



4-8-2017

Cognitive Constraints on Valuing Annuities


Jeffrey R. Brown

Arie Kapteyn

Erzo F. P. Luttmer

Olivia Mitchell

Follow this and additional works at: http://repository.upenn.edu/bepp_papers

 Part of the [Business Commons](#), [Economics Commons](#), and the [Public Affairs, Public Policy and Public Administration Commons](#)

Cognitive Constraints on Valuing Annuities

Abstract

This paper documents consumers' difficulty valuing life annuities. Using a purpose-built experiment in the American Life Panel, we show that the prices at which people are willing to buy annuities are substantially below the prices at which they are willing to sell them. We also find that buy values are *negatively* correlated with sell values and that the sell–buy valuation spread is negatively correlated with cognition. This spread is larger for those with less education, weaker numerical abilities, and lower levels of financial literacy. Our evidence contributes to the emerging literature on heterogeneity in financial decision-making abilities, particularly regarding retirement payouts.

Disciplines

Business | Economics | Public Affairs, Public Policy and Public Administration

NBER WORKING PAPER SERIES

COGNITIVE CONSTRAINTS ON VALUING ANNUITIES

Jeffrey R. Brown
Arie Kapteyn
Erzo F.P. Luttmer
Olivia S. Mitchell

Working Paper 19168
<http://www.nber.org/papers/w19168>

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
June 2013

Previously circulated as “Decision Complexity as a Barrier to Annuitization”. The research reported herein was performed pursuant to a grant from the U.S. Social Security Administration (SSA) funded as part of the Financial Literacy Consortium. The authors also acknowledge support provided by the Pension Research Council/Boettner Center at the Wharton School of the University of Pennsylvania, and the RAND Corporation. The authors thank Jonathan Li, Caroline Tassot and Yong Yu for superb research assistance, and Tim Colvin, Tania Gutsche, Bas Weerman, and participants of the Netspar 2012 Paris conference and the NBER PE program meetings for their invaluable comments and assistance on the project. Brown is a Trustee of TIAA and has served as a speaker, author, or consultant for a number of financial services organizations, some of which sell annuities and other retirement income products. Mitchell is a Trustee of the Wells Fargo Advantage Funds and has received research support from TIAA-CREF. The opinions and conclusions expressed herein are solely those of the authors and do not represent the opinions or policy of SSA, any agency of the Federal Government, any other institution with which the authors are affiliated, or the National Bureau of Economic Research.

At least one co-author has disclosed a financial relationship of potential relevance for this research. Further information is available online at <http://www.nber.org/papers/w19168.ack>

NBER working papers are circulated for discussion and comment purposes. They have not been peer-reviewed or been subject to the review by the NBER Board of Directors that accompanies official NBER publications.

© 2013 by Jeffrey R. Brown, Arie Kapteyn, Erzo F.P. Luttmer, and Olivia S. Mitchell. All rights reserved. Short sections of text, not to exceed two paragraphs, may be quoted without explicit permission provided that full credit, including © notice, is given to the source.

Cognitive Constraints on Valuing Annuities

Jeffrey R. Brown, Arie Kapteyn, Erzo F.P. Luttmer, and Olivia S. Mitchell

NBER Working Paper No. 19168

June 2013, Revised October 2014

JEL No. D03,G02,H55

ABSTRACT

We show that people have difficulty valuing annuities, and this, instead of a preference for lumpsums, helps explain observed low annuity demand. Although the median price at which people are willing to sell an annuity stream is close to the actuarial value, many responses diverge greatly from optimizing behavior. Moreover, people will pay substantially less to buy than to sell annuities. We conclude that boundedly rational consumers adopt “buy low, sell high” heuristics when confronting a complex trade-off. This suggests that many consumers do not make optimizing decisions, underscoring the difficulty of explaining cross-sectional annuity valuation differences using standard models.

Jeffrey R. Brown
Department of Finance
University of Illinois at Urbana-Champaign
515 East Gregory Drive
Champaign, IL 61820
and NBER
brownjr@illinois.edu

Erzo F.P. Luttmer
6106 Rockefeller Center, Room 305
Department of Economics
Dartmouth College
Hanover, NH 03755
and NBER
Erzo.FP.Luttmer@Dartmouth.Edu

Arie Kapteyn
University of Southern California
Center for Economic and Social Research
635 Downey Way
Los Angeles, CA 90089-3332
and NBER
kapteyn@usc.edu

Olivia S. Mitchell
University of Pennsylvania
Wharton School
3620 Locust Walk, St 3000 SH-DH
Philadelphia, PA 19104-6302
and NBER
mitchelo@wharton.upenn.edu

A data appendix is available at:
<http://www.nber.org/data-appendix/w19168>

Cognitive Constraints on Valuing Annuities

It is difficult for the average person contemplating retirement to determine how to draw down his wealth. Choosing a wealth decumulation and consumption strategy that maximizes lifetime utility is a highly complex problem that requires the ability to optimize intertemporally and under multiple sources of uncertainty, aspects that require substantial cognitive effort. Most formal models of annuitization, however, ignore the possibility that individuals may differ in their decision-making competency. The commonly used life-cycle model assumes a rational and fully-informed individual who knows the distribution of mortality rates, market returns, inflation, future expenditures, and income, and who can use this knowledge to optimally choose the mix of financial products to smooth his marginal utility of consumption across time and states-of-the-world at the least cost.

Recently, economists have been more actively documenting differences in individual decision-making abilities and heterogeneity in the extent to which their behavior can be described as economically rational (Choi, Kariv, Müller, and Silverman 2014). A key implication of this line of research is that there can be a gap between peoples' actual decisions and the decisions they would have made if they had perfectly developed decision-making abilities. This paper explores this idea in the important context of retirement income security, focusing specifically on the quality of decision making when valuing an annuitized income stream. Rather than assessing quality by comparing actual annuitization behavior to a theoretically optimal level of annuitization for each individual – a Herculean task given that existing research provides a wide range of results (literally from zero to full annuitization) depending on model assumptions – we instead focus on whether people are internally consistent in their valuations across different elicitation methods. Specifically, we present individuals with scenarios in which they are offered an opportunity to exchange an annuity for a lump sum (what we call “selling” the annuity), as well as scenarios in which the same individuals are offered an opportunity to exchange a lump sum for an annuity (what we call “buying” the annuity). In both cases, we also vary whether the trade-off is presented in a manner consistent with compensating or equivalent variation measures from welfare economics, allowing us to distinguish variation in valuations from endowment effects.

Our central hypothesis is that people differ in their ability to meaningfully value a stream of life annuity income relative to a lump sum, and that this ability is correlated with measures of

cognitive ability including education, financial literacy, and numeracy. There are four primary reasons that we have chosen to study the lump-sum versus annuity decision rather than other financial or economic decisions. First, the annuitization decision is important in its own right as an academic research topic. Indeed, there is a vast academic literature, dating back a half century, on the role that annuities should and do play in peoples' portfolios in later life.¹ Second, it is also important and timely as an issue of retirement policy in the U.S. and numerous other developed nations (e.g., US DOL 2010). Third, the annuitization decision is a natural place to look for variation in consumers' decision-making abilities. Valuing an annuity versus a lump sum is cognitively challenging because it requires that one wrestle with multiple sources of uncertainty (e.g., mortality, returns, inflation) and it also requires that one make a near-term choice with far-distant consequences, two characteristics known to render decision making difficult (Beshears, Choi, Laibson, and Madrian 2008). Fourth, because individuals are typically faced with an annuitization decision only one or two times in their lives, this is not a transaction that people learn about through repeated market interactions (Bernheim 2002). In such settings, individuals are known to be less likely to optimize and more likely to adopt rule-of-thumb behaviors (Kling, Phaneuf, and Zhao 2012), which can drive a wedge between true versus revealed preferences (Beshears, Choi, Laibson, and Madrian 2008).

We provide six pieces of evidence consistent with the hypothesis that individuals have difficulty valuing annuities and that the degree of difficulty is correlated with their cognitive abilities. First, we show that a non-trivial fraction of the population has implied annuity values that are difficult to reconcile with optimizing behavior under *any* plausible set of parameters. Second, we uncover a large divergence between the price at which individuals are willing to *buy* versus *sell* an annuity, a result that cannot be explained by liquidity constraints or endowment effects. Third, and even more striking, we find that the buy and sell valuations are *negatively* correlated. In other words, people who demand higher sell prices are also more likely to offer very low buy prices. Fourth, we show that the size of the buy-sell valuation discrepancy is strongly negatively correlated with cognitive ability as measured by education, financial literacy, and numeracy. This is consistent with less cognitively capable individuals having much greater difficulty valuing a stream of annuity payments. Fifth, we use additional experimental variation to show that the elicited valuations are sensitive to anchoring effects and thus they violate the "invariance" criterion

¹ Two reviews include Benartzi, Previtro, and Thaler (2011) and Poterba, Venti and Wise (2011).

of rational decision making. Finally, we argue that it is difficult to explain observed cross-sectional variation in the measured annuity valuation amounts with theoretically attractive measures. In other words, the pattern of significant marginal valuation predictors is more consistent with individuals using simple heuristics rather than full optimization to value the trade-offs.

Our evidence is drawn from a randomized experiment we conducted in the American Life Panel (ALP), where we presented respondents with hypothetical choices between a lump sum and a Social Security annuity. By varying whether the questions elicited a compensating variation (CV) or an equivalent variation (EV) value, whether the individual was buying or selling the annuity, the size of the increments, and the order of the questions, we directly examine the coherence and stability of subjective valuations placed by respondents on their Social Security annuities. We collected a number of additional variables to control for potentially confounding factors such as heterogeneity in liquidity constraints and beliefs about political risk.

Like most economists, we usually find evidence based on actual choices in natural settings more compelling than evidence based on hypothetical choices. We acknowledge important drawbacks of using hypothetical choices, such as the possibility that lower stakes could lead respondents to exert less effort and seek out fewer resources to assist with their decisions. However, although these considerations may make hypothetical choice behavior noisy, it would be surprising if they led to systematic patterns in hypothetical choice behavior that would be completely absent in actual choices.

Counterbalancing these drawbacks are three important benefits of using a hypothetical choice setting. First, the hypothetical setting allows us to observe an individual's annuitization decisions for a wide range of annuity prices, from which we obtain individual-specific annuity valuations without having to rely on functional form assumptions. In real world settings, annuitization decisions are typically made at a single price (and if there is price variation, it is generally not exogenous). Second, in a hypothetical setting, it is feasible to observe both the price at which an individual is willing to buy and the price at which he is willing to sell the annuity. Such within-person variation turns out to be extremely valuable in exploring cognitive constraints on individuals' abilities to value annuities. Third, the hypothetical setting allows us to elicit annuitization choices for a broadly representative sample of the U.S. population. As discussed in the literature overview below, actual annuitization decisions in natural settings are typically only observed for rather select populations.

In addition to advancing our academic understanding of consumer behavior in this area, our results also have considerable practical policy relevance. There is currently an active discussion on what role annuities should play in defined contribution (DC) or 401(k) pension plans, with much debate about whether and how life annuities ought to be encouraged in such settings (Gale, Iwry, John, and Walker 2008; Steverman 2012). This debate, in part, revolves around whether people can make optimal payout decisions using their accumulated retirement assets. Moreover, many countries, including the U.S., are grappling with fiscally unsustainable pay-as-you-go public pension systems. To the extent that households are poorly equipped to value the annuities offered by their public pensions, this has implications for the political feasibility of reforms changing the benefit structure, particularly if retirees were to be offered a choice between a lump sum and future annuity payments. The same point applies to state and local public defined benefit plans (DB) in the U.S., which also face substantial underfunding problems (Novy-Marx and Rauh 2011); indeed some reformers have called for a reduction in DB annuities in exchange for lump-sum contributions to defined contribution (DC) accounts (e.g., Kilgour 2006).

In what follows, we first summarize key prior studies on the demand for annuities from both the neoclassical and the behavioral economics literatures. Next, we describe the American Life Panel (ALP) internet survey, a broadly representative sample of the U.S. population, and we outline how we elicited lump-sum versus annuity preferences in this survey, using a purpose-built module. We then present our key empirical results, followed by a number of robustness checks and further analyses for subgroups that vary according to financial capabilities. We conclude with a discussion of possible policy implications and future research questions.

