Estimating Current Auto Collateral Values

Visible Equity Analytics

May 13, 2016

Abstract

In this document we discuss the methodology that Visible Equity employs in calculating current values for automobiles, often referred to as collateral values or auto values. The primary goal of the model is to estimate the current value for any vehicle in a sample of loans. Estimated values are obtained via index methods, a standard tool used in collateral valuation. The following paper will provide a formal description of the derivation of the Visible Equity index method, implementations of the index for collateral valuation, and an example to facilitate interpretation of results.

Section 1 Characterization of the Data Set

Visible Equity, where possible, provides an opinion on the current market value of a candidate vehicle utilizing a proprietary index method. The implementation of this index approach primarily depends on the availability of certain key data elements. In all cases, the underlying information used to create the applicable index comes directly from the Visible Equity database. Visible Equity houses approximately 6 million unique automobile appraisals taken at the time of loan origination. Pulling a sample of these records for each make-model combination, indexed by vehicle age, allows Visible Equity to estimate specific make-model depreciation rates. These estimates are then used to value any vehicle given its year, make, and model.

Section 2 Index Estimation

Section 2.1 Hedonic vs Index Methods.

In general, there are two types of automated valuation models/methods (AVMs) commonly used to estimate auto values; hedonic methods and index methods. Some approaches use a hybrid of these methods.

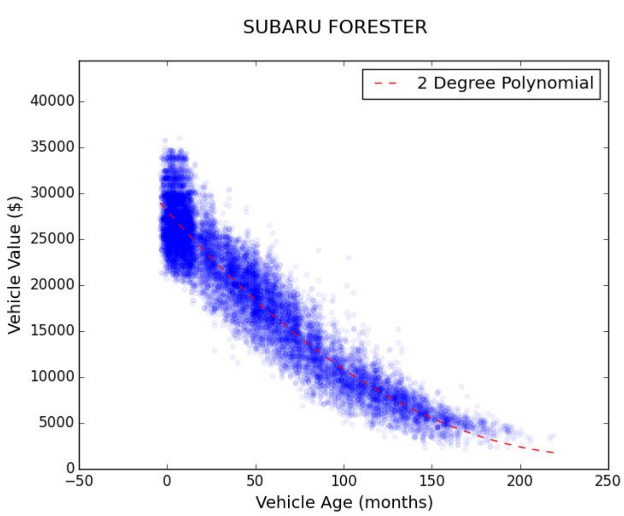
A hedonic method uses regression or similar statistical analyses to estimate the influence that each key feature of an automobile (e.g. make, model, mileage, age, etc.) has on the overall auto value. Hence, a subject automobile's key features are applied to the model and a current value is calculated.

An index method, on the other hand, utilizes changes in auto values as the basis for estimating auto prices. For any given make-model combination, all appraisal values are analyzed at each unit of vehicle age. Combining multiple vehicles at multiple vehicle age timestamps allows for the estimation of an "index", or a trend line that models depreciation over time. The typical index change between a subject vehicle's known value (from an appraisal or a sales transaction) and the date of estimation is then applied to the vehicle's known value. Where possible, Visible Equity utilizes this index method to value automobile collateral.

Section 2.2 Curve Fitting and Least Squares Regression Analysis.

In this section, we discuss the theory behind the creation of an index and fitting a least squares trend line to approximate typical depreciation for any make-model combination. As described above, data are aggregated for each automobile's make and model and a scatter plot is created to visualize the trend. A "best fit line" is then derived to represent the curvature of the data within the scatter plot. This curve fitting is the process of constructing a mathematical function that best fits the series of data points. Most commonly, a function of the form $y = f(x)$ is fitted to the data.

For example, starting with a first degree polynomial function would yield $y = ax + b$, a line with slope a that represents a constant depreciation through time. Such a function will fit a simple curve to two constraints. By nature of quadratic depreciation among auto classes, the majority of cases are derived to a second-degree polynomial of form $y = ax^2 + bx + c$. This will exactly fit a



simple curve to three constraints. An exact fit to all constraints however is not certain and therefore demands some method to evaluate each approximation within the data set.

