

Review of the Health and Retirement Survey

Subtopic: Psychology and economics.

November 4, 2002

Daniel Kahneman and David Laibson

We are highly impressed by the HRS. The survey coverage is outstanding and the questions are almost all well designed. All in all, we give the HRS very high marks. The resources dedicated to the HRS have been well spent. Our analysis yielded very few fundamental critiques.

We take this opportunity to focus on a list of ten issues that may merit consideration as the HRS is continuously amended in the years ahead.

1. Measures of life-satisfaction, affect, subjective quality of life, and psychosocial experiences.

Topics in these areas receive too little attention on the current survey. Indeed, this is the only set of topics that seems to us to be considerably underdeveloped in the current survey. One would expect such underdevelopment, since interest in and study of life-satisfaction and psychosocial experiences has expanded very rapidly since the HRS was first conceptualized in the 1980's. We propose that questions in these areas should be redesigned and significantly expanded. We believe that this expansion should be undertaken on a zero-based budgeting basis. Specifically, questions that are already in the survey should not necessarily be expected to appear in the redesigned survey. Instead, we suggest that instruments in these areas should be developed from scratch and only *then* compared to the pre-existing questions. Such a procedure will maximize the chances of useful innovation. We acknowledge that at the end of the day it is useful to adopt or adapt relevant pre-existing questions in these areas. But the survey redesign in these areas should not be locked into the pre-existing instruments and the redesign should not even take the preexisting instruments as a starting point.

2. Redundancy.

The survey may be getting too long, particularly in light of the need to incorporate new questions (see above). To address this problem we recommend reducing the number of questions by removing wholly or partially redundant questions. We recognize that this will slightly decrease the quality of the data --- since repetitive questions do force subjects to think harder about their answers --- but we believe that this reduction in data quality will be offset by the opportunity to incorporate new questions in areas that are insufficiently covered at the moment. We are particularly skeptical of partially repetitive questions that appear side-by-side or close together, since these questions will necessarily elicit similar answers from subjects because subjects want to appear consistent (both to the experimenter and to themselves). Hence, repetitive questions that are near each other will not generate “independent observations.”

Recent research by Kahneman, Krueger, Schkade, Schwarz and Stone¹ reports a new finding that can be used to eliminate some affect questions when evaluating situation-contingent quality of life (e.g., eating with friends, visiting family, shopping, waiting in line, commuting, cleaning, etc...). Kahneman et al report that if a situation elicits a particular negative affect (e.g., restlessness) in a population of respondents then that situation is highly likely to elicit other negative affects (e.g., anger and frustration) and is unlikely to elicit positive affects. Kahneman et al report associated correlation statistics (for population averages) around 0.8. This finding suggests that *for policy purposes* situational experiences can be placed on a single positive-negative affect continuum, potentially eliminating the need to measure many different specific types of affect (both positive and negative). Hence, evaluations of the social costs/benefits of an activity or situation can be made without measuring all of the different (and highly correlated) specific affects associated with the activity.² These results only apply to policy analysis, since the results apply to population averages and (naturally) do not apply to individual responses.³

¹ Reported at 2002 LSE conference on well-being and economics.

² The Postive Affect Negative Affect Scale (PANAS) measures 20 affects. The results of Kahneman et al imply that much simpler scales, which only measure three or four different affects, could be designed to answer policy questions about the social value of certain situations.

³ Individual responses are also influenced by individual fixed effects and measurement error. Both of these effects are “averaged out” in population averages.

3. Genetic measures.

Although genetic measures pose enormous logistical and ethical problems, we encourage the HRS to continue exploring this promising new area of research. The ongoing explosion in knowledge about the genome and the rapid drop in genotyping costs make genetic research a natural area of current and future inquiry. Perhaps the HRS should organize a conference that considers the feasibility and value of adding genotyping to the existing survey instruments. We believe that the addition of genetic information to the HRS will have spectacular payoffs over the next two decades. Such value derives from both traditional studies of biomedical pathology and new studies that relate genetic and environmental factors to socioeconomic phenotypes (savings, education, unemployment, etc...).

We encourage the HRS to consider the use of buccal swabs or other non-threatening technologies (e.g., finger pricks) for taking DNA samples. If possible, immortalized DNA lines should be created for respondents. We recognize that such research methods are quite costly, but we anticipate that these costs will continue to rapidly decline in coming years.

Genotyping also poses daunting ethical challenges. These ethical issues represent the most important counterarguments to the collection of genetic information; the administrators of the survey may feel ethically obligated to inform respondents about their genotyped disease predispositions, thereby giving the respondents information that will affect their survey participation and their behavior, thereby biasing the sample. The potential obligation to reveal such information to subjects should be evaluated by a panel of experts in genetics research.

