Retirement drawdown defaults: the role of implied endorsement

Jennifer Alonso-García* Hazel Bateman† Johan Bonekamp‡ Ralph Stevens§

February 26, 2018

Abstract

Implied endorsement is considered, together with inertia, as an explanation for the stickiness of defaults. This paper explores whether implied endorsement can serve as an explanation for the stickiness of defaults in the retirement decumulation phase. Using an experimental survey fielded in both the Netherlands and Australia, we analyze whether individuals perceive mandatory minimum withdrawals from their pension wealth set by the government as implicit advice from the government ("government knows best") or recommendations from peers ("what most people do"). We find that vulnerable groups, such as those with low financial resources and pension capabilities, are more likely to find implied endorsement important. However, those who are overconfident about their capabilities find it less important and are susceptible to adverse choices. As expected, the peer effect is reduced for those with less opportunity to participate in a social network.

JEL: D14, D90, J14, J26
Keywords: decumulation policy, defaults, peer effects

This project has received funding from the ARC Center of Excellence in Population Ageing Research (grant CE110001029) and ARC Linkage Grant 'Mandatory pre-funded retirement income schemes: Best policy and practice' (grant LP140100104). We thank Eduard Ponds and Arthur van Soest for feedback, and attendees at the Netspar Pension Day conference for useful comments. The authors are responsible for any errors.

*Email: j.alonsogarcia@unsw.edu.au; ARC Centre of Excellence in Population Ageing Research (CEPAR), UNSW Sydney, Australia; School of Risk and Actuarial Studies, UNSW Sydney, Australia.

†Email: h.bateman@unsw.edu.au; ARC Centre of Excellence in Population Ageing Research (CEPAR), UNSW Sydney, Australia; School of Risk and Actuarial Studies, UNSW Sydney, Australia.

‡Email: J.L.M.Bonekamp@uvt.nl; Department of Econometrics and Operations Research, Tilburg University, The Netherlands; Network for Studies on Pensions, Aging and Retirement (Netspar), The Netherlands.

§Email: ralph.stevens@unsw.edu.au; CPB Netherlands Bureau for Economic Policy Analysis, The Netherlands.

1 Introduction

The global move towards defined contributions (DC) from defined benefit (DB) pensions arrangements is changing the retirement benefit landscape. Retirees are less likely to be in receipt of lifelong benefits in the form of pensions or annuities, and more likely to be required to make decisions about retirement benefit products and regular drawdown patterns (OECD, 2015).

Irrespective of the institutional framework for pensions, deciding how much to spend and save in retirement is difficult for most people. The policy toolkit available to assist with retirement benefit and drawdown decisions include mandates (e.g., requiring full annuitization), economic incentives (e.g., providing concessions such as lower tax rates for the take-up of certain benefits or drawdown patterns) and nudges (which use choice architecture to alter behavior - see Thaler and Sunstein, 2008). Recent papers have substantiated the effectiveness of nudges as a policy tool (Benartzi et al., 2017; Chetty et al., 2014).

Our aim is to investigate the role that nudges may play in influencing retirement spending and saving decisions. Nudges are of many types and their number and variety are constantly growing (Sunstein, 2014). However, the most important for policy purposes can be classified into four different groups: 1) simplification and framing of information; 2) changes to default rules; 3) social norm comparisons; and 4) changes to the physical environment (Lehner et al., 2016). We focus on two most commonly used nudges - default rules and the use of social norm comparisons. That is, in the context of retirement drawdown decisions we study implied endorsement prompted by social norm comparisons, as a potential explanation for the stickiness of defaults.

A default is a prescribed alternative if an individual fails to make an active choice, or decides not to choose between alternatives (Johnson and Goldstein, 2003). Defaults are proven to be a powerful tool to achieve policy and company specific objectives. The power of defaults lies in their stickiness: that is, people rarely switch to an alternative and tend to stay with the default option. Sunstein (2013) argues that the three main underlying mechanisms for the stickiness of defaults are: inertia and procrastination; implied endorsement; and reference point and loss aversion. For instance, in the context of retirement saving, numerous studies have identified that switching the participation default in 401(k) savings plans from opt-in to opt-out led to very large increases in plan participation due to inertia and the perception of implicit advice (see for example, Madrian and Shea, 2001; Beshears et al., 2008).

Defaults have been found to be effective in guiding behavior in a variety of fields - including increasing organ donation intentions (Johnson and Goldstein, 2003), improving vaccination rates and therefore reducing the annual economic cost of flu (Chapman et al., 2016), improving the take-up of "green" electricity options (Pichert and Katsikopoulos, 2008), and increasing the coverage of comprehensive motor insurance (Johnson et al., 1993).

The effectiveness of defaults has also been proven in the field of retirement savings, particularly in the context of individual accounts where pension plan members can

be faced with a myriad of choices including whether to participate, which fund or plan to join, how much to contribute, which investment option to take, and how to decumulate at retirement. In the context of the United States 401(k) plans Madrian and Shea (2001) find that changing the participation default from opt-in to opt-out increased plan participation from around one third to over 80% for selected cohorts and that this stickiness extends to contribution rates and asset allocation. These findings are confirmed in Choi et al. (2002, 2004) and Choi et al. (2005). Research on Australia's funded income replacement pillar finds stickiness to the incumbent fund when fund choice is offered (Fry et al., 2007) and to the default DB plan when an alternative DC plan is available (Gerrans and Clark-Murphy, 2004), while Bateman et al. (2014) and Dobrescu et al. (2017) identify stickiness across all domains (default settings for plan type, contribution rate, asset allocation and life insurance) by members of UniSuper, the pension fund for staff in the university sector. Elsewhere, Hedesstrom et al. (2004) find stickiness to the default investment option in the Swedish premium pension (SPP) scheme. Following the wind down of information and marketing about choice options, take-up of the default investment fund by new entrants to the SSP increased from around one third in the first year of the scheme to over 90% by year five.

However, the stickiness of defaults can also be their downfall. Goda and Manchester (2013), Carroll et al. (2009) and Dobrescu et al. (2017) all find that if defaults are not carefully chosen, their characteristic stickiness can lead to sub-optimal outcomes in terms of lifetime wealth accumulation. Similarly, Beshears et al. (2016) in a study of the 401(k) plan behavior of ten large companies in the United States, find that defaults are more sticky for young and low income individuals. This, suggests that where identifiable groups are more affected by defaults, their needs should receive more weight in the "optimal" default.

In the retirement decumulation phase a series of papers have found that retiring plan members tend to stick to the default benefit option offered by their plan. This explains the predominance of annuities in the Swiss system, where the benefit default is generally a life annuity rather than a lump sum (Bütler and Teppa, 2007) and vice versa for US 401(k) plans (Benartzi et al., 2011) where a lump sum distribution is often the default. For retirees in the United Kingdom, Banks et al. (2015) find, using data from the English Longitudinal Study of Ageing (ELSA), that over 60% of individuals follow the "path of least resistance" and purchase an annuity from their pension provider, despite being informed, prior to retirement, about the possibilities of purchasing an annuity from an alternative provider. To our knowledge, there has been no research on the extent to which retirees stay with default or recommended drawdowns from their retirement accumulation[1].

In a similar vein, the use of social norm comparisons have been successful in influencing behavior across a number of domains including tax compliance, energy use and recycling. Here the strategy to nudge or influence behavior is to highlight what

[1]In Australia's superannuation scheme most retirees take their retirement benefits as phased withdrawal products (known in Australia as account-based pensions) (APRA, 2017) which is the effective default. However, the extent to which retirees stick to the regulated age-based minimum drawdown requirements is unclear.

"most people do". For example, in the case of tax compliance, people could be told "most people pay their taxes on time". Sunstein (2013) suggests that such information is often most powerful when it is as local and specific as possible, for example by qualifying the previous statement as "the overwhelming majority of people in your community pay their taxes on time". In a large-scale field experiment designed to learn about the factors that influenced the timely payment of taxes in the United Kingdom, Hallsworth et al. (2017) found that social norm comparisons were more effective than public good messaging (i.e., reminding people that paying taxes funds health, roads and schools). Evidence on the role of social norm comparisons (or peer effects), to assist with the complex lifetime spending and saving decisions is mixed. Duflo and Saez (2002, 2003) in two studies on the retirement saving behavior of university librarians found strong peer effects in employee pension plan decisions, and Brown et al. (2008) find similar effects for stock market participation. However, Beshears et al. (2015) found the opposite effect following the presentation of peer information about retirement savings. Similarly, individuals could interpret the institutional settings (such as the prescribed minimum drawdown requirements in Australia's superannuation system) as "advice" in terms of "what people should do".

Our study focuses on retirement drawdown decisions. We use an online experimental survey of retirement saving and spending decisions of soon to retire households in Australia and the Netherlands to analyze the effect of a default setting which prescribes minimum withdrawal rates from pension wealth. The implied endorsement associated with this default is communicated as either the "government knows best" or "what most people do". This allows us to test the effectiveness of two alternative prompts for implied endorsement communicated as implicit advice from the government ("government knows best") or as a social norm comparison or peer effect ("what most people do")[2].

In our survey we use vignettes to present short descriptions of hypothetical retiree households with given patterns of wealth and income (out of three alternatives), given expected future health status (out of four alternatives) and a default minimum drawdown set by government (or not). We then ask participants to advise the retiree household on a spending level and to rank the importance of a set of saving motives consistent with the spending advice that is given. The set of saving motives is drawn from economic theory and the behavioral and psychology literature and, for the default minimum drawdown treatment, includes implicit advice from government and a social norm comparison as motivators of implied endorsement (i.e., following the minimum drawdown pattern as set by government).

The Netherlands and Australia are ideal as settings for our study. Both countries are consistently ranked in the top three pension systems internationally by the Melbourne Mercer Global Pension Index (Mercer, 2017) and have a similar structure comprising a publicly funded state pension and a funded income replacement pillar supplemented by private saving. However, there is an important difference in

[2]Dinner et al. (2011), in a study of the effect of implied endorsement in the choice of light bulbs, labeled "government" implied endorsement as "direct" implied endorsement and the peer effect as "external" implied endorsement.

the decumulation phase. Dutch retirees are required to transfer the largest part of their retirement savings from the income replacement pillar into lifetime annuities while Australian retirees have considerable flexibility, with the discretion to take a lump sum, a phased withdrawal product (known as an account-based pension) or a term or lifetime annuity. Most Australian retirees take account-based pensions (the phased withdrawal option) which are associated with tax concessions for minimum annual withdrawals (APRA, 2017).

The experimental setup allows us to answer a number of important questions about the impact of nudges on retirement drawdown behavior, and the collection of information on demographics, preferences, psychological traits, and financial knowledge and competencies as well as country of residence enables us to identify the characteristics of those who are sensitive to the nudges we implement. First, we investigate the impact of default drawdown requirements by assessing whether individuals change their spending to match the default spending pattern suggested by the minimum withdrawal requirements in the experimental setup. We find that some do. Second, we are interested in the strength of two different prompts for implied endorsement - that the household wants to stick to the withdrawal amounts as suggested by the government as "government knows best" and that the household wants to stick to the withdrawal amounts as suggested by government as "that is what most people do", which we interpret as a "peer effect". We find that vulnerable groups, such as those with low financial resources and pension capabilities, are more likely to find implied endorsement important. However, those who are overconfident about their capabilities find it less important and are susceptible to adverse choices. As expected, the peer effect is reduced for those with less opportunity to participate in a social network.

The remainder of this paper is organized as follows. Section 2 describes the online experimental survey. In Section 3 we describe the data and report descriptive statistics. Section 4 discusses the model and the estimation results, while in Section 5 we discuss our findings and conclude.

2 Experimental survey

We designed and implemented and experimental survey to elicit spending patterns and associated saving motives of elderly Dutch and Australians[3].

The objective of the embedded experimental task was to investigate the effect of retirement income policy design, implicit advice (via a default minimum drawdown requirement) and future health and household shocks on spending patterns and saving motives of retirees. The effect of retirement income policy design was elicited by varying prescribed flexibility or annuitization from 100% annuitization to 100% flexibility. The importance of implied endorsement was assessed by setting a government prescribed drawdown pattern. Lastly, the impact of the health and household shocks

[3]An interactive link to the English version of the survey is available at http://survey.us.confirmit.com/wix/6/p3081554696.aspx. Static copies of the full survey are available in the supplementary material.

were elicited through a gradual deterioration of ADLs (Activities of Daily Living) or widowhood. We are also interested in the explanatory power of demographics, personality traits, financial competence, and other personal characteristics. In this paper we focus on implicit advice[4].

We used vignettes to present the choice sets (see Louviere et al., 2000). The vignettes consist of short descriptions of a hypothetical household that differs by retirement income policy design (which determine pairs of wealth and income), minimum government prescribed drawdown requirements (or not) and expected future health status. A clear advantage of the vignette methodology in cross-country analyses is that we can, to a certain extent, control for country specific factors. For example, in the Netherlands the state pension is universal, whereas in Australia it is means-tested[5]. Another advantage is that participants whose actual personal situation differs from the hypothetical scenarios presented, can still authentically participate in the experiment.

A drawback of stated preferences studies is that one cannot be fully certain that individuals would behave accordingly to their responses in the real world. However, this drawback does not outweigh the advantages of a stated preference study. First, in an experimental survey we can control various otherwise unobserved circumstances (e.g. state pension eligibility and amount, homeownership, etc.). Second, we can learn a lot about the participants through asking questions about personality traits, risk preferences, etc. Third, we can ask participants for the motivation for their decisions.

Eligible participants from the Netherlands and Australia were invited to participate in the experimental survey. Participants were eligible if they were aged 50-64 and not yet retired. In the Netherlands participants were recruited from the LISS and CentER panels. In Australia participants were recruited through the commercial web panel provider 'TEG rewards'. The Dutch sample comprised of 1,798 eligible participants, and the Australian sample comprised of 1,004 individuals.

2.1 Survey overview

The survey instrument comprised preliminary screening questions and four sections. The preliminary screening questions covered marital status, age of the participant and their partner (if applicable), employment status and gross household income. These questions are only included in the Australian version of the survey since this information is available in the LISS and CentER panels for the Dutch survey. The aim was to ensure the desired sample of individuals aged 50-64 who were not yet retired, or part of a household where at least one was not yet retired. The gross household income information was used to allocate participants into one of four income groups to prevent alienation as they proceed through the task. The remainder of the survey is (almost) the same in the two countries: the exception is

[4]The other two objectives are tackled in a companion paper by Alonso-García et al. (2017).

[5]An overview of the Dutch and Australian pension schemes is provided in Appendix A and discussed in detail in Bateman et al. (2016b).

questions collected as background questions in the Dutch version via the LISS and CentER panels.

The first section of the experimental survey is the experimental task which we describe in more detail in Section 2.2. The second section consists of questions on retirement planning and personality traits including time preference and planning horizon (Fisher and Montalto, 2011), risk attitude (Dohmen et al., 2011) and life expectancy (of the participant and partner - if applicable) (Bateman et al., 2016a; Teppa et al., 2015). To save survey time, we include the ten-item personality inventory (TIPI) (Gosling et al., 2003) instead of the much lengthier Big Five questionnaire in order to obtain a proxy for personality traits. The third section consists of questions on the pension arrangements and financial competence of the participant. Financial competence is measured by questions on financial literacy (Lusardi and Mitchell, 2011), numeracy (Lipkus et al., 2001) and pension knowledge (Agnew et al., 2013). The fourth section concludes with questions on demographics and personal characteristics. The median time for completion of the survey was 30 minutes for Australia and 31 minutes for the Netherlands.

