Can Survey Participation Alter Household Financial behavior?*

Thomas Crossley[†] Jochem de Bresser[‡] Liam Delaney[§] Joachim Winter[¶]

October 29, 2013

For review by the Netspar board of the international pension workshop, January 2014. Please do not quote or redistribute.

Abstract

Much empirical research in economics is based on data from household surveys. Panel surveys are particularly valuable for understanding dynamics and heterogeneity. A possible concern with panel surveys is that survey participation itself may alter subsequent behavior. We provide novel evidence of survey effects on household financial behavior in a developed country. We exploit randomized assignment to survey modules within the LISS-panel, an internet panel survey which is representative of the Dutch population. We show that households that respond to detailed questions on expenditures and needs in retirement subsequently reduced their non-housing savings by 1,700 euros (about 3.5 percent of income). Our saving measure is based on linked administrative wealth data, allowing us to distinguish changes in saving behavior from changes in reporting behavior. We document heterogeneity in the effect.

Key words: Survey effects, savings, administrative data, retirement

1 Introduction

Much research in empirical economics and finance relies on the analysis of survey data collected during interviews with individuals and households. Increasingly, models that account for individual or household heterogeneity make use of panel data where people are surveyed repeatedly over time. A wide range of potential survey effects need to be considered to ensure that data collected from such exercises is valid and reliable. A fundamental question is whether confronting respondents with detailed questions about their financial circumstances can alter their later behavior. The

*The authors thank Arthur van Soest, Marcel Das, Pierre-Car Michaud, Rob Alessie, Martin Salm and the participants in seminars at Tilburg University, the University of Groningen, the University of Stirling, Mannheim University, Ludwig Maximilian University of Munich and Statistics Canada for their useful comments.

†Essex University & Institute for Fiscal Studies, London

‡University of Groningen; email: J.de.Bresser@rug.nl

§University of Stirling

¶Ludwig Maximilian University of Munich

extent to which measurement alters the behaviors it is measuring is a core question for many disciplines and one that has not been confronted in depth in economics and finance. We provide the strongest empirical evidence to date on this important question in a central domain of economic behavior.

There are a number of intuitive reasons why survey-induced behavioral change may occur. One is the notion of limited attention, which is common in behavioral finance (see DellaVigna, 2009, for an overview of empirical evidence for limited attention in field experiments). Limited attention means that individuals tend to overlook some of the consequences of their decisions. If those unnoticed consequences materialize in the future, as do the benefits of saving today, this results in biases that are similar to those induced by limited self-control (Karlan et al., 2012). However, in contrast to self-control problems, limited attention suggests that behavior can be corrected by focusing individuals' attention on the aspects they are missing. For instance, Karlan et al. (2012) show that reminders are an effective means to increase savings. In the context of developing countries, they find that reminding people who enrolled in goal-specific savings programs of the benefits of saving increases the amount invested in those accounts and the likelihood of attaining the corresponding goal. Surveys may also serve as shocks to attention. Stango and Zinman (2011) show that survey respondents in the US are less likely to incur overdraft fees, penalties paid to banks in case the balance of a deposit account turns negative, after answering non-informative survey questions that ask for their attitude towards those fees. They find that more general questions on spending control that do not directly concern overdraft fees have the same effect. This suggests that individuals think associatively when it comes to household finance and that relatively subtle reminders of certain aspects of (spending) decisions affect subsequent behavior.

Another literature in psychology proposes that there may be a "question-behavior" effect whereby asking respondents to predict future behavior results in decisions that are in line with those predictions (Dholakia, 2010). One strand of literature has examined a "self-prophecy" effect, whereby committing to a normatively desirable action can increase the likelihood of that action occurring (Sherman, 1980). For example, Spangenberg (1997) shows that asking people to predict their workout behavior induces them to visit the gym more often. Several papers have also shown that surveying people about risky behaviors can increase the propensity to engage in risk behaviors (Fitzsimmons and Moore, 2008; Fitzsimmons and Shiv, 2001; Dholakia, 2010). This literature demonstrates that asking people questions can influence specific behaviors particularly when the question relates to defined time periods. There have even been recommendations in the public health literature to use self-recording of behavior as a potential behavioral change intervention (Michie et al., 2011a,b).

However, survey measurement could also influence reporting of behavior with respondents either becoming fatigued from regularly reporting information and therefore less accurate or becoming more sensitive to perceived appropriate behavior. A body of work on food diary surveys has demonstrated that respondents tend to report fewer episodes of food consumption toward the end of 1-week diary studies (Lillegaard et al., 2007; Biltoft-Jensen et al., 2009).

Furthermore, it is important to disentangle effects that arise from filling out the survey and effects that arise due to differential attrition. Panel conditioning has been discussed in the survey and economics literature for several years. Several studies have examined this potential effect in domains such as subjective well-being (Van Landeghem, 2012); marital satisfaction (Glenn, 1998) and a number of other areas. However, as pointed out by Das et al. (2011), disentangling panel conditioning from unobserved factors influencing attrition is a complex task. Das et al. (2011) find that, controlling for unobserved attrition factors, that there are significant panel conditioning effects on knowledge questions but not in other types of questions.

An important recent paper by Zwane et al. (2011) provides the most compelling evidence to date that surveys can directly impact on objectively observed economic behavior. Their study examined the effect of randomly assigning people to extra survey monitoring in the context of randomized controlled trials conducted across five developing country applications. A significant feature of their study was the ability to monitor objective outcomes, thus ensuring that any survey effects were indeed due to behavioral change rather than changes in reporting style. Strikingly, while they find no effect in two of the experiments, randomly assigning respondents to extra surveys significantly increased the probability of water treatment product usage, medical insurance usage and biased estimates of the health benefits of improved water source quality. As well as being statistically significant, the magnitude of the survey effect on water quality is large enough to obscure the effect of improved source quality on the incidence of diarrhea. Such effects demonstrate that measurement effects on behavior operate distinctively from effects of actually being observed.

We make a number of contributions in this paper. Firstly, we provide the first experimental estimates of the effect of being randomly assigned to a detailed financial survey on later financial behavior in a large representative population survey. In particular, we examine the effect of being asked questions about retirement provision on later savings behavior. Our methodology provides for an almost perfectly clean experimental estimate of the effect and we employ a number of recently developed techniques for dealing with treatment non-compliance to deal with any potential threats to validity.

We test the main hypothesis using the Dutch LISS panel, a household survey that has taken place in the Netherlands since 2007. LISS is particularly useful as it is conducted on a monthly basis assessing a range of financial, health and social outcomes. Crucially for this study, respondents are randomly assigned to be eligible for some of the modules. Thus, we have a strong measure of intention-to-treat when it comes to our main treatment. The treatment itself is a lengthy survey on retirement expectations taken by respondents in January 2008. The survey neither provides any information on the Dutch pension system in general nor on the pension entitlements of respondents. This was the first randomized module introduced on the LISS panel. Our main outcome variable is constructed through linkage, with respondent permission, to the Dutch national tax record system. This records savings and debt across different asset classes. We are thus able to construct a very robust measure of saving between 2007 and 2009 and examine whether the survey randomization influenced real behavior. Thus, as well as providing experimental evidence on a large representative sample, we also can be sure that any experimental effect is affecting behavior as opposed to survey reporting. Furthermore, the nature of the sample allows for exploration of heterogeneous treatment effects. Savings behavior is crucial to economics and finance and providing experimental evidence in such a central domain is particularly relevant.