I. Related Literature

A. Annuity Demand

There is a very large economics literature focused on modeling the optimal level of annuitization for life-cycle consumers under various assumptions.² That literature began with

² Rather than providing a comprehensive review here, we instead highlight those studies most germane to the research that follows. Readers interested in the broader literature on life annuities may consult Benartzi, Previtero, and Thaler (2011); Poterba, Venti, and Wise (2011); Brown (2008); Horneff, Maurer, Mitchell, and Dus (2008); and Mitchell, Poterba, Warshawsky, and Brown (1999). Note that we use the term “life annuity” because we are interested in products that guarantee income for life, as opposed to financial products such as “equity indexed annuities” that are mainly used as tax-advantaged wealth accumulation devices (and hence they are rarely converted into life-contingent income).

Yaari's (1965) paper in which he noted a set of conditions under which it would be optimal for an individual to annuitize all of his wealth.³ Extensions to the theory went on to show that full annuitization would be optimal under more general conditions,⁴ a puzzling prediction in light of very low annuitization rates in the real world (Mitchell, Piggott, and Takayama 2011). Extended life-cycle models have been constructed to measure consumer valuations of life annuities and to compute how optimal annuitization will vary with other factors including: pricing (Mitchell, Poterba, Warshawsky, and Brown 1999); pre-existing annuitization (Brown 2001; Dushi and Webb 2006); risk-sharing within families (Kotlikoff and Spivak 1981; Brown and Poterba 2000); uncertain health expenses (Turra and Mitchell 2008; Sinclair and Smetters 2004; Peijnenburg, Nijman, and Werker 2010a, 2010b); bequests (Brown 2001; Lockwood 2011); inflation (Brown, Mitchell, and Poterba 2001, 2002); the option value of learning about mortality (Milevsky and Young 2007); stochastic mortality processes (Reichling and Smetters 2012; Maurer, Mitchell, Rogalla, and Kartashov 2013); and broader portfolio issues including labor income and the types of assets on offer (Inkmann, Lopes, and Michaelides 2011; Koijen, Nijman, and Werker 2011; Chai, Horneff, Maurer, and Mitchell 2011; Horneff, Maurer, Mitchell, and Stamos 2009, 2010). Opinions differ on the extent to which this theoretical literature has been successful in resolving the annuity puzzle, even for marginal annuitization decisions (e.g., Shepard 2011).

There is also a sizable empirical literature on annuities, despite the small size of the voluntary life annuity market in most countries. Brown (2001) used the 1992 wave of the U.S. Health and Retirement Study (HRS) to show that expected annuitization from DC plans was correlated with the annuity valuations predicted by a life-cycle model based on demographic characteristics, though only for persons with sufficiently long (>1 year) planning horizons. Hurd and Panis (2006) explored payouts from DB plans in the HRS and found peoples' behavior was consistent with *status quo* bias. Büttler and Teppa (2007) used Swiss administrative data to track choices made by employees in ten different pension plans and concluded that annuitization was

³ The conditions include no bequest motives, time-separable utility, exponential discounting, and actuarially fair annuities (among others).

⁴ Davidoff, Brown, and Diamond (2005) showed that full annuitization is optimal under complete markets with no bequest motives. Peijnenburg, Nijman, and Werker (2010a; 2010b) found that if agents saved optimally out of annuity income, full annuitization can be optimal even in the presence of liquidity needs and precautionary motives. They further found that full annuitization is suboptimal only if agents risk substantial liquidity shocks early after annuitization and do not have liquid wealth to cover these expenses. This result was robust to the presence of significant loads.

higher in plans where an annuity was the default payout option. Chalmers and Reuter (2012) exploited exogenous variation in the price of annuities using Oregon public-sector workers; they (unexpectedly) found that worker demand for partial lump-sum payouts rose rather than fell as the value of the forgone life-annuity payments increased, leading them to conclude that the decisions were being made by unsophisticated individuals. Previtro (2014) showed that annuity demand was negatively correlated with the prior year's stock returns, consistent with consumers engaging in naïve trend chasing.

Several experimental papers have also suggested that annuitization decisions may be reflective of non-optimizing behavior. Agnew, Anderson, Gerlach, and Szykman (2008) showed that individuals could be steered toward or away from life annuities in an experimental setting, depending on whether the products were described using positive or negative frames. Brown, Kling, Mullainathan, and Wrobel (2008) used an internet survey to show that perceptions of annuity value relative to alternative financial products were heavily influenced by whether the products were described using “consumption” or “investment” frames. Beshears, Choi, Laibson, Madrian, and Zeldes (2014) also found evidence that framing affects annuity demand. Brown, Kapteyn, and Mitchell (2013) showed that Social Security claiming behavior (which is akin to making an annuitization decision) was influenced by framing changes. Accordingly, this small literature suggests that individuals behave in a manner at odds with optimizing models.

Overall, the difficulty explaining annuitization behavior using optimizing models, combined with the empirical and experimental work suggesting behavior is not fully consistent with rational decision making, motivates our research into whether individuals are making optimization errors in making decisions about annuities and whether the heterogeneity in decision-making quality is correlated with other observable factors.

B. Variation in Decision-making Abilities

In contrast to the neoclassical literature, which posits perfectly informed and rational individuals optimizing their consumption path over the life cycle, Lusardi and Mitchell (2014) provide a comprehensive review of the large and growing literature relating financial literacy to behavior, including the robust finding that many households lack basic financial knowledge. Research has documented that households make a range of financial mistakes when managing financial affairs (e.g., Calvet, Campbell, and Sodini 2007, 2009; Agarwal, Driscoll, Gabaix, and Laibson 2009) and that households making mistakes often lack day-to-day financial skills (Hilgert,

Hogarth, and Beverly 2003). It has also established that financial literacy is correlated with the propensity to participate in financial markets (Kimball and Shumway 2006; Christelis, Jappelli, and Padula 2010; van Rooij, Lusardi, and Alessie 2011; Almenberg and Dreber 2011; and Arrondel, Debbich, and Savignac 2013). Earlier work by Lusardi and Mitchell (2007, 2011) also showed that more financially literate individuals were more likely to engage in retirement planning and to accumulate retirement wealth.

A related literature has focused on variation in cognitive abilities and financial decision making. Fang, Keane, and Silverman (2008) found that cognitive functioning was a stronger predictor of Medigap purchase than risk preferences. Agarwal and Mazumder (2013) reported that performance on cognitive tests helped explain the quality of financial decisions related to the use of credit. A subset of this literature has also focused more specifically on retirement-preparedness among older individuals. For example, McArdle, Smith, and Willis (2011) and Banks, O’Dea, and Oldfield (2010) found that cognitive abilities helped explain retirement wealth accumulation.

Taken together, these and many other studies suggest that people differ in their financial decision-making abilities and that these differences are important correlates of financial well-being late in life. Taking this literature an important step further in establishing causality, Choi, Kariv, Müller, and Silverman (2014) conducted a large-scale experiment designed to directly test the extent to which individual decisions were consistent with the Generalized Axiom of Revealed Preference (GARP); they found substantial heterogeneity. Importantly for our work, they found that their measure of decision-making quality was higher among those with more education and among younger individuals, while they also showed that individuals with better decision-making skills accumulated more wealth. Behrman, Mitchell, Soo, and Bravo (2012) also reported that the more financially literate saved more in their pensions, controlling for the possible endogeneity of financial knowledge.

Our experiments contribute to this literature in two ways. First, we focus on a decision important in its own right – annuitization – but where heterogeneity in decision-making quality has not been studied. Second, we use an outcome that is a novel measure of decision-making ability; we take the spread of peoples’ responses across different approaches to eliciting stated valuations for life annuities as a measure of low decision-making quality. We find that this spread is strongly inversely related to cognition.

II. Methodology and Data

A. *The Social Security Context*

Our experiments, described below in greater detail, use Social Security benefits as the context rather than describing an unfamiliar hypothetical annuity product. This approach has several advantages. First, most workers have an understanding that Social Security pays benefits to retirees that last for as long as they live (Greenwald, Kapteyn, Mitchell, and Schneider 2010; Liebman and Luttmer 2011), which means that respondents will understand the nature of our “offer” to trade off annuities and lump sums. Second, our context provides a simple way to control for possible concerns about the private annuity market that might influence results such as the lack of inflation protection (our question makes it clear that Social Security is adjusted for inflation), or concerns about counterparty risk of the insurer providing the annuity.⁵ Third, given the ongoing debate about the U.S. long-term fiscal situation, our setting is highly policy-relevant. For example, past discussions of possible pension reforms around the world, as well as at the state and local levels in the U.S., have included proposals to partially “buyout” benefits by issuing government bonds to workers in exchange for a reduction in their annuitized benefits. Several U.S. corporations have also recently offered to buy back defined benefit pension annuities from retirees, in exchange for a lump sum (c.f., Wayland 2012).

B. *The American Life Panel*

To test how people value their Social Security annuity streams, we fielded a survey between June and August of 2011 using the RAND American Life Panel, a panel of U.S. households that regularly take surveys over the Internet. If, at the recruiting stage, households lacked Internet access, this was provided by RAND.⁶ By not requiring Internet access in the recruiting stage, the ALP has an advantage over most other Internet panels when it comes to generating a representative sample.⁷ The American Life Panel included about 4,000 active panel members at the time of our experiment. Our survey was conducted over two waves of the ALP to

⁵ Below we examine whether concerns about the fiscal sustainability of Social Security influences people’s valuation of the Social Security annuity. See Luttmer and Samwick (2012) for a detailed analysis of the effects of policy uncertainty on valuations of future Social Security benefits.

⁶ Initially these households received a WebTV allowing them to access the Internet. Since 2008, households lacking Internet access at the recruiting stage have received a laptop and broadband Internet access.

⁷ A more detailed explanation of the ALP is provided in Online Appendix A, along with a brief description of how we estimated Social Security benefits for survey respondents. Our survey instrument is included in Online Appendix B.

keep the length of each questionnaire within manageable bounds, and we invited ALP participants age 18 or older to take our survey. For the first wave, we selected 2,954 respondents, of whom 2,478 completed the survey for a response rate of 83.9%. Those who completed the first wave were invited to participate in a second survey at least two weeks later; of these, 2,355 respondents completed the second wave for a response rate of 95.0%. If a participant indicated he did not think he would be eligible to receive Social Security benefits (either on his own earnings records or on those of a current, late, or former spouse), he was asked to assume for the purposes of the survey that he would receive Social Security benefits equal to the average received by people with his age/education/sex characteristics (see Online Appendix B). Our sample comprised 2,112 complete responses for both waves 1 and 2.⁸

Table 1 compares our sample characteristics with those of the same age group in the Current Population Survey (CPS).⁹ Our sample is, on average, five years older, more female, more non-Hispanic white, better educated, slightly higher-income, and has a somewhat smaller household size than the CPS; the regional distribution is close to that of the CPS. The fact that our sample is more highly educated means that, if anything, our respondents should be in a better position to provide meaningful responses to complex annuity valuation questions, compared to a national sample. Despite the differences between the ALP and the CPS, our ALP sample does include respondents from a wide variety of backgrounds, so in this sense, we think of the ALP as broadly representative of the U.S. population.

Table 1 here

C. Eliciting Lump-Sum versus Annuity Preferences

To elicit preferences over annuitization, respondents were posed several questions of the following sort:

In this question, we are going to ask you to make a choice between two money amounts. Please click on the option that you would prefer. Suppose Social Security gave you a choice between:

(1) Receiving your expected Social Security benefit of \$SSB per month.

or

⁸ Of the 2,355 respondents who completed the second wave, we dropped 69 observations from the pilot version of wave 2 (where the questionnaire was slightly different). We further dropped 168 observations where survey instrument was incorrectly administered due to a technical glitch and we dropped 6 observations with missing information on basic demographics (age, education, or marital status).

⁹ Summary statistics of other key variables from our survey (such as annuity valuations, discussed below) are listed in Online Appendix Table A.1.

(2) Receiving a Social Security benefit of $\$(SSB-X)$ per month and receiving a one-time payment of $\$LS$ at age Z .

The variable SSB is an estimate of the individual's estimated monthly Social Security benefit; the variable LS refers to the lump-sum amount; and Z is the individual's self-reported expected claiming age. For those not currently receiving benefits, the trade-off was posed as a reduction in future monthly Social Security benefits in exchange for a lump sum to be received at that person's expected claiming age. For those currently receiving Social Security benefits, the questions were modified so as to compare a change in monthly benefits to the receipt of a lump sum in one year. In both cases, the receipt of the lump sum was in the future rather than immediately; we did this to avoid having the answers being affected by possible present-biased preferences. Before asking the annuity trade-off question, we instructed all respondents to "please assume that all amounts shown are after tax (i.e., you don't owe any tax on any of the amounts we will show you)" and "please think of any dollar amount mentioned in this survey in terms of what a dollar buys you today (because Social Security will adjust future dollar amounts for inflation)." In the trade-off question, we told married respondents, "Benefits paid to your spouse will stay the same for either choice." Thus, individuals were asked to value a single-life inflation-indexed annuity.

