ECEDHA Survey Takes Giant Leap Forward

This year ECEDHA is proud to announce the unveiling of its new interactive data system that will allow participating ECEDHA members to perform custom queries of the ECEDHA survey data. The annual survey has been a part of the ECEDHA program for more than a decade, collecting important comparative benchmark data on enrollments, salaries, graduate stipends, space allocations, graduation statistics, research expenditures, and much more. The new interactive data system is arguably a revolutionary upgrade to the ECEDHA survey that will allow members access to information like never before.

Evolution of the ECEDHA survey
In the years prior to 2007, survey results with school names omitted were distributed in spreadsheet form to the ECEDHA schools that participated in the survey. The practice of anonymous data sharing dates back to the origin of the survey and was enacted to encourage widespread participation and uninflated reporting.

One of the recognized shortcomings of presenting information in spreadsheet form with redacted school names is that reliable peer benchmarking can be difficult, since peer schools are not identified. In 2007, ECEDHA took a major step to address this problem by issuing an annual survey report. The report provided peer cohort comparisons in graphical form. As a rule of thumb, cohort sizes were chosen to be less than 20. Schools were grouped by Carnegie designation (as in previous ECEDHA surveys) and then subdivided, where applicable, according to US News and World Report ranking or, where not applicable, by program size. The grouping procedure resulted in 10 cohorts. While this formula for determining peer sets is far from ideal, it did provide chairs with documented comparisons that were useful. Feedback from ECEDHA members was overwhelmingly positive. But clearly opportunities for improvement remained.

A number of issues were identified after 2007 pertaining to data entry. The 2007 software was written under the assumption that chairs would, for the most part, enter data in one sitting. In practice, most chairs visit the entry portal multiple times, entering bits and pieces during each visit. With the 2007 software, chairs could occasionally lose their submitted work if they didn't save their entries using the save button. Fortunately, this problem (a relatively simple one to address) was fixed the year after it was reported.   A more fundamental and ongoing challenge, however, is data accuracy. With the relatively large number of data entry fields, typographical errors can and often do occur. Furthermore, we observed entry errors associated with misinterpretation of questions or misreading questions. These kinds of errors and inconsistencies showed up as anomalies when generating the survey report plots and had to be screened and corrected manually. This involved contacting the school chairs to obtain the correct information. While a best effort was made to correct these errors, it is likely some errors slipped by.

This year, ECEDHA has taken a giant leap forward, with upgraded software for data entry and the creation of an interactive data query system. As part of the data entry software upgrade, all entries are automatically saved and marked as such next to the entry field. Missing data fields are highlighted in pink to make it easy for chairs who complete the survey in multiple sessions to locate the questions remaining. To help reduce errors, the numerical entries from last year are

provided next to the entry box. In addition, there are preset bounds in place to help detect accidental entry errors. For example, if the survey question asks for an academic year salary and the chair enters a monthly salary by mistake, a courtesy message appears indicating that the entry is outside of the expected bounds and that the entry may be an error. This added safeguard has helped considerably.

Most exciting, however, is the new interactive data query system. In accordance with longstanding ECEDHA practice, we have taken great care to construct the interactive data system in a way that preserves confidentiality. After the data collection period ends and we open the system for interactive customized query, ECEDHA users will be able to specify their own set of peers and display individual peer responses visually as bar graphs. All names will be replaced by a letter designation (A, B, C ...) and the designated letters that are assigned will be randomized as a privacy safeguard every time a plot is generated. As an additional privacy measure, users will be required to specify a minimum of 12 peer institutions for data query. This helps protect school identities that might otherwise be apparent if very small sets of peers were allowable.

The system allows users to print out the custom plots they generate as well as download as a spreadsheet the underlying data being plotted. When viewing the bar graphs, users will have the ability to zoom in on the plots as well as the ability to see the exact entry value by holding the cursor over the bar of interest in the plot. In addition, the system allows users to normalize plots by total number of faculty. This feature allows you to plot benchmarks like research expenditures per faculty, space per faculty, and number of students per faculty. The system also provides and displays the peer set average associated with each question, a feature we expect many chairs to find useful.

We are still in the process of collecting survey data, but expect to close the entry portal in January. Shortly thereafter, the interactive data query system will be open to all participating chairs. We hope you will enjoy the new system and look forward to receiving your comments and suggestions for additional features we might add to enhance the system.