations that we have mentioned should be represented in the analysis, meaning that, ideally, the most complicated DEPS model discussed above should be used.

It should be kept in mind, however, that the choice of valuation system is a central strategic decision for a mortgage lender, very much a part of its basic competitive strategy. The costs of doing the evaluation of mortgage-valuation systems properly are likely to be small relative to the potential strategic advantages obtained by making the right choice of valuation systems. It is well worth the while of mortgage lenders to obtain a suitably large data set on competing property valuation systems and evaluate them with a method like our DEPS method.

The various simple measures of valuation system quality, such as the skewness of the distribution of valuation errors or the correlation of selling-price errors with valuation errors, probably cannot be used well in isolation without reference to the basic issues that the DEPS method attempts to quantify. On the other hand, since there is much latitude of choice in devising the expected profits model that underlies the DEPS method, those who evaluate valuation systems will probably want to understand the characteristics of the valuation systems as revealed by these simple measures.

Acknowledgment

The authors are indebted to John M. Clapp, Carmelo Giaccotto, Steven Grenadier, Terrance E. Loebs, and Douglas A. McManus for helpful comments.

Notes

1. The type I error costs may also tend to be excluded from analysis since they are less easily quantified.
2. In our analysis here, which is aimed at the fundamental usefulness of valuation systems, we assume that the approval criteria are decided by the mortgage lender, which does not defer to FNMA or FHLMC criteria.

3. Alternatively, the $(\nu - \hat{\nu})^2$ term in the above expression could be replaced by abs $(\nu - \hat{\nu})$, so that the measure would be mean absolute error. The $(\nu - \hat{\nu})^2$ term could also be replaced with a function $e(\nu - \hat{\nu})$ that is positive only for values of $\nu - \hat{\nu}$ that are less than a small number ϵ in absolute value and that is zero elsewhere. In the limit, as ϵ goes to zero, the expression would allow us to measure how well $\hat{\nu}$ represents the "most probable sales price."
4. Note that the shape of the π function resembles (the negative of) the shape of the familiar function relating the value of a put option to the price of the underlying. Indeed, the mortgage lender is in a sense writing a put option, an option to default. The shape of the π function would differ across mortgage lenders that pursue different business strategies. For example, with the B&C mortgage lending market, aimed at a substandard mortgagor credit spectrum, interest rates and points charged may be much richer, offsetting the high servicing and default-related costs. For these lenders, the π may be much higher than 1.5% for positive ν.
5. Skewness is measured as the third moment around the mean divided by the cube of the standard deviation.
6. Note that with such a step π function, taking C as a sunk cost, the expected cost of type I error is $\pi^+\text{Prob}(\nu > \nu^+$ and $\hat{\nu} < \hat{\nu}^*)$ while the expected cost of type II error is $-\pi^-\text{Prob}(\nu < \nu^-$ and $\hat{\nu} > \hat{\nu}^*)$. The first probability can be inferred from the probabilities discussed in the text if $\text{prob}(\nu > \nu^+)$ is given.
7. If we replaced $p - \hat{\nu}$ in the above expression with $\nu - \hat{\nu}$, so that the expression more nearly resembles expression (1), and if the variables p and ν are jointly normally distributed, then the optimal $\hat{\nu}$ that is a function of p would be the fitted value in a regression (with constant term) of ν on p. But running such a regression is not recommended since it maximizes an objective function that misrepresents lenders' true interests.
8. The problem of determining thresholds could be simplified if the two are constrained to be the same.
9. We have supposed here that data available for evaluation of valuation system are unforced sale prices. Of course, the loss of value caused by forced sale after default is a relevant factor for our analysis. We suppose that this factor is represented in the π function.
10. In fact, there is also an incentive to provide mortgage lenders as well with forecasts of the loss in value that occurs to properties securing loans that are defaulted on, therefore refining their knowledge of the π function.
11. As above, we are of course taking as constant here factors other than collateral value in the underwriting process, and we are assuming that the π function already reflects strategic decisions and investments already made in terms of lender reputation and market position.
12. Cole, Guilkey, and Miles (1987) addressed this issue with data on commercial appraisals.

References

Case, Karl E., and Robert J. Shiller. (1989). "The Efficiency of the Market for Single Family Homes," *American Economic Review* 79 (March), 125–137.

Cole, Rebel, David Guilkey, and Mike Miles. (1986). "Toward an Assessment of the Reliability of Commercial Appraisals," *Appraisal Journal* 54 (July), 422–432.

Cole, Rebel, David Guilkey, and Mike Miles. (1987). "Pension Fund Investment Managers' Unit Values Deserve Confidence," *Real Estate Review* 17 (Spring), 84–89.

Isakson, Hans R. (1986). "The Nearest Neighbors Appraisal Technique: An Alternative to the Adjustment Grid Method," *Journal of the American Real Estate and Urban Economics Association* 14 (Summer), 274–286.

Ito, Takatoshi, and Keiko Nosse Hirono. (1993). "Efficiency of the Tokyo Housing Market," *Bank of Japan Monetary and Economic Studies* 11, 1–32.

Lovell, Douglas D. (1992). "Improving your Appraisals by Avoiding Forecast Errors," *Appraisal Journal* 60 (July), 424–426.

Marchitelli, Richard. (1992). "Rethinking the Cost Approach," *Appraisal Journal* 60 (July), 424-426,

Poterba, James. (1991). "House Price Dynamics: The Role of Tax Policy and Demography," *Brookings Papers on Economic Activity* (1991), 143–99.

Révész, P. (1984). "Density Estimation." In P. R. Krishniah and P. K. Sen (eds.), *Handbook of Statistics*. Amsterdam: North Holland.

Shlaes, Jared. (1992). ''Real-Time Appraising: Bringing Value up to Date,'' *Appraisal Journal* 60 (July), 419–424.

Silverman, B. W. (1986). *Density Estimation for Statistics and Data Analysis*. London: Chapman and Hall.

Tapia, R. A., and J. R. Thompson. (1978). *Nonparametric Density Estimation*. Baltimore: Johns Hopkins University Press.

Webb, R. Brian. (1994). ''On the Reliability of Commercial Appraisals: Analysis of Properties Sold from the Russell-NCREIF Index (1978–1992),'' *Real Estate Finance* 11 (Spring), 62–65.