In order to probe the reliability of the valuations provided by respondents, we varied the question in a systematic way along two dimensions. First, we elicited how large a lump sum would be required to induce an individual to accept a reduction of (i.e., to *sell*) a portion of his Social Security income; below we refer to this version of the question with the shorthand "sell." We also elicited how much the individual would be willing to *pay* in order to increase his Social Security annuity (the "buy" condition). The difference in responses to these alternative solicitations is a central focus of what follows.

A second dimension along which we varied our questions depended on whether we elicited a compensating variation (CV) – the annuity/lump-sum trade that would keep people at their existing utility level – or an equivalent variation (EV) –the lump-sum amount that would be equivalent in utility terms to a given change in the monthly annuity amount. As we discuss in greater detail below, an analysis of the CV versus EV distinction should allow us to distinguish our findings from a simple *status quo* bias or endowment effect. This is because in the EV version of the questions, each individual had to choose an increment or decrement to his annuity. The *status quo* was not an option in this scenario.

In total, we elicited all four measures and designate them as **CV-Sell** (as in the example above), **CV-Buy**, **EV-Sell**, and **EV-Buy**. The chart below illustrates the essential differences across these four scenarios. We define *SSB* as the monthly Social Security benefit the individual was currently receiving (if the individual was a current recipient), or was expected to receive in the future (if the individual was not a recipient); *X* is the increment (or decrement, if subtracted) to that monthly Social Security benefit. Finally, we set *LS* as the lump-sum amount offered in exchange for the change in monthly benefits. In essence, this paper is about how individuals tradeoff a monthly benefit of $\$X$ for a lump-sum of amount $\$LS$.

Four Variants of the Annuity Valuation Tradeoff Question

	"Sell"-version		"Buy"-version	
	Choice A	Choice B	Choice A	Choice B
Compensating Variation (CV)	$[SSB-X] + LS$	$[SSB]$	$[SSB+X] - LS$	$[SSB]$
Equivalent Variation (EV)	$[SSB]+ LS$	$[SSB+X]$	$[SSB] - LS$	$[SSB-X]$

Note: *SSB* stands for current/expected monthly Social Security benefits, *X* is the amount by which monthly Social Security benefits would change and *LS* is a one-time, lump-sum payment. Positive amounts are received by the individual while negative amounts indicate a payment by the individuals. Amounts between square brackets are paid monthly for as long as the individual lives, whereas *LS* is a one-time payment. The individual is asked to elect Choice A or Choice B.

The CV-Sell scenario presented individuals with a choice between their current or expected Social Security benefits (*SSB*), versus a scenario in which their benefits would be reduced by $\$X$ per month in exchange for receiving a lump sum of $\$LS$. The EV-Sell scenario provided a choice between receiving a higher monthly benefit ($SSB+X$) or receiving $\$SSB$ plus a lump sum of $\$LS$. Note that within the Sell scenario, one can obtain EV simply by adding $\$X$ to each side of the CV trade-off. Given that $X=\$100$ per month in the baseline versions, the change in benefits is modest relative to total monthly income for most individuals. We would therefore expect CV and EV to be comparable, barring strong endowment effects that could be present in the CV formulation but not in the EV formulation (where the *status quo* was not an option).

Switching to the Buy scenarios, the CV-Buy question provided a choice between *SSB* and a benefit increased by $\$X$ in exchange for paying $\$LS$ to Social Security. EV-Buy provided a choice between receiving a lower monthly benefit ($SSB-X$) and paying a lump sum to maintain the existing benefit. Note that in these Buy scenarios, one can obtain CV simply by adding $\$X$ to each of the

EV scenarios. Again, it is worth noting that no *status quo* option was available in the EV case.

In order to converge on the subjective valuation resulting from any given measure above, the survey used a “branching” approach. For example, we started with a \$100 increment to the monthly annuity versus a \$20,000 lump sum. Then, based on the individual’s response, we either increased or decreased the amount of the lump-sum payment. By walking each respondent through a multi-stage branching process, we converged on a small range of lump-sum values that approximated the respondent’s implied subjective valuation of the change in the annuity stream.

As part of our experimental design, we chose one of our four approaches as a benchmark, permitting us to do additional sensitivity tests along other dimensions. While there is no theoretical basis for suggesting that one treatment would be preferred to the other three, we selected the CV-Sell option because it is more relevant to policy discussions. For example, offering retirees an opportunity to sell their annuities for lump sums is a transaction observed in recent years (e.g., GM offered retirees lump sums in lieu of their life annuities). The Sell measure is also less likely than the Buy measure to be bounded by people’s access to liquidity. Accordingly, all respondents were asked the CV-Sell question in one of the two waves, whereas the other three versions (CV-Buy, EV-Sell, and EV-Buy) were asked in a randomized order in the other wave. The placement of the CV-Sell question in the first or second wave was randomized across respondents. The two surveys were administered approximately two weeks apart. Below, we test whether responses to CV-Sell differ depending on whether this question was in the first or second wave.

Two other studies have employed a branching approach to elicit annuity valuations, although they were much more limited in focus.¹⁰ Cappelletti, Guazzarotti, and Tommasino (2013) used a national survey of Italian households in 2008 to ask whether people would give up half their monthly pension income (assumed to be €1000) in exchange for a lump sum of €60,000, paid immediately. Depending on their responses, individuals were branched to higher or lower lump-sum amounts. Liebman and Luttmer (2011) conducted a 2008 survey on the perceived labor supply

¹⁰ In addition to the two studies discussed in the text, a subset of the present authors has previously sought to elicit subjective annuity valuations in two HRS experimental modules. Errors in questions or response coding unfortunately interfered with analysis. One module in the 2004 HRS asked respondents their willingness to trade \$500 of a hypothetical \$1000 monthly Social Security benefit for a lump sum. Although the lump-sum amount offered to unmarried individuals was approximately actuarially fair, the amount offered to married couples was too low. A second module was fielded in the 2008 HRS, but coding instructions provided to field interviewers led to an inability to distinguish answers at the two extremes, i.e., those who placed a zero value on an annuity versus those who placed a very high value on annuities.

incentives in Social Security, and they included in their survey a question asking people for the equivalent variation of a \$100/month increase in their Social Security benefits (this is equivalent to our EV-Sell question.) Because each of those papers used only a single elicitation method, neither addresses the hypotheses tested in this paper.

D. Other Sources of Experimental Variation

We also randomized along a number of other dimensions for two reasons. First, we randomized the order of the options within a question, to test whether respondents were taking the survey seriously (as opposed to, say, always choosing option A). Second, we tested for anchoring effects in our benchmark question (CV-Sell) as well as whether responses varied with the magnitude of the change in the benefit, to provide a further assessment of the role of cognitive limitations. Furthermore, we asked a version of the questions designed to control for political risk, to ensure that our results were not driven by concern over the system's pending insolvency. Each of these factors is discussed in detail after we present our main results.

III. Evaluating Heterogeneity in Annuity Valuations

Figure 1 reports the cumulative distribution function (CDF) of the responses to the CV-Sell and CV-Buy questions, while Figure 2 provides a similar plot for EV-Sell and EV-Buy. Given our bracketing of responses, the two Figures plot both the upper and lower bounds for each respondent's annuity valuation.¹¹ In Figure 1, the midpoint of the upper and lower bounds for the CV-Sell question indicates a valuation of \$13,750 for a \$100-per-month change in Social Security benefits. The CV-Buy question midpoint valuation is only \$3,000. In Figure 2, the comparable valuations are \$12,500 for EV-Sell and \$3,000 for EV-Buy. By comparison, the median *actuarial value* of this annuity for respondents in our sample is \$16,855 (computed using Social Security's intermediate assumptions of a three percent interest rate and intermediate mortality).

Figures 1 and 2 here

Four patterns are evident in these two figures. First, median valuations are all substantially below the actuarial value. Second, there is substantial dispersion of valuations in all four valuation approaches. Third, the distributions of EV and CV valuations appear similar, holding fixed

¹¹ The CV-Sell figure plots valuations only for individuals who saw the \$100 increment first (the other three annuity valuation questions are only asked for \$100 increments). Other respondents saw higher annuity amounts first and, as we discuss below, this anchoring effect led to a higher valuation in their response to CV-Sell.

whether it is a Buy or Sell valuation, although we will see below that the correlation is far from perfect. Fourth, there is a very large difference between the Sell and Buy valuations, regardless of whether it is elicited in a CV or an EV setting. After briefly discussing each of these issues, we will then delve deeper into analyzing differences in valuations within and across individuals.

A. Median Valuations

If one simply pools the responses to our four valuation questions – CV-Sell, CV-Buy, EV-Sell and EV-Buy – 70% of the responses have an upper bound below the actuarially fair level and 64% of the responses have an upper bound that is at least \$5,000 below the actuarially fair level.¹² This finding is interesting given the ongoing discussion in the literature about the “annuity puzzle,” i.e., the fact that life-cycle models typically place a high utility value on annuitization, whereas most consumers appear to avoid purchasing them. Yet there are many reasons – some rational (e.g., bequest motives, a desire for liquidity), some behavioral (e.g., present-biased preferences) – for why people might value an annuity below its actuarially fair level and it is not the purpose of this paper to explain the annuity puzzle. Rather, our goal is to test whether valuations are consistent with optimizing behavior. We do this not by comparing elicited values with theoretical values, but rather by examining the stability and consistency of stated preferences. The remainder of our results should be viewed in light of this important distinction.

B. Dispersion of Valuations

The CDFs in Figures 1 and 2 reveal substantial heterogeneity in respondent valuations. For example, five percent of the sample reports upper-bound CV-Buy valuations of \$1,500 or less, levels so low that they are difficult to explain using any “rational” economic model, since the \$100 monthly annuity payments would provide more than this amount in only 16 months. The exception would be if some individual were virtually certain that he would die in that period, but we find that these outliers persist even when we control for respondents’ self-reported health status and survival probabilities. At the other extreme, 16 percent of the respondents reported lower-bound CV-Sell annuity values of \$60,000 or higher – nearly four times the actuarial value of the annuity. Moreover, over six percent of the respondents in the CV-Sell approach said they would not accept a lump sum of less than \$200,000. This is unexpected, since even if someone earned only a 60

¹²As in the figure, we limit the sample for the CV-Sell response to individuals who saw the \$100 increment first to avoid anchoring effects. If we double the weight on the remaining half of the CV-Sell responses (to compensate for the fact we dropped CV-Sell responses affected by anchoring), the percentages become 68% and 61%, respectively.

basis-point (0.60%) annual return on the \$200,000 lump sum, he could replace the \$100 per month he was giving up with this return *and still keep the lump sum* of \$200,000. As discussed in more detail below, these findings are not explained by subjective life expectancy, concerns about political risk, or other plausible “rational” explanations.¹³ In other words, many respondents appear to be having difficulty providing economically meaningful values for the Social Security annuity, at least in the tails of the CDF.

C. Comparing CV and EV

As noted above, we obtain the EV-Sell questions by simply adding \$100 to both of the options in the CV-Sell questions. Given the small magnitude of the shift (relative to mean estimated monthly benefits of \$1,395), we expect that a fully rational decision maker would provide quite similar assessments across these two ways of eliciting value. Although the distributions of CV-Sell and EV-Sell look similar in Figures 1 and 2 (as do the distributions of CV-Buy and EV-Buy), individual responses are not as correlated as one might infer. Table 2 reports the correlations across the four different measures.¹⁴ In column 1, we see that although CV-Sell and EV-Sell are significantly positively correlated, the correlation of +0.31 is far from perfect. It is notable that we asked the CV-Sell and the EV-Sell questions in different survey waves; thus each individual answered the two questions at least two weeks apart. Given this lag, it is unlikely that the correlation is driven by anchoring or memory effects that could arise if the questions had been asked within the same questionnaire. At +0.72, the correlation between CV-Buy and EV-Buy is substantially higher than the correlation between CV-Sell and EV-Sell.

