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

Statistical Analysis And Class Actions: Part 2

Law360, New York (May 28, 2015, 1:13 PM ET) --

Nathan D. Woods

This is the second of three pieces discussing statistical analysis of class certification topics in wage-and-hour class and collective actions. In part one, I reviewed problems associated with relying on overall averages in addressing class certification. In part two, below, I discuss rigorous analysis and statistical testing of whether common patterns are present in a proposed class. Finally, in part three, I will discuss the use of sampling in analysis of class certification.

Statistical Testing of Commonality in Wage-and-Hour Class and Collective Actions

Too often analyses advanced in support of certification in wage-and-hour class and collective actions fail to critically assess commonality issues in a statistically rigorous manner. Indeed, as discussed in an earlier piece, in many cases I see an overreliance on summary measures such as overall averages to determine whether class treatment is appropriate. On their own these summary measures are not instructive as to whether class treatment is suitable since, by construction, they do not allow for assessment of potential underlying variability.

For data analyses to be informative at the class certification stage analysts in wage-and-hour cases should take into account and explicitly test for the existence of statistically important variation between members of the same proposed class. The demonstration of important variation between different parts of the same proposed class would suggest class treatment is inappropriate.

Rigorous Analysis in Wage-and-Hour Class Actions

There are a number of possible methodologies which facilitate more rigorous analysis and explicit statistical testing of whether common issues predominate. Among them, multivariate approaches such as regression analysis are perhaps the most commonly used. Regression is a method used regularly in class action litigation to estimate the effect of a given explanatory factor on an outcome measure, while simultaneously "controlling" the effect of other factors.[1] In a wage-and-hour context, for example, regression could be used to estimate the extent to which working in a particular location has an important effect on how many minutes are recorded before an employee's scheduled start time, while controlling for the effect of other factors such as employee job.

The effect of the factor may be assessed through the regression coefficient, which measures the estimated size of the effect and the associated standard deviation value, which enables a "test" for whether the estimated effect is "statistically significant." Statistically significant simply means that the estimated effect is different from zero. In general, the larger the absolute value of an estimated effect, the smaller the associated variance, and the larger the sample size, the larger will be the number of associated standard deviations and the more likely the effect will be statistically distinguishable from zero. In labor and employment litigation, an explanatory factor is viewed as having a statistically significant (i.e., nonzero) influence on the outcome measure if the number of standard deviations associated with the estimated effect is at least two.

Just as an estimated effect for a factor is "tested" for whether the effect is statistically discernable from zero, the estimated effects of different factors may be tested against each other to ascertain whether different factors have statistically similar or different impacts relative to each other. In this way, the hypothesis of a "common" effect measured across different factors (e.g., different locations), can be tested.

Rigorous Analysis Applied to an Example Off-The-Clock Allegation

A typical class allegation is one involving a claim that unpaid work is occurring prior to the beginning of employees' shifts. Consider the following example for ACME Health, a company operating sprawling health care-related facilities featuring multiple entrances, numerous floors and time clocks located throughout.[2] Employees are paid based on a scheduled start time but use a time clock which functions essentially as an attendance clock. Plaintiffs allege the time clock stamp indicates the beginning of actual work time and claim 10 minutes of pre-shift time is worked off-the-clock. A preliminary analysis of the time clock punches relative to the scheduled start times reveals an overall average of 7.91 minutes elapse between clock in and the beginning of paid time for members of the putative class. The associated standard deviation is 7.77 minutes.

This relatively large standard deviation is an indication of large variance between employees underlying the overall average. By itself this measure provides evidence that undermines the likelihood that employees work pre-shift on a consistent classwide basis. Other factors may explain this wide variation and show whether the employees in the proposed class are or are not similarly situated. More rigorous analysis can shed light on this question.