If further cost-benefit analysis supports the addition of a genetic component on the HRS, we encourage the NIH to provide the necessary funds to make such research possible. We view genetic measurement as an important long-run consideration for all demographic surveys in the decades ahead. In our view, genetic information may prove to be as useful as standard demographic information like age and gender.

4. Situational determinants of wealth accumulation.

Recent research has shown that total retirement savings are extremely sensitive to largely arbitrary institutional factors like DB pensions, automatic 401(k) enrollment, 401(k) matches, and automatic IRA rollovers of 401(k) balances at separation (e.g., Choi, Laibson, Madrian, and Metrick 2001).⁴ We recommend that the HRS measure institutional features of retirement savings plans. Here is a list of the factors that have already appeared in the literature, most of which are not currently measured in the HRS:

- defined benefit pensions⁵
- automatic enrollment (automatic non-enrollment)
- automatic IRA rollover at separation (automatic cash distribution at separation)
- employer non-contingent 401(k) contribution
- employer matching contribution (match threshold and match percentage)
- company stock available in 401(k) plan
- employer match made in company stock
- 401(k) loans

The HRS should measure these and other institutional features of 401(k) plans, enabling researchers to evaluate the relative frequency of these plan rules and their impact on participant choices. Such measurements are particularly important for the younger, currently employed cohorts that will be sampled in future HRS waves.

5. Behavioral finance.

The past two decades a large and growing body of research in behavioral finance has revealed a lack of sophistication in the investment decisions of individual households.⁶

⁴ Choi, James J., David Laibson, Brigitte C. Madrian, and Andrew Metrick, [Defined Contribution Pensions: Plan Rules, Participant Decisions, and the Path of Least Resistance](#), *Tax Policy and the Economy*. 2001.

⁵ DB pensions are currently measured in the HRS.

⁶ Numerous papers by Brad Barber and Terry O'Dean and by Shlomo Benartzi and Richard Thaler are particularly influential in this literature. Here are a few examples: Barber, Brad and Terrance O'Dean, [Are Individual Investors Tax Savvy? Evidence from Retail and Discount Brokerage Accounts](#) with Brad Barber, forthcoming in the *Journal of Public Economics*. _____ and _____, ["Boys will be Boys: Gender, Overconfidence, and Common Stock Investment"](#), *Quarterly Journal of Economics*, February 2001, Vol. 116, No. 1, 261-292.

Many of these problematic investment choices have significant consequences for household welfare. The HRS should shed light on these issues by considering survey questions that measure:

- mutual fund management fees and household awareness of these fees (e.g., ask respondent for the names of their mutual funds so that beliefs about management fees can be compared to the actual management fees)
- portfolio churning, related tax and transaction cost implications, and household awareness of these issues
- underdiversification due to concentrated equity holdings (e.g., holding a large fraction of equity wealth in company stock)
- overoptimistic forecasts of stock returns (e.g., measure expected returns so they can be compared to actual returns)
- excessive credit card debt and credit card interest rates
- failures to refinance mortgages after interest rate declines (i.e., what is the interest rate on the homeowner's mortgage)
- predatory lending (particularly to the elderly)
- sources of financial advice and nature of relationship (e.g., commission based or fee for service)

The HRS should also measure financial literacy as well as financial expectations (first and second moments). As defined benefit pension plans become less and less common, households are increasingly being called upon to make their own financial decisions. These trends will accelerate if social security is partially "privatized." Hence, it is important to take stock of the financial literacy of US households. Many questions in

_____ and _____, "[Trading is Hazardous to Your Wealth: The Common Stock Investment Performance of Individual Investors](#)", *Journal of Finance*, Vol. LV, No. 2, April 2000, 773-806.

Benartzi, Shlomo, "[Excessive Extrapolation and the Allocation of 401\(k\) Accounts to Company Stock?](#)" *Journal of Finance*, October 2001, Vol. 56.5, pp. 1747-1764.

_____ and Richard Thaler, "[How Much Is Investor Autonomy Worth?](#)" *Journal of Finance*, August 2002, Vol. 57.4, pp. 1593-1616.

_____ and _____, "Naive Diversification Strategies in Retirement Saving Plans," *American Economic Review*, March 2001, Vol. 91.1, pp. 79-98.

O'Dean, Terrance, "[Are Investors Reluctant to Realize Their Losses?](#)", *Journal of Finance*, Vol. LIII, No. 5, October 1998, 1775-1798.

the HRS already do this, but it would be useful to get an even more fine-grained portrait of the financial knowledge and sophistication of US households.