¹³ We control for political risk in two ways in this study. First, we asked individuals about their confidence that the Social Security system will be able to provide them with the level of future benefits they are supposed to get under current law. Including responses to this question as a control variable in various analyses does not substantially affect our findings. Second, we asked a version of our CV-Sell annuity valuation question in which we explicitly instructed individuals not to consider political risk by stating: “From now on, please assume that you are absolutely certain that Social Security will make payments as promised, and that there is no chance at all of any benefit changes in the future other than the trade-offs discussed in the question below.” Using the most unbiased comparison available (i.e., comparing the response to the no-political-risk question to the baseline CV-Sell question for those for whom the two questions were asked in different waves of the survey), we find that the response to the no-political-risk question is a statistically insignificant 10 percent *lower* than the response to the baseline CV-Sell question. Taken literally, this implies a negative risk premium. We believe, however, the more likely explanation is that our question may have had the unintended effect of making political risk more salient, rather than less. Overall, our analysis suggests that the incorporation of political risk does not alter our main findings.

¹⁴ To control for correlations induced by common experimental manipulations, we regress the log midpoint value of the response on controls for the relevant manipulations and then correlate the residuals, which are reported in Table 2. Uncorrected correlations are similar and shown in Online Appendix Table A.2.

However, given that CV-Buy and EV-Buy were asked in the same wave, we cannot rule out that anchoring effects may have contributed to this higher correlation.

Table 2 here

D. Sell Prices Exceed Buy Prices

The most striking pattern that emerges from Figures 1 and 2 is that the distributions of annuity valuations from the Buy solicitations are substantially below those of the Sell solicitations. Recall that the Sell question asked how much a person would have to be compensated to *give up* part of his Social Security annuity, whereas the Buy question asked how much he would be *willing to pay* to increase his Social Security annuity. In Figure 1, the median midpoint response drops from \$13,750 for CV-Sell to only \$3,000 for CV-Buy, and responses at other points on the distribution drop as well.

If we observed this result only in the CV case, one might argue that this could result from *status quo* bias (Samuelson and Zeckhauser 1988) or an endowment effect (Kahneman, Knetsch, and Thaler 1991). Yet Figure 2 shows that an almost identical shift occurs when we use the EV-Sell and EV-Buy responses, where the *status quo* is not an option because both the annuity and the lump sum are represented as deviations away from the initial endowment.

To rule out the possibility that answers might be driven by liquidity constraints, we also asked respondents about their ability to come up with the money needed for the lump sum. The vast majority (91 percent) indicated that their choice was *not* due to liquidity constraints,¹⁵ and the clear divergence in valuations persists even when we focus on the non-liquidity constrained subsample.

Rather than *status quo* bias, endowment effects, or liquidity constraints, we conjecture that this wedge is the outcome of valuation difficulties on the part of respondents. This conjecture has two testable implications. First, individuals who have difficulty valuing annuities may seek to

¹⁵ Specifically, we asked whether each respondent could come up with \$5,000 “if you had to” and, separately, whether he could come up with the lump sum needed to purchase the higher annuity. The time frame for accessing the money was the same time frame as in the annuity valuation question, namely one year from now or the respondent’s expected claim date, whichever was later. About two-thirds of the respondents answered that they were certain they could come up with \$5,000, and over 90 percent responded that they could come up with the amount probably or certainly. About 82 percent of respondents indicated that they could come up with the lowest lump-sum amount that they declined to pay. Of the 18 percent that indicated they could not come up with this amount, half said that even if they had the money, they would decline to pay the lump sum. Thus, for 91 percent of the respondents, liquidity constraints were not the reason for the low reported annuity valuation in the CV-Buy trade-off question.

protect themselves by agreeing to an annuity transaction only if the annuity is priced very attractively, which would lead them to demand a high price to sell, but offer a low price to buy. We refer to this as a “reluctance to exchange,” which would imply that buy valuations will be *negatively* correlated with sell valuations. Second, it implies that the size of the wedge between buy and sell valuations will be decreasing with cognitive abilities.

E. Negative Buy-Sell Correlations

Although Figures 1 and 2 reveal large differences in the distributions of responses between Sell and Buy valuations, they do not reveal whether *within-person* responses to these alternative valuation measures are correlated. Hence we cannot yet conclude that the entire distribution is shifting to the left, or whether the same individuals are also changing their positions in the distribution depending on whether they see a Sell or Buy measure. This is addressed in Table 2, where we find a striking negative correlation between Buy and Sell valuations. Specifically, the correlation between CV-Sell and CV-Buy is -0.11, whereas the correlation between EV-Sell and EV-Buy is -0.15; both are highly statistically significant. These negative correlations suggest that individuals who place above-average values on a \$100 monthly annuity when asked to sell it tend to be willing to pay less than average to purchase a 100/month annuity flow. The negative correlation also suggests substantial movement within the distributions, rather than just a downward shift for everyone when we move from a Sell to a Buy elicitation method. This pattern is consistent with individuals who have difficulty valuing annuities being reluctant to exchange annuities because they wish to ensure that they are not losers in a transaction they have difficulty evaluating.

F. The Role of Cognition and the Sell-Buy Spread

If the Sell-Buy differential is the result of differences in decision-making abilities, then the size of the spread should be correlated with measures of cognition. To explore this, we construct a measure of the “Spread” that is equal to the absolute value of the difference between the log CV-Sell and the log CV-Buy valuation of a \$100 change in monthly Social Security benefits. The spread is defined as an absolute value because a discrepancy between Sell and Buy valuations in either direction is indicative of a lack of optimizing behavior. In practice, the spread is dominated by the 80 percent of individuals who place a higher value on CV-Sell than on CV-Buy. Because the spread is measured as the difference in log points, this variable reflects the relative value of Sell and Buy solicitations. The mean of the Spread variable is 2.58 and its median is 2.30,

indicating that the median individual reports Sell and Buy valuations that differ by a factor of 10. In the large majority that places a higher value on the Sell than the Buy valuation, the mean and median of the Spread variable are respectively 2.63 and 2.35, which indicates that the median person in this sample demands a price to sell a \$100 per month annuity that is 10.5 times higher than the price that he is willing to pay to purchase this same annuity.

In Figure 3, we show the bivariate relation between the CV Sell-Buy Spread and various measures of cognition. Panel A groups respondents according to an index of financial literacy, measured as the sum of correct answers to the three questions devised for the Health and Retirement Study and used in the ALP to rate respondents' financial literacy levels.¹⁶ Consistent with our hypothesis that the discrepancy between Sell and Buy is driven by differences in decision-making ability, we find that the wedge between the responses falls sharply with measured increases in financial literacy. In Panel B, we make use of the scores on a number scoring test, where respondents are shown incomplete sequences of numbers and are asked to complete the sequence (e.g.: 7, 8, ..., 10). Based on six such sequences and a RASCH scoring algorithm (Linacre 1999), each respondent is assigned a score. We find a sharp and monotonic decline in the Sell-Buy Spread across the quintiles of this numeracy measure. In Panel C, we split the Spread measure by level of education and once again we find a sharply declining pattern. Panel D uses an overall cognition index, which is the first principal component of the three measures of cognition.¹⁷ Not surprisingly, given the patterns in the earlier panels and the fact that the weight on each of the three factors is roughly equal, we find a strong, monotonic, negative relation between the Sell-Buy Spread and our cognition index.¹⁸

Figure 3 here

We repeat this analysis in a regression framework in Table 3, along with controls for other factors. In column 1, we regress the Spread against age dummies and controls for experimental

¹⁶ The three questions test for an understanding of inflation, compound interest, and risk diversification (for a complete listing of the questions see Lusardi and Mitchell 2014).

¹⁷ Although we use principal components analysis to construct the weights in the cognition index, the contributions of each of the three components turn out to be very close to equal. Thus, we obtain very similar results if we construct an index in which we simply give each of the three components equal weight.

¹⁸ Online Appendix Figure A.1 shows that findings of Figure 3 (monotonically declining spreads in each of the three measures of cognition as well as in the index) also hold when the Sell-Buy spread is based on EV valuations rather than CV valuations. Online Appendix Figure A.2 shows that we obtain declining spreads (but not always monotonically so) when we use the spread between CV and EV valuations (both using the Sell condition).

variations (to be discussed in the next section). We note that the coefficient on age 65+ is positive and significant; these older individuals have a 44 percentage point larger absolute difference between their Sell and Buy valuations than the reference age category (age 34 or younger) and the difference remains large (39 percentage points) and significant if we compare them to the age 50-64 category. This finding is important for two reasons. First, it addresses concerns that younger individuals may be less attuned to our decision context because they are further removed from making the comparable real-world decision and also because they may be less likely to be familiar with how Social Security benefits are paid as monthly income for life. Yet we find the opposite: younger individuals have more consistent valuations than do older individuals. Second, this evidence fits with our hypothesis regarding cognitive functioning. Agarwal, Driscoll, Gabaix, and Laibson (2009) show that cognitive functioning declines at older ages and may contribute to a decline in financial decision making. As we will see below, adding direct controls for cognition strengthens, rather than mitigates, this relation.

Table 3 here

In column 2, we add in our three direct measures of cognition. All three of them – financial literacy, education, and numeracy – are significantly negatively correlated with the Sell-Buy Spread. Each additional correct answer on the financial literacy questions reduces the spread by 0.32 log points. Moving up one education category reduces the spread by 0.24 log points. A one standard deviation increase in the score on the number series questions reduces the spread by 0.31 log points. In column 3, we combine these measures into our cognition index, and we find a strongly significant relation. A one standard deviation increase in cognition corresponds to a 0.59 log point reduction in the Sell-Buy Spread. In column 4, we add additional controls for demographics and preferences, including sex, marital status, race and ethnicity, family income, home ownership, self-reported health, children, risk aversion, return expectations, and political risk, among others.¹⁹ Even with this rich set of additional controls, the coefficient on the cognition index remains highly significant, with a coefficient of -0.42. The results of Table 3 are similar if we use the Sell-Buy Spread based on EV valuations rather than on CV valuations (see Online Appendix Table A.4). If we take the spread between CV-Sell and EV-Sell, which have similar values on average because they are both Sell measures, we continue to find significant negative

¹⁹ Definitions and summary statistics of control variables are given in Online Appendix Table A.1. The coefficients on the control variables can be found in Column 1 of Online Appendix Table A.3.

effects of the cognition index on the spread, though the magnitude of the coefficient drops by half (see Online Appendix Table A.5). This indicates that individuals with higher values of the cognition index give more internally consistent answers even in settings where the “reluctance-to-exchange” motive should not be prominent.²⁰

Thus far, we have shown that an important fraction of the population has implied annuity values that are difficult to reconcile with optimizing behavior under *any* plausible set of parameters. We have also documented a large divergence between the price at which individuals are willing to *buy* an annuity and the price at which they are willing to *sell* an annuity, and we showed that this cannot be explained by liquidity constraints or endowment effects. Moreover, buy and sell valuations are *negatively* correlated, and the size of the sell-buy disparity is highly correlated with cognitive ability, as measured by education, financial literacy, and numeracy. The next section extends our analysis in several additional directions.

IV. Robustness and Further Results

A. Are the Responses Meaningful?

In view of the implausible values in the tails of the distributions, the negative correlation across Sell and Buy valuations, and the large Sell-Buy Spread, some might surmise that a subset of respondents might not have taken the survey seriously (or perhaps did not understand it).²¹ Nevertheless, we have already shown that there is information contained in the elicited valuations: respondents provide consistent responses to similarly constructed offers (e.g., CV-Sell and EV-Sell) despite being asked in different waves two weeks apart. Additionally, as part of our experimental design, we included two additional sources of variation solely designed to test whether responses were meaningful. First, we randomized the order of the scenarios to which

²⁰ In Online Appendix Figure A.3 and Table A.6, we show that the negative correlation between Sell and Buy valuations decreases in absolute value for those with higher levels of cognition. In addition, the positive correlation between CV and EV valuations tends to increase with cognition. Hence, cognition also has effects in the expected direction on our correlations. However, because correlations cannot be measured at the individual level (but only for subsamples), these results have less statistical precision. We therefore placed them in the Online Appendix and present the effect of cognition on the Sell-Buy Spread given that the Sell-Buy Spread is measured at the individual level.

²¹ In the mechanism design literature on contingent valuation, concerns of this type are often referred to as being about whether the choices are consequential. The concern is that if respondents do not believe their survey responses are consequential, they may not dedicate effort to the survey.

people were exposed.²² Second, we also randomized the order of the options *within* a question (i.e., whether the lump-sum increment was the first or the second response). If the order of the questions or the order of the options within the questions mattered, this would suggest that individuals had difficulty with the survey itself. Therefore we test for sensitivity to “asked in wave 1” and “lump-sum option shown first” at the same time we test for sensitivity to anchoring and starting values in the next sub-section. As we shall see, the evidence is consistent with respondents reading and understanding the questions.

