Session Data Preservation and Migration - Problems and Solutions

Jason Ostergren and Helena Stolyarova

The Health and Retirement Study (HRS) at the University of Michigan Institute for Social Research is a longitudinal study that originated in 1992, switched to conducting interviews using Blaise 4 in 2002, and Blaise 5 in 2018. Among the challenges HRS has faced since the switch to Blaise 5 is how to handle Blaise 5 Session Data when an interview is not completed in one sitting. Session Data is the working database that maintains the state of the instrument, including the values of temporary data and properties of fields. If that data is lost in a case where an interview was interrupted and has to be resumed later, significant problems can result if the instrument is dependent on session data to resume correctly, as is the case with HRS. There are a number of aspects to this problem and HRS has tackled new and different ones in each of the three (2018, 2020, and 2022) waves of interviews since adopting Blaise 5. Data migration has been at the center of some of these issues - for example, HRS has had to incorporate careful attention to harmless changes into processes for updating instruments in the field. Most problems have been solved, sometimes with significant help from CBS, but with varying degrees of completeness. For example, it was thought that data migration issues had been solved in 2020, but it turned out that mode switches rendered that solution incomplete. Finally, HRS has determined that it would be useful to retain and preserve session data after interview completion, which has been partially solved as well. This paper will break down why session data has been important to HRS as well as the various issues and solutions to the problems that have arisen, including testing tools, and what remains to be done.