A least squares polynomial function is then fit to minimize error. The method of least squares is an approach typical to regression analysis and attempts to approximate the solution of overdetermined systems, or systems in which there are more equations than unknowns. The best fit in the least-squares sense minimizes the sum of squared residuals, a residual being the difference between an observed value and the fitted value provided by a model. Thus, the objective consists of adjusting the parameters of a model function to best fit a data set.

Each auto class dataset consists of n points, or data pairs, (x_i, y_i), $i = 1, \ldots, n$ where x_i is the independent variable, vehicle age, and y_i is the dependent variable, vehicle value. The model function then has the form $f(x, \beta)$ where m adjustable parameters are held in the vector β. The objective is to find the parameter values for the model which "best" fits the data. The least squares method finds its ideal when the sum of squared residuals, denoted

$$S = \sum_{i=1}^{n} r_i^2$$

is a minimum. Note, a residual is defined as the difference between the actual vehicle value and the value predicted by the model: $r_i = y_i - f(x_i, \beta)$. The minimum of the sum of squares is found by setting the gradient, or slope, equal to zero. Since the model contains m parameters, there are m gradient equations. Thus:

$$\frac{\partial S}{\partial \beta_j} = 2 \sum_i r_i \frac{\partial r_i}{\partial \beta_j} = 0 \text{ for } j = 1, \ldots, m .$$

Since $r_i = y_i - f(x_i, \beta)$, the gradient equations become:

$$-2 \sum_i r_i \frac{\partial f(x_i, \beta)}{\partial \beta_j} = 0 \text{ for } j = 1, \ldots, m .$$

Thus each particular auto class index requires unique expressions for the model and its partial derivatives in order to minimize the total squared error and find a true "best fit" trend line. By default, we attempt to fit a second order polynomial function to the data. However, if an auto class index does not meet certain requirements[1], a simple first order polynomial function is fit to the data. As previously stated, for every vehicle's make and model, sample data is collected and a best fit line is derived. If data are sparse[2], Visible Equity fits a trend line only to the make of a vehicle.

Section 3 Implementing the Index Method

With trend lines created for each auto class, Visible Equity then implements the index in one of two ways: a true index method, and a true average method. The selection of these methods is largely a function of the availability of certain key data elements.

The primary implementation is a true index approach and is employed when data can be collected on a vehicle's make, model, year and original value (i.e., the value at loan origination). First, we determine the appropriate index to use. This is accomplished by matching a subject vehicle to its specific make and model index. Second, we note the original value of the vehicle, the vehicle age at origination, and the vehicle age at the date of estimation. We then calculate the percent change in index value from these two points in time, and apply the change to the original value. This yields an estimate of the value as of the date of estimation.

---

[1] We first attempt to fit a 2nd order function to vehicle values of the same make and model. If the $\square_1$ coefficient is negative (indicating concavity), we attempt to fit a 1st order polynomial. This, along with other controls, eliminate the possibility of vehicle appreciation.

[2] If the number unique values for a given make/model is small (less than an arbitrary cutoff), then we estimate an index with vehicle values of the same make using the process described above.

If the client is unable to provide a known value, Visible Equity implements the index as a true average. In other words, the current value is estimated purely as a function of vehicle age[3]. Admittedly, this approach is less precise, but it is effective at providing a naïve, baseline value in the absence of an original appraisal.

Section 4 Auto Valuation Example

An example will serve to make this methodology more concrete. Consider the index for Honda Civics as seen in Figure 2. First, a sample of known values for Honda Civics is collected from Visible Equity's database. Second, a least squares, second-degree polynomial function is fit to the data. Now consider two new Honda Civics: Sample Car 1, which has an original value of $27,000 and Sample Car 2, which has an original value of $21,000. Suppose that we want to estimate the value when each vehicle is six years old (72 months). We would simply calculate the change[4] in the index value from 0 to 72 months, and apply this change to the value of Sample Car 1 and Sample Car 2. Using this method, the adjusted appraisal after 72 months for Sample Car 1 would be $14,079. Similarly, the adjusted appraisal for Sample Car 2 would be $10,950.

