Authors’ Reply to the Commentary

Willard L. Rodgers and Mary Beth Ofstedal

Survey Research Center, University of Michigan, Ann Arbor.

We thank Freedman and Martin (2003) for their comments and for stimulating discussion on this important topic. Their comments speak to a fundamental tenet of survey methodology: the superiority of repeated cross-sectional data for studying aggregate population changes and the drawbacks associated with using panel data for such studies. In general, however, the research community tends to be more interested in studying individual-level transitions rather than population trends, and for this reason greater priority has been placed on conducting long-term panel studies rather than repeated cross-sectional studies. Panel surveys, particularly those using cohort replacement, can be used to study cohort trends, but researchers must be aware of the potential caveats and address them carefully in the analysis. We did this to the best of our understanding of the issues, but we realize that researchers will differ in their views and approaches. Indeed, we feel that such differences are critical to the advancement of science.

In response to Freedman and Martin’s (2003) appeal for additional sensitivity analysis, we replicated the analyses reported in Rodgers, Ofstedal, and Herzog (2003), first by using a cruder method of imputations for missing data on the cognitive measures (the same method used in their analysis), and second by using different cutoffs to identify those with severe impairment. The crude method of imputation involved assigning a score of zero to most of the cognitive items that the respondents refused to answer (except for the immediate recall and Serial 7s tests, for which we assigned scores of 2 and 1, respectively). The primary difference between the regression analyses based on those imputations and those shown in Table 2 of Rodgers and colleagues (2003) is that there was a statistically significant decline on the total cognitive scores after we controlled for age and prior interview status. For the second type of sensitivity analysis, we used total cognitive scores of 8 and 12 to identify those with severe impairment and compared the findings from logistic regressions based on those cutoffs with those shown in Table 3, which are based on a cutoff score of 10. Again, the differences were small and none of the logistic regression coefficients for Waves 2, 3, or 4 is statistically significant for Models 2–4, regardless of which cutoff score is used. These modifications do not change our conclusion that the data fail to provide evidence of improvement in cognitive functioning across the most recent cohorts of Americans aged 70 years and older.

We did one final analysis: To avoid any ambiguity that may be introduced by adjusting for prior experience with the cognitive tests in the regression models and because of differences in some of the measures and their administration introduced following Wave 1 (the word lists were changed, and the interviewer instructions on the Serial 7s task were modified), we looked only at those who were interviewed for the second time at Wave 4 and compared them with their counterparts.
at Wave 2. This also had the effect of restricting the analysis primarily to those in the rather narrow age range of 72–76. The unadjusted mean scores on both the immediate and delayed word recall tasks fell by about 0.2 words between Waves 2 and 4, and scores on the Mental Status and Serial 7s did not change significantly. On the other hand, the proportion with scores of 10 or less also declined, from 3% down to 2% ($p < .05$).

Controlling on design and respondent characteristics somewhat increased the estimated decline in mean scores and reduced (generally to nonsignificance) the estimated decline in the proportion with very low scores.

We share Freedman and Martin’s (2003) concern about the extent to which differences in measurement, methodological, and analytic approaches cloud the underlying truth and impede understanding of a particular phenomenon. In this regard, Freedman, Martin, and Schoeni (2002) have conducted a careful review of studies of disability trends in the United States and initiated an effort in conjunction with the investigators responsible for those studies to reconcile differences in findings. We applaud this effort and feel that a similar effort on cognitive trends would be valuable. We are somewhat less sanguine about their call for a coordination of analytic approaches and harmonization of measures across surveys and settings. This presumes that researchers know how best to measure and model a particular phenomenon, and we are skeptical that researchers have reached that point, particularly with regard to the measurement of cognitive impairment in a large-scale prospective survey. Furthermore, it is our view that the most compelling findings are those that hold up under varied survey designs and analytic approaches.

Again we thank Freedman and Martin (2003) for opening this dialogue, and we look forward to future exchanges with them and others on this topic.

ACKNOWLEDGMENTS

Address correspondence to Dr. Willard L. Rodgers, Institute for Social Research, University of Michigan, Ann Arbor, MI 48104-1248. E-mail: wrodgers@umich.edu

REFERENCES


Received April 30, 2003
Accepted April 30, 2003