B. Sensitivity to Anchoring and Starting Values

We also incorporated two sources of experimental variation designed to further test for the consistency of valuations with optimizing behavior. First, we varied the starting values of the size of the lump sum, randomizing across \$10,000, \$20,000 and \$30,000; we refer to this as “log of starting value.” Second, in the CV-Sell case, we varied the order of size of the increment of the monthly benefit. Specifically, we presented the CV-Sell version multiple times to each respondent for $X=\$100$, $X=\$500$, $X=\$SSB$ (i.e., the entire amount of the respondent’s Social Security benefits), and for a random X that was a multiple of \$100 (less than $\min(\$SSB-100, 2000)$, and not equal to 100 or 500). We also randomized whether we asked the CV-Sell with the X s arranged in increasing order or with the X s arranged in decreasing order. We control for this randomization in the regressions (i.e., whether people were shown values from small-to-large or large-to-small). We refer to this in our tables as “asked after larger version.”

All four of these randomizations (two used to test for meaningfulness of responses and two to test for consistency with optimizing behavior) were conducted independently. A simple correlation analysis (not detailed here) confirmed that this randomization was indeed done correctly, such that variation along each dimension was orthogonal to the variation along the other dimensions.

C. Results of these Extensions

If individuals have difficulty valuing annuities, then we would expect to find that people would be sensitive to irrelevant cues such as starting values and the whether asked after a larger version. Conversely, we do not necessarily expect that the order of the scenarios or the options

²² We first randomized at the individual level whether CV-Sell was asked in the first or second wave of our survey; CV-Buy, EV-Sell, and EV-Buy were asked in the other wave of the survey. Within the wave in which CV-Buy, EV-Sell, and EV-Buy were asked, we randomized their order over each of the six possible orderings.

would matter for valuation decisions, as long as the respondent tried to answer the questions. Our results appear in the first column of Table 4, where we regress the log midpoint of our baseline CV-Sell variable (using a \$100 variation in Social Security benefits) against the four variables capturing all sources of randomization.²³

Table 4 here

Several findings stand out. First, there is no evidence that individuals simply elected the first option shown (i.e., there is no effect of “Lump sum shown last”), giving some comfort that the respondents *did* take care in answering the survey questions. Relatedly, it does not matter whether the question was asked in the first or second wave (i.e., “Asked in wave 1” has a small and insignificant coefficient estimate). Second, there is bias with respect to both of the other measures, as would be expected if individuals had difficulty valuing annuities. Specifically, the starting value had a statistically significant coefficient of +0.37. Because both the annuity valuation and the starting value are measured in logs, this means that increasing the first lump-sum amount shown by 10% raised respondents’ valuations by an average of around 3.7%. Furthermore, when the CV-Sell question was shown after a CV-Sell question with a larger change in Social Security benefits (so the order was large-to-small), respondents reported a 0.7 log-point higher average valuation of the annuity than if the baseline CV-Sell question was seen first.

We next re-run this regression on two respondent sub-samples: those in the top (col. 2) and those in the bottom quintile (col. 3) of the cognition index. We find the effect of the log of the starting value is insignificant for individuals with higher cognitive abilities, whereas it is substantially more important (the coefficient is 0.92) for those in the bottom quintile of cognition. This suggests that less cognitively capable people are more sensitive to anchoring effects. Interestingly, however, the effect of “asked after larger version” is significant and similar in magnitude across the cognition quintiles. Moreover, the “asked after larger version” effect is extremely persistent across any cut of the data by measures of cognition. Whether the lump-sum option was shown as the first or second option continues to be insignificant for all quintiles, although we do now find that whether the question was asked in wave 1 or wave 2 is significant in the lowest cognition quintile. That is, respondents in the bottom quintile appear to provide a

²³ We do this analysis on the CV-Sell version because only the CV-Sell version asks for different increment sizes of the Social Security amount. This means that we can randomize the order in which the increment sizes are shown only for the CV-Sell version.

significantly higher valuation in the second wave, though why this is so is not evident.

In column 4, we return to the full sample and include interactions between the cognition index and each of the four survey manipulations. Results tell a similar story: people are sensitive to starting values and even more so if they have lower cognitive abilities. “Asked after larger version” is highly significant, but is not mitigated by cognition. “Asked in wave 1” and “lump-sum option shown last” continue to be statistically insignificant.

D. Explaining Annuity Valuations

A key reason that life annuities play such an important role in life-cycle economic models is that they provide a cost-effective way to smooth consumption by insuring against longevity risk. Although numerous authors have calculated the welfare gains associated with annuitization, there is conflicting evidence on the extent to which real-world individuals actually value the insurance. Brown (2001) showed that a utility-based measure of annuity valuation was correlated with a binary measure of intended annuitization of asset balances. Büttler and Teppa (2007) documented similar findings in the Swiss system. Nevertheless, such measures explain only a very small fraction of the variation in the annuitization decision. Brown, Kling, Mullainathan, and Wrobel (2008) suggested that the ubiquitous framing of retirement planning in terms of wealth accumulation has conditioned individuals to ignore the insurance aspects of annuities and view them through an investment lens, consistent with individuals resorting to simplified decision-making heuristics in the face of complexity (Benartzi, Previtro, and Thaler 2011).

We hypothesize that when individuals confront trade-offs of the type presented in our survey, they find it difficult to sort through the lifetime utility implications and instead resort to thinking in simpler financial terms. To test this hypothesis, we regress annuity valuations against various determinants of annuity demand in our data. Column 1 of Table 5 regresses the average of CV-Sell and CV-Buy valuations against the actuarial value of the annuity offer presented (which varies by cohort, age at annuitization, and sex; it also assumes a real interest rate of three percent).²⁴ The actuarial value term has a coefficient of 1.02, suggesting that there is approximately a one-for-one correspondence between the annuity’s actuarial value and individuals’ subjective valuations of the annuity. Column 2 replaces the actuarial value with a utility-based measure.

²⁴ We use the CV versions because, unlike the EV versions, they were asked in different waves of the survey. We take the average of CV-Sell and CV-Buy because there is no a priori reason to consider one more credible than the other.

Following Brown (2001), this value is derived from a parameterized life-cycle model with variation coming from age at annuitization, mortality differences by cohort and sex, marital status (which determines whether it is a single or joint life optimization), risk aversion, current levels of non-annuitized wealth, current annuitized wealth, and interactions of these variables through the utility-maximizing model. We find that the coefficient on this theoretical, utility-based annuity value in column 2 is not significantly different from zero, though it is highly significantly different from one.²⁵ In columns 3 and 4, we repeat this analysis using even more control variables and we obtain very similar results.²⁶

Table 5 here

Overall, we view these results as being consistent with individuals using simpler financial decision rules (e.g., “How long will it take me to break even?”) rather than taking into account the more complex consumption-smoothing and insurance considerations. We also note that the R-squared values are very low: around 0.05 to 0.06 depending on the specification. These findings are consistent with prior studies (e.g., Brown 2008) which also found it difficult to account for the observed variation in annuitization decisions.

Table 6 reports coefficient estimates on the annuity valuations for respondents based on our cognition index. Here we provide both estimated coefficients and the root mean squared error (MSE) for each specification. Although the coefficients differ non-monotonically across the cognition quintiles, in all cases they are not significantly different from 1.²⁷ We do, however, find that the root MSE is monotonically declining with the level of cognition. Recalling that our dependent variable is in logs, these differences are economically meaningful. For example, the root mean squared distance from the regression line declines by 0.24 log-points when one moves

²⁵ In results not detailed here, when we include both the actuarial value and the utility-based measure, we continue to find that the coefficient on the actuarial value is approximately one and that the utility-based measure is insignificant.

²⁶ In Online Appendix Table A.7, we repeat the regressions in columns 1 and 3 of Table 5, but now we use each of our four separate valuation measures (CV-Sell, CV-Buy, EV-Sell and EV-Buy). In seven of the eight additional specifications, we cannot reject the hypothesis that the coefficient on the actuarial value is equal to one, but we can reject that it is equal to zero. In unreported results, we also find that the utility-based measure is not significant using these alternative dependent variables.

²⁷ We also run a single regression in which the cognition index is interacted with the actuarial value (including the same controls as in Table 6, the direct effect of the cognition index and the direct effect of actuarial value). We find that the interaction term has a coefficient estimate of 0.001 (s.e.: 0.146), which confirms that the effect of the actuarial value on the annuity valuation does not differ significantly by cognition.

from the bottom to the second cognition quintile and a decline of a full 0.60 log-points when one moves from the bottom to the top cognition quintile. In other words, decisions made by less cognitively capable individuals are substantially noisier than those made by the more cognitively able.

Table 6 here

E. Robustness to Outliers and to Age of Sample

To verify that possible outliers have no major effects on the results, we replicated Figure 3 and Tables 2 through 6 using only annuity valuation measures that have been winsorized at the 10th and 90th percentiles. The results for these winsorized variables are very similar (see Online Appendix Figure A.4 and Tables A.8 through A.12). We also examine the possibility that our results might be driven by a lack of interest in our questions by younger respondents, who may not have given retirement planning much thought. We therefore replicated Figure 3 and Tables 2 through 6 using only respondents age 50 and above. The results for this sample are very similar (see Online Appendix Figure A.5 and Tables A.13 through A.17).

V. Discussion and Conclusions

This paper provides support for the hypothesis that there is important cross-sectional variation in individuals' abilities to value a stream of life annuity payments. Furthermore, this variation is highly correlated with an index of cognition we built using measures of education, financial literacy, and numeracy. On average, we find that consumers tend to value annuities *less* when given the opportunity to *buy* more, but they value them *more* when given the opportunity to *sell* annuities in exchange for a lump sum. Because this finding holds even when no *status quo* option is available, this finding is not driven by standard *status quo* or endowment effects. Instead, our results are consistent with consumers having difficulty ascertaining the value of an annuity. As a result, many people are only willing to buy or sell an annuity when it appears to be an exceptionally good deal and this tendency is strongest among the least cognitively able. We also show that people are sensitive to framing and starting values, which provides further evidence against optimizing behavior. Finally, we show that the cross-sectional variation in subjective annuity valuations is correlated with the relatively simple actuarial value, but not with amore cognitively-challenging utility-based value, and that decisions made by less cognitively adept individuals are substantially noisier than those of the more cognitively able.

Our findings raise questions about whether consumers are able to make utility-maximizing choices when confronted with a decision about whether to buy longevity protection. Although our results do not speak directly to why average annuity values are so low (and thus do not speak directly to what the literature calls the “annuity puzzle”), our findings do imply that one should not necessarily infer from the observed lack of annuitization that optimizing agents prefer lump sums over annuities. For example, the fact that so few people annuitize their defined contribution pension balances when given the opportunity to do so should not be interpreted as conclusive revealed preference evidence that they do not value longevity protection.

In addition to advancing our intellectual understanding, our results also have considerable policy relevance. For example, in early 2014, the UK eliminated its mandatory annuitization requirement, thus leaving the decision of whether or not to annuitize retirement account balances in the hands of individuals (Donaldson and Hutton 2014). By contrast, in Singapore, the government now requires retirees to purchase life annuities sufficient to cover subsistence needs (Fong, Mitchell, and Koh 2011). In the U.S., policymakers are increasingly expressing interest in ways to further encourage annuitization in 401(k) and other defined contribution plans (Gale, Iwry, John, and Walker 2008). There is also an emerging debate in the U.S. about whether to encourage or discourage a particular form of “de-risking” of corporate defined benefit pension plans, by offering retirees a lump sum in lieu of lifelong pension benefits. In these and other instances, our results suggest that many individuals lack the cognitive ability to make appropriate judgments about how to manage their money in retirement. Accordingly, policymakers would do well to recognize this substantial heterogeneity in people’s optimization capabilities when individuals make crucially important financial decisions.

