Predicting Response Mode During Data Collection in the National Survey of College Graduates

Chandra Erdman*
U.S. Census Bureau, Washington, DC, USA – chandraerdman@gmail.com

Survey organizations routinely manage trade-offs among cost, timeliness, and coverage when selecting a mode of data collection. Cost and time are typical limitations for face-to-face surveys; coverage is a common concern in telephone and Internet surveys; and, decreasing response rates are common to all surveys. To overcome these limitations, many surveys have increased the number of mode choices offered to sample members. This is true for most surveys conducted by the U.S. Census Bureau, including the National Survey of College Graduates (NSCG). The NSCG starts most sample members in Internet mode and, if no response is received after a pre-specified period of time, a subset of non-respondents are offered mail mode. Finally, if no response is received after another pre-specified of time, another subset of non-respondents are placed in computer-assisted telephone interviewing. This fixed sequential mode design minimizes costs by starting sample members in the least expensive mode and only offering more expensive modes to subsets of non-respondents. However, such a fixed design may lessen the timeliness of data products. This paper outlines research on adaptive design for the NSCG. Our primary goal is to place respondents in their preferred mode as quickly as possible to shorten the data-collection period for the NSCG while maintaining response rates and representativeness. Using paradata, historical survey responses, and a rich set of frame data, we examine a set of competing risks models for estimating response propensities and time to response by mode, throughout data collection. We discuss model development and evaluation along with applications in the 2015 National Survey of College Graduates.

Keywords: adaptive design; responsive design; paradata; competing risks.