Our findings demonstrate a very strong, robust and negative effect of the treatment on savings rates in the sample. We demonstrate this through the use of a quantile IV procedure that uses the survey offer as an intention-to-treat variable. There is no conceivable sample selection effect that drives our results. Attrition from the survey and permission to grant access to tax records were both unaffected by the randomization. Furthermore, a falsification test examining the effect of the survey treatment on pre-survey savings behavior yields completely null results. The Dutch retirement system is among the highest ranked in the world in terms of covering post-retirement income, with state and occupational pensions alone covering a median of 70 per cent of highest income post-retirement (Bovenberg and Meijdam, 2001). Therefore, in that institutional context surveying leads mainly to dissaving. To the extent that there are heterogeneous treatment effects, the treatment largely affects respondents who are closer to retirement and better educated. Such heterogeneity is plausible given that older and better educated households have the highest pension entitlements (De Bresser and Knoef, 2013). It should be noted that these are not Hawthorne effects, whereby respondents might change their behavior if aware directly they were being monitored. There is no incentive for respondents to change their behavior in terms of relationships between the participants and the survey agency. Therefore, what we are measuring is a pure survey measurement effect that is particularly concerning for panel studies of financial behavior.

The overall reduction in savings and the effect heterogeneity along the lines of age and education are consistent with limited attention. According to that interpretation, the survey on pensions and retirement increased the salience of those subjects, reminding individuals of issues in the future that they tend to overlook. After reflecting on their expenditure needs in retirement, older and highly educated households concluded that they can afford to save less while the young and poorly educated marginally increased their savings. This reflection was aided by the availability of information on individual-level pension entitlements provided by Universal Pensions Overviews (UPOs), which financial institutions were obliged to provide to all pension holders from 2008 onward. Those UPOs give members of occupational and private pension funds standardized information on current entitlements and projects for age 65 (the statutory retirement age for public pensions during the period covered by the sample).

Our findings are the most robust evidence to date of survey effects on behavior. One implication is that targeted informational interventions may work in the field of household finance as they do in other fields (examples of information affecting household finance are given in Stango and Zinman, 2011, and Karlan et al., 2012). Methodologically, survey effects need to be given far greater consideration in terms of potential biases they introduce to models that require the use of panel data. The results confirm the notion by Zwane et al. (2011) that a survey design which infrequently fields surveys to a large panel may be preferable to one that surveys a smaller sample intensively. Moreover, our research design illustrates the use of randomization of survey modules as a credible identification strategy for the estimation of causal effects. Such randomized modules can be used to study the updating of expectations and the effectiveness of various types of financial education and their effects on financial behavior. In general, randomization in panel surveys provides a relatively cost efficient way to evaluate interventions and test theories, outside the confines of the lab.

The rest of this paper is structured as follows. Section 2 outlines the research design in the study. Section 3 outlines the LISS panel data and administrative data being used. Section 4 describes in details the experimental results, robustness checks and falsification tests. Section 5 concludes with implications for future research.

2 Research design

2.1 Overview

The aim of this paper is to identify the effect of survey participation on household savings. We use survey data from the LISS panel, which is a representative random sample from the Dutch population that was initiated during the autumn of 2007.

The LISS panel is administered by CentERdata, a survey research institute affiliated with Tilburg University, and follows close to 8,000 individuals from 5,000 households. Surveys are distributed over the internet every week. Though the Netherlands has a high rate of internet access, more than 80% of Dutch households are connected, CentERdata safeguards representativeness by providing selected households with an internet subscription and a simple computer when necessary.

In addition to core surveys on subjects such as demographics, income and assets, researchers have the possibility to design their own questionnaires. Those customized modules are usually distributed only once and, to keep costs down, are often limited to a random subsample of the eligible sample. This distribution of modules to random subsets of the LISS panel generates exogenous variation in survey participation, which we exploit to estimate the effect of participation on household financial behavior.

2.2 The treatment

We define treatment as participation in the survey entitled "What is an adequate old age income?". This questionnaire was constructed by Johannes Binswager and Daniel Schunk and it was the first randomized module to be fielded in the LISS panel in January 2008. It consists of around 60 items that primarily concern minimal and desired expenditure levels in retirement, the tradeoff between current and future consumption and risk preferences with respect to income after retirement. Moreover, the questionnaire asks about respondents' willingness to cut back on housing expenditures and the extent to which they have thought about retirement. The questionnaire did not provide respondents with any information about the Dutch system of retirement income provision in general or the respondent's personal entitlements in particular. For our purpose of cleanly identifying a survey effect, it is important to note that the survey did not include any questions on predicted or intended savings. Therefore, our design rules out question behavior: respondents are not asked to predict their own behavior.

Eligibility for the survey was limited to all LISS members that were 25 years or older, who had a net household income of at least 800 euro per month and who were either the head of the household or his/her partner (children or other household members were excluded from participation). This lead to a total eligible sample of 5,435 individuals, 2,755 of which were selected at random and were offered the survey. The take-up rate among those that received the offer was 74%; see Binswanger and Schunk (2012) for more information on the questionnaire and an analysis of the answers it elicited.

2.3 Outcome measures

We investigate the effect of survey participation on household savings and on satisfaction with the state of the Dutch economy. Though the LISS data include elaborate biannual surveys on assets and debt, we prefer to use administrative wealth records for two reasons. Firstly, there is the general concern about the quality of self-reported survey data on assets (Bound et al., 2001). Exploratory analysis reveals that the self-reports of wealth in the LISS data are no exception. Comparing administrative records with self reports from the same households, we find that the reported ownership of asset classes such as stocks and bonds is approximately 10 percentage points lower in the LISS assets module. Furthermore, conditional on ownership panel members tend to understate the value of their assets. Therefore, measurement error is an important reason to prefer tax-derived administrative records over self-reports when it comes to savings. Secondly, we want to rule out the confounding influence of differences in reporting behavior between those who were and were not offered the retirement expenditures survey. If we would find an effect of survey participation on self-reported savings, one could argue that the survey changed reporting styles rather than behavior. Deriving our outcome measures from administrative data mitigates that concern.

Informed consent for the match of LISS data with administrative records was elicited in September of 2011. Unfortunately, panel attrition limits the number of households for which we could obtain a match: out of the 3,125 households that contain at least 1 member that was eligible for the retirement expenditures survey, we could match only 1,602. De Bresser and Knoef (2013) show that this loss of data is mostly due to panel attrition rather than objections: only 10% of the respondents to the retirement expenditures survey that were still in the LISS panel in 2011 objected against the match.

The administrative assets data are taken from the Complete Asset Data of the Netherlands-dataset (Integraal Vermogensbestand, CAD), which was constructed by Statistics Netherlands. The CAD contains a detailed decomposition of household-level wealth for the entire Dutch population. The categories of assets that we observe are checking and saving accounts, bonds, stocks, property, other real estate, business capital and other tangibles. For debt the CAD distinguishes between mortgage and other debt. It measures assets on the first of January for the years 2007, 2008 and 2009 (data for more recent years are not yet released at the time of writing). Available records thus allow us to compute yearly savings during 2007 and 2008 as the differences between wealth stocks in consecutive years. We compute these wealth stocks net of the value of the primary residence, because we want to focus on pure savings and housing has an important consumption component.

The CAD is based on tax records, which are supplemented with information from banks. Though the records provide a measure of assets that is likely to be more accurate than survey data, the fact that they are mostly derived from taxes means they are not complete. We miss savings held in small accounts, because banks are not obliged to report accounts with a balance less than 500 euro or 15 euros in interest payments. We also do not observe debt for households without capital income, which means that we miss most short-term debt. Finally, we miss savings held in tax-exempt 3rd pillar pensions. Such accounts are taxed only during the payout phase and are therefore invisible in tax records up to retirement. Given that the treatment-survey concerns expenditures after retirement, one may expect the effect of participation on savings to be particularly strong for those savings vehicles aimed at generating additional income after retirement. We cover these third pillar accounts by means of specific items from the LISS assets module that ask for ownership and balances of such accounts. Because of the high likelihood of substantial measurement error, we do not add the self-reported private pensions to the tax records. Instead, we analyze them separately and report the findings in section 4.5.

