


ARTICLE

Introduction to the Special Issue on New Longitudinal Data for Retirement Analysis and Policy

Marco Angrisani¹, Anya Samek^{2*} and Arie Kapteyn¹ 

¹University of Southern California, Dana and David Dornsife College of Letters Arts and Sciences, Center for Economic and Social Research, Los Angeles, California 90089-3332, USA and ²University of California, San Diego, Rady School of Management, La Jolla, CA 92093, USA

*Corresponding author. Email: anysamek@gmail.com

(Received 15 August 2020; revised 20 November 2020; accepted 13 January 2021)

The number of data sources available for academic research on retirement economics and policy has increased rapidly in the past two decades. Data quality and comparability across studies have also improved considerably, with survey questionnaires progressively converging towards common ways of eliciting the same measurable concepts. Probability-based Internet panels have become a more accepted and recognized tool to obtain research data, allowing for fast, flexible, and cost-effective data collection compared to more traditional modes such as in-person and phone interviews. In an era of big data, academic research has also increasingly been able to access administrative records (e.g., Kostøl and Mogstad, 2014; Cesarini *et al.*, 2016), private-sector financial records (e.g., Gelman *et al.*, 2014), and administrative data married with surveys (Ameriks *et al.*, 2020), to answer questions that could not be successfully tackled otherwise.

The availability of more detailed and better quality data has also opened up new economics research opportunities. The growth of surveys and initiatives to collect innovative data in the area of demography and economics of aging has been the most remarkable. This is due, in part, to the aging of the population worldwide and the need to better understand the impact of aging on wellbeing, public health, and social security programs. Issues central to the policy debate in developed economies are the determinants of individuals' financial preparedness for retirement, the factors driving labor force attachment at older ages, and the effects of public and private pension incentives in shaping observed retirement paths. In this Special Issue, we collect nine papers that use innovative data sources to explore these and related questions from a new perspective. These papers are notable examples of how recent advances in data collection can promote research on retirement and indicate ways to enhance data content and quality, thereby allowing the pursuit of promising research avenues in the future.

The organization of this Special Issue reflects the types of data sources used by the featured papers. Our issue begins with a study that builds on newly harmonized data on older adults across countries, to construct measures of retirement incentives. It continues with a series of papers investigating various retirement-related issues through the lenses of rich and comprehensive data collected via Internet panels. It ends with an investigation of how automatic and voluntary pension plan enrollments affect retirement wealth accumulation at job separation using administrative data. Below, we highlight the original contribution of featured papers and emphasize the crucial roles played by these new data in inspiring and informing their analyses.

The Health and Retirement Study (HRS) is the most prominent example of a survey focused on older adults in the USA. The HRS provides rich and longitudinal information to study transitions from work into retirement, changes in health status over time, and economic wellbeing at older

ages. Launched in 1992, the HRS has recently enriched its content with genetic data, biomarkers, various administrative linkages, and a harmonized cognitive assessment protocol, all of which enable researchers to comprehensively examine the challenges faced and posed by an aging population.

Following the HRS, similar surveys have been administered in over 30 countries on five continents. Designed to be harmonized with the HRS, these surveys provide a unique opportunity for cross-national studies based on comparable data. This work, therefore, allows academics and policymakers to draw lessons about the role of different programs and institutions in different economic and socio-cultural contexts. To this end, the Gateway to Global Aging Data project at the University of Southern California provides user-friendly harmonized datasets of 10 HRS-family surveys, greatly facilitating cross-country studies on the aging population. Exploiting this unique collection of comparable datasets, Knapp *et al.* (2020) (this volume) evaluate several alternative approaches to computing prospective pension benefits for current workers using common survey questions and validate them against matched administrative data. The authors document that survey-based measures of pension benefit growth capture the financial incentives of the US Social Security rules adequately. Since survey-based measures of prospective pension benefits rely on individual characteristics collected by all HRS-sister surveys, the results of this study indicate that reliable harmonized retirement incentives can be computed for different countries and made available to the research community. The Gateway has embarked on this ambitious project, which will promote future comparative research on public pensions and the specific retirement incentives they provide in different contexts.