2.2 Experimental Task

The experimental task consists of eight vignettes that describe a hypothetical house hold of two individuals age 65 who have just retired and own the house they live in. The base vignette and experimental design is summarized in Appendix B, Figure B.1 and Table B.1. To answer our research questions, the vignettes have varying levels of liquid wealth and lifetime income (linked to the given alternative for retirement income policy design), implied endorsement (the minimum government prescribed drawdown requirements), and expected health status. Vignettes 1-3 present three levels of retirement income policy design. Vignette 1 presents a high liquidity of wealth and low annuitization, similar to the Australia's flexible drawdown pension system, whereas vignette 3 has a low liquidity of wealth and high annuitization, similar to the Dutch pension system where annuitization is mandatory. Vignette 2 presents the average between vignette 1 and 3 (reflecting partial annuitization and some flexibility) as an intermediate situation and possible future policy direction for both Australia and The Netherlands. The hypothetical household in vignette 4 has the same liquidity and lifetime income combination as in vignette 1, but is required by the government to annually withdraw a part of their liquid wealth - that is, the default or implied endorsement. Vignettes 5 to 8 differ with respect to expected future health status, represented by declining ADLs and/or death of a spouse for the hypothetical households. In the experiment, participants are presented with vignettes 1-3 in random order, followed by vignette 4 and then vignettes 5-8. For vignettes 5-8 participants are allocated randomly to a high liquidity treatment (high wealth, low income - a proxy for the Australian institutional arrangements), or a low liquidity treatment (low wealth, high income - a proxy for the Dutch institutional arrangements).

Vignette 4 presents the default or implied endorsement alternative in which the hypothetical household is faced with a minimum government prescribed drawdown

requirement. This mimics the tax driven minimum wealth withdrawals in the Australian pension system. In vignette 4 the hypothetical household is obliged to withdraw a fixed amount of wealth each year equal to five percent of their initial wealth at the age of retirement (as determined by the high wealth treatment - vignette 1). The wording of vignette 4 differs only by the following sentence placed before the five spending alternatives: *"Government regulations require that they withdraw a part of their wealth each year to supplement their income. This corresponds to <<income category dependent amount>>. The household is not obliged to (fully) spend this supplementary income."* (see the bold text in the representative screen shot shown in Figure B.2, Appendix B). This regulated drawdown is the default drawdown set by the government, which we interpret as "implied endorsement".

Minimum withdrawal rates are used in a number of countries including the United States (Internal Revenue Service, 2017), Canada (Canada Revenue Agency, 2017) and Australia (MoneySmart, 2016). For example, in Australia, retirees who follow age-based minimum withdrawals from their pension accounts ranging from 5% of the account balance at ages 65-74 to 14% for individuals aged 95 and older receive tax concessions (MoneySmart, 2016). However, this minimum withdrawal requirement does not equate to a minimum spending requirement. The minimum withdrawals can be used to increase savings in non-pension accounts. We capture this feature in vignette 4 by explicitly stating that the hypothetical household is not forced to consume all this (additional) income, as shown in Figure B.2, Appendix B.

For each vignette, participants are required to complete two tasks. First, the participant is asked to advise a spending pattern out of five alternatives for the hypothetical household based on their own preferences. In order to help the participant, for each spending pattern, information on remaining wealth of the hypothetical household at the age of 65, 75, 85, and 95 is presented (see Figure B.4, Appendix B for a screen shot of the first task in the context of vignette 1). Second, the participant is presented with five motives for the chosen spending (and consequent saving) pattern and asked to rank these in terms of consistency with the chosen spending pattern, using two rounds of best/worst (see Figure B.5, Appendix B for a representative screen shot).

For vignettes 1-3 and 5-8 the five saving motives are drawn from a list of ten. These ten are the result of a pre-test (using best/worst scaling) to identify the most important saving motives for people approaching or in retirement from an exhaustive list of 19 drawn from economic theory and the behavioral economics and psychology literature (see supplementary materials for details). The ten short listed saving motives are the first ten listed in Table 1. For each of vignettes 1-3 and 5-8 the five saving motives are randomly selected, subject to three from the six "rational" motives and two from the four "psychological" motives[6].

Vignette 4 is designed to test the role of the default drawdown as a form of implied endorsement. To do so, we supplement the short list of ten saving motives with two additional motives designed to be prompts for implied endorsement. The two prompts are communicated as implicit advice from the government ("government

[6]Further explanation of the experimental design can be found in Appendix B.

knows best") or as a social norm comparison or peer effect ("what most people do"). The full text of these additional saving motives as used in the experimental task associated with vignette 4 can be found in the bottom two rows of Table 1. The first is written as "*The household wants to stick to the withdrawal amounts as suggested by the government as government knows best.*" and is designed to capture implicit advice from the government. The second is written as "*The household wants to stick to the withdrawal amounts as suggested by the government as that is what most people do.*" is a social norms comparison designed to capture the peer effect. For vignette 4, the five randomly selected saving motives are selected from the subset of 12 listed in Table 1, using the decision rule that included the three most important (highest ranked) saving motives in vignette 1 (as a comparator for vignette 4), supplemented by the two implied endorsement motives "government knows best" and "what most people do" (see Figure B.3, Appendix B for a screen shot of the second task in the context of the implied endorsement vignette)[7].

Table 1: Saving motives used in the vignettes

Name	Text in vignette
Rational	
precautionary	wants to ensure that they will be able to finance any unforeseen expenditures other than health and aged care expenditures
precautionary health	wants to ensure that they will be able to finance unforeseen health and aged care expenditures
life-span risk	wants to ensure that they will not outlive their wealth
intended bequest	wants to ensure that they will be able to leave a bequest to their dependents or estate
liquidity	wants to ensure that they have enough cash on hand at any time
intra-household bequest	wants to ensure that if they die, their partner is able to maintain his/her standard of living
Psychological	
autonomy	wants to ensure that they remain financially independent
security	wants to ensure that they have enough money to have peace of mind
self-gratification	wants to ensure that they are able to enjoy life now as well as later
political risk	wants to ensure that they are protected against a change in the superannuation / pension rules
Implied Endorsement	
government advice	wants to stick to the withdrawal amounts as suggested by the government as government knows best.
peer effect	wants to stick to the withdrawal amounts as suggested by the government as that is what most people do.

In this paper our focus is on the role of implied endorsement (the default mini-

[7]Contrary to Beshears et al. (2015), we do not specify whether the "people" in the "peer effect" prompt are individuals who belong to the same 5- or 10-year age bracket as the participants. In their field experiment, they study the effect of providing peer information in saving decisions for individuals between the age of 20 and 69. To assess the effect of peer information and increase the power of social comparison, they use 5- or 10-year age brackets of the recipient's age (e.g. individuals aged 20-29) and compare this with not providing peer information. Their findings suggest that providing peer information matters, but it hardly matters whether this is presented in 5- or 10-year age brackets. Note that our participants are aged between 50 and 64 and not yet retired, thereby making an age-specific "peer effect" saving motives redundant.

mum drawdown requirement) and the relative strength of two alternative prompts summarized as "implicit advice from the government" and "peer effect". As such we consider only the stated preferences from vignette 1 (high liquidity of wealth, no default drawdown, no expected decline in health status) and vignette 4 (default drawdown/implied endorsement, no expected decline in health status)[8].

3 Data and descriptive statistics

Almost 2,700 eligible participants were invited to participate in the experimental survey. Around 1,000 individuals were recruited in Australia from the web panel provider 'TEG rewards' and almost 1,700 individuals were recruited from the Dutch LISS and CentER panels. For Australian participants, we asked a question on their gross household income at the beginning of the survey to be able to allocate them into representative income groups. For most Dutch individuals from the two panels we had information on their gross household income as background variables. The Dutch individuals for whom information on gross household income was missing were assigned to an income group at random when entering the survey. Given that we asked a question on gross household income later in the survey, we could verify the accuracy of this randomization procedure. Individuals who were severely mismatched[9] were excluded from the analysis sample (138 out of 1,669 Dutch individuals) as preliminary analysis suggest that their survey responses differed statistically significantly from those who were not severely mismatched. In addition, participants who afterwards turned out not to be eligible to participate were excluded together with those from whom information on relevant covariates was missing. Our analysis sample consists of 2,420 individuals of whom 1,437 are Dutch and 983 are Australian. The key covariates we collect are defined in Table 2.

Our study focuses on the impact of nudges on retirement drawdown behavior in an experimental setting. We are interested in investigating the impact of a default drawdown by the government on advised spending patterns. Therefore, we construct the variable *changed advised spending pattern*, that equals 1 if the participant changed their advised spending pattern in vignette 4 (with a default drawdown set by the government) from vignette 1 (without a default drawdown), and 0 otherwise. In the Australian sample over 30% of the individuals changed their advised spending pattern compared to slightly over 25% in the Dutch sample (see Table 3)[10].

The implicit advice communicated can either have an effect because it can be in-

[8] An analysis of stated preferences in vignettes 1-3, 5-8 can be found in Alonso-García et al. (2017).

[9] Individuals are flagged as 'severely mismatched' if their self-reported income category differed by at least two positions from their randomly assigned income category.

[10] Our nudge was constructed in such a way that the participants should advise 'spending plan 1' to the hypothetical household. Around 10% (15%) of the Australian (Dutch) participants advised 'spending plan 1' in vignette 4. However, the word "require" in vignette 4 might have induced some unforeseen responses. If the participants interpreted the nudge such that the hypothetical households should always be able to withdraw the amount of wealth as suggested by the government, this could lead to some more conservative behavior. Therefore, we focus in this paper on changing the advised spending pattern, as opposed to advising 'spending plan 1'.

Table 2: Description of the relevant covariates

Covariate	Explanation
Financial resources	
Household income (Q3 and Q4)	1 if participant belongs to the upper quantiles of (current) gross household income, 0 otherwise
Homeowner	1 if participant owns (potentially with a mortgage) the house (s)he lives in, 0 otherwise
Future orientation	
Time horizon: more than 5 years from now	1 if participant answered a time horizon of at least five year to the question "People use different time-horizons when they decide which part of their income to spend, and which part to save. Which of the time-horizons mentioned below is in your household most important with regard to planning expenditures and savings for you?", 0 otherwise
Pension capabilities	
Objectively measured	1 if participant made less mistakes than the average of the population in financial literacy (Lusardi and Mitchell, 2011), numeracy (Lipkus et al., 2001), and pension system knowledge questions (Bateman et al., 2016a), 0 otherwise
Self-assessed capabilities[b]	1 if participant is more confident in his/her own pension related capabilities than the average of the population, 0 otherwise
Overconfidence	1 if *objectively measured* equals 0 and *self-assessed capabilities* equals 1, 0 otherwise
Social network	
Employed[a]	1 if participant indicates his current work status is employed, 0 otherwise
Employed x (gender = female)	Interaction between *Employed* and *gender = female*. Thus, 1 if participant is employed and is female, 0 otherwise
Religious / Member of a church community	1 if participant considers himself / herself as member of a religion or church community, 0 otherwise
Additional control variables	
gender = female	1 if female, 0 if male
marital status = partner	1 if participant lives together with partner, 0 otherwise
children living at home	1 if participant has at least one child living at home, 0 otherwise
age: 50-54	1 if participant is in the 50 to 54 age group, 0 otherwise
age: 55-59	1 if participant is in the 55 to 59 age group, 0 otherwise
age: 60-64	1 if participant is in the 60 to 64 age group, 0 otherwise
willingness to take risk	standardized measure comprised of the following question: "How do you see yourself: Are you generally a person who is fully prepared to take risks or do you try to avoid taking risks?"
born in the country they are currently living in	1 if participant is born in the country (s)he lives in, 0 otherwise
personality: TIPI conscientiousness	standardized measure for the personality trait conscientiousness, comprising two conscientiousness related questions of the ten-item personality inventory (TIPI) questionnaire (Gosling et al., 2003)
personality: TIPI extraversion	standardized measure for the personality trait extraversion, comprising two extraversion related questions of the TIPI questionnaire
personality: TIPI neuroticism	standardized measure for the personality trait neuroticism, comprising two neuroticism related questions of the TIPI questionnaire
Country of residence	
Australia	1 if participant took part in the Australian questionnaire, 0 otherwise

Notes: Standardized measures are standardized (mean 0 and standard deviation of 1) using the full analysis sample.

[a] Based on (1)-(3) from *"Which of the following best describes our current work status?"* (1) paid employment; (2) works or assists in family business; (3) autonomous professional, freelancer, or self-employed; (4) job seeker following job loss; (5) first-time job seeker; (6) exempted from job seeking following job loss; (7) attends school or is studying; (8) takes care of the housekeeping; (9) has (partial) work disability; (10) performs unpaid work while retaining unemployment benefit; (11) performs voluntary work; (12) does something else.

[b] Measure is comprised of the answers to the following questions (scale 1: strongly disagree to 7: strongly agree): (1) *"I am very knowledgeable about financial planning for retirement"* (2) *"I know more than most people about retirement planning"* (3) *"I am very confident in my ability to do retirement planning"*.

Table 3: Descriptive statistics for the dependent variables

	Analysis Sample		Australia		The Netherlands	
	Mean	Sd	Mean	Std. Err.	Mean	Std. Err.
Dependent variables						
Changed advised spending pattern	0.29	0.45	0.32	0.015	0.27	0.012
Importance: Government prompt	0.06	0.23	0.06	0.008	0.05	0.006
Importance: Peers effect prompt	0.08	0.27	0.10	0.010	0.06	0.006
Observations	2,420		983		1,437	

Notes: Difference between "changed advised spending pattern" for Australia and for The Netherlands are significant at the 5% level (two-sample test of proportions). "importance: Peers effect prompt" at the 1% level (two sample test of proportions). The difference between Australia and the Netherlands for the "Importance: Government prompt" is not statistically significant.

terpreted as "government knows best" or it provides information on "what most people do" (a peer effect). Therefore, we construct two variables that capture the importance of these implied endorsement motives. As the experimental design for vignette 4 includes (for a given participant) the two implied endorsement saving motives together with the top three saving motives from vignette 1, an intuitive way is to model their importance as follows. If the government (or peers) saving motive is the first or second most important in the best/worst choice task, we assign value 1 and 0 otherwise. In our analysis sample 6.2% (10.2%) out of 983 Australian participants rank the government (peers) saving motive as the first or second most important in the best/worst choice task (see Table 3). For the Dutch this occurs for 5.4% (6.2%) out of the 1437 participants. Although these numbers look small initially, recall that this is relative to their (imputed) top three saving motives.

Table 4 shows the descriptive statistics for the main covariates in the analysis sample. Our measures of financial resources are household income and home ownership. The proportion of participants with a high income is similar in both the Dutch sample and the Australian sample. However, reflecting country differences, Australians are more likely to be homeowners than the Dutch. In Australia most older individuals own their home and only a minority rent. In the Netherlands there is a large social housing pool with substantial renter's protection and low rents.