In our analysis of savings we look both at *levels*, in euro per year, and *rates*, which are levels divided by yearly disposable income. The data on the yearly disposable income of households are also taken from tax records. We use the Complete Household Income Data of the Netherlands-dataset (Integraal Huishoudens Inkomstenbestand, CHID), assembled by Statistics Netherlands. The measure for primary income in the CHID is quite complete: in addition to labor income it includes income from entrepreneurship and from assets (interest payments and imputed rent for homeowners). Disposable income is defined as primary income plus government transfers that the household received minus the transfers and taxes paid by the household. The administrative income measure that we use is likely to be more accurate than survey measures of income, since information about the various income streams is provided electronically by employers and financial institutions to the tax authority. Hence, our income data is unlikely to suffer from reporting errors.

Though our focus is on the effect of survey participation on household savings, we also look at the impact of the survey on subjective outcomes that may be relevant to savings. The LISS data provide a rich set of relevant outcome measures. In particular, the income module that was fielded during the summer of 2008 asked respondents about their satisfaction with the economic situation in the Netherlands. Such subjective perceptions may drive savings and therefore we also investigate whether they were affected by the survey.

2.4 Institutional context

In order to understand how a survey about expenditures in retirement might affect savings, we need to describe briefly the institutional context of Dutch pensions. The Dutch system of income provision during retirement is easily understood in terms of 4 categories or "pillars". The first pillar is that of the public pension, which provides everybody who lived in the Netherlands between the ages of 15 and 65 with a subsistence income. Coverage of the public pension is close to universal, since uninterrupted residence in the country is the only criterion (benefits are cut by two percent for each year spent abroad).[1] The level of the public pension is set in reference to the minimum wage. Since public pensions only provide a minimum income, almost all employees accumulate additional entitlements in occupational pensions (the second pillar). Such arrangements cover 90% of all employees and are usually organized at the level of the sector or of the company (Bovenberg and Meijdam, 2001). Participation in the first two pillars is mandatory and together they replace 70 percent of gross last earned income on average, which translates to replacement rates net of taxes above 80 percent (Kapeyn and De Vos, 2008; Bovenberg and Meijdam, 2001). The third pillar contains all private savings vehicles that are aimed specifically at retirement, such as life annuities. Such voluntary arrangements are especially important for individuals that are not included in occupational pensions, such as the self employed. Finally, all other forms of wealth that can be drawn down to generate additional income after retirement, such as savings accounts, investments and real estate, make up the fourth pillar. As explained in the previous section, our assets data allow us to look at the effect of survey participation on the accumulation of assets in this fourth pillar.

It is important to stress that the Dutch pension system in 2008 was characterized by arrangements that were almost universal and provided extremely generous income replacement. In this institutional environment it is not surprising that the first two pillars, public old age pensions and occupational pensions, together provided 95% of income in retirement (Kapeyn and De Vos, 2008). The final 5% was accounted for by private pension products and other savings.

Starting from 2008, individuals can find detailed information on their own pension entitlement in occupational and private funds in their Uniform Pension Overview (UPO). These UPOs pro-

[1]Technically, one is covered by the public pension if one's income is subject to Dutch income taxes. Residence abroad does not affect the accumulation of entitlements as long as your income is taxed within the Dutch system.

vide all members of pension funds, both in the second and third pillar, with yearly updates on their current entitlements and projected future entitlements at age 65. UPOs are mandatory for all financial institutions in the Netherlands since january 1st of 2008.

2.5 Threats to validity

Our analysis faces two threats to internal validity. The first problem is the issue of incomplete compliance with the treatment: not everybody who was offered the survey responded. We apply two remedies. First, we do an intention-to-treat (ITT) analysis that compares those who did receive the offer with those who did not (instead of comparing those who were treated with those who were not). Second, we do instrumental variables (IV) analyses in which we use the random offer of treatment as an instrument for being treated. Both methods allow us to obtain estimates of treatment effects that are not affected by endogenous sample selection as a result of non-response, since they rely on exogenous variation in the survey-offer. In addition to IV regressions for the conditional mean of the savings distribution, we also estimate unconditional decile treatment effects in order to establish the robustness of our results. We use the estimator proposed in Frölich and Melly (2013) (for Stata code implementing this estimator, see Frölich and Melly (2010)). Our estimates can be interpreted as local average treatment effects, since our research design imposes the monotonicity requirement for the effect of the instrument on the likelihood of being treated explained in Angrist et al. (1996): nobody who wasn't offered the survey could participate in it.

The second threat to internal validity is that of selection into outcome measurement due to the substantial loss of observations when we match LISS observations with administrative records. As mentioned above, we could only link 1,602 out of 3,125 eligible households in the LISS panel to administrative records because of attrition in the period between the survey (January 2008) and the match (September 2011). Therefore, it is important that we verify that sample selection is not related to the offer of the retirement expenditures survey. We check this by testing for mean independence of the instrument from sample selection.

3 Data

3.1 Matching LISS and administrative data

The basic unit of our analysis of savings is the household, since we measure both wealth and income at the household level. We classify a household as being offered the survey if at least one household member that was eligible for the survey received the offer of participation. Likewise, we classify all households as treated in which at least one member that was offered the survey actually filled it out. The construction of our estimation sample starts with 3,125 households that contain at least one member that was eligible for the retirement expenditures survey according to the criteria mentioned in section 2.2. After matching the LISS respondents to administrative data, we obtain wealth records for 1,429, 1,437 and 1,449 households in the years 2007-2009 respectively. We drop those households for which all eligible members were retired in 2008, reducing the sample to 1,275 households. Finally, we trim all households for which 2008 savings rates relative to after-tax household income were larger than 50% in absolute value, leaving us with an estimation sample of 999 households.[2]

Table 1 presents descriptive statistics for the full sample and for the estimation sample, separately for couples and singles (note that the full sample consists of 2,816 rather than 3,125 households, because we exclude households for which all eligible members were retired in 2008). For couples, defined as households in which two partners live together irrespective of their marital status, individual-specific attributes are reported for the head of the household. For instance, about 12% of couples have a female head of the household, while around 60% of the singles are females. The mean age is 47 for couples and 45 for singles and couples have more children on average than do singles (1.15 compared with 0.40). Homeownership is prevalent among couples but not among singles: 82% of the former own their home while only 45% of the latter do. Among those living with a partner, over 80% are married compared to 5% among people living without a partner. Around half of the singles were never married, while a little over a third are widowed. Three education levels stand out that account for about 25% of the sample each: intermediate secondary education, intermediate vocational and higher vocational training. Household heads and singles alike are mostly engaged in some form of employment (75% of the former and 70% of the latter are).

In addition to describing the complete LISS data, Table 1 allows one to compare the characteristics of all eligible LISS households with those of the final estimation sample. Both for couples and singles the samples are found to be similar.

One difference between the full sample and the estimation sample that is not in Table 1 is that of compliance to the survey offer. As mentioned in Section 2.2, 74% of the individuals in the complete sample who were offered the survey participated. Household-compliance is around 5 percentage points higher: among those households for which at least one eligible member was offered the survey at least one member filled it out in 79% of the cases. In the estimation sample the corresponding compliance rates are 82% for individuals and 87% for households, which is 8 percentage points higher than in the complete sample. It is not surprising that compliance is related to being observed in our final dataset, since non-compliers to the survey offer are more likely to attrit from the LISS-panel altogether. As a result, non-compliers were less likely to give their permission for the match to administrative records and are lost from our estimation sample. However, this does not compromise our research design, so long as the instrument is orthogonal to this selection process. In Section 4.1 we show that the instrument is not correlated with selection

[2] We also tried trimming the sample at savings rates equal to 75% and 100% of net household income and found that our results are robust. Estimates are reported in Appendix B.