As Internet interviewing is becoming an increasingly popular mode of data collection, several authors in this Special Issue discuss data from web surveys of probability-based, representative samples. For research purposes, probability-based Internet panels are greatly preferred to convenience samples, which do not have an underlying sampling frame, and, hence, make it difficult to draw proper inferences for a population larger and more interesting than just the volunteer survey respondents (Butz and Torrey, 2006).

Compared to other common modes like in-person and phone interviews, online surveys are self-administered, thereby reducing interviewer effects and social desirability biases when eliciting individuals' traits, preferences, and beliefs, and when confronting respondents with sensitive issues such as politics or religion. In addition, online surveys allow more user-friendly, interactive interfaces along with visual and audio aids that may help better guide and engage respondents throughout the interview process. Such features are likely to improve data quality and response rates relative to in-person, phone, and paper-based surveys. Online surveys are also amenable to a variety of experiments that would be more difficult (and costly) to implement with traditional interview modes. For instance, researchers can easily explore different question wording, randomize response order within a question to control for primacy and recency effects, and test for framing effects by presenting identical scenarios in different ways. Implementing such experiments in paper-based, phone, or in-person interviews would likely involve higher questionnaire costs and measurement error. Online surveys can also be successfully used to assess the effect of information treatments, where respondents are provided with different pieces of information before making hypothetical choices. In such instances, online surveys allow the respondent to read and understand relevant information at his or her own pace, eliminating the implicit pressure created by the presence of an interviewer in a phone or in-person interview. Relative to self-administered paper-based surveys, online surveys also often provide *para-data* (e.g., how much time the respondent spent on a page), which can be used to proxy for the 'degree of treatment exposure.' Moreover, this approach greatly reduces questionnaire costs and data entry mistakes.

Growing Internet access in the population has reduced coverage concerns surrounding online surveys, though it has not eliminated them completely since Internet access is not homogeneous across socioeconomic and demographic groups. For this reason, non-convenience samples typically used for research purposes such as the Knowledge Panel (GfK), the RAND American Life Panel (ALP), and the Understanding American Study (UAS) at the University of Southern California, provide households lacking Internet access with the tools to gain access. Selection bias may still arise, though it can often

be satisfactorily corrected in most cases with appropriate weighting (Schonlau *et al.*, 2009). Overall, online surveys represent the most promising and exciting means of collecting new data for empirical research in the social sciences.

In their paper, Barcellos and Zamarro (2020) (this volume) designed and administered a survey in the ALP to examine the use of formal and alternative financial services among minority groups. Minority groups are less likely to have a bank account and more likely to rely on payday loans, car title loans, and rent-to-own agreements than the general population, which results in lower financial security. This is a cause of national concern given the increasing presence of Hispanic and Black households in the USA and the rapid aging of the population. Leveraging the rich background information available for ALP panel members and exploiting primary data, the authors show how financial inclusion among minorities is shaped by financial literacy, trust in financial institutions, social networks, and time preferences. While the white-minority gap in bank account ownership is driven by differences in socioeconomic status and circumstances, nevertheless the higher use of alternative financial services among minorities remains largely unexplained.

Five papers in this Special Issue use the UAS to investigate retirement-related issues. Notably, all UAS respondents are invited to take the entire HRS questionnaire every 2 years. To reduce the response burden, the UAS-HRS questionnaire has been split into six separate surveys, with some adjustments to accommodate the self-administered web survey mode. Unlike the HRS, the UAS administers these surveys to every participant, not just those age 50+. Differences between the HRS and its online version in the UAS allow for an examination of the evolution of physical and cognitive functioning as well as employment, assets, and pension wealth over the complete life cycle. Moreover, one can also test the effect of online (UAS) and in-person/phone (HRS) interview modes on data quality and survey outcomes (see Angrisani *et al.*, 2019, for an example of such a study).