Related to future orientation, we observe from our sample that Australians tend to be more future oriented. An explanation for this is that the Australian retirement system puts more responsibility on individuals, making it more important to plan long term. The Dutch retirement system has little choice and individual responsibility, making it less important to plan long term, as long-term retirement planning is already taken care of. Related to pension capabilities, we find that Australians made less mistakes in the financial literacy, numeracy and pension literacy questions and thus score higher on the objective measure of the pension capabilities. This might be because they are more exposed to financial decisions as the retirement income system puts more emphasis on individual choice. However, there is no different in self assessed pension capabilities on average between the Australian and Dutch samples, although the Dutch are more likely to be overconfident. For social networks, we observe from our sample that, on average, the Dutch are less likely to

Table 4: Descriptive statistics for the control variables

	Analysis Sample				Australia				The Netherlands			
	Mean	Sd	Min	Max	Mean	Sd	Min	Max	Mean	Sd	Min	Max
Covariates												
Financial resources												
Household income (Q3 and Q4)	0.32	0.47	0.00	1.00	0.34	0.47	0.00	1.00	0.31	0.46	0.00	1.00
Homeowner	0.78	0.42	0.00	1.00	0.83	0.37	0.00	1.00	0.74	0.44	0.00	1.00
Future orientation												
Time horizon: more than 5 years from now	0.44	0.50	0.00	1.00	0.52	0.50	0.00	1.00	0.38	0.49	0.00	1.00
Pension capabilities												
Objectively measured	0.24	0.43	0.00	1.00	0.34	0.47	0.00	1.00	0.18	0.39	0.00	1.00
Self-assessed capabilities	0.54	0.50	0.00	1.00	0.55	0.50	0.00	1.00	0.54	0.50	0.00	1.00
Overconfidence	0.37	0.48	0.00	1.00	0.31	0.46	0.00	1.00	0.41	0.49	0.00	1.00
Social network												
Employed	0.77	0.42	0.00	1.00	0.89	0.31	0.00	1.00	0.68	0.47	0.00	1.00
Employed x (gender = female)	0.33	0.47	0.00	1.00	0.42	0.49	0.00	1.00	0.27	0.45	0.00	1.00
Religious / Member of a church community	0.32	0.47	0.00	1.00	0.30	0.46	0.00	1.00	0.33	0.47	0.00	1.00
Additional control variables												
gender = female	0.50	0.50	0.00	1.00	0.50	0.50	0.00	1.00	0.50	0.50	0.00	1.00
marital status = partner	0.72	0.45	0.00	1.00	0.74	0.44	0.00	1.00	0.71	0.45	0.00	1.00
children living at home	0.40	0.49	0.00	1.00	0.45	0.50	0.00	1.00	0.36	0.48	0.00	1.00
age: 50-54	0.34	0.47	0.00	1.00	0.38	0.49	0.00	1.00	0.31	0.46	0.00	1.00
age: 55-59	0.34	0.47	0.00	1.00	0.39	0.49	0.00	1.00	0.31	0.46	0.00	1.00
age: 60-64	0.32	0.46	0.00	1.00	0.23	0.42	0.00	1.00	0.37	0.48	0.00	1.00
willingness to take risk	0.00	1.00	-1.97	2.40	0.06	0.98	-1.97	2.40	-0.04	1.01	-1.97	2.40
born in the country they are currently living in	0.85	0.36	0.00	1.00	0.74	0.44	0.00	1.00	0.92	0.27	0.00	1.00
personality: TIPI conscientiousness	0.00	1.00	-4.11	1.52	0.17	0.94	-4.11	1.52	-0.12	1.02	-4.11	1.52
personality: TIPI extraversion	0.00	1.00	-2.24	2.09	-0.21	0.98	-2.24	2.09	0.14	0.99	-2.24	2.09
personality: TIPI neuroticism	0.00	1.00	-3.41	1.64	-0.21	0.98	-3.41	1.64	0.14	0.99	-3.41	1.64
Country of residence:												
Australia	0.41	0.49	0.00	1.00	1.00	0.00	1.00	1.00	0.00	0.00	0.00	0.00
Observations	2,420				983				1,437			

Notes: Objectively measured: is a combined measure that equals 1 if respondent made less mistakes than the average of the population in the financial literacy, numeracy, and pension system knowledge related questions, and 0 otherwise. Self-assessed: is a combined measure that equals 1 if respondent is more confident in his/her own pension related capabilities than the average of the population, and 0 otherwise. Measure is comprised of the answers to the following questions (scale 1: strongly disagree to 7: strongly agree): (1) I am very knowledgeable about financial planning for retirement (2): I know more than most people about retirement planning (3): I am very confident in my ability to do retirement planning. Overconfidence: equals 1 if Objectively measured equals 0 and Self-assessed equals 1, and 0 otherwise.

be employed[11] and are slightly more likely to consider themselves as members of a religion or church community.

Finally, for our samples we also observe that the Australian participants are on average slightly less likely to be single, slightly younger than the Dutch, and more likely to have at least one child living at home. Moreover, compared to the Dutch, they tend to be more risk loving. The Dutch are more likely to be born in the country they live in, and score higher for the personality traits extraversion and neuroticism, but lower on the conscientiousness measure.

4 Model description and estimation results

Out experimental survey is designed such that we can investigate the impact of a nudge on drawdown, spending and saving behavior in retirement. Our drawdown nudge consists of the inclusion of the following two sentences to an otherwise identical vignette: *"Government regulations require that they withdraw part of their wealth each year to supplement their income. This corresponds to (...)"*. This allows us to explore the effect of the default drawdown requirement on spending and saving behavior in retirement. The participant's ranking of five saving motives is used to assess the importance of the prompts for implied endorsement ("government advice" and "peer effect") on their spending and saving preferences. The rich set of information we collect for each participant on demographics, preferences, psychological traits and financial knowledge and competencies, allows us to identify the characteristics of those who are more sensitive to the nudges we implement.

To address our research questions, we make use of a (Multinomial) Logit model (see, for example, Cameron and Trivedi, 2005). Our main specification also includes a set of nuisance parameters to account for the different sets five saving motives presented to participants[12]. In Section 4.2 a set of robustness checks analyses the sensitivity of our results to some design elements (importance of the realized ordering of the varying levels of liquid wealth vignettes, and cognitive effort) of the experiment[13].

4.1 Results

We have two main aims. The first aim is to investigate the relative effectiveness of the "peer effect" prompt compared to the "government advice" prompt. A prompt is more effective when it alters the advised spending pattern more often and the adjustment is in the appropriate direction. This aim is investigated in Section 4.1.1. The

[11] This difference is most likely driven by the lower labor force participation rates for women in the Netherlands - see Euwals et al. (2011).

[12] Nuisance parameters are modeled as binary variables to account for the different saving motives presented to the participant, as a result of the randomization procedure. Our results appear robust to a specification without nuisance parameters (cf. column (5) and (11) in Table 7 and Table 8).

[13] As an additional robustness check, we look at the sensitivity of our results to different distributional assumptions (a Probit model and a Linear Probability Model). Our results appear robust to a specification with a different distributional assumption (cf. Table C.3 and Table C.4 in Appendix C).

second aim is to identify which personal characteristics lead individuals to be more sensitive to implied endorsement. Typically, the government's aim when designing a nudge is to protect the vulnerable in society from making decisions which can substantially reduce their utility. Therefore, it is important to investigate whether those who are more sensitive to implied endorsement cues are those who would be in the target group. This aim is investigated in Section 4.1.2. To check whether results can be generalized across countries we present our results by country[14]. Our analyses indicates that, even after controlling for a rich set of covariates, there remain country differences in the effect of the nudge on the advised spending and saving behavior between the Australians and the Dutch (see Table C.1 in Appendix C.1).

4.1.1 For whom is implied endorsement more effective?

A small nudge, provided in two additional sentences in vignette 4, has a substantial impact on the behavior of participants. The descriptive statistics (cf. Section 3) show that around 25% of participants altered their advised spending pattern after presentation of the nudge. Importantly, a substantial proportion indicated that they advised the spending pattern in vignette 4 as a result of the implied endorsement motive (that is, the "government advice" and "peer effect" prompts were highly ranked). Related to the propensity to change the spending pattern, there are two hypotheses we test. First, we hypothesize that the "peer effect" prompt is more likely to alter the advised spending pattern than the "government advice" prompt. We would expect individuals to place greater value on advice from their closest relatives and friends who are more aware of their personal circumstances than from the government, who can only design and communicate one-size-fits-all advice. In addition, we observed from the descriptive statistics in Section 3 that the "peer effect" prompt is likely to be more important than the "government advice" prompt. Second, we hypothesize that the "government advice" prompt is more effective in increasing spending. In the nudge we implement, the government sets the minimum withdrawal level which is equal to the highest advised spending level, therefore the cue from the government is to spend a relatively large amount of retirement wealth. In this subsection we formally test these two hypotheses using a Multinomial Logit regression.

To test these conjectures, we introduce two additional variables. First, we introduce the additional variable *spending pattern* with three outcomes: -1 (1) if the participant advises a higher (lower) spending pattern in vignette 4 (with the implied endorsement) than in vignette 1 (without implied endorsement), and 0 otherwise. Our nudge is constructed in such a way that individuals who follow the implied endorsement given in the additional cue, should be inclined to advise the highest spending pattern. Second, we introduce a *peer ranked above government* dummy variable that equals 1 if the participant ranked the "peer effect" prompt higher than the "government advice" prompt and either the "peer effect" or "government ad-

[14]We now standardize the relevant variables per country, whereas in Table 4 standardization took place based on the analysis sample. Thus, rather than reported in Table 4, the mean (standard deviation) for, for example, "personality: TIPI conscientiousness" in The Netherlands would now equal 0 (1).

vice" motive is ranked in third place or higher, and 0 otherwise. The additional restriction is necessary to reduce noise in this variable. If participants do not want to alter their top three saving motives from vignette 1, there is no clear incentive in vignette 4 to rank the "peer effect" motive over the "government advice" prompt, or vice versa[15].

A Multinomial Logit regression is performed on the propensity to change the advised spending level on a set of explanatory variables, including our variable of interest, the *peer ranked above government* dummy variable. Table 5 presents the average marginal effects per outcome for the key explanatory variables. Columns (1)-(3) presents the results for the regression without additional control variables, whereas columns (4)-(6) presents the results including additional control variables.

First, from Table 5 we observe that the "peer effect" prompt is more likely to alter the advised spending pattern than the "government advice" prompt, as the variable *peer ranked above government* is negative and significant in columns (2) and (5). This holds for the Australian sample as well as the Dutch sample. For the Dutch sample, the size (in absolute value) of the effects of the key variable, that is *peer ranked above government*, is smaller, as expected. Outside the experimental setup, Australians have to make an active choice about the decumulation of their retirement wealth (which is mainly taken as phased withdrawal products known as account-based pensions). In contrast, the Dutch are required to fully annuitize their pension wealth. Hence, it is more likely that (in the real world) Australians discuss their choice options with their peers (e.g. whether or not to follow the government prescribed minimum withdrawal rates) thereby leading to more sizable effects for the *peer ranked above government* dummy.

Second, we do not obtain statistically significant evidence that the "government advice" prompt is more effective at increasing the spending level than the "peer effect" prompt. The estimate of the *peer ranked above government* variable for the increased spending outcome (see columns (3) and (6) of Table 5) is of the expected sign, but not statistically significant, even at a significance level of 90%. We have identified two potential explanations for the lack of significance. First, we hypothesize that the lack of statistically significant results is due to the word "require" in the vignette, thereby potentially suggesting to the participants that the hypothetical households should always be able to withdraw the amount of wealth as suggested by the government, even after they run out of money at advanced age. Unfortunately, we are not able to formally test this possibility with the variables available in our dataset. Second, the lack of statistically significant results may simply be driven by the limited number of observations[16].

[15]Note that the estimation results reported in Table 5 act as a lower bound for the increased / decreased outcome as individuals who advise the highest spending pattern in both vignette 1 and vignette 4 are still included in the country specific analysis sample of Table 5. This concerns 77 Australian participants and 157 Dutch participants. Table C.2 in Appendix C.2 presents the estimation results excluding these individuals.

[16]115 Australian participants increased the advised spending pattern of whom 39 also had the *peer ranked above government* dummy variable equal to 1. In total 211 individuals (implicitly) ranked the "peer effect" prompt higher than the "government advice" prompt and had either the "peer effect" or "government advice" prompt (implicitly) in their top 3.

Controlling for a large set of observable characteristics slightly decreases the size of the effect of the *peer ranked above government* variable on holding advised spending constant, while hardly affecting the propensity to change the advised spending pattern[17] (compare columns (1)-(3) with columns (4)-(6) in Table 5).

Table 5: Propensity to change advised spending pattern

Propensity to change advised spending pattern						
	Australian sample					
	Spending pattern:			Spending pattern:		
	Decreased	Constant	Increased	Decreased	Constant	Increased
	(1)	(2)	(3)	(4)	(5)	(6)
peer ranked above *government*	0.0568**	-0.0948***	0.0380	0.0528**	-0.0773**	0.0245
	(2.14)	(-2.74)	(1.39)	(1.97)	(-2.22)	(0.90)
Control variables				Yes	Yes	Yes
Nuisance parameters	Yes	Yes	Yes	Yes	Yes	Yes
Pseudo R2		0.0163			0.0428	
Observations		983			983	
Log likelihood		-814.1			-792.2	
	Dutch sample					
	Spending pattern:			Spending pattern:		
	Decreased	Constant	Increased	Decreased	Constant	Increased
	(1)	(2)	(3)	(4)	(5)	(6)
peer ranked above *government*	0.0501**	-0.0849**	0.0348	0.0432*	-0.0636*	0.0204
	(2.00)	(-2.55)	(1.38)	(1.70)	(-1.90)	(0.81)
Control variables				Yes	Yes	Yes
Nuisance parameters	Yes	Yes	Yes	Yes	Yes	Yes
Pseudo R2		0.00959			0.0386	
Observations		1437			1437	
Log likelihood		-1104.8			-1072.4	

Notes: *, **, and *** denote significance at 90%, 95%, and 99% respectively. The t-statistics of the average marginal effects are in parentheses. Control variables consists of: Income / wealth related, social network, future orientation, pension capabilities (objective and subjective), and additional control variables.

[17]Note that, by construction, the average marginal effects for a single variable have to sum up to 0 when aggregated over the outcomes.

4.1.2 Who is more sensitive to the implied endorsement prompt?

Libertarian paternalism allows individuals to make their own decisions while providing guidance to prevent the more vulnerable from making decisions which would be disadvantageous to them (that is, would diminish their utility). Therefore, it is important to investigate who is more sensitive to the implied endorsement provided by the cue (that is, the government required minimum drawdown). To do so, we perform a Logit regression with the dependent variable being whether the "peer effect" and "government advice" saving motives respectively are among the top two savings motives nominated by the participant in vignette 4. We undertake regressions for the Australian and Dutch participants separately. Table 6 presents the average marginal effects for the importance of the "peer effect" and "government advice" prompt for two alternative model specifications. The model specifications differ in the measurement of pension capabilities, by including different variables related to this set of covariates[18]. The table displays the results for covariates that are proxies for groups who are more vulnerable to making choices which negatively impact their utility as measured by financial resources, future orientation and pension capabilities. In addition, we investigate the impact of the social network of the participant, as this might explain the effectiveness of the peer effect. Other control variables and nuisance parameters are included in the estimation, but the parameter estimates are not shown in this table.