References

- Agarwal, Sumit, John Driscoll, Xavier Gabaix, and David Laibson. 2009. "The Age of Reason: Financial Decisions over the Life-Cycle and Implications for Regulation" *Brookings Papers on Economic Activity*, (2): 51–117.
- Agarwal, Sumit, and Bhashkar Mazumder. 2013. "Cognitive Abilities and Household Financial Decision Making." *American Economic Journal: Applied Economics*. 5(1): 193–207.
- Agnew, Julie R., Lisa R. Anderson, Jeffrey R. Gerlach, and Lisa R. Szykman. 2008. "Who Chooses Annuities? An Experimental Investigation of the Role of Gender, Framing and Defaults." *American Economic Review*. 98(2): 418–422.
- Almenberg, Johann, and Anna Dreber. 2011. "Gender, Financial Literacy, and Stock Market Participation." Stockholm School of Economics Working Paper.
- Arrondel, Luc, Majdi Debbich, and Frédérique Savignac. 2013. "Financial Literacy and Planning in France." *Numeracy*. 6(2): 1–17.
- Banks, James, Cormac O’Dea, and Zoe Oldfield. 2010. "Cognitive Function, Numeracy and Retirement Saving Trajectories." *The Economic Journal*. 120(8): F381–F410.
- Behrman, Jere, Olivia S. Mitchell, Cindy Soo, and David Bravo. 2012. "Financial Literacy, Schooling, and Wealth Accumulation." *American Economic Review*. 102(3): 300–304.
- Benartzi, Shlomo, Alessandro Previtro, and Richard H. Thaler. 2011. "Annuitization Puzzles." *Journal of Economic Perspectives*. 25(4): 143–64.
- Bernheim, B. Douglas. 2002. "Taxation and Saving." In *Handbook of Public Economics*. Alan J. Auerbach and Martin Feldstein, Eds. New York: Elsevier: 1173–1249.
- Beshears, John, James J. Choi, David Laibson, and Brigitte C. Madrian. 2008. "How Are Preferences Revealed?" *Journal of Public Economics*. 92(8-9): 1787–1794.
- Beshears, John, James J. Choi, David Laibson, Brigitte C. Madrian, and Stephen P. Zeldes. 2014. "What Makes Annuitization More Appealing?" *Journal of Public Economics*. 116: 2–16.
- Brown, Jeffrey R. 2001. "Private Pensions, Mortality Risk, and the Decision to Annuitize." *Journal of Public Economics*. 82(1): 29–62.
- Brown, Jeffrey R. 2008. "Understanding the Role of Annuities in Retirement Planning." In *Overcoming the Savings Slump: How to Increase the Effectiveness of Financial Education and Saving Programs*. Annamaria Lusardi, Ed. Chicago: University of Chicago Press: 178–206.
- Brown, Jeffrey R., Arie Kapteyn, and Olivia S. Mitchell. 2013. "Framing and Claiming: How Information-Framing Affects Expected Social Security Claiming Behavior." *Journal of Risk and Insurance*. Forthcoming.
- Brown, Jeffrey R., Jeffrey R. Kling, Sendhil Mullainathan, and Marian Wrobel. 2008. "Why Don't People

- Insure Late Life Consumption? A Framing Explanation of the Under-Annuitization Puzzle.” *American Economic Review*. 98(2): 304–309.
- Brown, Jeffrey R., Olivia S. Mitchell, and James M. Poterba. 2001. “The Role of Real Annuities and Indexed Bonds in an Individual Accounts Retirement Program.” In *Risk Aspects of Investment-Based Social Security Reform*. John Campbell and Martin Feldstein, Eds. Chicago: University of Chicago Press: 321–360.
- Brown, Jeffrey R., Olivia S. Mitchell, and James M. Poterba. 2002. “Mortality Risk, Inflation Risk, and Annuity Products.” In *Innovations in Retirement Financing*. Olivia S. Mitchell, Zvi Bodie, Brett Hammond, and Steve Zeldes, Eds. Philadelphia: University of Pennsylvania Press: 175–197.
- Brown, Jeffrey R., and James M. Poterba. 2000. “Joint Life Annuities and the Demand for Annuities by Married Couples.” *Journal of Risk and Insurance*. 67(4): 527–553.
- Bütler, Monika, and Federica Teppa. 2007. “The Choice between an Annuity and a Lump Sum: Results from Swiss Pension Funds.” *Journal of Public Economics*. 91(10): 1944–1966.
- Calvet, Laurent, John Campbell, and Paolo Sodini. 2007. “Down or Out: Assessing the Welfare Costs of Household Investment Mistakes.” *Journal of Political Economy*. 115(5): 707–747.
- Calvet, Laurent, John Campbell, and Paolo Sodini. 2009. “Measuring the Financial Sophistication of Households.” *American Economic Review*. 99(2): 393–398.
- Cappelletti, Giuseppe, Giovanni Guazzarotti, and Pietro Tommasino. 2013. “What Determines Annuity Demand at Retirement?” *The Geneva Papers*. 38: 777–802.
- Chalmers, John, and Jonathan Reuter. 2012. “How Do Retirees Value Life Annuities? Evidence From Public Employees.” *Review of Financial Studies*. 25(8): 2601–2634.
- Chai, Jingjing, Wolfram Horneff, Raimond Maurer, and Olivia S. Mitchell. 2011. “Optimal Portfolio Choice over the Life Cycle with Flexible Work, Endogenous Retirement, and Lifetime Payouts.” *Review of Finance*. 15(4): 875–907.
- Choi, Syngjoo, Shachar Kariv, Wieland Müller, and Dan Silverman. 2014. “Who is (More) Rational?” *American Economic Review*. 104(6): 1518–1550.
- Christelis, Dimitris, Tullio Jappelli, and Mario Padula. 2010. “Cognitive Abilities and Portfolio Choice.” *European Economic Review*. 54 (1): 18–38.
- Davidoff, Thomas, Jeffrey R. Brown, and Peter A. Diamond. 2005. “Annuities and Individual Welfare.” *American Economic Review*. 95(5): 1573–1590.
- Donaldson, Kitty and Robert Hutton. 2014. “Osborne Scraps Annuity Rule in U.K. Budget Aimed at Savers.” Bloomberg.com, March 19. <http://www.bloomberg.com/news/2014-03-19/osborne-scraps-annuity-rule-in-u-k-budget-aimed-at-savers.html>
- Dushi, Irena, and Anthony Webb. 2006. “Rethinking the Sources of Adverse Selection in the Annuity

- Market.” In *Competitive Failures in Insurance Markets: Theory and Policy Implications*. Pierre-André Chiappori and Christian Gollier, Eds. Cambridge: MIT Press: 185–212.
- Fang, Hanming, Michael P. Keane, and Dan Silverman 2008. “Sources of Advantageous Selection: Evidence from the Medigap Insurance Market.” *Journal of Political Economy*. 226(2): 303–350.
- Fong, Joelle H.Y., Olivia S. Mitchell, and Benedict S. K. Koh. 2011. “Longevity Risk Management in Singapore’s National Pension System.” *Journal of Risk and Insurance*. 78(4): 961–981.
- Gale, William G., J. Mark Iwry, David C. John, and Lina Walker. 2008. *Increasing Annuitization of 401(k) Plans with Automatic Trial Income*. Retirement Security Project Report. Washington, D.C.: Brookings Institution.
- Greenwald, Mathew, Arie Kapteyn, Olivia S. Mitchell, and Lisa Schneider. 2010. “What Do People Know about Social Security?” Financial Literacy Consortium Report to the SSA, September.
- Hilgert, Marianne, Jeanne Hogarth, and Sondra Beverly. 2003. “Household Financial Management: The Connection between Knowledge and Behavior.” *Federal Reserve Bulletin*: 309–322.
- Horneff, Wolfram J., Raimond H. Maurer, Olivia S. Mitchell, and Ivica Dus. 2008. “Following the Rules: Integrating Asset Allocation and Annuitization in Retirement Portfolios.” *Insurance: Mathematics and Economics*. 42(1): 396–408.
- Horneff, Wolfram J., Raimond H. Maurer, Olivia S. Mitchell, and Michael Z. Stamos. 2009. “Asset Allocation and Location over the Life Cycle with Survival-Contingent Payouts.” *Journal of Banking and Finance*. 33(9): 1688–1699.
- Horneff, Wolfram J., Raimond H. Maurer, Olivia S. Mitchell, and Michael Z. Stamos. 2010. “Variable Payout Annuities and Dynamic Portfolio Choice in Retirement.” *Journal of Pension Economics and Finance*. 9(2): 163–183.
- Hurd, Michael, and Stan Panis. 2006. “The Choice to Cash out Pension Rights at Job Change or Retirement.” *Journal of Public Economics*. 90(12): 2213–2227.
- Inkmann, Joachim, Paula Lopes, and Alexander Michaelides. 2011. “How Deep is the Annuity Market Participation Puzzle?” *Review of Financial Studies*. 24(1): 279–319.
- Kahneman, Daniel, Jack L. Knetsch, and Richard H. Thaler. 1991. “Anomalies: The Endowment Effect, Loss Aversion, and Status Quo Bias.” *Journal of Economic Perspectives*. 5(1): 193–206.
- Kilgour, John G. 2006. “Public Sector Pension Plans: Defined Benefit Versus Defined Contribution.” *Compensation Benefits Review*. 38(1): 20–28.
- Kling, Catherine L., Daniel J. Phaneuf, and Jinhua Zhao. 2012. “From Exxon to BP: Has Some Number Become Better than No Number?” *Journal of Economic Perspectives*. 26(4): 3–26.
- Kimball, Miles, and Tyler Shumway. 2006. “Investor Sophistication and the Participation, Home Bias, Diversification, and Employer Stock Puzzles.” University of Michigan Working Paper.

- Koijen, Ralph S.J., Theo E. Nijman, and Bas J.M. Werker. 2011. "Optimal Annuity Risk Management." *Review of Finance*. 15(4): 799–833.
- Kotlikoff, Laurence J., and Avia Spivak. 1981. "The Family as an Incomplete Annuities Market." *Journal of Political Economy*. 89(2): 372–391.
- Liebman, Jeffrey B., and Erzo F.P. Luttmer. 2011. "The Perception of Social Security Incentives for Labor Supply and Retirement: The Median Voter Knows More Than You'd Think." Working Paper, Dartmouth College.
- Linacre, J. M. (1999). "Understanding Rasch Measurement: Estimation Methods for Rasch Measures." *Journal of Outcome Measurement*. 3(4): 382–405.
- Lockwood, Lee. 2011. "Bequest Motives and the Annuity Puzzle." *Review of Economic Dynamics*. 15(2):226–243.
- Lusardi, Annamaria, and Olivia S. Mitchell. 2007. "Baby Boomer Retirement Security: The Roles of Planning, Financial Literacy, and Housing Wealth." *Journal of Monetary Economics*. 54(1): 205–224.
- Lusardi, Annamaria, and Olivia S. Mitchell, Eds. 2011. *Financial Literacy: Implications for Retirement Security and the Financial Marketplace*. Oxford, Oxford University Press.
- Lusardi, Annamaria, and Olivia S. Mitchell. 2014. "The Economic Importance of Financial Literacy: Theory and Evidence." *Journal of Economic Literature*, 52(1): 5-44.
- Luttmer, Erzo F.P., and Andrew Samwick. 2012. "The Welfare Cost of Perceived Policy Uncertainty: Evidence from Social Security." Unpublished Manuscript. Dartmouth College.
- Maurer, Raimond, Olivia S. Mitchell, Ralph Rogalla, and Vasily Kartashov. 2013. "Lifecycle Portfolio Choice with Stochastic and Systematic Longevity Risk, and Variable Investment-Linked Deferred Annuities." *Journal of Risk and Insurance*. 80(3): 649–676.
- McArdle, John J., James P. Smith, and Robert Willis. 2011. "Cognition and Economic Outcomes in the Health and Retirement Survey." In *Explorations in the Economics of Aging*. David A. Wise, Ed. Chicago: University of Chicago Press: 209–233.
- Milevsky, Moshe, and Virginia R. Young. 2007. "Annuitization and Asset Allocation." *Journal of Economic Dynamics and Control*. 31(9): 3138–3177.
- Mitchell, Olivia S., James M. Poterba, Mark J. Warshawsky, and Jeffrey R. Brown. 1999. "New Evidence on the Money's Worth of Individual Annuities." *American Economic Review*. 89(5): 1299–1318.
- Mitchell, Olivia S., John Piggott, and Noriyuke Takayama, Eds. 2011. *Securing Lifelong Retirement Income*. Oxford: Oxford University Press.
- Novy-Marx, Robert, and Joshua Rauh. 2011. "The Liabilities and Risks of State-Sponsored Pension Plans." *Journal of Economic Perspectives*. 23(4): 191–210.