The growing popularity of online panels is partly explained by their flexibility in accommodating user-friendly and interactive interfaces. By making questionnaires more intuitive and easier to navigate, they have the potential to greatly increase respondents' understanding of complex questions and reduce response biases. Online studies are particularly amenable to survey experiments, allowing researchers to combine experiments' causal power with the generalizability of population-based samples. The UAS has been at the forefront of such data collection efforts, using both newly designed interfaces and experiments.

Another study using the UAS is by Perez-Arce *et al.* (2020b) (this volume), where the authors inform respondents of Social Security's pending shortfalls. Respondents are then confronted with hypothetical scenarios where alternative policy changes are offered to counteract the shortfalls. The scenarios include an increase in the payroll tax rate, a reduction in benefits, an increase in the wage ceiling, and an income-tax increase. Scenarios can include a single policy change in isolation or a combination of multiple changes, and they are randomly assigned to survey participants. In each case, the authors measure respondents' expectations about the benefits they will receive and potential behavioral responses that may follow in terms of labor supply, retirement path, and savings.

In related work by Perez-Arce and Rabinovich (2020) (this volume), the authors implement an experiment to test whether simpler and more concise information about the Retirement Earnings Test (RET) than currently offered can enhance individuals' understanding of RET rules. The authors also evaluate whether better knowledge of the short- and long-term tradeoffs involved in the RET affect retirement intentions. Similarly, in the study by Perez-Arce *et al.* (2020a), the authors use vignettes to provide UAS respondents with alternative ways to think about Social Security spousal benefits. The authors then elicit respondents' advice about when each vignette character should claim Social Security benefits, along with respondents' own claiming intentions. Taken together, these studies provide important insights regarding the potential effects of information campaigns about retirement-related concepts and Social Security rules. They also highlight heterogeneity in such effects across population segments, exploiting the rich background information available for UAS members.

In the paper by Angrisani and Casanova (2020) (this volume), differences in retirement preparedness are explored in terms of peoples' different levels of subjective and objective retirement financial

knowledge. For this purpose, the authors rely on the UAS Comprehensive File (CF), which merges core UAS surveys repeated every 2 years. The CF contains an extensive array of demographics, cognitive, financial literacy, and personality scores, along with self-assessments of retirement preparedness, and knowledge of Social Security rules; additionally it includes a complete household balance sheet for each UAS panel member. The authors find that overconfident individuals (those with relatively high self-rated but low objective financial knowledge) exhibit a level of retirement preparedness no different from others who are less confident, yet they are the least interested in learning more about retirement planning. Underconfident individuals (those with relatively low self-rated but high objective financial knowledge) have worse economic outcomes than their counterparts and express a clear interest in learning more about retirement-related financial issues. These findings suggest that it is the *combination* of actual and perceived financial knowledge that shapes behaviors. Hence educational programs should not only enhance financial literacy but also boost awareness of actual knowledge, to more effectively influence financial decision-making.

Financial, and more generally, cognitive skills are not the only determinants of retirement preparedness: non-cognitive skills and personality may play a crucial role too, though such skills are often overlooked in the policy debate. In her paper, Zamarro (2020) (this volume) uses a novel approach enabled by the combination of UAS survey data and paradata (i.e., statistics about the data collection process) to delve into this topic. She analyzes a series of UAS surveys, quantifies respondents' effort as measured by item non-response and careless answering, and shows that respondents' efforts on the surveys correlates strongly with self-reported personality traits. She also explores how constructed and self-reported measures of character skills relate to individuals' financial capability and retirement readiness.

Big data have been increasingly used for academic research purposes in part because large administrative and proprietary private sector datasets can offer more accurate measures to track and describe economic activity. Additionally, they can enable better research designs to assess the effects of different policy interventions, as illustrated in the paper by Hung *et al.* (2020) (this volume). The authors use an administrative dataset on retirement plan savings accounts to explore how plan enrollment design and default distribution rules impact leakage at job separation. This question could not have been tackled without information plan details, while heterogeneity across different groups of workers could not have been fully explored without access to a large sample.