Financial resources

In line with our expectations, those with less financial resources find the implied endorsement more important in their spending and savings decisions. This is the typical group for whom nudges are designed. The proxies for having high financial resources - belonging to the highest two income categories and owning a house - decrease the likelihood of finding the "government advice" and "peer effect" prompts important. For Australian participants, the effect of homeownership is negligible, as most Australians close to retirement are homeowners. These results do not contradict the findings of Beshears et al. (2016), albeit in a slightly different setting. They find that individuals who were younger than the age of 34 when they were first hired, as well as those with low incomes, are more likely to follow the default, even after controlling for contribution rate preferences.

Future orientation

Another group who are more vulnerable to choose an option which substantially reduces their utility are individuals who do not plan for the future. Such behavior might induce individuals to spend too little, as they might be concerned about outliving their wealth. Indeed, the minimum drawdown rates for retirement wealth in Australia are designed to ensure people spend the savings in their retirement account rather than continue to save and consume less in retirement than during their working life.

[18]The difference in the average marginal effects for the other covariates related to financial resources, social network, and future orientation between the two specifications (compare odd columns (1), (3), (5), (7) with even columns (2), (4), (6) and (8) respectively) is negligible.

This aligns with our survey setting where following the government advice implies the highest spending level during retirement, which reduces savings accordingly. We find that those who plan less are more inclined to follow the implied endorsement (that is, the drawdown required by government) although the effect is not significant. The association of future orientation to finding either the government or peer motive important is negative but not significant. That we do not observe a significant effect might be because being future oriented might have two opposite effects on the saving and spending decision. On the one hand, those who are future oriented probably have thought about their financial arrangements during retirement, and thus are less inclined to find the implied advice and the related saving motive important. On the other hand, those who have a higher present bias, and are thus less future oriented, were already more likely to prefer a higher spending level before the nudge.

Pension capabilities

Another group of individuals are those who have limited capabilities to make the cognitively difficult decision. Therefore, in the odd columns (1), (3), (5) and (7), we include two measures of pension capabilities. The first measure is one where we test knowledge of financial literacy, numeracy and pensions system knowledge and construct a variable based on whether participants answers these questions correctly or not. Table 6 shows that those scoring higher than the average of the (country specific) analysis sample on objectively measured pension capabilities, which include financial literacy, numeracy and pension system knowledge, are less likely to find the "government advice" and "peer effect" prompts important. In a related context, Agnew and Szykman (2005) find that individuals with a low knowledge in financial matters more often opt for a default asset allocation in 401(k) plans. Our findings suggest that financial knowledge has a similar relationship with the importance of "government advice" and "peer effect" prompts.

Similarly, for the Dutch sample, the self-assessed pension knowledge question, although not statistically significant, is inversely associated with the importance of the "government advice" and "peer effect" prompts. This is in line with our expectations, as those who are less confident about their financial capabilities and pension knowledge, are less likely to indicate that either the "government advice" or "peer effect" prompts are among the two most important saving motives. However, Australian participants scoring higher than the average of the analysis sample in the self-assessed pension capabilities measure are associated with a higher likelihood of finding the implied endorsement motive important. For the government advice motive the measure is significant at a 5% significance level.

To test what drives the positive effect for the self-assessed capabilities we re-estimate the model with an overconfidence measure instead (see columns (2), (4), (6) and (8) of Table 6) of the objective and self-assessed measures. This overconfidence measure equals 1 if the participant did not have fewer mistakes in the financial literacy, numeracy and pension knowledge questions than the average of the (country-specific) sample, but is more confident in his or her own pension capabilities than the (country-specific) sample average. The re-estimation leads to sizeable and statistically significant estimates for the overconfidence measure for the Australian

sample[19], although for the Dutch sample the marginal effects are very close to zero. Our nudge implies to advise the highest spending pattern, which exhausts the wealth at age 95 for all income categories. Following the nudge may therefore also be interpreted as risk-seeking behavior[20]. Those who are overconfident are more likely to engage in more risky behavior Indeed, related literature finds that individuals who report being financial literate, are more likely to engage in risk seeking behavior. For example, self-assessed financial literacy is associated with lower allocation to a product offering longevity protection (Bateman et al., 2017).

The country difference in the effect of the overconfidence measure (cf. even columns (2), (4), (6) and (8)) or the difference in sign between countries for the self-assessed capabilities (cf. odd columns (1), (3), (5) and (7)) could relate to differences between countries with respect to their pension system. In the Netherlands, the government mandates a lifelong retirement income payment and does not require a minimum withdrawal rate. However, in Australia, those purchasing phased withdrawal products are required to withdraw a minimum amount to be able to access tax concessions. As Australians are more familiar with phased withdrawal strategies, overconfident individuals might be more inclined to think that they can manage their finances well enough to sustain a high spending level. Therefore, our experimental setting, and the familiarity of the Australian participants to it, might be driving the result of the overconfidence and self-assessed financial literacy measure.

Social network

Having a good social network is a prerequisite for the "peer effect" savings motive. Having more, and closer, connections would potentially increase the likelihood that people discuss financial matters with peers and follow their advice. The social network effect could be captured through several channels, for example marital status, having children living at home, employment status and religion[21]. To control for part-time work, which is the norm for Dutch women (Euwals et al., 2011), but less so for Australian women, we include an interaction term between employment status and gender.

The size and quality of the social network should not impact the importance of the government advice saving motive. Indeed, the variables related to the social network of participants do not have a statistically significant effect on the importance of the "government advice" prompt at conventional levels.

We find that having a social network is important for the "peer effect" savings motive for the Dutch participants, but not those resident in Australia. For the Dutch, the parameter estimates indicate that that being employed, being a female employee,

[19]The difference in the financial resources, social network, and future orientation related variables between the two estimations is negligible.

[20]Note that in the additional control variables we also included a self-reported risk attitude variable. One would expect this to capture the risk attitude rather than the overconfidence measure. However, for our risk attitude variable we find estimates which are not significant at conventional levels and very close to 0.

[21]The channels of the social network through marital status and having children living at home are included as control variables as we consider them as second order effects. These control variables would require additional connections through the partner or children to have an effect on spending and saving decisions. For these variables we do not observe statistically significant effects.

and belonging to a church community significantly affects the importance of the "peer effect" savings motive. First, being employed implies that the participant has a social network including their colleagues, with whom they can discuss financial strategies for retirement. However, since Dutch women are more likely to work part-time, their connection to their colleagues is likely to be less strong and provide less opportunity for discussion of financial matters. For the Dutch sample, although being employed has a positive effect, this effect is reduced for females, as the parameter estimate for being employed and being female is negative[22]. A potential explanation of the significant interaction term between female and employment status using the Dutch sample could relate to social interactions. Peer effects are found to influence (retirement) savings outcomes in various studies (Duflo and Saez, 2003; Beshears et al., 2015; Breza and Chandrasekhar, 2015). These studies, however, only control for gender and do not analyze potential gender specific effects.

4.2 Robustness checks

In this section we discuss what drives the pension capabilities measure. We also analyze the sensitivity of our results to some design elements of the experiment: the importance of the realized ordering of the varying levels of liquid wealth vignettes, and the effect of cognitive effort. Moreover, we assess the sensitivity of our results to the inclusion of the additional control variables and nuisance parameters.

4.2.1 Decomposing objectively measured pension capabilities

Table 6 in Subsection 4.1.2 highlighted that objectively measured pension capabilities are not only significant for both prompts and countries, but they have a sizeable effect. To assess the drivers of this effect we decompose the measure into financial literacy, numeracy and knowledge of the pension system and re-estimate the model. The results are presented in columns (1) and (7) in Table 7 for Australia and Table 8 for the Netherlands.

The results indicate that financial literacy explains the most (except for the "government advise" motive for the Dutch) of the importance of the pension capabilities measure the most. Note that the decomposition leads to a 50% higher marginal effect for the financial literacy measure for the "peer effect" motive in the Dutch sample, than the aggregate measure presented in Table 6. In the Netherlands, those who are more financial literate than the Dutch sample are, even at a 1% significance level, less likely to find the "peer effect" motive important.

The size of the effect of financial literacy is closely followed by the size of the effect

[22] A more formal interpretation of the marginal effects is as follows. For example, on average, a one-unit increase in "Employed", that is the participant is employed as opposed to not employed, increases the probability of the participant to indicate that the "peer effect" prompt is most important in either the first or the second round of best worst by almost 3.74%. Note that this "one-unit" interpretation does not generally hold for continuous variables. However, as we have standardized our continuous variables, this interpretation should therefore remain (more or less) valid.

Table 6: Main results for the importance of the "government advice" and "peer effect" prompts. Average marginal effects based on Logit estimates per country

Covariates associated with the importance of implied endorsement motives (Logit)								
	government advice				peer effect			
	Australia		The Netherlands		Australia		The Netherlands	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Financial resources								
Household income (Q3 and Q4)	-0.0741***	-0.0758***	-0.0115	-0.0188	-0.0597**	-0.0617***	-0.0302*	-0.0404**
	(-3.15)	(-3.20)	(-0.73)	(-1.21)	(-2.52)	(-2.60)	(-1.78)	(-2.39)
Homeowner	-0.00847	-0.0105	-0.0325**	-0.0361***	-0.00425	-0.0130	-0.0226	-0.0268*
	(-0.44)	(-0.55)	(-2.34)	(-2.59)	(-0.17)	(-0.53)	(-1.55)	(-1.84)
Future orientation								
Time horizon: more than 5 years from now	-0.0160	-0.0160	-0.0103	-0.0123	-0.0271	-0.0273	-0.0121	-0.0141
	(-1.07)	(-1.06)	(-0.81)	(-0.97)	(-1.45)	(-1.45)	(-0.91)	(-1.06)
Pension capabilities								
Objectively measured	-0.0531**		-0.0436**		-0.0952***		-0.0509**	
	(-2.51)		(-2.11)		(-3.75)		(-2.42)	
Self-assessed capabilities	0.0348**		-0.00265		0.0258		-0.0119	
	(2.22)		(-0.22)		(1.33)		(-0.94)	
Overconfidence		0.0353**		0.00417		0.0515***		-0.000293
		(2.29)		(0.34)		(2.67)		(-0.02)
Social network								
Employed	0.0285	0.0253	0.00613	0.00515	0.0683	0.0649	0.0374*	0.0360*
	(0.50)	(0.44)	(0.35)	(0.29)	(0.99)	(0.95)	(1.83)	(1.75)
Employed x (gender = female)	-0.0385	-0.0330	-0.00930	-0.00738	-0.0376	-0.0313	-0.0883***	-0.0864***
	(-0.61)	(-0.52)	(-0.35)	(-0.27)	(-0.47)	(-0.39)	(-2.82)	(-2.73)
Religious / Member of a church community	-0.00666	-0.00762	0.00551	0.00539	0.0187	0.0148	0.0273**	0.0265**
	(-0.40)	(-0.45)	(0.44)	(0.43)	(0.93)	(0.73)	(2.16)	(2.10)
Additional control variables	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Nuisance parameters	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Pseudo R2	0.170	0.158	0.136	0.127	0.129	0.113	0.186	0.174
Observations	983	983	1437	1437	983	983	1437	1437
Log likelihood	-189.7	-192.6	-261.8	-264.5	-281.7	-286.7	-271.6	-275.7

Notes: *, **, and *** denote significance at 90%, 95%, and 99% respectively. The t-statistics of the average marginal effects are in parentheses. Control variables consists of: gender = female, marital status, age brackets (50-54; 55-59; 60-64), having children living at home, risk measure (standardized), born in the country they are currently living in, personality (measures of TIPI conscientiousness, extraversion, and neuroticism - standardized). Objectively measured: is a combined measure that equals 1 if respondent made less mistakes than the average of the population in the financial literacy, numeracy, and pension system knowledge related questions, and 0 otherwise. Self-assessed: is a combined measure that equals 1 if respondent is more confident in his/her own pension related capabilities than the average of the population, and 0 otherwise. Measure is comprised of the answers to the following questions (scale 1: strongly disagree to 7: strongly agree): (1) I am very knowledgeable about financial planning for retirement (2): I know more than most people about retirement planning (3): I am very confident in my ability to do retirement planning. Overconfidence: equals 1 if Objectively measured equals 0 and Self-assessed equals 1, and 0 otherwise.

of numeracy, especially for the government advice prompt. The size of the pension system knowledge is typically small. The variable is not statistically significant in any of the regressions.

4.2.2 Design elements of the implied endorsement vignette

Although we have randomized the first part of the experimental task (cf. Section 2.2), some of our estimation results could be driven by ordering effects and cognitive exhaustion. Our survey design allows us to test three conjectures. The results of the conjectures we test are presented in Table 7 and Table 8.

First, if the participant first advised a spending pattern in a familiar liquidity setting, this could lead to different outcomes as opposed to a (similar) participant in an unfamiliar first liquidity setting. To test the first conjecture, we construct an indicator variable that equals one if the first choice task by the Australian (Dutch) participants consisted of the high (low) liquidity vignette, and zero otherwise. From columns (2) and (8) in Table 7 and Table 8 we observe that this variable is not significant and that our results are robust to the first vignette the participant has answered.

Second, participants who were presented with the high liquidity vignette prior to the implied endorsement effect, could respond differently in the implied endorsement vignette as they are more likely to be aware of what they answered in the previous vignette. To test the second conjecture, we construct an indicator variable that equals one if the third vignette presented consisted of the high liquidity vignette, and zero otherwise. From columns (3) and (9) in Table 7 and Table 8 we observe that this variable is not significant and that our results are robust to the prior vignette the participant has answered.

Third, another confounding factor might be cognitive exhaustion. To control for this, we include a variable that captures the respondents' total survey response time. Note that individuals were not required to complete the survey in a fixed period of time[23]. Hence, our distribution of the survey response time is initially skewed towards (very) long response time. To prevent that the estimate for this variable is driven by these outliers, we winsorize the survey duration from the top to 75 minutes. This affects 121 out of 983 Australian participants, and 298 out of 1437 Dutch participants. The estimation results for the main variables of interest are unaffected by the inclusion of this proxy of cognitive exhaustion (see columns (4) and (10) in Table 7 and Table 8). However, in half of the cases the estimate of this proxy for cognitive exhaustion is statistically significant at conventional levels. The statistical negative parameter implies that participants who take more time to answer the survey are less likely to find the implied endorsement savings motive important. This could indicate that they do not suffer from cognitive exhaustion, but take the time to make a thoughtful choice. We interpret this as evidence against the conjecture of cognitive exhaustion affecting our results, as otherwise we would

[23]For example, a respondent could start answering the survey for 30 minutes, then decide to close the browser and finish the survey in an additional 30 minutes a day later. This would be recorded as a survey duration of 1500 (= 30 + 30 + 24*60) minutes.

have expected a similar effect size between the different prompts.