Table 1: Descriptive statistics

	Couples				Singles			
	Full sample (LISS)		Estimation sample		Full sample (LISS)		Estimation sample	
	Mean	(SD)	Mean	(SD)	Mean	(SD)	Mean	(SD)
Female	0.12	(0.32)	0.11	(0.32)	0.58	(0.49)	0.62	(0.49)
Age	47.4	(11.8)	46.9	(11.5)	45.4	(12.1)	44.6	(11.6)
Children	1.16	(1.15)	1.15	(1.16)	0.41	(0.81)	0.36	(0.78)
Homeowner	0.83	(0.38)	0.81	(0.40)	0.49	0.50)	0.42	(0.50)
Marital status								
Married	0.81	(0.39)	0.83	(0.38)	0.06	(0.23)	0.04	(0.19)
Separated/divorced	0.05	(0.23)	0.04	(0.20)	0.33	(0.47)	0.38	(0.49)
Widowed	0.002	(0.04)	0.001	(0.04)	0.10	(0.30)	0.07	(0.25)
Never married	0.13	(0.34)	0.11	(0.32)	0.51	(0.50)	0.52	(0.50)
Education								
Primary	0.08	(0.28)	0.08	(0.28)	0.09	(0.28)	0.09	(0.29)
Int. Secondary	0.23	(0.42)	0.25	(0.43)	0.24	(0.43)	0.26	(0.44)
Higher secondary	0.07	(0.26)	0.08	(0.28)	0.08	(0.27)	0.07	(0.25)
Int. vocational	0.25	(0.43)	0.25	(0.43)	0.23	(0.42)	0.24	(0.43)
Higher vocational	0.25	(0.43)	0.25	(0.43)	0.25	(0.43)	0.28	(0.45)
University	0.11	(0.32)	0.09	(0.29)	0.10	(0.30)	0.06	(0.25)
Most imporant activity								
Employed	0.72	(0.45)	0.75	(0.43)	0.69	(0.46)	0.70	(0.46)
Self employed	0.11	(0.31)	0.07	(0.26)	0.09	(0.29)	0.08	(0.28)
HH work	0.01	(0.11)	0.01	(0.11)	0.05	(0.21)	0.06	(0.23
Retired	0.11	(0.31)	0.11	(0.32)	0.00	(0.00)	0.00	(0.00)
Disabled	0.03	(0.16)	0.03	(0.17)	0.07	(0.26)	0.07	(0.25)
Other	0.03	(0.16)	0.02	(0.15)	0.10	(0.30)	0.09	(0.29)
N	2167 (77.0%)		768 (76.9%)		649 (23.0%)		231 (23.1%)	

For couples all individual-specific variables refer to the head of the household.

into the estimation sample. Hence, the comparison of households based on the random survey offer is as valid there as it would be in the complete sample.

3.2 Descriptive statistics

Table 2 describes our administrative assets records for the years 2007, 2008 and 2009 (all in 2008 euros). The single most important category of assets is that of the primary residence, with an average value of around 200,000 euro. Savings accounts follow at great distance as the second most important type both in terms of mean (27,000 euro) and median (13,000 euro) value. Real estate other than the primary residence is also important, but only for a small minority: the mean value is around 7,000 euro though only 8% of the sample has any non-residential real estate. The mean value of risky assets, stocks and bonds, drops from 7,210 euro in 2007 to 4,857 euro in 2009 (median holdings are zero in all years). Business wealth and other wealth are the least important categories of assets with a mean value below 1,500 euro in all years.

Table 2: Descriptives of assets and debt

	2007			2008			2009		
	Mean	Median	SD	Mean	Median	SD	Mean	Median	SD
Assets									
Saving accounts	25,551	12,799	40,870	26,728	12,728	42,568	28,008	13,196	44,963
Risky assets	7,210	0	23,974	6,627	0	22,560	4,857	0	17,468
Property	196,713	205,571	162,020	201,325	212,589	161,463	199,616	212,181	155,063
Real estate	9,808	0	53,750	6,906	0	41,722	7,689	0	44,442
Business	1,202	0	12,588	1,212	0	13,842	1,459	0	15,485
Other	861	0	9,914	959	0	10,871	1,014	0	10,779
Debt									
Mortgage	105,119	86,091	105,553	104,079	84,827	103,787	108,243	91,412	106,892
Non-mortgage debt	2,375	0	13,796	1,992	0	12,621	2,432	0	18,823
Net worth	133,852	78,760	172,072	137,687	81,551	170,305	131,968	77,021	167,938
Net housing wealth	91,594	40,984	131,904	97,246	47,128	128,481	91,374	43,949	124,627
Net worth excl. housing	42,258	16,292	82,970	40,440	15,643	76,797	40,594	15,232	77,325
N		983			999			999	

All assets are reported in 2008 euros.

As mentioned in section 2.3, we observe both mortgage and non-mortgage debt. On average households have about 105,000-110,000 euro in mortgage debt and around 2,000-2,500 euro of non-mortgage debt. Non-mortgage debt is concentrated in a small minority of 6% of the sample, among which the mean non-mortgage debt is around 20,000 euro.

Taking assets and debt together, the mean net worth of the households in the sample is around 135,000 euro. Unsurprisingly, net worth is concentrated in the primary residence, which has a mean value net of mortgage of around 95,000 euro. Because of the consumption value of housing, we compute savings based on the remaining 40,000 euro of non-housing savings.

Table 3 presents summary statistics of net income and for the outcome variables for savings used in the analysis. Mean household income is 38,165 euro in 2008 and the median is 35,699 euro, both of which are slightly higher than the average of 33,100 euro for the Dutch population at large (Centraal Bureau voor de Statistiek (CBS), 2012). 2008 non-housing savings, computed as the difference between the non-housing wealth stocks of January 1st 2008 and 2009, has a sample mean of 154 euro and a median of 2 euro, showing that the distribution of savings is centered around zero. There is, however, considerable variation in savings: the standard deviation is 9,411 euro. We compute savings rates as the level of savings divided by after-tax income. The distribution of savings rates is centered around zero, but there is considerable variation: the standard deviation of the savings rate is 19 percentage points.

Table 3: Descriptive statistics of outcomes

			Percentiles				
	Mean	Std. dev.	0.05	0.25	0.5	0.75	0.95
HH income[a]	38,165	17,649	16,289	27,056	35,699	46,107	67,474
Non-housing savings[b]							
Levels (2008 euros)	154	9,411	-13,632	-3,221	2	3,084	14,860
Savings rates	-0.01	0.19	-0.40	-0.09	0.00	0.09	0.33
N				999			

[a] HH income net of taxes.
[b] Savings corrected for inflation and net of property value and mortgages.

4 Results

4.1 Validity of the instrument

As explained in section 3.1, we lose about half of our sample when we match LISS records with administrative data. This loss of data compromises the internal validity of our empirical strategy if attrition from the LISS panel is related to our instrument, the offer of the retirement expenditures survey. Therefore, it is important to establish that the offer of the survey is not related to sample selection. Table 4 shows estimates of a linear model that uses our instrument, called ”offer”, to explain an indicator of sample retention. We find that sample selection is not correlated with the offer of the retirement expenditures survey, so the substantial loss of data that comes from matching survey participants to administrative data does not affect our identification.

In the regression in Table 4 and in all other models reported below we control for the presence of multiple household members. This is necessary, because the randomization of the offer was

Table 4: Exogeneity of the instrument w.r.t. sample selection

	Dependent variable: indicator for estimation sample
Offered	-0.0209
	(0.0201)
Multiple eligibles	0.00700
	(0.0215)
Constant	0.364***
	(0.0207)
Number of selected HHs	999
N	2,816

Robust standard errors in parentheses.
*significant at 10%; **significant at 5%; ***significant at 1%

done at the level of the individual while the outcome variables we analyze are measured at the household level. We classify a household as being offered the survey if at least one eligible member received the offer, so by construction households with multiple eligible members are more likely to receive the offer. Conditional on the number of eligible household members, however, randomization across individuals ensures that the offer is random at the household level. We checked whether the randomization was successful by regressing our instrument, an indicator for being offered the survey, on all socio-demographic variables listed in Table 1 (controlling for the presence of multiple household members). All the variables from Table 1 are jointly insignificant, with a P-value equal to 0.901. Hence, there is no evidence to suggest that the randomization failed.