In sum, this Special Issue offers readers a glimpse into many exciting new data sources rapidly becoming available for research on retirement preparedness. Harmonization of these different data sources across countries, the availability of data from probability-based Internet panels, and growing access to big administrative datasets from firms represent exciting developments for empirical researchers. For the first time, these data sources are enabling research spanning many populations (including cross-cultural research and research with under-represented minorities), approaches (including use of Internet-based questionnaires that simplify and explain concepts using online tools, and experiments that use randomization to assess the causal impact of interventions), and topical areas (including using paradata to study associations of non-cognitive skills with retirement preparedness). And most importantly, much of the data is now publicly available. For example, harmonized cross-national data for HRS-family surveys can be obtained at the Gateway to Global Aging Repository (<https://g2aging.org/>). Additionally, data from over 150 UAS surveys are available free of charge to any researcher, subject to a data user agreement (<https://uasdata.usc.edu/index.php>). Researchers across different institutions are also able to conduct their own surveys using the UAS.

References

- Ameriks J, Briggs J, Caplin A, Shapiro MD and Tonetti C (2020) Long-term-care utility and late-in-life saving. *Journal of Political Economy* 128(6), 2375–2451.
- Angrisani M and Casanova M (2020) What you think you know can hurt you: under/over confidence in financial knowledge and preparedness for retirement. *Journal of Pension Economics & Finance*.

- Angrisani M, Finley B and Kapteyn A** (2019) Can internet match high quality traditional surveys? Comparing the health and retirement study and its online version. In Tripathi G, Jacho-Chavez DT and Hyunh KP (eds), *The Econometrics of Complex Survey Data. Advances in Econometrics no. 39*. Bingley, UK: Emerald Publishing Limited, pp. 3–33.
- Barcellos SH and Zamarro G** (2020) Unbanked status and use of alternative financial services among minority populations. *Journal of Pension Economics & Finance*.
- Butz WP and Torrey BB** (2006) Some frontiers in social science. *Science (New York, N.Y.)* **312**(5782), 1898–1900.
- Cesarini D, Lindqvist E, Östling R and Wallace B** (2016) Wealth, health and child development: evidence from administrative data on Swedish lottery players. *Quarterly Journal of Economics* **131**, 687–738.
- Gelman M, Kariv S, Shapiro MD, Silverman D and Tadelis S** (2014) Harnessing naturally occurring data to measure the response of spending to income. *Science (New York, N.Y.)* **345**, 212–215.
- Hung A, Luoto J, Burke J, Utkus S and Young J** (2020) Automatic enrollment and job market turnover. *Journal of Pension Economics & Finance*.
- Knapp D, Lis M, Lee J and Phillips D** (2020) Evaluating alternative approaches for predicting pension benefits and incentives. *Journal of Pension Economics & Finance*.
- Kostol AR and Mogstad M** (2014) How financial incentives induce disability insurance recipients to return to work. *American Economic Review* **104**(2), 624–655.
- Perez-Arce F and Rabinovich L** (2020) Improving understanding of the retirement earnings test. *Journal of Pension Economics & Finance*.
- Perez-Arce F, Rabinovich L, Samek A and Yoong J** (2020a) The effect of informational prompts about spousal benefits on social security claim intentions. *Journal of Pension Economics & Finance*.
- Perez-Arce F, Rabinovich L and Yoong J** (2020b) Policy levers to reduce the social security shortfalls. Considering their likely impact on benefit and behavior expectations. *Journal of Pension Economics & Finance*.
- Schonlau M, van Soest A, Kapteyn A and Couper M** (2009) Selection bias in web surveys and the use of propensity scores. *Sociological Methods & Research* **37**(3), 291–318.
- Zamarro G** (2020) Alternative measures of Non-cognitive skills and their effect on retirement preparation and financial capability. *Journal of Pension Economics & Finance*.