4.2.3 Sensitivity to additional controls and nuisance parameters

To assess whether our results are robust to the inclusion of the nuisance parameters and additional control variables, we present the marginal effects of the re-estimations without it in columns (5) and (6) for the "government advice" motive and columns (11) and (12) for the "peer effect" motive in Table 7 for Australia and Table 8 for the Netherlands.

From columns (5) and (11) we observe that the inclusion of the nuisance parameters does enhance the fit of the model. However, it does not affect the interpretation and significance of the financial resources, social network, future orientation and pension capabilities variables substantially.

From columns (6) and (12) we observe that the non-inclusion of other control variables, besides the clear decrease in overall fit, does affect the significance of our main control variables. Indeed, the size of the objective and self-assessed pension capabilities increases slightly, and more importantly, the significance increases in all cases for the Australian and Dutch sample. On the other hand, the size and significance of the financial resources related measure decreases. For the social network related variables we observe a stark change. Indeed, the non-inclusion of other variables which include the female control, leads to a change in the size and significance of the employed and female employed control. As earlier discussed, this may be due to the tendency of females to work part-time, more than their male counterparts.

Table 7: Robustness using the Australian data (experimental design; decomposition objectively measured pension capabilities; nuisance parameters and additional control variables): covariates associated with the importance of implied endorsement motives (Logit)

	Australia											
	government advice						peer effect					
	(1)	(1)	(2)	(3)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Income / wealth related												
Household income (Q3 and Q4)	-0.0683***	-0.0728***	-0.0756***	-0.0764***	-0.0781***	-0.0788***	-0.0577**	-0.0608**	-0.0594**	-0.0599**	-0.0659***	-0.0620***
	(-2.93)	(-3.11)	(-3.21)	(-3.27)	(-3.26)	(-3.25)	(-2.42)	(-2.57)	(-2.51)	(-2.53)	(-2.73)	(-2.67)
Homeowner	-0.00843	-0.00857	-0.00856	-0.000934	-0.00639	-0.0139	-0.00708	-0.00492	-0.00420	-0.00250	-0.0124	-0.00860
	(-0.45)	(-0.45)	(-0.45)	(-0.05)	(-0.34)	(-0.75)	(-0.29)	(-0.20)	(-0.17)	(-0.10)	(-0.51)	(-0.37)
Future orientation												
Time horizon: more than 5 years from now	-0.0140	-0.0152	-0.0166	-0.0125	-0.0184	-0.0205	-0.0260	-0.0276	-0.0268	-0.0268	-0.0256	-0.0279
	(-0.93)	(-1.01)	(-1.11)	(-0.84)	(-1.21)	(-1.34)	(-1.38)	(-1.48)	(-1.43)	(-1.43)	(-1.34)	(-1.50)
Pension capabilities												
Objectively measured		-0.0519**	-0.0533**	-0.0492**	-0.0546**	-0.0616***		-0.0958***	-0.0951***	-0.0937***	-0.105***	-0.0943***
		(-2.45)	(-2.52)	(-2.34)	(-2.54)	(-2.85)		(-3.78)	(-3.75)	(-3.69)	(-3.94)	(-3.76)
Obj. measured: Financial literacy	-0.0295*						-0.0371*					
	(-1.70)						(-1.73)					
Obj. measured: Numeracy	-0.0289*						-0.0214					
	(-1.75)						(-1.04)					
Obj. measured: Pension system knowledge	-0.0162						-0.0280					
	(-1.01)						(-1.38)					
Self-assessed capabilities	0.0360**	0.0342**	0.0358**	0.0259*	0.0338**	0.0329**	0.0235	0.0276	0.0250	0.0240	0.0283	0.0273
	(2.29)	(2.17)	(2.28)	(1.67)	(2.15)	(2.08)	(1.20)	(1.42)	(1.28)	(1.23)	(1.42)	(1.44)
Social network												
Employed	0.0268	0.0239	0.0331	0.0255	0.0259	0.0295	0.0628	0.0707	0.0671	0.0683	0.0647	0.0848**
	(0.48)	(0.42)	(0.59)	(0.46)	(0.45)	(1.10)	(0.92)	(1.03)	(0.97)	(0.99)	(0.94)	(2.28)
Employed x (gender = female)	-0.0348	-0.0336	-0.0437	-0.0367	-0.0338	-0.0234	-0.0249	-0.0402	-0.0363	-0.0369	-0.0331	-0.0574***
	(-0.55)	(-0.53)	(-0.69)	(-0.59)	(-0.53)	(-1.40)	(-0.31)	(-0.50)	(-0.45)	(-0.46)	(-0.41)	(-2.75)
Religious / Member of a church community	-0.00633	-0.00795	-0.00626	-0.00384	-0.00355	-0.00441	0.0162	0.0196	0.0185	0.0195	0.0236	0.0191
	(-0.38)	(-0.47)	(-0.37)	(-0.23)	(-0.21)	(-0.26)	(0.80)	(0.97)	(0.92)	(0.97)	(1.15)	(0.96)
Robustness: Design elements vign.												
Vignette 3: High liquidity of wealth		0.0218						-0.0238				
		(1.39)						(-1.14)				
Vignette 1: High liquidity of wealth			0.0168						-0.00797			
			(1.08)						(-0.40)			
Survey duration (standardized)				-0.0324***						-0.00594		
				(-3.19)						(-0.59)		
Additional control variables	Yes	Yes	Yes	Yes	Yes	No	Yes	Yes	Yes	Yes	Yes	No
Nuisance parameters	Yes	Yes	Yes	Yes	No	Yes	Yes	Yes	Yes	Yes	No	Yes
Pseudo R2	0.179	0.174	0.173	0.199	0.131	0.116	0.117	0.131	0.129	0.129	0.0698	0.122
Observations	983	983	983	983	983	983	983	983	983	983	983	983
Log likelihood	-187.6	-188.8	-189.1	-183.1	-198.6	-202.1	-285.4	-281.0	-281.6	-281.5	-300.7	-283.8

Notes: *, **, and *** denote significance at 90%, 95%, and 99% respectively. The t-statistics of the average marginal effects are in parentheses. Control variables consists of: gender = female, marital status, age brackets (50-54; 55-59; 60-64), having children living at home, risk measure (standardized), born in the country they are currently living in, personality (measures of TIPI conscientiousness, extraversion, and neuroticism - standardized). Objectively measured: is a combined measure that equals 1 if respondent made less mistakes than the average of the population in the financial literacy, numeracy, and pension system knowledge related questions, and 0 otherwise. Self-assessed: is a combined measure that equals 1 if respondent is more confident in his/her own pension related capabilities than the average of the population, and 0 otherwise. Measure is comprised of the answers to the following questions (scale 1: strongly disagree to 7: strongly agree): (1) I am very knowledgeable about financial planning for retirement (2): I know more than most people about retirement planning (3): I am very confident in my ability to do retirement planning. Overconfidence: equals 1 if Objectively measured equals 0 and Self-assessed equals 1, and 0 otherwise.

Table 8: Robustness using the Dutch data (experimental design; decomposition objectively measured pension capabilities; nuisance parameters and additional control variables): covariates associated with the importance of implied endorsement motives (Logit)

	The Netherlands											
	government advice						peer effect					
	(1)	(1)	(2)	(3)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Income / wealth related												
Household income (Q3 and Q4)	-0.00362	-0.0113	-0.0115	-0.0115	-0.0210	-0.0115	-0.00952	-0.0308*	-0.0300*	-0.0307*	-0.0397**	-0.0318*
	(-0.22)	(-0.72)	(-0.73)	(-0.73)	(-1.32)	(-0.73)	(-0.56)	(-1.81)	(-1.77)	(-1.81)	(-2.28)	(-1.85)
Homeowner	-0.0267*	-0.0326**	-0.0325**	-0.0324**	-0.0348**	-0.0305**	-0.0126	-0.0224	-0.0226	-0.0231	-0.0233	-0.0226*
	(-1.93)	(-2.34)	(-2.34)	(-2.34)	(-2.48)	(-2.33)	(-0.90)	(-1.55)	(-1.56)	(-1.60)	(-1.59)	(-1.67)
Future orientation												
Time horizon: more than 5 years from now	-0.00735	-0.0103	-0.0103	-0.00930	-0.00861	-0.00873	-0.00478	-0.0122	-0.0120	-0.00752	-0.0138	-0.0123
	(-0.58)	(-0.81)	(-0.81)	(-0.73)	(-0.67)	(-0.69)	(-0.37)	(-0.91)	(-0.90)	(-0.56)	(-1.01)	(-0.91)
Pension capabilities												
Objectively measured		-0.0434**	-0.0436**	-0.0426**	-0.0421**	-0.0429**		-0.0511**	-0.0509**	-0.0483**	-0.0475**	-0.0481**
		(-2.11)	(-2.12)	(-2.06)	(-2.00)	(-2.08)		(-2.43)	(-2.42)	(-2.31)	(-2.19)	(-2.27)
Obj. measured: Financial literacy	-0.0250						-0.0721***					
	(-1.60)						(-3.95)					
Obj. measured: Numeracy	-0.0301**						-0.0319**					
	(-2.27)						(-2.39)					
Obj. measured: Pension system knowledge	-0.0101						-0.0153					
	(-0.77)						(-1.10)					
Self-assessed capabilities	-0.00217	-0.00227	-0.00271	-0.00324	-0.000262	-0.00595	-0.0115	-0.0124	-0.0119	-0.0139	-0.00846	-0.0164
	(-0.18)	(-0.18)	(-0.22)	(-0.26)	(-0.02)	(-0.49)	(-0.92)	(-0.97)	(-0.93)	(-1.09)	(-0.65)	(-1.29)
Social network												
Employed	0.00769	0.00654	0.00616	0.00532	0.00637	0.0330**	0.0448**	0.0368*	0.0373*	0.0339*	0.0378*	0.0503***
	(0.44)	(0.37)	(0.35)	(0.30)	(0.36)	(2.37)	(2.23)	(1.81)	(1.82)	(1.67)	(1.80)	(3.50)
Employed x (gender = female)	-0.00555	-0.0104	-0.00926	-0.00775	-0.0123	-0.0599***	-0.0877***	-0.0868***	-0.0883***	-0.0812***	-0.0877***	-0.106***
	(-0.21)	(-0.39)	(-0.35)	(-0.29)	(-0.45)	(-3.27)	(-2.88)	(-2.78)	(-2.82)	(-2.61)	(-2.72)	(-4.60)
Religious / Member of a church community	0.00330	0.00548	0.00546	0.00550	0.00459	0.00478	0.0245**	0.0277**	0.0274**	0.0278**	0.0289**	0.0290**
	(0.27)	(0.44)	(0.44)	(0.44)	(0.36)	(0.39)	(1.98)	(2.19)	(2.17)	(2.22)	(2.26)	(2.32)
Robustness: Design elements vign.												
Vignette 3: High liquidity of wealth		-0.00693						0.0103				
		(-0.55)						(0.81)				
Vignette 1: Low liquidity of wealth			0.00125						-0.00175			
			(0.10)						(-0.13)			
Survey duration (standardized)				-0.00410						-0.0155**		
				(-0.72)						(-2.43)		
Additional control variables	Yes	Yes	Yes	Yes	Yes	No	Yes	Yes	Yes	Yes	Yes	No
Nuisance parameters	Yes	Yes	Yes	Yes	No	Yes	Yes	Yes	Yes	Yes	No	Yes
Pseudo R2	0.148	0.137	0.136	0.137	0.0821	0.111	0.233	0.187	0.186	0.196	0.131	0.147
Observations	1437	1437	1437	1437	1437	1437	1437	1437	1437	1437	1437	1437
Log likelihood	-258.2	-261.7	-261.8	-261.6	-278.2	-269.3	-256.1	-271.3	-271.6	-268.4	-290.0	-284.6

Notes: *, **, and *** denote significance at 90%, 95%, and 99% respectively. The t-statistics of the average marginal effects are in parentheses. Control variables consists of: gender = female, marital status, age brackets (50-54; 55-59; 60-64), having children living at home, risk measure (standardized), born in the country they are currently living in, personality (measures of TIPI conscientiousness, extraversion, and neuroticism - standardized). Objectively measured: is a combined measure that equals 1 if respondent made less mistakes than the average of the population in the financial literacy, numeracy, and pension system knowledge related questions, and 0 otherwise. Self-assessed: is a combined measure that equals 1 if respondent is more confident in his/her own pension related capabilities than the average of the population, and 0 otherwise. Measure is comprised of the answers to the following questions (scale 1: strongly disagree to 7: strongly agree): (1) I am very knowledgeable about financial planning for retirement (2): I know more than most people about retirement planning (3): I am very confident in my ability to do retirement planning. Overconfidence: equals 1 if Objectively measured equals 0 and Self-assessed equals 1, and 0 otherwise.

5 Conclusion

In this paper we design and implement an online experimental survey to investigate the role that two commonly used nudges, implied government endorsement and social norm comparisons, could play in influencing retirement spending and saving decisions. We investigate whether people are influenced by government regulated minimum withdrawals from their pension wealth, whether they perceive such regulations as advice from the government (implied government endorsement) or peers (social norm comparison) and whether the default setting actually influences the vulnerable groups it is designed to target.

Our experimental setup uses two vignettes to present short descriptions of hypothetical retiree households with a given income similar to the state pension level, with accrued retirement savings and given expected to be in good health at age 70. One vignette includes a required minimum pension drawdown set by government whereas the other one excludes it. Participants are asked to advise a hypothetical retiree household on a spending pattern and rank the importance of a set of saving motives consistent with that spending advice. This design enables us to compare retirement spending and saving behavior and motivations with, and in the absence of, the implied endorsement (i.e., the required minimum pension drawdown). The saving motives presented to participants include two prompts for the implied endorsement, communicated as implicit advice from the government ("government knows best") and as a social norm comparison or peer effect ("what most people do"). To enable us to identify the characteristics of those who are sensitive to the nudges we also collect background information on the participant. This includes demographics, preferences, psychological traits and financial knowledge and competencies and country of residence (Australia or the Netherlands) as a proxy for the institutional retirement income system.

Our key findings can be summarized as follows:

- We find that the nudges we implement are effective, with 30% of the Australian participants and 25% of the Dutch participants changing their spending pattern to match the default drawdown pattern. A significant proportion of these indicate that they did so as a result of the "implied endorsement" prompts. Of the two prompts we test, we find peer effects communicated as "what most people do" to be more influential than implicit government advice, communicated as "government knows best" and that these results hold for both the Australian and Dutch participants.

- Our findings indicate that the default setting was followed more by the more vulnerable participants in our sample. As freedom of choice advocates allowing the less vulnerable to make their own decisions, while the aim of a nudge, such as default setting, would be to assist those who are less able to make welfare enhancing choices. Our findings indicate that the default setting was successful in this regard, and was followed by the more vulnerable participants in our sample, specifically those with fewer financial resources, and lower (objectively measured) pension capability. However, we also find a positive association

between following the default and higher "self-confidence". We interpret this as risk-seeking behavior.

- We also find that that peer effects are important, but for the Dutch and not the Australian participants in our experimental survey. The main collected characteristics in the experimental survey associated with the influence of peer effects are being employed and belonging to a church community. Importantly, peer effects were significantly less important for those with less opportunity to build social connections (such as employed Dutch women, who have a greater tendency to work part-time). We conclude that defaults are less likely to be successful for those with a small social network.