As a final example, consider a third sample car with an unknown original value. After determining the correct index, the age of the vehicle would simply be plotted on the red best-fit line and followed as a true average for the desired time period, arriving at an average estimated current value. At 72 months, this vehicle would have an adjusted appraisal of $11,519.

[3] Consider a fictitious vehicle with an index modeled by the following function: $.5x^2-100x+30000$. In this case, simply plugging vehicle age into x, would yield an estimated value.

[4] Change is calculated as a percent change in index value from time 1 to time 2

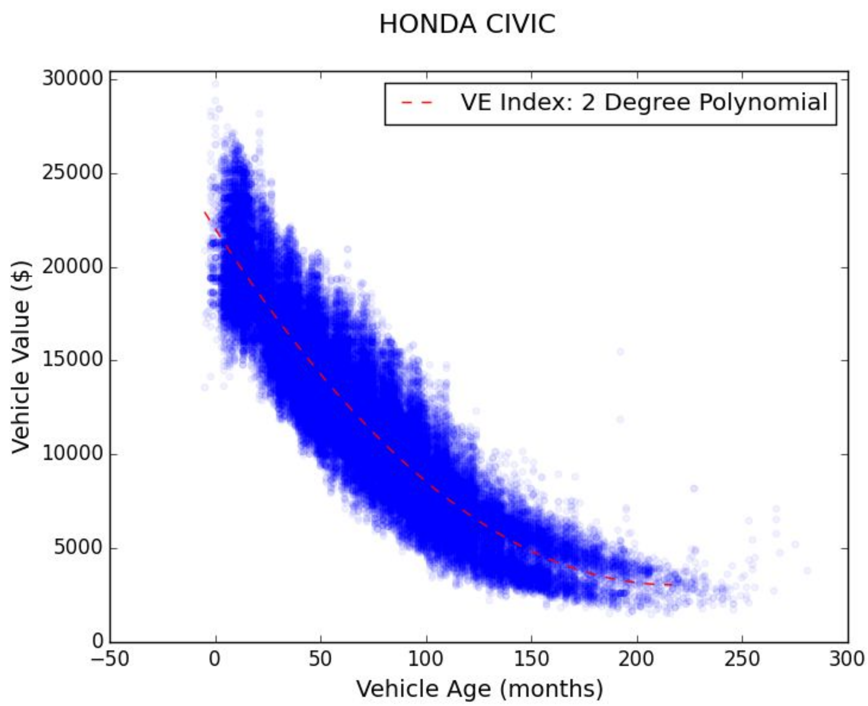


Figure 2

Section 5 Conclusion

In this paper we have described the derivation and implementation of Visible Equity's auto valuation index. This index method proves to be an effective approach to estimate updated values for large pools of automobiles. Financial institutions that engage in this type of lending will find these methods of particular use as they also allow for continuous estimates of loan-to-value ratios.

The precision of statistical models, by their nature, are clearly dependent on the sample employed. For this reason, Visible Equity will continue to evaluate the performance of the model as data availability increases and as economic conditions change. While indices for vehicle models with robust histories are not expected to change significantly between model iterations, those of newer vehicle models might. Thus, Visible Equity expects model performance to improve and calibrate to new conditions over time.

Appendix I Out-of-Sample Prediction

The methods described above offer portfolio managers an effective tool to update values of collateralized auto loans. By definition, empirical models are based on data, and are therefore subject to the constraints of the data employed. In the case of the Visible Equity auto valuation index, one important constraint to consider is the depth of historical data available for each vehicle model. Because these indices model collateral value as a function of vehicle age, they can only offer an estimated value for a subject vehicle if its age falls within the vehicle age range of the given data set.