There may be many different possible factors, but in this example the location of the time clocks used by the members of the proposed class could be important. In the ACME Health facilities, each time clock can be mapped relative to an employee's workstation and the distance between where an employee clocks in and where the employee actually works can be determined.

For simplicity's sake, in this example consider three distances: one ("TC1") is closest to an employee's work station, another ("TC2") is further away near a break room and a third ("TC3") is furthest away, located at the lobby entrance to a building.[3] The combined data across all three time clocks results in the average difference of 7.91 minutes between the clock-in time and scheduled start of shift. However, examining the data separately for each of the three time clocks reveals that the employees clocking in at the time clock closest to their work station, TC1, show an average difference of 1.5 minutes with a standard deviation of 0.41, the time clock near the break room, TC2, gives an average of 12 minutes with a standard deviation of 4.16, and the time clock near the building entrance, TC3, gives an average of 17.5 minutes with a standard deviation of 8.24.

At 1.5, 12 and 17.5 minutes, respectively, the averages for each of the three time clocks appear to be considerably different from one another. This pattern suggests that differences between clock-in time and the scheduled start of shift are likely related to which time clock an employee uses and how close that clock is to her workstation.

More concrete evidence countering the possibility of a common class would be the existence of systematic differences between subparts of the same proposed class. As described above, regression methodology can be used to establish whether these apparent differences are statistically meaningful (which depends in part on how much variance underlies the average for each time clock and the size of the samples analyzed). If the measures are statistically significantly different from one another then the data exhibit systematic differences within the proposed class.

Here the model would use pre-shift minutes as an outcome measure with separate explanatory measures for the different time clocks and perhaps other control measures as well, like time frame, department or job title.[4] This model could measure what effect the use of different time clocks has on the number of minutes indicated pre-shift. Based on the estimates produced by the model, the hypothesis of a common effect (or lack thereof) between the three time clocks can be tested. In this example, this test would show statistically significant differences between time clocks. In my view, establishing these systematic differences within the proposed class means a common policy or practice could not be operating and common class treatment would be inappropriate.

Implications for Rigorous Analysis

In the above example, the wide variation between employee-shifts indicates differences across the proposed class that appear to be inconsistent with the alleged classwide pattern. Furthermore, statistical testing between the time clocks indicates that a factor other than a presumed classwide policy or practice may explain much of the variation in pre-shift minutes.

While this example for ACME Health is meant only as an illustration, it demonstrates some of the questions that should be analyzed when studying the question of class certification in wage-and-hour matters. What the data in these cases have in common is that the patterns uncovered through more rigorous analysis present a sometimes starkly different perspective than could be gained through reliance on aggregate summary measures alone. Given recent court decisions, this more rigorous analysis is now more than just sound analytical practice — it is likely the standard for analysis in wage-and-hour class and collective actions.

In the next piece I will discuss sampling strategies if use of a sample is contemplated in assessment of class certification and liability in wage and hour class and collective actions.

—By Nathan D. Woods, Edgeworth Economics LLC

Nathan Woods is a partner in Edgeworth Economics' Washington, D.C., office.

[1] This method is often referred to as “multiple” regression when more than one explanatory factor is

included in the estimate. For my purposes here in testing for possible differences within a proposed class, this terminology can be extended to include other multivariate approaches, including logit or probit. These methods also allow for controlling the effects of several explanatory factors but are intended for estimates of binary outcome variables, such as those measuring whether an outcome occur or not (e.g., was overtime paid in a work week?; yes = 1 and no = 0).

[2] Examples for ACME Hotel/Resort or ACME Manufacturing might be similar in that the buildings are often large, with multiple entrances and a variety of time clocks from which to choose for employees to clock in and out.

[3] Depending on the size of the facility, the number of entry points and whether the facility operates as part of a larger cluster of buildings, we might see time clocks used adjacent to parking, in separate facilities, or on different floors of a multifloor building.

[4] Depending on the case, many other factors may also be included as controls or studied explicitly for important variation themselves.