- Peijnenburg, Kim, Theo Nijman, and Bas J.M. Werker. 2010a. "Optimal Annuitization with Incomplete Annuity Markets and Background Risk during Retirement." Netspar Discussion Paper.
- Peijnenburg, Kim, Theo Nijman, and Bas J.M. Werker. 2010b. "Health Cost Risk and Optimal Retirement Provision: A Simple Rule for Annuity Demand." Wharton School Pension Research Council Working Paper No. WPS 2010-08.
- Poterba, James, Steven Venti, and David Wise. 2011. "The Composition and Drawdown of Wealth in Retirement." *Journal of Economic Perspectives*. 25(4): 95–118.
- Previtero, Alessandro. 2014. "Stock Market Returns and Annuitization." *Journal of Financial Economics* 113(2): 202–214.
- Reichling, Felix, and Kent Smetters. 2012. "Optimal Annuitization with Stochastic Mortality Probabilities." BYU Computational Public Economic Conference Presentation, December.
- Samuelson, William, and Richard J. Zeckhauser. 1988. "Status-quo Bias in Decision Making." *Journal of Risk and Uncertainty*. 1: 7–59.
- Shepard, Mark. 2011. "Social Security Claiming and the Life-Cycle Model." Unpublished Manuscript. Harvard University.
- Sinclair, Sven H., and Kent A. Smetters. 2004. "Health Shocks and the Demand for Annuities." Technical Paper Series; Congressional Budget Office. Washington, DC: GPO, July.
- Steverman, Ben. 2012. "Mark Iwry: Bringing Annuities to 401(k)s." Bloomberg.com <http://www.bloomberg.com/news/2012-04-17/mark-iwry-bringing-annuities-to-401-k-s.html>
- Turra, Cassio, and Olivia S. Mitchell. 2008. "The Impact of Health Status and Out-of-Pocket Medical Expenditures on Annuity Valuation." In *Recalibrating Retirement Spending and Saving*. John Ameriks and Olivia S. Mitchell, Eds. Oxford: Oxford University Press: 227–250.
- US Department of Labor (USDOL). 2010. "U.S. Labor Department Seeks Public Comments on Lifetime Income Options for Retirement Plans." February 2. <http://www.dol.gov/ebsa/newsroom/2010/09-1571-NAT.html>
- van Rooij, Maarten, Annamaria Lusardi, and Rob Alessie. 2011. "Financial Literacy and Stock Market Participation." *Journal of Financial Economics*. 101(2): 449–472.
- Wayland, Michael. 2012. "GM Pensions: 13, 200 White Collar Retirees Taking Buyouts Makes Sense, Analysts Say." Nov. 1. Mlive.com. http://www.mlive.com/auto/index.ssf/2012/11/gm_pensions_12600_white-collar.html.
- Yaari, Menahem. 1965. "Uncertain Lifetime, Life Insurance, and the Theory of the Consumer." *Review of Economic Studies*. 32(2): 137–150.

Figure 1: Cumulative Distribution Functions of CV-Sell and CV-Buy Annuity Valuations



Figure 2: Cumulative Distribution Functions of EV-Sell and EV-Buy Annuity Valuations

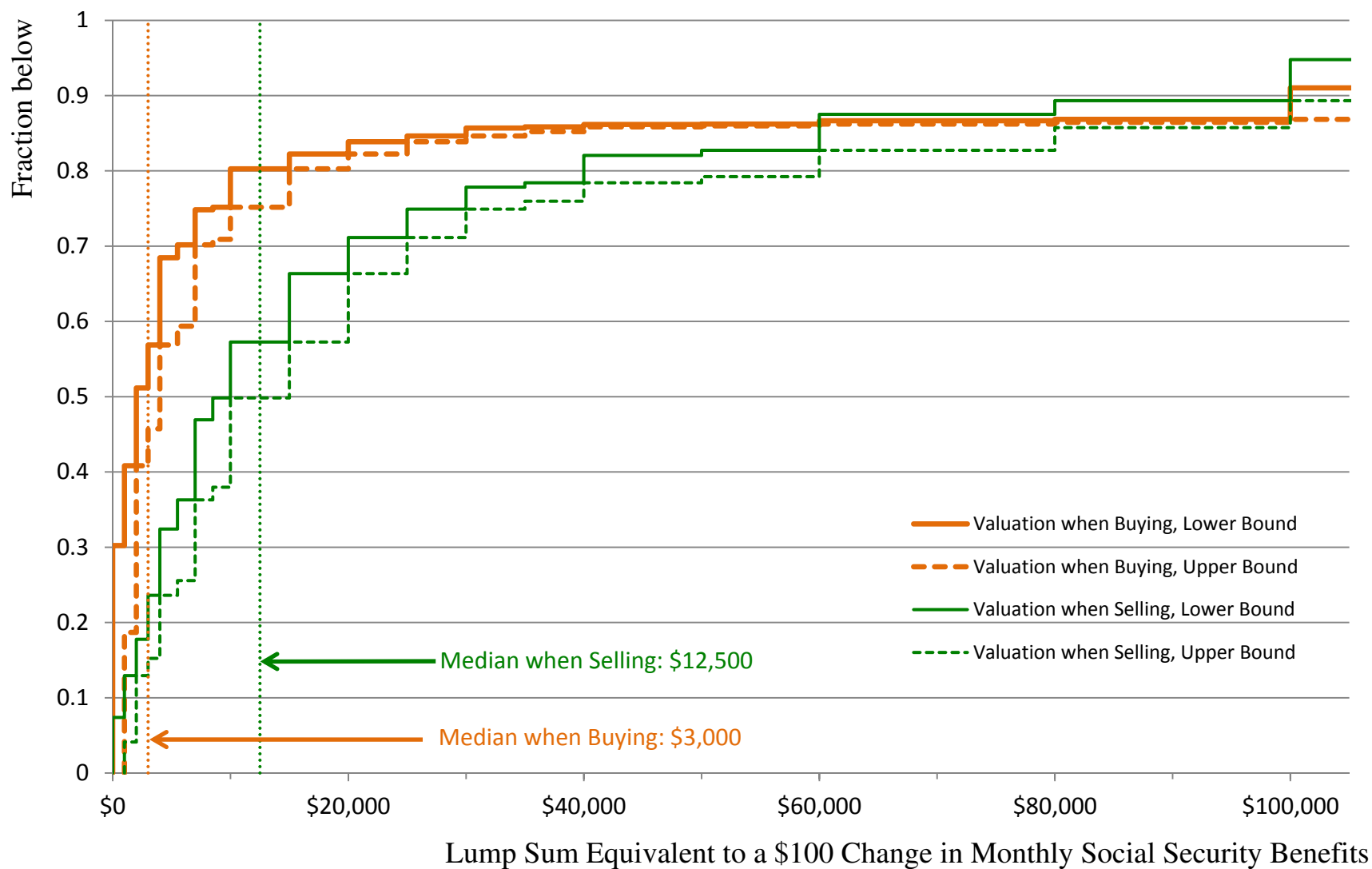
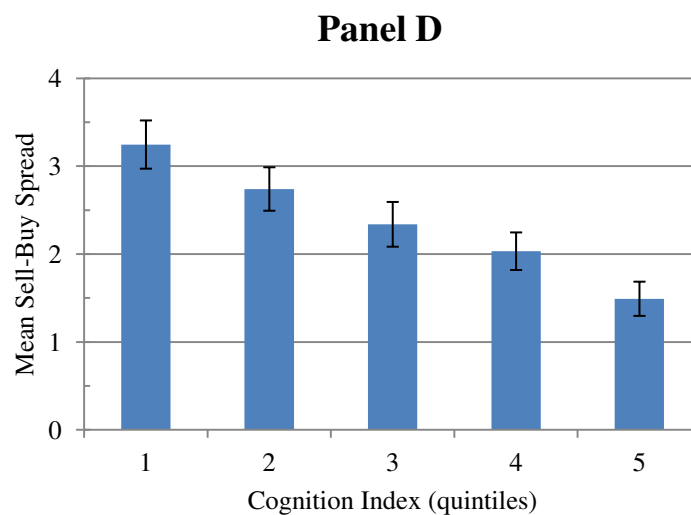
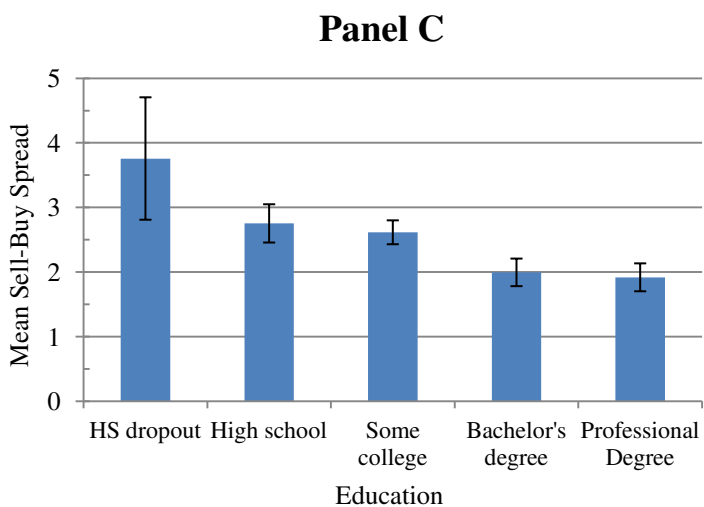
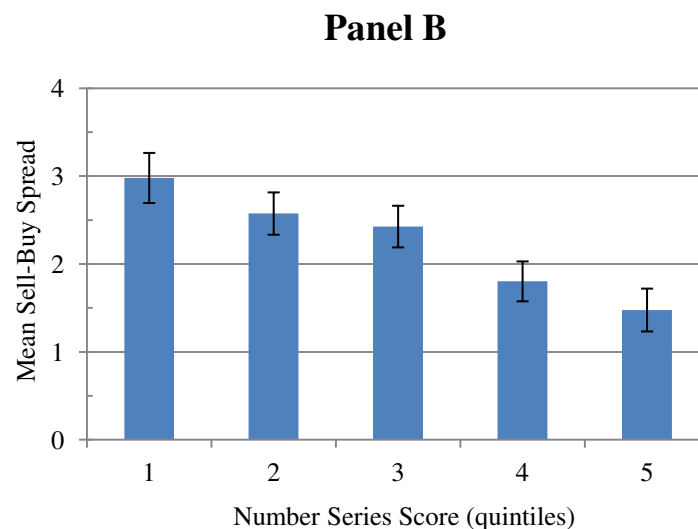
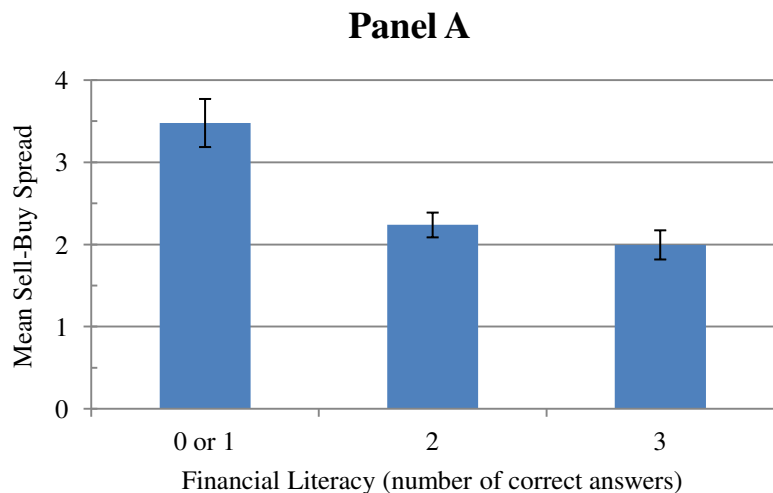


Figure 3: Sell-Buy Spread by Measures of Decision-Making Ability



Note: The whiskers represent 95% confidence intervals. The Sell-Buy Spread is measured as the absolute value of the difference between the log CV-Sell valuation and the log CV-Buy valuation of a \$100 change in monthly Social Security benefits. For the Financial Literacy measure, we grouped those with no correct answers with those with one correct answer because there are very few observations (4%) with no correct answers.

