

## Regression to Compare Process Costs

Mississippi Department of Human Service's Child Support Enforcement wanted to know the cost of individual processes as part of considering outsourcing. Two methods of outsourcing are common: outsource whole offices, or outsource individual process. Both are frequent: in Mississippi previous outsourcing had been done at the process level: collections, call center and legal services. In order to consider further functional outsourcing, MDHS leadership needed to know what CSE processes really cost—in total, including those things workers forget to put on a process map.

Cost allocation, process mapping and Monte Carlo simulation are other examples of methods of calculating costs from the bottom up, using field workers' estimation of process time. However, these estimations are always just a guess. Regression overcomes the inherent shortcomings of all bottoms-up methods of establish process cost:

- Individuals estimates are only approximate
- Workers often comingle work on several processes making it impossible divide total labor by output
- Bottoms-up methods always leave out part of the time spent during a day: breaks, down time, conversations with co-workers. These are all an important part of working, and crucial to include in the assessment of process cost

Each CSE case goes through two phases: set-up and maintenance. During the set-up phase, the caseworker collects information from the mother and father, researches outside records and prepares court documents and supports the court. This all happens over the months during which a case is initially prepared for the court hearing to establish an order. After a court order is obtained, a case takes on a whole different life: requiring little effort from the caseworker on a day-to-day basis. Thus, the "caseload" is nearly irrelevant for the purpose of understanding workload. At the same time, considering only new case load would ignore the cost of maintaining the case.

Regression provides a method for calculating the impact both new set-up and maintenance have on the size of the workforce required in an office. Regression links the amount of work product resulting from several processes (independent variables) to the number of workers required to complete the work done in an office (dependent variable). Thus, the volume of each type of output for each office is regressed against the number of workers in an office (or alternatively, the dollar amount of payroll).

In Mississippi, CSE is organized into 83 county offices providing 83 data points for the regression. The regression used a work year, expressed in hours as the dependent variable. Thus, the number of workers was multiplied by 2,000 hours per year. Cases created was one independent variable, tacked by case worker and summed by office. The other independent variable was the sum of all types of case modifications and other activity recorded in the case system. It would be possible to have more independent variables, but not necessary for MDHS purposes.

Regression was useful to MDHS in two ways:

1. Adjusted R-Squared tell leadership how much of the variation in office staffing is explained by the volume of the work they complete. The R-Squared was .94, indicating that nearly all of the

variation in the number of workers across offices was explained by the number of cases they set up and managed.

2. The regression coefficients tell leadership the cost of each process (in hours or dollars). The regression told MDHS leadership that new cases account for 12 hours of a case worker's time. This is very different from the bottom up time. A small part of this difference is the time some case workers spend actually in court. Most of the difference shows the inadequacy of bottoms-up costing as a tool for explaining all of a worker's time. In addition, workers spend 30 minutes per year maintaining an average case.

Regression as a tool for process costing requires some care. The dependent variable must have variance, so regression cannot be performed at the level of an individual worker – since case workers are salaried and their hours do not vary enough worker to worker. Regression results should always be developed on one set of data, then tested on another. This can be done by using two sets of offices, or two year's data. This is an important way to find out whether the results are truly indicative of workload, or the agency is merely assigning a set of workers, and not adjusting staffing for changes in workload.

The regression findings told MDHS leadership a great deal about improving managing CSE:

- Managing by the number of open cases is misleading—management should be based on both new cases and case load
- The time to set up a case is much greater than workers think. This is part of the operation that requires more management attention
- While the regression produced strong results for a single year, back-casting the results to prior years broke down. This tells management that it is not adjusting office staffing to meet changes in workload
- Using the regression coefficients to recalculate office staffing, MDHS leadership has a method for managing in an environment of changing workload—one that better reflects the work done than existing cases.
- The approach provided dollar values against which MDHS could evaluate outsourcing proposals for the volume of new and maintenance cases in the offices to be considered as part of a pilot