Having established that the instrument is uncorrelated with sample selection and that the randomization succeeded in creating a valid instrument, it remains to be shown that the instrument is relevant for our treatment. The first-stage regression shows that the instrument is highly relevant: the F-statistic for the coefficient of the instrument in a model that controls for the presence of multiple eligibles is 4,818. Complete estimation results for the first stage are given in Appendix A.

4.2 Main results on saving

Figure 1 presents box-and-whiskers diagrams of savings levels (left panel) and savings rates (right panel) by offer-status, separately for couples and singles. This comparison of savings between households that were offered the survey with savings of those that were not constitutes an ITT-analysis, which is valid because of random assignment to the groups. The left figure shows that the median level of savings in households with multiple eligibles is similar regardless of being offered the survey, but that both the 25th and the 75th percentiles are lower for households that did receive the offer. The dispersion for savings is much less among singles and the differences between offered and non-offered households are smaller. In the right panel of Figure 1 we find the same patterns for savings rates: the median savings rate is similar for all households but both

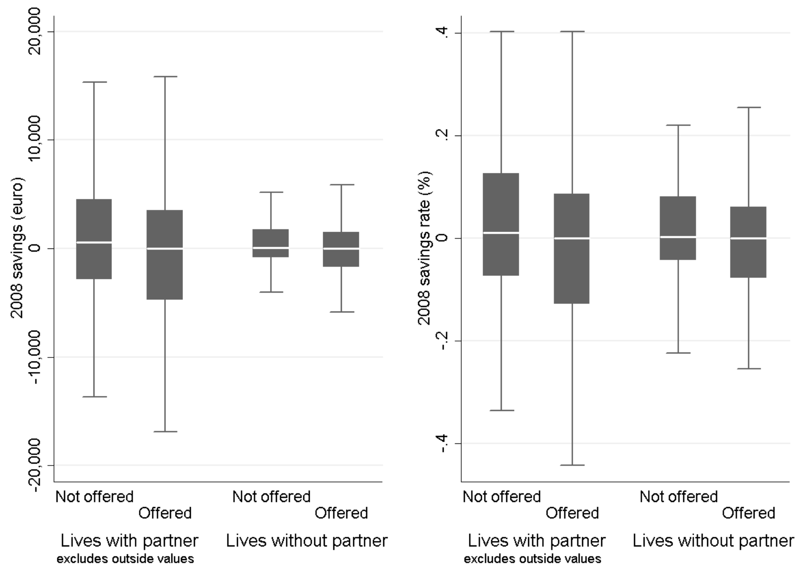

Figure 1: Graphical intention-to-treat analysis.

the 25th and 75th percentiles are substantially lower for those households that were offered the retirement expenditures survey. In contrast to the levels, for savings rates this difference is present both for couples and singles. Our ITT analysis suggests that the random offer of the survey reduced household savings on both sides of the median.

Table 5 presents our main estimation results for the effect of participation in the retirement expenditures survey on savings. The top panel uses 2008 non-housing savings as outcome measure, while the bottom panel explains the 2008 savings rate (non-housing savings divided by household income). The leftmost column shows the estimated coefficients and accompanying standard errors for the treatment dummy in 2SLS models where we instrument survey participation with the random offer of the survey. Participation in the survey caused households to save 1,683 euro less on average during 2008. This is a large effect, especially considering the sample average of 154 euro and the standard deviation of 9,411. When we express savings relative to household income, we also find a significant and negative effect. Survey participation caused households to save 3.5 percentage points less on average, compared with a sample average of -1% and standard deviation of 19 percentage points.

In order to get a better feel for the nature of the effect and its robustness, we check which parts of the savings distribution were affected most strongly by means of IV quantile models. We report estimates for the second up to the eight decile in the remaining columns of Table 5. For the level of savings, we find significant and large effects for the third, sixth and seventh deciles. The estimated coefficients for the other deciles are also all negative. For the non-housing savings rate, we find

Table 5: The effect of survey participation on savings

		Deciles[a]						
	Mean	0.20	0.30	0.40	0.50	0.60	0.70	0.80
	Dependent variable: 2008 non-housing savings (thousands of euros)							
Treated	-1.683**	-1.792	-1.193**	-0.644	-0.474	-0.955**	-1.085**	-0.784
	(0.764)	(1.119)	(0.552)	(0.458)	(0.438)	(0.461)	(0.530)	(0.709)
Sample statistics	0.154	-4.935	-2.061	-0.583	0.002	1.060	2.245	4.393
Proportion compliers	0.875							
N	999							
	Dependent variable: 2008 non-housing savings rate (1 = 100%)							
Treated	-0.0351**	-0.0519*	-0.0337**	-0.0224*	-0.00922	-0.0352**	-0.0247	-0.0317
	(0.0144)	(0.0286)	(0.0166)	(0.0135)	(0.0130)	(0.0141)	(0.0153)	(0.0208)
Sample statistics	-0.01	-0.14	-0.07	-0.02	0.00	0.03	0.07	0.13
Proportion compliers	0.875							
N	999							

[a] For decile models we report unconditional treatment effects.
We control for the presence of multiple eligibles.
Standard errors in parentheses.
*significant at 10%; **significant at 5%; ***significant at 1%

strongly significant effects at the third and sixth decile as well as marginally significant effects at the second and fourth deciles. These estimates show that large parts of the savings distribution were shifted by the survey, with similar effect sizes below and above the median.[3]

The finding that participation in a survey about retirement expenditures reduced savings may seem counterintuitive. One might expect that such survey induces respondents to think about their consumption needs during retirement, which would underline the importance of building a nest egg. However, we believe that the finding makes sense in the context of the Dutch pension system. Remember that in the Netherlands 95% of the entitlement to income during retirement is accumulated in public and occupational pension accounts. Savings in those accounts are mandatory for those who are covered and together they provide generous income replacement relative to the final earned wage. Hence, it is reasonable that respondents give the issue more thought and conclude that there is really no need for the accumulation of additional resources to finance their desired expenditures. De Bresser and Knoef (2013) show that around over 60% of the Dutch accumulate enough resources in non-voluntary pensions alone to meet their minimal level of expenditures (the median household exceeds their consumption floor by 23%). In such an institutional context it is

[3]Assets held in bank accounts and risky investments are provided directly by banks to the tax authority and as a result they are probably measured most accurately. Therefore, we also tried yearly savings in those categories as alternative outcome variables. Appendix C presents the estimates and shows that our results are robust, especially for the savings rates. Moreover, we find no evidence that the survey led people to change their ownership status with respect to risky assets.

not surprising that thinking more about one's consumption needs later in life causes a decline in savings.

4.3 Falsification tests

Our identification is based on the randomized distribution of the retirement expenditures survey to a subset of the eligible panel members. This reliance on an actual randomization allows us to cleanly measure the causal effect of interest. In order to establish the credibility of our identification strategy, we ran the same models reported in Table 5, but now controlling for all covariates described in Table 1. All estimated effects are robust to changing the specification.[4]

As an additional check, we run the same models on 2007 savings. The estimation results are shown in Table 6. We find no evidence for any systematic difference in savings behavior prior to the time of the survey, neither in terms of the average level of savings nor any of the deciles. Note that, in contrast to Table 5, the coefficients of the various deciles are not even of the same sign.