While very unlikely, it is possible that a subject vehicle's age is greater than the range of vehicle ages from the training data set. To solve this issue, a separate out-of-sample valuation index is employed. The relationship between the Visible Equity and out-of-sample indices can be described as:

$$v(t : t \leq t_{max}) = at^2 + bt + c$$

$$v(t : t > t_{max}) = g(t) = v(t_{max})\delta^{(t - t_{max})}$$

Where *v(t)* is the subject-specific Visible Equity index, *g(t)* is the out-of-sample index, and t_{max} is the maximum observed time, *t,* given the sample of auto values. The function *g(t)* is a value depreciation function with decay parameter δ, which is modelled by:

$$\delta(t_{max}) = .0002t_{max} + .905$$

In other words, δ is scaled by the size of t_{max} and the behavior of the decay is a function of the depth of the vehicle history in the data sample. The parameter δ is scaled to mimic the end-behavior of the Visible Equity valuation index. Examples can be found in figure 3 and figure 4.

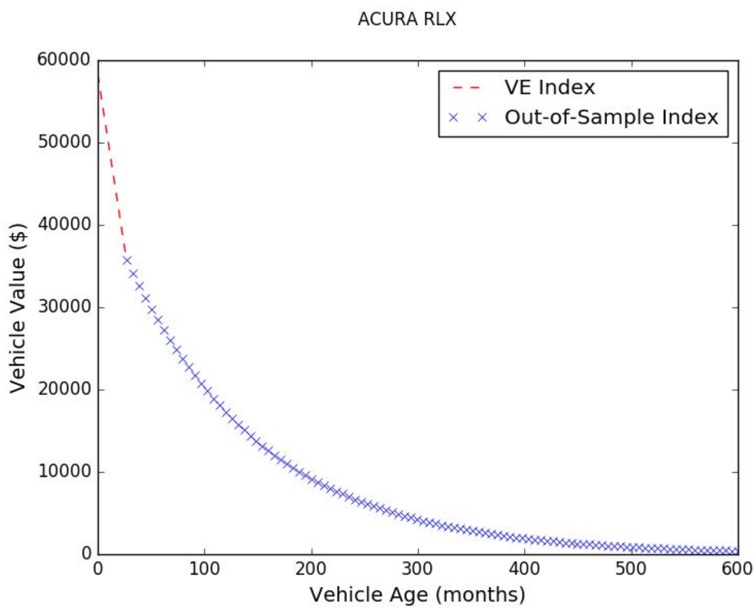


Figure 3

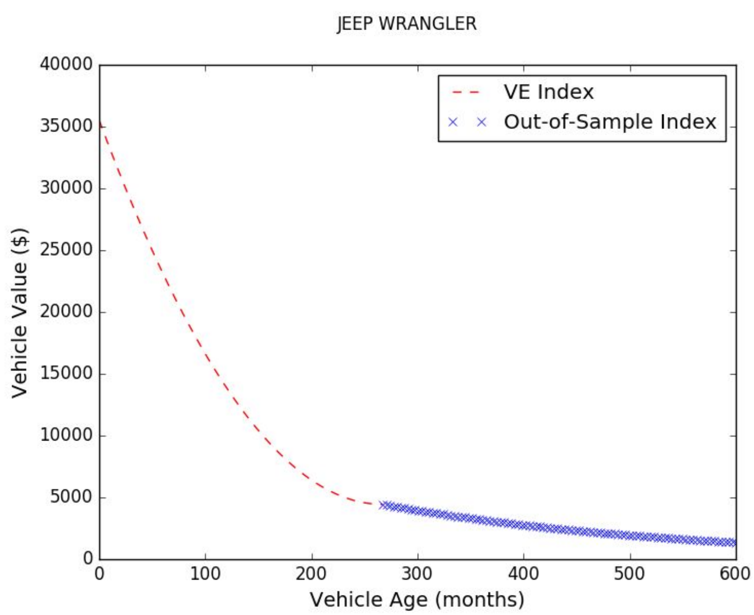


Figure 4