6. Widowhood and widowerhood.

The death of a spouse is a critical event because of its effects on household finances and the subjective well-being of the surviving spouse. These effects can be unusually profound if the deceased spouse managed the household's finances or was a caregiver.

Older adults are particularly vulnerable when faced with the death of a spouse because their social networks may be weak already. Their social interactions are adversely affected by separation from job-related social networks, movements out of old neighborhoods, deaths of non-spousal family/friends, physical infirmity that makes it difficult to visit family/friends, and the challenges of making new friendships late in the lifecycle. Ongoing research on subjective well-being identifies social networks, friendship and companionship as key determinants of subjective well-being.

We recommend a workshop or workshops to discuss these issues and develop related survey instruments. The workshops should focus on the issue of widowhood and widowerhood, but should also address the related question of elderly poverty (particularly of widows) and the challenges of measuring critical social opportunities for older adults.

7. Consumption measure.

The current survey has only a limited measure of consumption. But consumption is an extremely important piece of information for evaluating the quality of life of older adults. In addition, savings adequacy can be inferred from the consumption path over the lifecycle. Consumption paths that rapidly decline in retirement suggest inadequate savings and/or inadequate annuitization.

Wealth holdings can also be used to infer savings adequacy, but wealth measures have the drawback that they only represent financial wealth. It is quite hard to measure claims to pension benefits, which represent roughly half of the total wealth of the typical

_____, "[Do Investors Trade Too Much?](#)", *American Economic Review*, Vol. 89, December 1999, 1279-1298.

US retiree. Consumption measures avoid these problems, since consumption measures are relatively insensitive to the source of the income.

Consumption measures are also useful, since they reveal important lifecycle dynamics in the quality of life. For example, if consumption is disaggregated into specific types of expenditure (e.g., housing, food, vacation, prescription drugs, etc....), the researcher can gain a clearer picture of how the quality of life changes with age. For example, an age-related rise in expenditure on prescription drugs sends a different message than an age-related rise in vacation spending. Hence, an ideal measure of consumption would reveal not just the total magnitude of consumption, but also a breakdown by consumption category.

A workshop that invites the designers and some prominent users⁷ of the Consumption Expenditure Survey (CEX) might provide a useful way of thinking about these issues. However, it is important to note that the goal is not to measure consumption as carefully as the CEX. Instead, the goal of the HRS should be to design a comparatively short instrument that measures consumption in a way that facilitates comparison to the CEX and exploits the large amounts of experience that has been gained by the designers and users of the CEX.

8. Compliance

Most people have a hard time following through on their best intentions and on medical advice to take medications, exercise regularly, eat healthfully, etc... It may be worthwhile trying to measure this compliance gap. Are older adults following their doctor's advice? What fraction of their pills do they forget/fail to take? Are they following their doctors' recommendations to exercise and cut down on cholesterol and salt intake? Are diabetics reducing sugar intake appropriately? Do older adults have a hard time complying with their doctors' advice? And if so why?

9. Risk aversion.

Measuring risk aversion is difficult, particularly because economic models of risk aversion do not correspond to consumer attitudes toward risk. For example, loss aversion

⁷ A list of CEX users might include Chris Carroll, Jonathan Parker, and Nicholas Soulelos.

(“first order risk aversion”), which has been measured in hundreds of laboratory and field studies, should not exist if standard *economic* models of risk aversion are descriptively accurate. We are concerned that the HRS relies too heavily on standard economic models of risk aversion in the design of the instruments that measure attitudes toward risk.

With these issues in mind, we suggest that the HRS reevaluate its current risk aversion measures. There is a growing body of research by financial industry practitioners that attempts to measure attitudes toward risk with innovative measures. Much of this new knowledge has accumulated among people who do 401(k) consulting and advising. Daniel Kahneman has some experience with these new measures and would be happy to talk with the HRS about the design of these measures.

10. Implementation

We believe that all of the HRS reviews should be interpreted as suggestions only. Demand for change ultimately has to come from the people who are actually involved in the design and use of the survey. The external reviewers may not have an insider’s view of which parts of the HRS are running well and which parts can easily be revised without disrupting ongoing research projects.

For example, some HRS questions are currently being actively used in ongoing and valuable time series analysis. Naturally, these instruments should probably not be changed. More generally, particular researchers may have an important stake in certain questions in the survey. We do not want to encourage the HRS to overlook the interests of these “stakeholders.” We believe that our suggestions are most useful in designing new areas of emphasis and bringing new ideas to bear on the future redesign of the survey. We encourage the HRS designers to consult active users of the existing instrument during future efforts to amend or abandon existing questions.