Table 6: Falsification tests

		Deciles[a]						
	Mean	0.20	0.30	0.40	0.50	0.60	0.70	0.80
	Dependent variable: 2007 non-housing savings (thousands of euros)							
Treated	-0.406	-0.582	0.130	0.158	-0.090	-0.393	-0.684	-1.792
	(0.749)	(0.889)	(0.460)	(0.414)	(0.417)	(0.478)	(0.832)	(1.251)
Proportion compliers	0.866							
N	1,014							
	Dependent variable: 2007 non-housing savings rate (1 = 100%)							
Treated	-0.0136	-0.00961	-0.00279	0.00197	-0.00198	-0.00296	-0.0211	-0.0257
	(0.0147)	(0.0249)	(0.0133)	(0.0121)	(0.0122)	(0.0136)	(0.0218)	(0.0298)
Proportion compliers	0.866							
N	1,014							

[a] For decile models we report unconditional treatment effects.
We control for the presence of multiple eligibles.
Standard errors in parentheses.
*significant at 10%; **significant at 5%; ***significant at 1%

4.4 Effect heterogeneity

So far we have shown that participation in the retirement expenditures survey reduced average household non-housing saving, both in absolute terms and relative to household income. Next we investigate effect heterogeneity. One approach would be to run IV analyses on subsamples, but

[4]Estimates available on request.

many variables that could be used for interesting splits of the sample are correlated. Examples are income and education and income and age. Therefore, we prefer a regression approach, where we regress savings on an indicator equal to one if a household was offered the survey (our instrument in the preceding analysis); all covariates shown in Table 1; and interaction terms of covariates with the offer. Hence, we can interpret the results in this section as (heterogeneous) intention-to-treat effects. We investigate heterogeneity along the lines of income, education and age. Note that, for reasons of sample size, the specification does not contain dummies for all cells shown in Table 7, but only interactions of the separate variables with the offer indicator. The upper panel of Table 7 displays coefficient estimates for the main effect and interaction terms of the model with the level of savings as the dependent variable. According to these estimates, the offer of the survey increased average savings of young, income-poor households that are poorly educated by 2,910 euro per year. We find strong evidence for differential effects along the lines of age and education: households with more highly educated or older heads reduced their savings more after filling out the survey. The lower panel of Table 7 shows the differences in savings between offered and non-offered households for subsamples defined along age, education and income categories. Offered households with poorly educated heads who are younger than 40 actually saved 2,900-3,000 euro more than non-offered households with the same education and age. Households in the highest education category saved less regardless of the age of the head and their household income, but we find the strongest effects for older households: the intention-to-treat-effect is -2,400 euro for households below 40; -3,500/-3,600 for age 40-54; and -6,800/-7,000 for households aged 55 or older.

The magnitudes of the effects are similar for households above and below median income, the interaction term of being offered the survey with income is not significant, so we should find smaller differences for high-income households if we divide savings by household income. Table 8 shows exactly that pattern: highly educated households with incomes below the overall median reduced savings by 8-15 percentage points depending on age, while high income households only reduced savings by 3-10 percentage points.

In their analysis of the differences in pension entitlements between socio-economic groups, De Bresser and Knoef (2013) show that education and age are positively correlated with forecasted (occupational) pension entitlements at age 65. Households in which at least one partner has finished a university degree have 37% higher entitlements on average compared to households in which neither spouse finished secondary school. Our analysis of heterogeneous effects indicates that those groups that can look forward to the most generous pensions, namely older and more highly educated households, cut their savings the most after participating in the survey on retirement expenditures. The pension-poor young and poorly educated increased their savings marginally as a result of the survey. These patterns suggest that the survey made respondents reflect on their preparedness for retirement, after which some groups concluded that they could afford to save a little less while others reached the opposite conclusion.

Table 7: Heterogeneous intention-to-treat effects – level of savings

	Dependent variable: 2008 savings (thousands of euros)
Offered	2.910**
	(1.321)
Offered × HH inc. high	0.160
	(1.225)
Offered × educ. middle	-1.027
	(1.195)
Offered × educ. high	-5.315***
	(1.475)
Offered × age 40-54	-1.243
	(1.360)
Offered × age 55+	-4.624***
	(1.779)
Controls	Yes
R-squared	0.0749
N	999

	Heterogeneous effects					
	Income below median			Income above median		
	Education			Education		
	Low	Middle	High	Low	Middle	High
Age <40	2.910**	1.883*	-2.405**	3.070**	2.043	-2.245
	(1.321)	(1.056)	(1.157)	(1.535)	(1.339)	(1.601)
Age 40-54	1.667	0.640	-3.648***	1.827	0.800	-3.488**
	(1.388)	(1.252)	(1.346)	(1.397)	(1.288)	(1.565)
Age 55+	-1.714	-2.741*	-7.029***	-1.554	-2.581	-6.869***
	(1.149)	(1.538)	(1.813)	(1.755)	(2.047)	(2.379)

Standard errors in parentheses.
*significant at 10%; **significant at 5%; ***significant at 1%

Table 8: Heterogeneous intention-to-treat effects – savings rate

	Dependent variable: 2008 savings rate (1 = 100%)
Offered	0.0213
	(0.0300)
Offered × HH inc. high	0.0482*
	(0.0268)
Offered × educ. middle	-0.0150
	(0.0319)
Offered × educ. high	-0.100***
	(0.0319)
Offered × age 40-54	-0.0200
	(0.0294)
Offered × age 55+	-0.0706**
	(0.0330)
Controls	Yes
R-squared	0.0574
N	999

	Heterogeneous effects					
	Income below median			Income above median		
	Education			Education		
	Low	Middle	High	Low	Middle	High
Age <40	0.0213	0.00626	-0.0789***	0.0695*	0.0545*	-0.0307
	(0.0300)	(0.0260)	(0.0290)	(0.0377)	(0.0304)	(0.0298)
Age 40-54	0.00129	-0.0138	-0.0989***	0.0495	0.0344	-0.0507*
	(0.0300)	(0.0308)	(0.0330)	(0.0319)	(0.0280)	(0.0270)
Age 55+	-0.0493*	-0.0643*	-0.150***	-0.00111	-0.0162	-0.101***
	(0.0280)	(0.0348)	(0.0342)	(0.0349)	(0.0371)	(0.0336)

Standard errors in parentheses.
*significant at 10%; **significant at 5%; ***significant at 1%

In addition to the effect heterogeneity reported in Tables 7 and 8, we also checked whether the intention-to-treat effect of the survey offer differs depending on whether the individual(-s) who received the offer is a husband; a wife; both husband and wife; a single male; or a single female. However, we find no evidence for different intention-to-treat effects depending on whether men or women were offered the survey. Moreover, we investigated whether the intensity of the effect was different for respondents that took longer than the median time to answer the questions, but found no evidence to suggest that completion time affected the effect size. Estimates are available on request.

4.5 Evidence from survey data

Though our use of administrative data to reduce measurement error adds to the reliability of the analysis, it also limits our scope. As explained in section 2.3, we do not observe investments in private pension schemes since those funds are tax-exempt during the accumulation phase. However, one might expect that the negative effect of the survey on household non-housing savings is indicative of a re-allocation of assets to accounts that are linked explicitly to retirement. In order to investigate that possibility, we used the LISS assets module to look specifically at assets invested in private pensions. We looked for an effect on ownership, changes in ownership, balance conditional on ownership and unconditional balances and carried out the analysis at the level of the individual and of the household. However, we do not find any evidence for a change in savings in voluntary pension accounts.[5] Unfortunately, we cannot rule out the possibility of measurement error obscuring an actual effect.

We also used self-reports in the LISS to investigate whether participation in the retirement expenditures survey altered the subjective outlook of respondents in ways that are relevant to savings. We find some evidence that survey participation made respondents more satisfied with the economic situation in the Netherlands. However, that effect is not robust to limiting the sample to those individuals that could be matched with administrative records.

5 Conclusion

In this paper we show that participating in a non-informative survey on expenditures after retirement led Dutch households to save significantly less. Our analysis uses administrative wealth data to calculate clean measures of savings that are not contaminated by the reporting styles of survey respondents. Participation in the survey is instrumented by randomized assignment to treatment conditions, so our estimates are unaffected by endogenous compliance. Estimated effects are large: the survey caused households to save 1,700 euro, or 3.5 percentage points relative to disposable income, less during 2008 (sample means: 154 euro and -1%). These effects are robust to various

[5] Estimates available on request.

trimming rules and decile IV models show that they occur on both sides of the median of the distribution of savings. Furthermore, falsification checks reveal no effect on savings prior to the survey, supporting the validity of our identification strategy. We find evidence for heterogenous effects, with the strongest impact among highly educated and older households.