Table 1: Characteristics of the ALP Sample

	(1)	(2)	(3)
	ALP Sample Mean	CPS Mean	Difference ALP-CPS
Age	51.4	46.2	5.19***
Age: 18-34	0.16	0.31	-0.15***
Age: 35-49	0.25	0.27	-0.02
Age: 50-64	0.41	0.25	0.16***
Age: 65+	0.18	0.17	0.01
Female	0.58	0.51	0.07***
Married	0.60	0.54	0.07***
Race/Ethnicity			
Non-Hispanic White	0.80	0.68	0.12***
Non-Hispanic Black	0.08	0.12	-0.04***
Hispanic	0.09	0.14	-0.05***
Other Race/Ethnicity	0.03	0.07	-0.03***
Education			
High School Dropout	0.02	0.13	-0.11***
High School	0.16	0.30	-0.14***
Some College	0.37	0.28	0.09***
Bachelor's Degree	0.25	0.18	0.07***
Professional Degree	0.19	0.10	0.10***
Ln Family Income	10.89	10.77	0.13***
Family Income: Below 25k	0.18	0.24	-0.07***
Family Income: 25k-50k	0.27	0.24	0.04***
Family Income: 50k-75k	0.21	0.18	0.03***
Family Income: 75k-100k	0.13	0.12	0.01**
Family Income: Above 100k	0.21	0.23	-0.02*
Household size	2.68	3.00	-0.33***
Household size of one	0.22	0.14	0.08***
Household size of two	0.36	0.33	0.03***
Household size of three	0.15	0.19	-0.04***
Household size of four +	0.27	0.33	-0.06***
Region			
Northeast	0.17	0.18	-0.02*
Midwest	0.24	0.22	0.02**
South	0.35	0.37	-0.01
West	0.24	0.23	0.01
Observations	2,112	146,785	

Notes: * significant at 10%, ** significant at 5%, *** significant at 1%. In both the ALP and the CPS the sample is restricted to those aged 18 and older. The ALP sample was collected between June and August of 2011. The CPS data are from March 2011 and use CPS person weights; the ALP data are unweighted.

Table 2: Correlations between Annuity Valuation Measures

Pairwise correlations	CV-Sell (in logs)	EV-Sell (in logs)	CV-Buy (in logs)	EV-Buy (in logs)
CV-Sell (in logs)	1			
EV-Sell (in logs)	0.31***	1		
CV-Buy (in logs)	-0.11***	-0.17***	1	
EV-Buy (in logs)	-0.11***	-0.15***	0.72***	1

Notes: * significant at 10%, ** significant at 5%, *** significant at 1%. Each entry gives the pairwise correlation between the variable listed in the column and in the row. All four variables are a measures of the valuation that an individual places on \$100 change in monthly Social Security benefits. See the text for exact definitions. All measures are expressed in logs of the midpoint between the upper and lower bounds. To correct for correlations induced by common experimental manipulations (such as the starting value) across the four variables, we regress each variable on the relevant experimental manipulations and take the residual. The correlations between the resulting residuals are shown in the Table; uncorrected pairwise correlations are very similar and provided in Online Appendix Table A.2.

Table 3: Explaining the Sell-Buy Spread

Explanatory Variables	Dependent Variable: Absolute Value of Difference between Log CV-Sell and Log CV-Buy			
	(1)	(2)	(3)	(4)
Age 35 to 49	-0.10 (0.13)	0.08 (0.13)	0.09 (0.13)	0.22 (0.13)
Age 50 to 64	0.05 (0.12)	0.33*** (0.12)	0.34*** (0.12)	0.42*** (0.13)
Age 65 and older	0.44*** (0.14)	0.66*** (0.14)	0.68*** (0.14)	0.66*** (0.16)
Cognition index, standardized			-0.59*** (0.04)	-0.42*** (0.07)
Financial literacy index, 0-3 scale		-0.32*** (0.06)		
Education index, 1-5 scale		-0.24*** (0.04)		
Number series score, standardized		-0.31*** (0.05)		
Controls for demographics and preferences	No	No	No	Yes
Controls for experimental variation	Yes	Yes	Yes	Yes
Adjusted R ²	0.0279	0.1230	0.1233	0.1677
Number of observations	2065	2065	2065	2065
Mean of dependent variable	2.58	2.58	2.58	2.58
Standard deviation of dependent variable	1.85	1.85	1.85	1.85

Notes: Robust standard errors between parentheses. * significant at 10%, ** significant at 5%, *** significant at 1%. Each column contains an OLS regression of the Sell-Buy Spread (absolute value of the difference between log CV-Sell and log CV-Buy) on the explanatory variables listed in the rows. CV-Sell is the lump-sum amount given to the individual that would exactly compensate the individual for a \$100 decrease in monthly Social Security benefits. CV-Buy is the lump-sum amount the individual is just willing to pay for a \$100 increase in monthly Social Security benefits. All regressions also include controls for missing values of explanatory variables and controls for experimental manipulations, namely: log of starting value, asked after larger version, asked in wave 1, lump-sum option shown last. The financial literacy index is equal to the number of correct answers to three financial literacy questions (mean: 2.12 s.d.: 0.80). The education index equals 1 for high school dropouts, 2 for high school graduates, 3 for some college, 4 for bachelor's degree, and 5 for professional degree (mean: 3.42 s.d.: 1.05). The number series score is a standardized measure of performance on a number of questions that involve completing number series. The cognition index is the first principal component of the financial literacy index, the education index, and the number series score. The coefficients on the demographic and preference variables of the regression in column 4 are shown in Online Appendix Table A.3 column 1.

Table 4: Effects of Randomizations and Interactions with the Cognition Index

Explanatory Variables	(1)	(2)	(3)	(4)
	Dependent Variable: log CV-Sell			
	Entire sample	Top quintile of cognition index	Bottom quintile of cognition index	Entire sample
Log of starting value	0.37*** (0.07)	0.17 (0.13)	0.92*** (0.21)	0.39*** (0.07)
Asked after larger version	0.70*** (0.07)	0.70*** (0.12)	0.77*** (0.19)	0.69*** (0.07)
Asked in wave 1	0.04 (0.07)	0.01 (0.12)	0.38** (0.19)	0.05 (0.07)
Lump-sum option shown last	0.09 (0.07)	0.01 (0.12)	-0.03 (0.19)	0.08 (0.07)
Log of starting value × Cognition index				-0.20** (0.08)
Asked after larger version × Cognition index				-0.09 (0.07)
Asked in wave 1 × Cognition index				-0.03 (0.07)
Lump-sum option shown last × Cognition index				0.03 (0.07)
Cognition index				-0.17*** (0.04)
Adjusted R ²	0.0600	0.0832	0.0827	0.0737
N	2,090	385	412	2,090
Mean of dependent variable	10.02	9.82	10.27	10.02
Standard deviation of dependent variable	1.56	1.19	1.96	1.56

Notes: Robust standard errors in parentheses. * significant at 10%, ** significant at 5%, *** significant at 1%. Each column contains an OLS regression of the baseline CV-Sell measure on the explanatory variables listed in the rows. The baseline CV-Sell measure is the lump-sum amount given to the individual that would exactly compensate the individual for a \$100 decrease in monthly Social Security benefits. CV-Sell is expressed in logs of the midpoint between the upper and lower bounds. The starting value for the annuity valuation was randomized at \$10,000, \$20,000, or \$30,000. "Asked after larger version" equals one if the baseline CV-Sell measure was asked after a CV-Sell question in which Social Security benefits were varied by more than \$100. Whether this occurred was randomized. "Asked in wave 1" is a dummy variable that equals one if the CV-Sell question was asked in the first wave, and "Lump sum option shown last" is a dummy variable that equals one if the option involving the lump-sum amount was shown after the alternative option. Both dummy variables were randomized. The cognition index is the first principal component of the financial literacy index, the education index, and the number series score. All variables interacted with the cognition index are demeaned so that the coefficient on the cognition index can be interpreted as the effect of the cognition literacy index when the interaction variables are equal to their sample means.

Table 5: Explaining Annuity Valuations

Explanatory Variables	Dependent Variable: Mean of log CV-Sell and log CV-Buy			
	(1)	(2)	(3)	(4)
Log actuarial value	1.02*** (0.25)		0.84*** (0.26)	
Log theoretical utility-based annuity value		0.04 (0.04)		0.18 (0.13)
Age	-0.05*** (0.01)	-0.02 (0.01)	-0.04** (0.02)	-0.01 (0.01)
Age squared/100	0.06*** (0.02)	0.02 (0.01)	0.04*** (0.02)	0.01 (0.01)
Female	-0.08 (0.06)	0.01 (0.05)	-0.03 (0.06)	0.01 (0.06)
Married	0.05 (0.06)	0.04 (0.06)	0.08 (0.06)	0.13* (0.07)
Black	0.05 (0.12)	0.08 (0.12)	0.03 (0.12)	0.04 (0.12)
Hispanic	0.34*** (0.11)	0.36*** (0.11)	1.02*** (0.12)	0.32*** (0.12)
Other	-0.08 (0.13)	-0.08 (0.13)	-0.08 (0.12)	-0.09 (0.13)
Education index, 1-5 scale			-0.03 (0.03)	-0.03 (0.03)
Log family income			0.03 (0.04)	0.02 (0.04)
Owns an annuity			-0.07 (0.06)	-0.04 (0.06)
Owns home			-0.16* (0.08)	-0.16*** (0.08)
Log financial wealth			0.01 (0.02)	-0.01 (0.03)
Self-reported health index, 1-5 scale			-0.03 (0.03)	-0.03 (0.03)
Ever had kids			-0.03 (0.06)	-0.04 (0.06)
Risk aversion (standardized)			0.02 (0.03)	0.02 (0.03)
Precaution (standardized)			-0.07** (0.03)	-0.07** (0.03)
Expects returns greater than 3% p.a.			0.10* (0.06)	0.10* (0.06)
Confident SS will pay promised benefits, 1-4 scale			1.02*** (0.03)	0.14*** (0.03)
Controls for experimental variation	Yes	Yes	Yes	Yes
Adjusted R ²	0.0543	0.0473	0.0649	0.0607
Number of observations	2065	2065	2065	2065
Mean of dependent variable	9.18	9.18	9.18	9.18
Standard deviation of dependent variable	1.22	1.22	1.22	1.22

Notes: Robust standard errors between parentheses. * significant at 10%, ** significant at 5%, *** significant at 1%. Each column contains an OLS regression of annuity valuation (mean of log CV-Sell and log CV-Buy) on the explanatory variables listed in the rows. CV-Sell is the lump-sum amount given to the individual that would exactly compensate the individual for a \$100 decrease in monthly Social Security benefits. CV-Buy is the lump-sum amount the individual is just willing to pay for a \$100 increase in monthly Social Security benefits. All regressions also include controls for missing values of explanatory variables and controls for experimental variation, namely: log of starting value, asked after larger version, asked in wave 1, lump-sum option shown last. To calculate the theoretical utility-based annuity value, we solve the lifecycle dynamic programming problem for a household that matches the respondent on age, gender, marital status, spousal age (if married), start date of the annuity, financial wealth, existing annuity wealth, and coefficient of risk aversion, assuming a real discount rate of three percent per year. We solve this lifecycle dynamic programming problem twice: once for the CV-Sell equivalent wealth and once for the CV-Buy equivalent wealth. We take the log of both amounts and average them. The education index equals 1 for high school dropouts, 2 for high school graduates, 3 for some college, 4 for bachelor's degree, and 5 for professional degree. Summary statistics and sources for the explanatory variables are provided in Online Appendix Table A.1.

Table 6: Predictive Power of Actuarial Value by Quintile of the Cognition Index

	(1)	(2)	(3)	(4)	(5)
Dependent Variable:	Coefficient on log	p-value on	Root	Adjusted	
Mean of log CV-Sell and log CV-Buy	actuarial value	coefficient=1	MSE	R ²	N
Sample split by quintiles of the cognition index					
1. <i>Bottom quintile</i>	0.46 (0.77)	0.483	1.488	0.0922	404
2. <i>Second quintile</i>	0.76 (0.59)	0.686	1.246	0.0259	451
3. <i>Third quintile</i>	1.24** (0.49)	0.618	1.163	0.0204	392
4. <i>Fourth quintile</i>	0.77 (0.50)	0.650	1.034	0.0498	433
5. <i>Fifth quintile</i>	1.49*** (0.51)	0.340	0.889	0.0677	385

Notes: Robust standard errors in parentheses. * significant at 10%, ** significant at 5%, *** significant at 1%. Here we estimate specification 1 of Table 5 by subsample. Each row contains an OLS regression of the log annuity valuation (mean of CV-Sell and CV-Buy) on the log actuarial value and additional controls. Additional controls are those in specification 1 of Table 5. The cognition index is the first principal component of the financial literacy index, the education index, and the number series score.