Portfolio Media. Inc. | 860 Broadway, 6th Floor | New York, NY 10003 | www.law360.com
Phone: +1 646 783 7100 | Fax: +1 646 783 7161 | customerservice@law360.com

Use The Right Model For Auto Defect Suits

Law360, New York (July 09, 2013, 12:56 PM ET) -- Automobile companies continue to be the focus of consumer class actions alleging that they sold defective vehicles. Recent lawsuits have claimed significant flaws in braking systems, on-board navigation systems, power windows and convertible tops. One of the damages theories plaintiffs commonly propose in these matters is that the accused vehicles have diminished in value.

To evaluate these claims, attorneys often engage economic experts who rely on regression analysis to assess whether the depreciation rate of the accused vehicles is higher than that of a benchmark set of vehicles. But how can litigation teams be sure that the results of those analyses are accurate? After all, there are multiple regression approaches experts can take, using different sets of data, each of which may yield different results.

In fact, a common mistake experts make when assessing claims of diminished value in auto defect lawsuits is to perform something called an ordinary least squares (OLS) regression using average vehicle prices at the trim level from Kelley Blue Book or a similar source. (The trim level refers to the potential bundle of options included in a vehicle — for instance, engine size, number of doors, manual or automatic transmission and different combinations of accessories.) So rather than using a single-observed average price per period for the regression model, the expert relies on multiple observations of average price, one for each trim version.

If the goal is to obtain statistically significant results, this may seem like an attractive approach because it increases the number of observations the expert can analyze, thereby increasing the power of the regression. But from a statistical and damages perspective, there is a problem with this technique: It assumes that the individual observations in the expert's analysis are independent, which is not likely to be the case if there are multiple trim observations for the same car model because trim levels for the same model tend to depreciate at similar rates.

In short, an OLS regression using trim-level data could indicate that a vehicle is experiencing excess depreciation when, in reality, it is not, thereby supporting an incorrect conclusion that class members may have been damaged.

Attorneys may not be performing these regressions directly, but they should be aware that the expert they engage has a number of statistical techniques at their disposal — techniques that will allow him to analyze trim-level data with the benefit of increased sample sizes but without the risk of drawing false conclusions.

How Regressions Work

Regression analysis is a statistical technique for estimating relationships between variables. For example, an expert can assess how the price of gasoline in the United States, the dependent variable, is affected by factors such as the price of crude oil and demand for gasoline in China, which are the independent variables.

For each independent variable, the regression will produce two numbers: the coefficient and the t-statistic. The coefficient is an estimate of how large an effect the independent variable has on the dependent variable.

The t-statistic provides information about how certain the expert should be that the estimated coefficient indicates a true relationship and not just one that occurs by chance. The higher the t-statistic, the greater certainty the expert has about the relationship between the variables being assessed.

In the context of auto defect class actions and plaintiffs' claims of diminished value, the dependent variable would be the value of a vehicle at a specific point in time. The independent variables would include an identifier for the vehicle at issue (for example, a number that is equal to one for every observation for the vehicle at issue and zero for all other models), the age of the vehicle and an interaction term that is the product of these variables.

The ability of a regression to identify a relationship between the dependent and independent variables depends in part on the amount of data used. Generally, the more data analyzed, the more accurate the regression should be.

Suppose an expert is analyzing the depreciation rate of the 2010 Volkswagen Jetta and wants to use the Toyota Corolla, Hyundai Elantra and Subaru Impreza as benchmarks. If the expert has quarterly data from 2005 to the end of 2012, he would have 32 observations per model or 128 observations total.

While this may be sufficient, the number of observations could be greatly increased by analyzing trim-level data instead of model-level data. Assuming there were five trim levels per vehicle model, this would increase the number of observations to 640. If he were to estimate the OLS regression using model-level data, the coefficient on the interaction term would be -0.17 percent, and the t-statistic would be -0.4, which indicates the relationship is not significant.

But if he estimated the regression using trim-level data, the coefficient changes to -1.04 percent, and the t-statistic indicates that the relationship is significant. And because the coefficient is negative, the expert would conclude that the Jetta depreciates more quickly than the benchmark, which the plaintiffs would argue supports a claim for diminished value.

This conclusion, however, would be wrong.

An Alternate Approach

Regression analysis is based on a number of assumptions about the data being analyzed — one of which is that each observation is independent of all other observations. When this assumption is violated, the results estimated by the regression are not correct.

Specifically, the regression attempts to extract information about the relationship between the dependent and independent variables from each observation. If the observations are related, then they contain some of the same information, which should not be counted twice.

The regression, however, assumes that each observation contains different information and places equal weight on each. Therefore, the regression assumes the data contain more information than they actually do, which is why the results of the regression are inaccurately precise.

There are alternate techniques that do not require independent observations. Hierarchical linear models (HLM), for instance, are commonly used in research applications in which the data are nested — that is, when they must be accounted for as subsets of a larger universe of data.

For example, a researcher who is analyzing test scores for a sample of students would have to account for the potential effects of the students' classroom on those scores, as well as the effects of the students' school. Auto depreciation data are similarly nested; the trim-level observations are all for the same vehicle model.

Going back to the earlier example, when the expert uses HLM to analyze the trim-level Jetta data, the coefficient on the interaction term is -0.64 percent, but now the t-statistic is only -1.76, which is not significant. So even though the coefficient is negative, there is not a statistically significant relationship between the dependent and independent variables.

This analytical approach provides a more accurate reading on the depreciation rate associated with the vehicle at issue compared with the benchmark vehicles, thereby allowing the expert to reach the correct conclusion regarding whether the vehicle at issue suffers from a higher level of depreciation.

When reviewing your own expert's work or preparing to depose an opposing expert's diminished value analysis, there are two questions you should keep in mind.

First, what type of regression model did the expert use: OLS, HLM or something else? You can learn this by reviewing the computer code provided by the opposing expert or, more likely, by consulting your own expert.

Second, did the expert use model-level or trim-level data, and is this consistent with the type of regression model used?

When conducting any regression, experts on both sides must be careful that the regression is properly defined and does not violate the statistical assumptions on which it is based. Otherwise, the results of the analysis will be inaccurate, prompting experts to possibly conclude that the accused vehicle experienced accelerated depreciation when it did not.

--By Mark Gustafson and Bruce Strombom, Ph.D., Analysis Group Inc.

Mark Gustafson is a vice president in the Los Angeles office of Analysis Group Inc. He specializes in applied econometrics and has assessed data in several large class actions involving U.S. auto manufacturers. Bruce Strombom is a managing principal in the Los Angeles office of Analysis Group.