The decline in savings after participating in a survey on consumption during retirement makes sense in the particular institutional context of the Netherlands in 2008, which was characterized by mandatory pensions that were generous, replacing 70% of final income on average, and covered nearly the entire population. Furthermore, De Bresser and Knoef (2013) show that older and more educated households can look forward to higher (occupational) pensions, which fits with the effect heterogeneity we document.

In terms of the mechanism behind these survey effects, limited attention seems to provide a coherent interpretation for our results. The heterogeneity in treatment effects we document is consistent with the survey acting as an attention shock. After being reminded of the tradeoff between current and future consumption, most individuals conclude that their pensions alone will provide them with an adequate income in retirement, resulting in a negative effect on average non-pension savings. However, some groups, namely the young and poorly educated, foresee that they may not be so lucky and do not cut back on savings (or even save a little more). In that sense, our results are comparable to those of Stango and Zinman (2011), who show that financial behavior can be influenced even by questions that do not directly concern that behavior.

Though we do not expect to find similar effects in different institutional contexts, the general point of household financial behavior being affected by participation in household surveys is of considerable importance to empirical economists. It may lead panels that are representative in terms of demographics to behave in non-representative ways. As a result the external validity of any study based on that data would be compromised. If future research confirms that financial decisions of households are susceptible to survey effects, that would be an important reason for researchers to prefer administrative data whenever available. Also, as noted by Zwane et al. (2011), the presence of survey effects suggests that large panels that participate in surveys infrequently may be a better way to get statistical power than small samples that fill out questionnaires on a weekly basis.

Our use of a randomized survey module to identify a causal effect highlights the potential of household panels as laboratories for economic experiments. Randomized questionnaires can be used to investigate how individuals update their expectations and how financial education affects perceptions and behavior. Researchers can use those surveys to identify causal effects of interventions and test economic theories.

One limitation of the present study is the fact that our wealth data only allow us to compute savings during 2007 and 2008. Therefore, we do not know how long-lived the survey effect is. Naturally we will investigate the durability of the effects when more data becomes available. Documenting the range of behaviors that are affected by surveys and the subpopulations that are most sensitive are fruitful areas for future research. Our evidence suggests that economists should follow psychologists and other social scientists and acknowledge the relevance of survey effects, even if that means that our favorite method of data collection is simultaneously more versatile and less innocent than we would like to believe.

References

Angrist, J., Imbens, G., and Rubin, D. (1996). Identification of causal effects using instrumental variables. *Journal of the American Statistical Association*, 91(434):444–455.

Biltoft-Jensen, A., Matthiesen, J., Rasmussen, L., Fagt, S., Groth, M., and Hels, O. (2009). Validation of the Danish 7-day pre-coded food diary among adults: energy intake v. energy expenditure and recording length. *British Journal of Nutrition*, 102:1838–1846.

Binswanger, J. and Schunk, D. (2012). What is an adequate standard of living during retirement? *Journal of Pension Economics and Finance*, 11(2):203–222.

Bound, J., Brown, C., and Mathiowetz, N. (2001). Measurement error in survey data. In Heckman, J. and Leamer, E., editors, *Handbook of Econometrics*, volume 5, pages 3705–3843. Elsevier, North-Holland, Amsterdam.

Bovenberg, A. and Meijdam, L. (2001). The Dutch pension system. In Borsch-Supan, A. H. and Miegel, M., editors, *Pension Reform in Six Countries*. Springer, New York.

Centraal Bureau voor de Statistiek (CBS), D. (2012). Welvaart in nederland - inkomen, vermogen en bestedingen van huishoudens en personen.

Das, M., Toepoel, V., and Van Soest, A. (2011). Nonparametric tests of panel conditioning and attrition bias in panel surveys. *Sociological Methods & Research*, 40(1):32–56.

De Bresser, J. and Knoef, M. (2013). Can the Dutch meet their own retirement expenditure goals? Mimeo, Tilburg University.

DellaVigna, S. (2009). Psychology and economics: evidence from the field. *Journal of Economic Literature*, 47(2):315–372.

Dholakia, U. (2010). A critical review of question-behavior effect research. *Review of Marketing Research*, 7(7):145–197.

Fitzsimmons, G. and Moore, S. (2008). Should we ask our children about sex, drugs and rock & roll? potentially harmful effects of asking questions about risky behaviors. *Journal of Consumer Psychology*, 18(2):82–95.

Fitzsimmons, G. and Shiv, B. (2001). Nonconscious and contaminative effects of hypothetical questions on subsequent decision making. *Journal of Consumer Research*, 28(2):224–238.

Frölich, M. and Melly, B. (2010). Estimation of quantile treatment effects with stata. *The Stata Journal*, 10(3):423–457.

Frölich, M. and Melly, B. (2013). Unconditional quantile treatment effects under endogeneity. *Journal of Business and Economic Statistics*, just-accepted.

Glenn, N. (1998). The course of marital success and failure in five American 10-year marriage cohorts. *Journal of Marriage and Family*, 60(3):569–576.

Kapeyn, A. and De Vos, K. (2008). Social security and retirement in the Netherlands. In Gruber, J. and Wise, D. A., editors, *Social Security and Retirement Around the World*, pages 269–304. University of Chicago Press.

Karlan, D., McConnell, M., Mullainathan, S., and Zinman, J. (2012). Getting to the top of mind: How reminders increase saving. No. w16205. National Bureau of Economic Research.

Lillegaard, I., Løken, E., and Andersen, L. (2007). Relative validation of a pre-coded food diary among children, under-reporting varies with reporting day and time of the day. *European Journal of Clinical Nutrition*, 61(1):61–68.

Michie, S., Churchill, S., and West, R. (2011a). Identifying evidence-based competences required to deliver behavioural support for smoking cessation. *Annals of Behavioral Medicine*, 41(1):59–70.

Michie, S., Hyder, N., Walia, A., and West, R. (2011b). Development of a taxonomy of behaviour change techniques used in individual behavioural support for smoking cessation. *Addictive Behaviors*, 36(4):315–319.

Sherman, S. (1980). On the self-erasing nature of errors of prediction. *Journal of Personality and Social Psychology*, 39(2):211–222.

Spangenberg, E. (1997). Increasing health club attendance through self-prophecy. *Marketing Letters*, 8(1):23–31.

Stango, V. and Zinman, J. (2011). Limited and varying consumer attention: Evidence from shocks to the salience of bank overdraft fees. No. w17028. National Bureau of Economic Research.

Van Landeghem, B. (2012). Panel conditioning and self-reported satisfaction: evidence from international panel data and repeated cross-sections. SOEPpapers on Multidisciplinary Panel Data Research.

Zwane, A., Zinman, J., Dusen, E. V., Pariente, W., Null, C., Miguel, E., Kremer, M., Karlan, D., Hornbeck, R., Gine, X., Duflo, E., Devoto, F., Crepon, B., and Banerjee, A. (2011). Being surveyed can change later behavior and related parameter estimates. *Proceedings of the National Academy of Sciences*, 108(1):1821–1826.