Overall, we have provided evidence of the implied endorsement mechanism of the stickiness of default settings in the context of the decumulation of pension wealth. We find that those most influenced by the default are the target vulnerable groups and effects are enhanced those with a strong social network. Given these findings it is very important that default settings and other forms of implied endorsement are well designed to ensure that they are not detrimental to the target groups.

References

Agnew, J. R., Bateman, H., and Thorp, S. (2013). Financial Literacy and Retirement Planning in Australia. *Numeracy*, 6(2):1–25.

Agnew, J. R. and Szykman, L. R. (2005). Asset allocation and information overload: The influence of information display, asset choice, and investor experience. *The Journal of Behavioral Finance*, 6(2):57–70.

Alonso-García, J., Bateman, H., Bonekamp, J., van Soest, A., and Stevens, R. (2017). Saving preferences in retirement: the impact of pension policy design and health status. (CEPAR Working Paper Series No. 2017/19). Available at: https://www.cepar.edu.au/publications/working-papers.

APRA (2017). Annual superannuation bulletin. (APRA Annual Superannuation Bulletin). Available at: http://www.apra.gov.au/Super/Publications/Pages/annual-superannuation-publication.aspx.

Banks, J., Crawford, R., and Tetlow, G. (2015). Annuity choices and income drawdown: evidence from the decumulation phase of defined contribution pensions in England. *Journal of Pension Economics & Finance*, 14(4):412–438.

Bateman, H., Deetlefs, J., Dobrescu, L. I., Newell, B. R., Ortmann, A., and Thorp, S. (2014). Just Interested or Getting Involved? An Analysis of Superannuation Attitudes and Actions. *Economic Record*, 90(289):160–178.

Bateman, H., Eckert, C., Iskhakov, F., Louviere, J., Satchell, S., and Thorp, S. (2016a). Individual Capability and Effort in Retirement Benefit Choice. *Journal of Risk and Insurance*.

Bateman, H., Eckert, C., Iskhakov, F., Louviere, J., Satchell, S., and Thorp, S. (2017). Default and naive diversification heuristics in annuity choice. *Australian Journal of Management*, 42(1):32–57.

Bateman, H., Piggott, J., and Stevens, R. (2016b). *Sustainable retirement income policies in aging society: Lessons from the UK, Australia and the Netherlands*. Mimeo.

Benartzi, S., Beshears, J., Milkman, K. L., Sunstein, C. R., Thaler, R. H., Shankar, M., Tucker-Ray, W., Congdon, W. J., and Galing, S. (2017). Should Governments Invest More in Nudging? *Psychological Science*, 28(8):1041–1055.

Benartzi, S., Previtero, A., and Thaler, R. H. (2011). Annuitization Puzzles. *The Journal of Economic Perspectives*, 25(4):143–164.

Beshears, J., Choi, J., Laibson, D., and Madrian, B. (2008). *The Importance of Default Options for Retirement Saving Outcomes: Evidence from the United States*, pages 59–87. Oxford University Press.

Beshears, J., Choi, J. J., Laibson, D., Madrian, B. C., and Milkman, K. L. (2015). The Effect of Providing Peer Information on Retirement Savings Decisions. *The Journal of Finance*, 70(3):1161–1201.

Beshears, J., Choi, J. J., Laibson, D., Madrian, B. C., and Wang, S. Y. (2016). *Who Is Easier to Nudge?* (Working Paper). Available at: https://scholar.harvard.edu/laibson/publications/who-easier-nudge.

Breza, E. and Chandrasekhar, A. G. (2015). *Social networks, reputation and commitment: Evidence from a savings monitors experiment*. (NBER Working Paper No. 21169). Available at: http://www.nber.org/papers/w21169.

Brown, J. R., Ivković, Z., Smith, P. A., and Weisbenner, S. (2008). Neighbors matter: Causal community effects and stock market participation. *The Journal of Finance*, 63(3):1509–1531.

Bütler, M. and Teppa, F. (2007). The choice between an annuity and a lump sum: Results from Swiss pension funds. *Journal of Public Economics*, 91(10):1944–1966.

Cameron, A. C. and Trivedi, P. K. (2005). *Microeconometrics: Methods and Applications*. Cambridge University Press.

Canada Revenue Agency (2017). Minimum withdrawal factors for registered retirement income funds.

Carroll, G. D., Choi, J. J., Laibson, D., Madrian, B. C., and Metrick, A. (2009). Optimal Defaults and Active Decisions. *The Quarterly Journal of Economics*, 124(4):1639–1674.

Chapman, G. B., Li, M., Leventhal, H., and Leventhal, E. A. (2016). Default clinic appointments promote influenza vaccination uptake without a displacement effect. *Behavioral Science & Policy*, 2(2):40–50.

Chetty, R., Friedman, J. N., Leth-Petersen, S., Nielsen, T. H., and Olsen, T. (2014). Active vs. Passive Decisions and Crowd-Out in Retirement Savings Accounts: Evidence from Denmark. *The Quarterly Journal of Economics*, 129(3):1141–1219.

Choi, J. J., Laibson, D., Madrian, B. C., and Metrick, A. (2002). Defined contribution pensions: Plan rules, participant choices, and the path of least resistance. *Tax Policy and the Economy*, 16:67–113.

Choi, J. J., Laibson, D., Madrian, B. C., and Metrick, A. (2004). For better or for worse: Default effects and 401 (k) savings behavior. In *Perspectives on the Economics of Aging*, pages 81–126. University of Chicago Press.

Choi, J. J., Laibson, D. I., and Madrian, B. C. (2005). Are empowerment and education enough? underdiversification in 401(k) plans. *Brookings Papers on Economic Activity*, 2005(2):151–213.

Dinner, I., Johnson, E. J., Goldstein, D. G., and Liu, K. (2011). Partitioning default effects: why people choose not to choose. *Journal of Experimental Psychology: Applied*, 17(4):332.

Dobrescu, L. I., Fan, X., Bateman, H., Newell, B. R., Ortmann, A., and Thorp, S. (2017). Retirement savings: A tale of decisions and defaults. *The Economic Journal*.

Dohmen, T., Falk, A., Huffman, D., Sunde, U., Schupp, J., and Wagner, G. G. (2011). Individual risk attitudes: Measurement, determinants, and behavioral consequences. *Journal of the European Economic Association*, 9(3):522–550.

Duflo, E. and Saez, E. (2002). Participation and investment decisions in a retirement plan: The influence of colleagues' choices. *Journal of public Economics*, 85(1):121–148.

Duflo, E. and Saez, E. (2003). The Role of Informtaion and Social Interactions in Retirement Plan Decisions: Evidence from a Randomizd Experiment. *Quarterly Journal of Economics*, 118(3):815–843.

Euwals, R., Knoef, M., and Van Vuuren, D. (2011). The trend in female labour force participation: what can be expected for the future? *Empirical Economics*, 40(3):729–753.

Fisher, P. J. and Montalto, C. P. (2011). Loss Aversion and Saving Behavior: Evidence from the 2007 U.S. Survey of Consumer Finances. *Journal of Family and Economic Issues*, 32(1):4–14.

Fry, T., Heaney, R., and McKeown, W. (2007). Will investors change their superannuation fund given the choice? *Accounting & Finance*, 47(2):267–283.

Gallagher, P. (2012). Treasury measurement of retirement income adequacy. Presentation to FEAL Pre-Forum Discussion, February 9, 2011.

Gerrans, P. and Clark-Murphy, M. (2004). Gender differences in retirement savings decisions. *Journal of Pension Economics & Finance*, 3(2):145–164.

Goda, G. S. and Manchester, C. F. (2013). Incorporating employee heterogeneity into default rules for retirement plan selection. *Journal of Human Resources*, 48(1):198–235.

Gosling, S. D., Rentfrow, P. J., and Swann, W. B. (2003). A very brief measure of the Big-Five personality domains. *Journal of Research in Personality*, 37(6):504–528.

Hallsworth, M., List, J. A., Metcalfe, R. D., and Vlaev, I. (2017). The behavioralist as tax collector: Using natural field experiments to enhance tax compliance. *Journal of Public Economics*, 148:14–31.

Hedesstrom, T. M., Svedsater, H., and Garling, T. (2004). Identifying heuristic choice rules in the swedish premium pension scheme. *The Journal of Behavioral Finance*, 5(1):32–42.

Internal Revenue Service (2017). Retirement topics - required minimum distributions (rmds).

Johnson, E. J. and Goldstein, D. (2003). Do Defaults Save Lives? *Science*, 302(5649):1338–1339.

Johnson, E. J., Hershey, J., Meszaros, J., and Kunreuther, H. (1993). Framing, probability distortions, and insurance decisions. *Journal of Risk and Uncertainty*, 7(1):35–51.

Knoef, M., Been, J., Alessie, R., Caminada, K., Goudswaard, K., and Kalwij, A. (2016). Measuring retirement savings adequacy: developing a multi-pillar approach in the netherlands. *Journal of Pension Economics & Finance*, 15(1):55–89.

Lehner, M., Mont, O., and Heiskanen, E. (2016). Nudging–a promising tool for sustainable consumption behaviour? *Journal of Cleaner Production*, 134:166–177.

Lipkus, I. M., Samsa, G., and Rimer, B. K. (2001). General Performance on a Numeracy Scale among Highly Educated Samples. *Medical Decision Making*, 21(1):37–44.

Lusardi, A. and Mitchell, O. S. (2011). Financial literacy around the world: an overview. *Journal of Pension Economics & Finance*, 10(4):497–508.

Madrian, B. C. and Shea, D. F. (2001). The power of suggestion: Inertia in 401 (k) participation and savings behavior. *The Quarterly Journal of Economics*, 116(4):1149–1187.

Mercer (2017). Melbourne mercer global pension index.

MoneySmart (2016). Account-based pensions.

OECD (2015). *Pensions at a Glance 2015 - OECD and G20 indicators*. OECD Publishing, Paris.

Pichert, D. and Katsikopoulos, K. V. (2008). Green defaults: Information presentation and pro-environmental behaviour. *Journal of Environmental Psychology*, 28(1):63–73.

Sunstein, C. R. (2013). Deciding by default. *University of Pennsylvania Law Review*, 162(1):1–57.

Sunstein, C. R. (2014). Nudging: a very short guide. *Journal of Consumer Policy*, 37(4):583–588.

Teppa, F., Thorp, S., and Bateman, H. (2015). *Family, Friends and Framing: A Cross-Country Study of Subjective Survival Expectations*. (De Nederlandsche Bank Working Paper No. 491). Available at: https://ssrn.com/abstract=2710670.

Thaler, R. H. and Sunstein, C. R. (2008). *Nudge: Improving Decisions about Health, Wealth, and Happiness*. New Haven, CT: Yale University Press.

A Dutch and the Australian pension systems

Dutch pension system
The first pillar of the Dutch pension system is a Beveridge type public pension which provides a flat-rate pension to all residents. Its level is linked to the minimum wage and payments are indexed to wages growth. The eligibility age is increasing from 65 years to 67 years over the period 2013 and 2021, and thereafter commencement age will be linked to the remaining life expectancy at age 65.

The second pillar of the Dutch pension system are occupational pension schemes, which are typically part of the labor contract negotiated between unions and employers in collective labor agreements. Employees are obliged to participate and have limited choice in the accrual of pensions. The most popular schemes are defined benefits arrangements where pension accrual and information communication is as income. Despite being referred to as defined benefit, employers have progressively withdrawn from their role as sponsors, leaving the participants bearing the investment and longevity risk. Indexation of occupational pensions is conditional on the funding ratio of the fund, and nominal pension cuts are possible. The typical pension product is a life annuity plus a 70% reversionary annuity for a surviving spouse. There is limited, but increasing, choice in pension plans. On a plan by plan basis, participants can choose to retire early and take the annuity income earlier, and to change the payment schedule to start with higher or lower income than the scheduled amount (reverting to lower/higher income in later years). Around 90% of workers are covered by the second pillar pensions. Most pension plans aim for a gross replacement rate of 70% of average career salary (including first pillar pension benefits) for an individual with 40 years of (full-time) employment (Knoef et al., 2016).

Australian pension system
The first pillar in the Australian pension system is also a Beveridge type public system which provides a means-tested flat pension to all residents (known as the Age Pension). The level (for a single pension) is set at 27% of male average earnings and payments are indexed to the maximum of price and wage inflation. The means-test is comprehensively defined with both an income test and an asset test, but excludes housing wealth. The eligibility age is increasing from 65 years to 67 years between 2017 and 2023.

The second pillar in the Australian pension system is an earnings-related, defined contribution scheme known as the "superannuation guarantee". Employers are required to contribute at least 9.5% of an employee's income into a pension account. Although participation is mandatory, fund members have plethora of choice. They can choose the pension provider, the pension plan, whether to make voluntary contributions (in excess of the minimum 9.5%) and the investment portfolio. In retirement, subject to an access age of 60, the participant can choose at which age to commence decumulation and in what form benefits are taken: a lump sum and/or a phased withdrawal product and/or an annuity. Most people take non-annuitized phased withdrawal products at retirement. Under current policy settings a person on average weekly earnings working for 40 years could expect a replacement rate of

65-70% from an annuitized superannuation accumulation and a part Age Pension (Gallagher, 2012).

B Summary of the experimental design

This Appendix summarizes the key features of the experimental design, including the base text for each vignette (Figure B.1), the characteristics of the eight vignettes and key screen shots.

Figure B.1: Base text for each vignette

The household consists of two individuals currently 65 years old who have just retired. [**INSERT FUTURE HEALTH EXPECTATIONS**].

Each household has a net of tax lifetime income of [**INSERT INCOME**] and their wealth at retirement is [**INSERT WEALTH**]. The household owns the house they live in, without a mortgage. They don't want to move or sell their house. If one member of the household dies, the survivor will receive less income but also spend less. The reduction in income is roughly equivalent to the reduction in spending.

At retirement the household has to plan how much they expect to save and spend, based on their income and current wealth. The following table shows five different spending plans together with income and wealth at different ages (if they survive). If their wealth is exhausted then the household has to adapt their spending to their income. [**INSERT IMPLIED ENDORSEMENT or not**]

Finally, you can assume that prices do not change over time.

Part A:
What spending plan do you advise the household to choose, based on your preferences?
<< Show five different **SPENDING PLANS**, accompanied by a reminder of annual and fortnightly/monthly income, and information about remaining wealth at ages 65, 75, 85, 95 >>

Part B:
Below you see five possible reasons to choose a specific spending plan.

Please indicate which reason is the most important for this household, based on your own preferences, and which saving motive is the least important. Then indicate which saving motive is the 2nd most important and the 2nd least important.
<< Show five different **SAVING MOTIVES** in each choice set, randomly selected from 10 (subject to category restrictions)>>

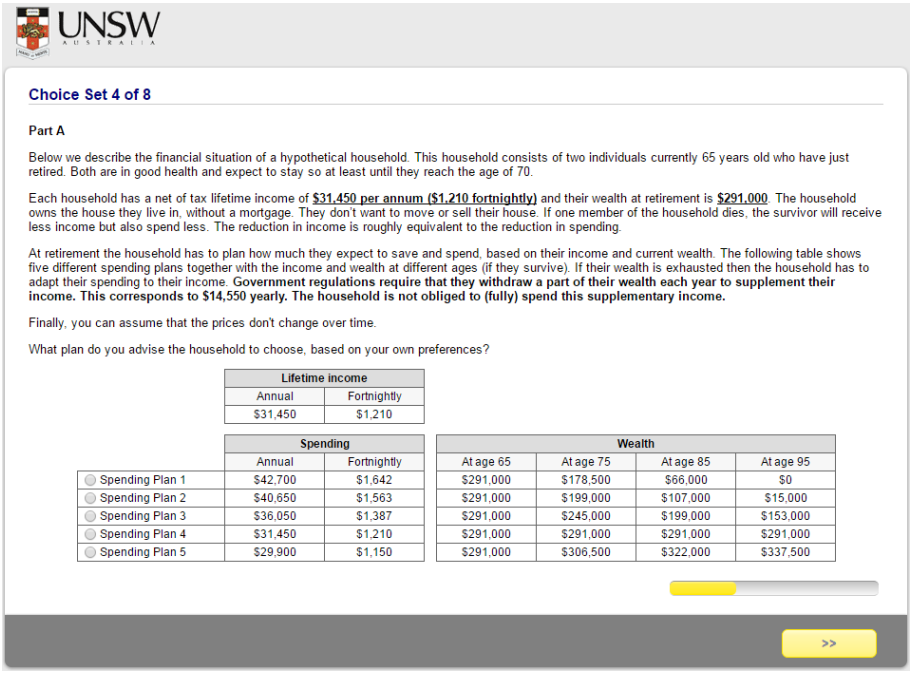
The features of each of the eight vignettes are summarized in Table B.1 below.

Table B.1: Characteristics of the eight vignettes

	Retirement income policy design			**Implied Endorsement**	**Expected future health status**				
	High wealth / Low income	Middle wealth / Middle income	Low wealth / High income	Government required minimum drawdown	Both healthy to 70	Both healthy to 75	One has ADL limitations	Widowed and healthy	Widowed and ADL limitations
Vignette 1	Yes				Yes				
Vignette 2		Yes			Yes				
Vignette 3			Yes		Yes				
Vignette 4	Yes			Yes	Yes				
Vignette 5	Randomly assigned to High wealth / Low income or High wealth / Low incomer					Yes			
Vignette 6	Randomly assigned to High wealth / Low income or High wealth / Low incomer						Yes		
Vignette 7	Randomly assigned to High wealth / Low income or High wealth / Low incomer							Yes	
Vignette 8	Randomly assigned to High wealth / Low income or High wealth / Low incomer								Yes

Notes: For Retirement income policy design, High wealth/Low income is a high liquidity treatment and represents the Australian institutional framework, Low wealth/high income is a low liquidity treatment and represents the Dutch institutional framework, Middle income/middle wealth represents the future policy direction for both Australia and The Netherlands. The exact wording of the alternatives for 'Expected future health status' in each of the vignettes is as follows: *Both healthy to age 70*: Both are in good health and expect to stay so at least until they reach the age of 70. *Both healthy to age 75*: They expect to be in good health until they reach the age of 75. *One has ADL limitations*: They expect that one of them will have difficulties with activities of daily living such as climbing one flight of stairs, driving the car of doing light housework (such as cleaning or dusting) by themselves within 10 years. *Widowed and healthy*: They expect that one of them will die within the next 10 years. Furthermore, they think that the partner will stay in good health until they reach the age of 75. *Widowed and ADL limitations*: They expect that one of them will die within the next 10 years. Furthermore, they think that the partner will have difficulties with activities of daily living such as climbing one flight of stairs, driving the car of doing light housework (such as cleaning or dusting) by themselves within 10 years.

Figure B.2: Screen shot of task 1 (Part A) for the default/implied endorsement vignette ('vignette 4') for the lowest income group in the Australian version of the survey.



UNSW AUSTRALIA

Choice Set 4 of 8

Part A

Below we describe the financial situation of a hypothetical household. This household consists of two individuals currently 65 years old who have just retired. Both are in good health and expect to stay so at least until they reach the age of 70.

Each household has a net of tax lifetime income of **$31,450 per annum ($1,210 fortnightly)** and their wealth at retirement is **$291,000**. The household owns the house they live in, without a mortgage. They don't want to move or sell their house. If one member of the household dies, the survivor will receive less income but also spend less. The reduction in income is roughly equivalent to the reduction in spending.

At retirement the household has to plan how much they expect to save and spend, based on their income and current wealth. The following table shows five different spending plans together with the income and wealth at different ages (if they survive). If their wealth is exhausted then the household has to adapt their spending to their income. **Government regulations require that they withdraw a part of their wealth each year to supplement their income. This corresponds to $14,550 yearly. The household is not obliged to (fully) spend this supplementary income.**

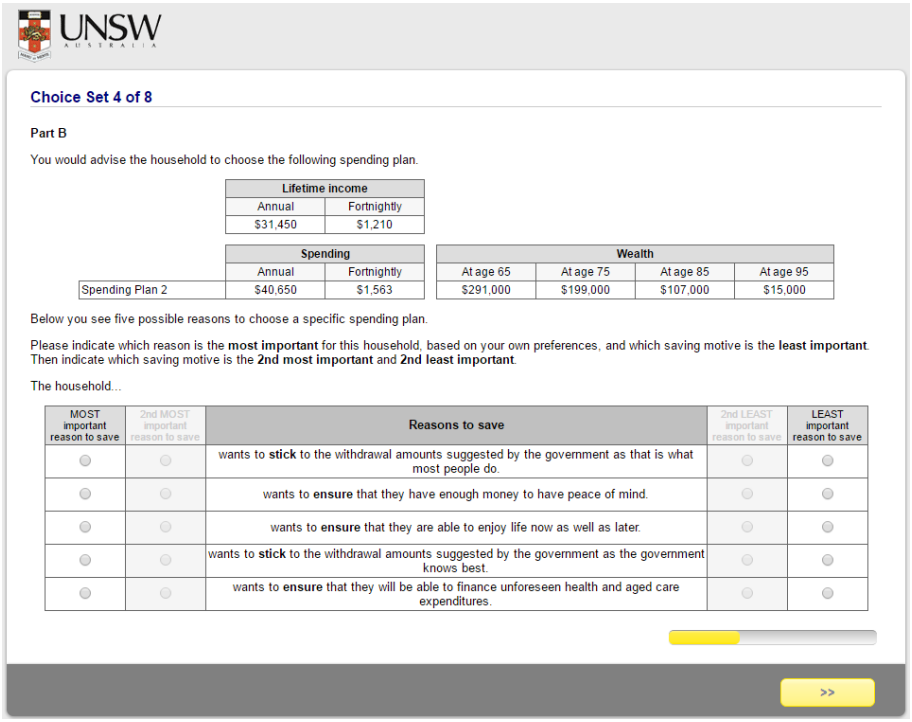
Finally, you can assume that the prices don't change over time.

What plan do you advise the household to choose, based on your own preferences?

	Lifetime income		Spending		Wealth			
	Annual	Fortnightly	Annual	Fortnightly	At age 65	At age 75	At age 85	At age 95
○ Spending Plan 1	$31,450	$1,210	$42,700	$1,642	$291,000	$178,500	$66,000	$0
○ Spending Plan 2			$40,650	$1,563	$291,000	$199,000	$107,000	$15,000
○ Spending Plan 3			$36,050	$1,387	$291,000	$245,000	$199,000	$153,000
○ Spending Plan 4			$31,450	$1,210	$291,000	$291,000	$291,000	$291,000
○ Spending Plan 5			$29,900	$1,150	$291,000	$306,500	$322,000	$337,500

>>

Figure B.3: Screen shot of task 2 (Part B) for the default/implied endorsement vignette ('vignette 4') for the lowest income group in the Australian version of the survey.



UNSW AUSTRALIA

Choice Set 4 of 8

Part B

You would advise the household to choose the following spending plan.

	Lifetime income		Spending		Wealth			
	Annual	Fortnightly	Annual	Fortnightly	At age 65	At age 75	At age 85	At age 95
Spending Plan 2	$31,450	$1,210	$40,650	$1,563	$291,000	$199,000	$107,000	$15,000

Below you see five possible reasons to choose a specific spending plan.

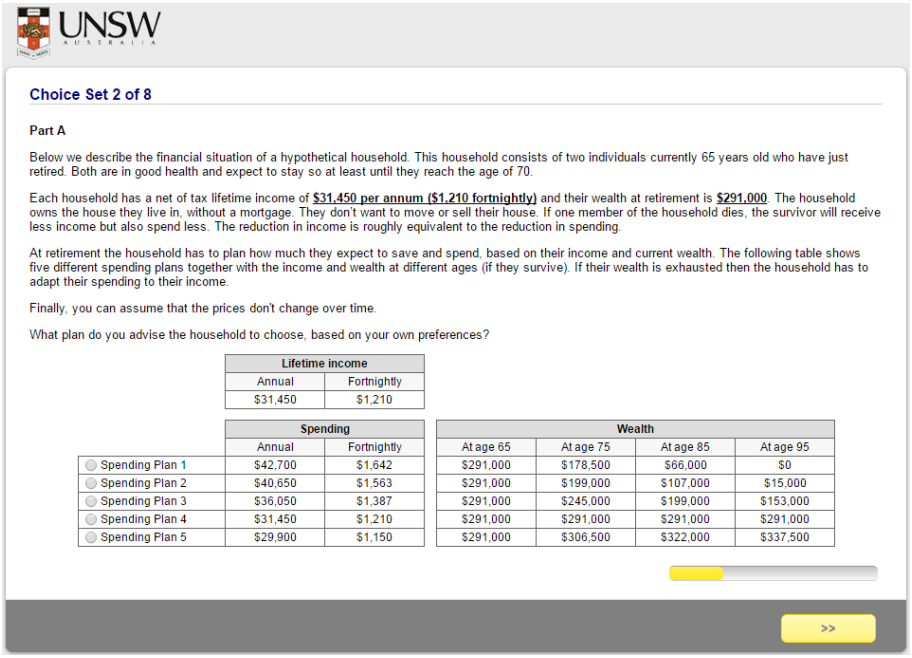
Please indicate which reason is the **most important** for this household, based on your own preferences, and which saving motive is the **least important**. Then indicate which saving motive is the **2nd most important** and **2nd least important**.

The household...

MOST important reason to save	2nd MOST important reason to save	Reasons to save	2nd LEAST important reason to save	LEAST important reason to save
○	○	wants to **stick** to the withdrawal amounts suggested by the government as that is what most people do.	○	○
○	○	wants to **ensure** that they have enough money to have peace of mind.	○	○
○	○	wants to **ensure** that they are able to enjoy life now as well as later.	○	○
○	○	wants to **stick** to the withdrawal amounts suggested by the government as the government knows best.	○	○
○	○	wants to **ensure** that they will be able to finance unforeseen health and aged care expenditures.	○	○

>>

Figure B.4: Screen shot of task 1 (Part A) for the high liquidity vignette ('vignette 1') for the lowest income group in the Australian version of the survey.



UNSW AUSTRALIA

Choice Set 2 of 8

Part A

Below we describe the financial situation of a hypothetical household. This household consists of two individuals currently 65 years old who have just retired. Both are in good health and expect to stay so at least until they reach the age of 70.

Each household has a net of tax lifetime income of **$31,450 per annum ($1,210 fortnightly)** and their wealth at retirement is **$291,000**. The household owns the house they live in, without a mortgage. They don't want to move or sell their house. If one member of the household dies, the survivor will receive less income but also spend less. The reduction in income is roughly equivalent to the reduction in spending.

At retirement the household has to plan how much they expect to save and spend, based on their income and current wealth. The following table shows five different spending plans together with the income and wealth at different ages (if they survive). If their wealth is exhausted then the household has to adapt their spending to their income.

Finally, you can assume that the prices don't change over time.

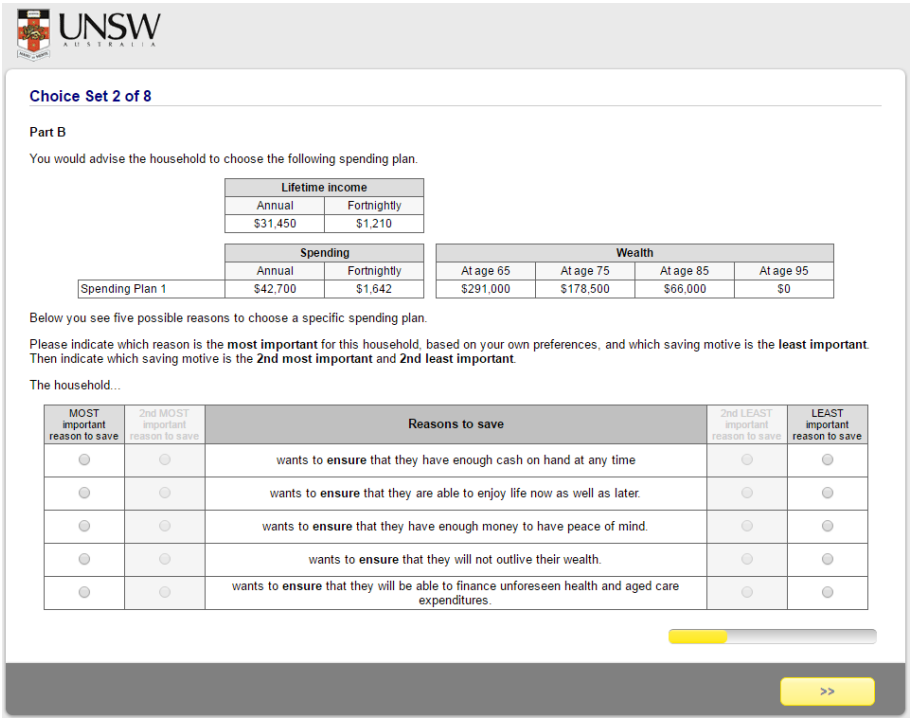
What plan do you advise the household to choose, based on your own preferences?

Lifetime income	
Annual	Fortnightly
$31,450	$1,210

	Spending		Wealth			
	Annual	Fortnightly	At age 65	At age 75	At age 85	At age 95
Spending Plan 1	$42,700	$1,642	$291,000	$178,500	$66,000	$0
Spending Plan 2	$40,650	$1,563	$291,000	$199,000	$107,000	$15,000
Spending Plan 3	$36,050	$1,387	$291,000	$245,000	$199,000	$153,000
Spending Plan 4	$31,450	$1,210	$291,000	$291,000	$291,000	$291,000
Spending Plan 5	$29,900	$1,150	$291,000	$306,500	$322,000	$337,500

>>

Figure B.5: Screen shot of task 2 (Part B) in the high liquidity vignette ('vignette 1') for the lowest income group in the Australian version of the survey.



UNSW AUSTRALIA

Choice Set 2 of 8

Part B

You would advise the household to choose the following spending plan.

Lifetime income	
Annual	Fortnightly
$31,450	$1,210

	Spending		Wealth			
	Annual	Fortnightly	At age 65	At age 75	At age 85	At age 95
Spending Plan 1	$42,700	$1,642	$291,000	$178,500	$66,000	$0

Below you see five possible reasons to choose a specific spending plan.