Appendix A First Stage

Table 9: First stage

	Dependent variable: HH treated
HH offered	0.879***
	(0.0127)
Multiple eligibles	-0.0376**
	(0.0159)
Constant	0.0231**
	(0.00988)
R squared	0.688
F(1, n-(k+1))	4,818.37***
N	999

Robust standard errors in parentheses.
*significant at 10%; **significant at 5%; ***significant at 1%

Appendix B Estimates Under Different Trimming Rules

Table 10: Robustness checks with different trimming rules

		Deciles[a]						
	Mean	0.20	0.30	0.40	0.50	0.60	0.70	0.80
					Sample trimmed at savings rates between -100% and +100%			
					Dependent variable: 2008 non-housing savings (thousands of euros)			
Treated	-1.862*	-1.282	-1.498*	-1.017*	-0.124	-0.913*	-1.085*	-0.510
	(1.085)	(1.440)	(0.809)	(0.544)	(0.507)	(0.495)	(0.575)	(0.829)
Proportion compliers					0.873			
N					1,124			
					Dependent variable: 2008 non-housing savings rate (1 = 100%)			
Treated	-0.0343	-0.0803*	-0.0464**	-0.0284*	-0.00260	-0.0281*	-0.0243	-0.0220
	(0.0216)	(0.0418)	(0.0217)	(0.0157)	(0.0147)	(0.0149)	(0.0162)	(0.0234)
Proportion compliers					0.873			
N					1,124			
					Sample trimmed at savings rates between -75% and +75%			
					Dependent variable: 2008 non-housing savings (thousands of euros)			
Treated	-1.900**	-0.828	-1.306*	-0.828	-0.0301	-0.975**	-1.284**	-0.771
	(0.868)	(1.378)	(0.722)	(0.510)	(0.483)	(0.481)	(0.581)	(0.804)
Proportion compliers					0.88			
N					1,080			
					Dependent variable: 2008 non-housing savings rate (1 = 100%)			
Treated	-0.0356*	-0.0615*	-0.0399**	-0.0270*	-0.00139	-0.0332**	-0.0277*	-0.0273
	(0.0182)	(0.0364)	(0.0200)	(0.0151)	(0.0141)	(0.0147)	(0.161)	(0.0226)
Proportion compliers					0.88			
N					1,080			

[a] For decile models we report unconditional treatment effects.
We control for the presence of multiple eligibles.
Standard errors in parentheses.
*significant at 10%; **significant at 5%; ***significant at 1%

Appendix C Financial Savings (Savings Accounts and Risky Assets)

In this appendix we present the results of an analogous analysis to that in the main text, which uses a different outcome variable. Here we analyze yearly financial savings, defined as the yearly differences of the sum of assets in savings accounts and risky assets. Information on savings accounts and risky assets holdings is provided directly by banks to the tax authority, whereas assets in other categories are reported by households. Though misreporting those other assets is a criminal offense, direct provision of account information by banks means that financial assets are probably measured even more precisely. Therefore, we report complementary results in Table 11 to substantiate our main findings.

Table 11: Alternative outcome variable: financial savings (savings accounts and risky assets)

		Deciles[a]						
	Mean	0.20	0.30	0.40	0.50	0.60	0.70	0.80
	2008 financial savings[b]							
	Dependent variable: financial savings (thousands of euros)							
Treated	-1.055	-1.818	-0.836*	-0.528	-0.0768	-0.852*	-0.833	-0.751
	(0.646)	(1.119)	(0.508)	(0.430)	(0.419)	(0.439)	(0.521)	(0.677)
Proportion compliers	0.865							
N	1,043							
	Dependent variable: financial savings rate (1 = 100%)							
Treated	-0.0248*	-0.0534**	-0.0308**	-0.0190	-0.00207	-0.0269**	-0.0217	-0.0209
	(0.0139)	(0.0273)	(0.0158)	(0.0128)	(0.0125)	(0.0129)	(0.0142)	(0.0177)
Proportion compliers	0.865							
N	1,043							
	Falsification check: 2007 financial savings[b]							
	Dependent variable: financial savings (thousands of euros)							
Treated	-0.040	-1.305*	-0.234	0.000	-0.0236	-0.198	-0.797	-1.005
	(0.639)	(0.738)	(0.416)	(0.373)	(0.378)	(0.431)	(0.767)	(1.081)
Proportion compliers	0.862							
N	1,066							
	Dependent variable: financial savings rate (1 = 100%)							
Treated	-0.0110	-0.0254	-0.0107	0.00135	9.92e-06	-0.00138	-0.0164	-0.0283
	(0.0137)	(0.0222)	(0.0124)	(0.0109)	(0.0109)	(0.0121)	(0.0192)	(0.0261)
Proportion compliers	0.862							
N	1,066							

[a] For decile models we report unconditional treatment effects.
[b] Savings are trimmed at -50% and +50% of disposable household income.
We control for the presence of multiple eligibles.
Standard errors in parentheses.
*significant at 10%; **significant at 5%; ***significant at 1%

Table 11 contains estimates for 2008 financial savings (top panel) and the corresponding falsification check using 2007 financial savings (bottom panel). Analogously to our analysis of non-housing savings in the main text, we trim households with financial savings below -50% or above 50% of disposable household income. When we look at the *level* of financial savings, we find smaller, less significant results than for non-housing savings: only the third and sixth decile show marginally significant effects of survey participation around -800/-850 euro per year. However, for the financial savings *rate*, defined as yearly savings divided by yearly household disposable income, there is stronger evidence of a survey effect. The estimated mean effect is marginally significant and equal to -2.48 percentage points. Moreover, the decile treatment effects indicate that this survey effect is most pronounced in the bottom of the savings rates distribution: the second and third deciles are reduced by 3-5 percentage points. The bottom panel of Table 11 shows that financial savings, like non-housing savings, pass the falsification test: the estimated treatment effects prior to the treatment are generally small compared to those after the treatment and not significantly different from zero.

We also checked whether there is any evidence that survey participation caused respondents to reallocate their portfolio towards or away from risky assets, but we find no evidence to support this possibility at the extensive margin. Instrumental variables models show that there is no significant effect of survey participation on the likelihood of owning risky assets during any of the years 2007-2009. Because it is impossible to verify to what extent changes in the value of the stock of risky assets result from the decision to buy or sell rather than from variation in stock and bond prices, we ran the analysis on changes in the balance of bank accounts. Table 12 contains the corresponding estimates and shows that the result of a negative effect of the survey on savings is corroborated by the quantile models ($p = 0.103$ in the model for the mean savings rate). However, the part of the distribution of savings that is affected is different compared to the combined bank accounts/risky assets outcome analyzed above. When we combined assets from those categories, we found the largest effects below the median, at the third and fourth decile. If we focus exclusively on funds in bank accounts, the largest effects occur at the median and sixth and seventh deciles.

Table 12: Alternative outcome variable: savings in bank accounts (without risky assets)

	Mean	Deciles[a] 0.20	0.30	0.40	0.50	0.60	0.70	0.80
					2008 savings in bank accounts[b]			
					Dependent variable: savings in bank accounts (thousands of euros)			
Treated	-0.840	-0.412	-0.712	-0.429	-0.774*	-0.939**	-0.989*	-0.218
	(0.635)	(1.131)	(0.479)	(0.401)	(0.398)	(0.427)	(0.506)	(0.689)
Proportion compliers					0.866			
N					1,078			
					Dependent variable: savings rate for bank accounts (1 = 100%)			
Treated	-0.0219	7.81e-06	-0.0220	-0.0149	-0.0253**	-0.0311**	-0.0227*	-0.0199
	(0.0134)	(0.0272)	(0.0142)	(0.0118)	(0.0116)	(0.0122)	(0.0133)	(0.0174)
Proportion compliers					0.866			
N					1,078			
					Falsification check: 2007 savings in bank accounts[b]			
					Dependent variable: savings in bank accounts (thousands of euros)			
Treated	-0.00686	-0.533	0.138	0.263	0.115	1.76e-04	-0.156	-1.047
	(0.673)	(0.699)	(0.389)	(0.351)	(0.356)	(0.410)	(0.716)	(1.094)
Proportion compliers					0.857			
N					1,076			
					Dependent variable: savings rate for bank accounts (1 = 100%)			
Treated	-0.00271	0.0116	0.00292	0.00836	0.00264	0.00401	-0.00831	-0.0174
	(0.0136)	(0.0230)	(0.0113)	(0.0102)	(0.0104)	(0.0117)	(0.0183)	(0.0283)
Proportion compliers					0.857			
N					1,076			

[a] For decile models we report unconditional treatment effects.
[b] Savings are trimmed at -50% and +50% of disposable household income.
We control for the presence of multiple eligibles.
Standard errors in parentheses.
*significant at 10%; **significant at 5%; ***significant at 1%