Please indicate which reason is the **most important** for this household, based on your own preferences, and which saving motive is the **least important**. Then indicate which saving motive is the **2nd most important** and **2nd least important**.

The household...

MOST important reason to save	2nd MOST important reason to save	Reasons to save	2nd LEAST important reason to save	LEAST important reason to save
○	○	wants to **ensure** that they have enough cash on hand at any time	○	○
○	○	wants to **ensure** that they are able to enjoy life now as well as later.	○	○
○	○	wants to **ensure** that they have enough money to have peace of mind.	○	○
○	○	wants to **ensure** that they will not outlive their wealth.	○	○
○	○	wants to **ensure** that they will be able to finance unforeseen health and aged care expenditures.	○	○

>>

C Additional estimation results

C.1 Country specific effect in the importance of the prompts

The average marginal effects of both the "government advice" and the "peer effect" prompts under a Logit specification, using the full sample (both countries) are presented in Table C.1. A full list of the controls can be found in Table 2 in the main text. Even after controlling for a large set of covariates, there remain country specific effects for the importance of the "peer effect" prompt, as indicated by the statistically significant country dummy.

C.2 Propensity to change the advised spending pattern (subsample)

Similar to Section 4.1.1, we investigate the importance of "peer effect" prompt as opposed to the "government advice" prompt in the propensity to change the advised spending pattern. The only difference between Table 5 in Section 4.1.1 and Table C.2 in the Appendix, is the country specific analysis sample. In the Appendix we have excluded individuals who advised "spending pattern 1" both in the high liquidity vignette (vignette 1) and implied endorsement vignette (vignette 4). This reduced the number of participants in the Australian (Dutch) sample by 77 (157) individuals.

Comparing Table 5 and C.2, we observe that the interpretation of the results presented in Section 4.1.1 is robust to this alternative specification. However, it is clear that the exclusion of those who consequently choose the highest spending pattern in the previous specification, of whom its ambiguous on whether they responded to the nudge, increases both the size and significance of the variable *peer ranked above government* slightly.

C.3 Sensitivity to the underlying distribution assumptions

In general, the underlying distribution of the estimation procedure to analyze binary outcome variables, hardly affects the resulting average marginal effects. This is a direct consequence the similar cumulative distribution functions for most of the (same) random variable. However, the tails of these distribution differ. Hence, as our estimation results for the importance of the prompts for implied endorsement ("government advice" and "peer effect") are driven by a relative small number of individuals in our analysis sample, this potentially could affect our results. Therefore, we re-estimate the importance of the cue for implied endorsement using a Probit model and using a Linear Probability Model (LPM) (see, for example, Cameron and Trivedi, 2005).

The result of this procedure for the "government advice" and "peer effect" prompt are presented in Table C.3 and C.4 respectively. We compare these results with

Table C.1: Results for the importance of the "government advice" and "peer effect" prompts using the whole sample. Average marginal effects based on Logit estimates

Covariates associated with the importance of implied endorsement motives (Logit)				
	government advice		peer effect	
	Australia and The Netherlands			
	(1)	(2)	(3)	(4)
Country of Residence				
Australia	0.0159	0.0129	0.0404***	0.0344***
	(1.45)	(1.17)	(3.24)	(2.76)
Financial resources				
Household income (Q3 and Q4)	-0.0465***	-0.0495***	-0.0532***	-0.0584***
	(-3.53)	(-3.76)	(-3.77)	(-4.13)
Homeowner	-0.0201*	-0.0233**	-0.0139	-0.0202
	(-1.84)	(-2.13)	(-1.10)	(-1.58)
Future orientation				
Time horizon: more than 5 years from now	-0.0150	-0.0161	-0.0224**	-0.0243**
	(-1.53)	(-1.63)	(-2.02)	(-2.18)
Pension capabilities				
Objectively measured	-0.0496***		-0.0701***	
	(-3.35)		(-4.35)	
Self-assessed capabilities	0.0166*		0.00426	
	(1.73)		(0.39)	
Overconfidence		0.0212**		0.0209*
		(2.22)		(1.92)
Social network				
Employed	0.00508	0.00394	0.0423*	0.0394*
	(0.30)	(0.23)	(1.88)	(1.74)
Employed x (gender = female)	0.000768	0.00232	-0.0507*	-0.0464
	(0.03)	(0.10)	(-1.79)	(-1.63)
Religious / Member of a church community	0.00201	0.00101	0.0266**	0.0243**
	(0.20)	(0.10)	(2.40)	(2.19)
Additional control variables	Yes	Yes	Yes	Yes
Nuisance parameters	Yes	Yes	Yes	Yes
Pseudo R2	0.114	0.104	0.122	0.108
Observations	2420	2420	2420	2420
Log likelihood	-471.3	-476.7	-582.2	-591.8

Notes: *, **, and *** denote significance at 90%, 95%, and 99% respectively. The t-statistics of the average marginal effects are in parentheses. Control variables consists of: gender = female, marital status, age brackets (50-54; 55-59; 60-64), having children living at home, risk measure (standardized), born in the country they are currently living in, personality (measures of TIPI conscientiousness, extraversion, and neuroticism - standardized). Objectively measured: is a combined measure that equals 1 if respondent made less mistakes than the average of the population in the financial literacy, numeracy, and pension system knowledge related questions, and 0 otherwise. Self-assessed: is a combined measure that equals 1 if respondent is more confident in his/her own pension related capabilities than the average of the population, and 0 otherwise. Measure is comprised of the answers to the following questions (scale 1: strongly disagree to 7: strongly agree): (1) I am very knowledgeable about financial planning for retirement (2): I know more than most people about retirement planning (3): I am very confident in my ability to do retirement planning. Overconfidence: equals 1 if Objectively measured equals 0 and Self-assessed equals 1, and 0 otherwise.

Table C.2: Propensity to change advised spending pattern (reduced sample)

Propensity to change advised spending pattern (reduced sample)						
	Australian sample					
	Spending pattern:			Spending pattern:		
	Decreased	Constant	Increased	Decreased	Constant	Increased
	(1)	(2)	(3)	(4)	(5)	(6)
peer ranked above *government*	0.0686**	-0.119***	0.0506*	0.0617**	-0.0966***	0.0349
	(2.40)	(-3.23)	(1.73)	(2.15)	(-2.61)	(1.19)
Control variables				Yes	Yes	Yes
Nuisance parameters	Yes	Yes	Yes	Yes	Yes	Yes
Pseudo R2		0.0197			0.0515	
Observations		906			906	
Log likelihood		-781.2			-755.9	
	Dutch sample					
	Spending pattern:			Spending pattern:		
	Decreased	Constant	Increased	Decreased	Constant	Increased
	(1)	(2)	(3)	(4)	(5)	(6)
peer ranked above *government*	0.0612**	-0.105***	0.0435	0.0512*	-0.0763**	0.0252
	(2.18)	(-2.83)	(1.54)	(1.81)	(-2.06)	(0.89)
Control variables				Yes	Yes	Yes
Nuisance parameters	Yes	Yes	Yes	Yes	Yes	Yes
Pseudo R2		0.0114			0.0423	
Observations		1280			1280	
Log likelihood		-1049.7			-1016.9	

Notes: *, **, and *** denote significance at 90%, 95%, and 99% respectively. The t-statistics of the average marginal effects are in parentheses. Control variables consists of: Income / wealth related, social network, future orientation, pension capabilities (objective and subjective), and additional control variables.

the findings under the Logit specification presented in Table 6. Despite the slight variation in significance and size of the control variables, the interpretation of the results is robust to the alternative distributional assumptions for the Australian and the Dutch sample for the two studied prompts.

Table C.3: Robustness with respect to the underlying distribution ("government advice" prompt)

Covariates associated with the importance of the "government advice" prompt (probit and LPM)								
	government advice							
	Australia				The Netherlands			
	probit	probit	LPM	LPM	probit	probit	LPM	LPM
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Financial resources								
Household income (Q3 and Q4)	-0.0688***	-0.0708***	-0.0595***	-0.0614***	-0.0101	-0.0167	-0.0112	-0.0169
	(-3.30)	(-3.36)	(-3.82)	(-3.93)	(-0.67)	(-1.13)	(-0.81)	(-1.27)
Homeowner	-0.00815	-0.00985	-0.00926	-0.0101	-0.0309**	-0.0349**	-0.0371**	-0.0403**
	(-0.43)	(-0.51)	(-0.39)	(-0.43)	(-2.19)	(-2.47)	(-2.26)	(-2.43)
Future orientation								
Time horizon: more than 5 years from now	-0.0147	-0.0145	-0.0189	-0.0175	-0.00766	-0.00935	-0.00616	-0.00817
	(-1.00)	(-0.98)	(-1.20)	(-1.13)	(-0.62)	(-0.75)	(-0.52)	(-0.70)
Pension capabilities								
Objectively measured	-0.0529***		-0.0437***		-0.0422**		-0.0383***	
	(-2.70)		(-3.05)		(-2.21)		(-2.84)	
Self-assessed capabilities	0.0349**		0.0421**		-0.00383		0.000193	
	(2.26)		(2.47)		(-0.31)		(0.01)	
Overconfidence		0.0355**		0.0440**		0.00322		0.00754
		(2.33)		(2.40)		(0.27)		(0.57)
Social network								
Employed	0.0314	0.0296	0.0288	0.0260	0.00277	0.00207	-0.0000320	-0.00100
	(0.58)	(0.53)	(0.67)	(0.60)	(0.15)	(0.11)	(-0.00)	(-0.04)
Employed x (gender = female)	-0.0401	-0.0363	-0.0347	-0.0294	-0.00587	-0.00372	0.00259	0.00499
	(-0.65)	(-0.58)	(-0.68)	(-0.57)	(-0.22)	(-0.14)	(0.09)	(0.17)
Religious / Member of a church community	-0.00631	-0.00716	-0.00587	-0.00606	0.00303	0.00348	0.00329	0.00427
	(-0.38)	(-0.44)	(-0.35)	(-0.36)	(0.24)	(0.28)	(0.26)	(0.33)
Additional control variables	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Nuisance parameters	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
(Pseudo) R2	0.175	0.162	0.0763	0.0720	0.136	0.127	0.0637	0.0603
Observations	983	983	983	983	1437	1437	1437	1437
Log likelihood	-188.5	-191.6	41.91	39.63	-261.9	-264.7	141.8	139.2

Notes: *, **, and *** denote significance at 90%, 95%, and 99% respectively. The t-statistics of the average marginal effects are in parentheses. Control variables consists of: gender = female, marital status, age brackets (50-54; 55-59; 60-64), having children living at home, risk measure (standardized), born in the country they are currently living in, personality (measures of TIPI conscientiousness, extraversion, and neuroticism - standardized). Objectively measured: is a combined measure that equals 1 if respondent made less mistakes than the average of the population in the financial literacy, numeracy, and pension system knowledge related questions, and 0 otherwise. Self-assessed: is a combined measure that equals 1 if respondent is more confident in his/her own pension related capabilities than the average of the population, and 0 otherwise. Measure is comprised of the answers to the following questions (scale 1: strongly disagree to 7: strongly agree): (1) I am very knowledgeable about financial planning for retirement (2): I know more than most people about retirement planning (3): I am very confident in my ability to do retirement planning. Overconfidence: equals 1 if Objectively measured equals 0 and Self-assessed equals 1, and 0 otherwise.

Table C.4: Robustness with respect to the underlying distribution ("peer effect" prompt)

Covariates associated with the importance of the "peer effect" prompt (probit and LPM)								
	peer effect							
	Australia				The Netherlands			
	probit	probit	LPM	LPM	probit	probit	LPM	LPM
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Financial resources								
Household income (Q3 and Q4)	-0.0554**	-0.0585**	-0.0529**	-0.0575***	-0.0311*	-0.0408**	-0.0293**	-0.0370***
	(-2.44)	(-2.56)	(-2.47)	(-2.68)	(-1.89)	(-2.50)	(-2.17)	(-2.74)
Homeowner	-0.00696	-0.0164	-0.00672	-0.0155	-0.0230	-0.0275*	-0.0243	-0.0286
	(-0.29)	(-0.67)	(-0.22)	(-0.52)	(-1.59)	(-1.90)	(-1.32)	(-1.55)
Future orientation								
Time horizon: more than 5 years from now	-0.0240	-0.0265	-0.0262	-0.0273	-0.0109	-0.0129	-0.00745	-0.0101
	(-1.30)	(-1.43)	(-1.37)	(-1.43)	(-0.83)	(-0.98)	(-0.61)	(-0.82)
Pension capabilities								
Objectively measured	-0.0972***		-0.0816***		-0.0457**		-0.0444***	
	(-4.08)		(-4.33)		(-2.34)		(-3.04)	
Self-assessed capabilities	0.0252		0.0261		-0.0137		-0.00700	
	(1.31)		(1.29)		(-1.08)		(-0.49)	
Overconfidence		0.0524***		0.0560**		-0.00378		0.00529
		(2.71)		(2.53)		(-0.30)		(0.39)
Social network								
Employed	0.0702	0.0678	0.0498	0.0482	0.0335*	0.0342*	0.0434*	0.0418*
	(1.03)	(0.99)	(0.90)	(0.87)	(1.65)	(1.67)	(1.70)	(1.65)
Employed x (gender = female)	-0.0394	-0.0364	-0.0277	-0.0237	-0.0808***	-0.0800***	-0.0700**	-0.0666**
	(-0.50)	(-0.46)	(-0.44)	(-0.37)	(-2.74)	(-2.69)	(-2.42)	(-2.33)
Religious / Member of a church community	0.0204	0.0167	0.0167	0.0152	0.0263**	0.0269**	0.0295**	0.0306**
	(1.02)	(0.83)	(0.79)	(0.71)	(2.08)	(2.12)	(2.12)	(2.18)
Additional control variables	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Nuisance parameters	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Pseudo R2	0.133	0.115	0.0870	0.0796	0.184	0.172	0.0895	0.0850
Observations	983	983	983	983	1437	1437	1437	1437
Log likelihood	-280.4	-286.2	-174.1	-178.0	-272.4	-276.2	72.96	69.40

Notes: *, **, and *** denote significance at 90%, 95%, and 99% respectively. The t-statistics of the average marginal effects are in parentheses. Control variables consists of: gender = female, marital status, age brackets (50-54; 55-59; 60-64), having children living at home, risk measure (standardized), born in the country they are currently living in, personality (measures of TIPI conscientiousness, extraversion, and neuroticism - standardized). Objectively measured: is a combined measure that equals 1 if respondent made less mistakes than the average of the population in the financial literacy, numeracy, and pension system knowledge related questions, and 0 otherwise. Self-assessed: is a combined measure that equals 1 if respondent is more confident in his/her own pension related capabilities than the average of the population, and 0 otherwise. Measure is comprised of the answers to the following questions (scale 1: strongly disagree to 7: strongly agree): (1) I am very knowledgeable about financial planning for retirement (2): I know more than most people about retirement planning (3): I am very confident in my ability to do retirement planning. Overconfidence: equals 1 if Objectively measured equals 0 and Self-assessed equals 1, and 0 otherwise.