November 7, 2022

Appendix

to

Communicating Social Security Reform

Andrew Caplin Eungik Lee

Søren Leth-Petersen Johan Sæverud

A Social Security Benefits in Denmark

The policy investigated in this paper changed the eligibility age, but did not affect the benefits conditional on being eligible, i.e., the benefit structure was unaffected by the policy change. Social security benefits consist of a basic benefit and a supplement.

Table A1: *Yearly Social Security Benefits (DKK), 2021*

	Single	In couple
Basic	78,216	78,216
Supplement	88,022	44,484
Total	166,238	122,700

The basic benefit is awarded to everyone who are eligible to social security benefits subject to not having earned income above a threshold. The basic social security benefit is tested against earnings such that benefits are reduced with 30 percent of the income above 344,600 DKK. This means that no social security benefits are paid out if earned income exceeds 605,300 DKK per year. Payments from pension schemes (be it private or labor market pension schemes) have no impact on the basic benefits.

The supplement is tested against household taxable income net of social pensions, including income from retirement accounts and income of the spouse. The supplement is reduced by 32 percent of (household) income exceeding a threshold. The threshold is 76,100 DKK for singles and 152,500 DKK for couples.

Social Security can be postponed with a bonus. For every two months that benefit payments are postponed a bonus of 1 percent is earned, meaning that payments will be 1 percent higher when benefit payments begin and will remain 1 percent higher for the remaining life time. It is possible to postpone benefit payments for up to ten years after which the benefit level will be 160 percent of the basic rate.

B Summary Statistics

Table A2: *Balance Table*

	Baseline survey 2021			*Follow-up survey 2022*		
	Participants	Non-participants	Difference	Participants	Non-participants	Difference
female	0.492	0.493	-0.001	0.476	0.502	-0.026
	(0.5)	(0.5)	(0.006)	(0.499)	(0.5)	(0.011)
age	48.389	43.601	4.788	52.68	47.472	5.208
	(12.179)	(12.8)	(0.135)	(11.218)	(12.308)	(0.248)
college	0.488	0.341	0.148	0.501	0.481	0.021
	(0.5)	(0.474)	(0.005)	(0.5)	(0.5)	(0.011)
employed	0.842	0.785	0.057	0.839	0.844	-0.004
	(0.365)	(0.411)	(0.004)	(0.367)	(0.363)	(0.008)
earnings	408,197	325,309	82,888	428,503	396,367	32,136
	(311,201)	(311,782)	(3,431)	(335,858)	(295,277)	(6,842)
wealth	443,926	248,848	195,078	519,329	399,997	119,331
	(4,792,147)	(2,620,078)	(50,299)	(1,673,064)	(5,891,539)	(81,160)
pension	1,579,293	993,810	585,483	1,894,543	1,395,629	498,914
	(1,874,468)	(1,490,113)	(20,150)	(1,944,205)	(1,807,606)	(40,326)
N	9,572	62,594		3,540	6,032	

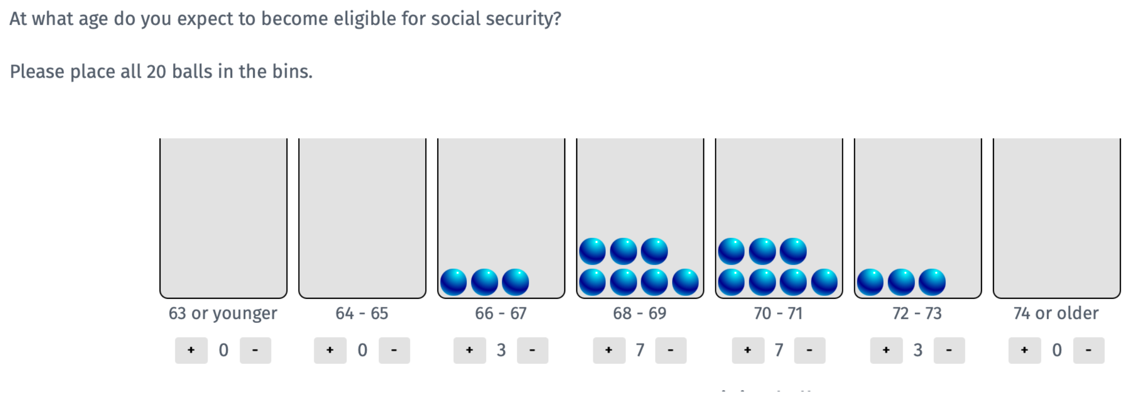
Table A3: *Tabulation of observations by social security eligibility age*

	Baseline survey 2021		*Follow-up survey 2022*			
Eligibility age	Control	Treatment	CC	CT	TC	TT
67.00	1,248	1,297	313	319	320	298
68.00	614	591	152	131	136	126
69.00	586	520	142	104	108	95
70.00	496	488	94	79	79	80
71.00	384	421	62	48	69	63
72.00	331	344	43	53	48	43
73.00	293	322	38	43	43	49
73.50	306	318	30	40	39	29
74.00	289	263	30	26	33	28
74.50	227	234	32	28	25	22

C The "Balls-in-Bins" Survey Instrument

Figure A1 show the graphical "balls-in-bins" interface that respondents meet in the survey. Along with the question, respondents are shown a graphical representation of seven bins into which 20 balls should be distributed, by pressing a "±" button underneath the bins, such that bins with more balls represent eligibility ages that they believe are more likely.

Figure A1: *"Balls-in-Bins"*



Notes: The graphical user interface where the respondents place 20 balls in seven bins to reflect their subjective beliefs.

Some respondents allocated many balls into the bin labelled "74 or older". The elicited subjective distribution, for respondents who associate substantial mass to this category, thus become censored, i.e. exhibit excess mass, at this category. To correct for this, we impose an underlying symmetric triangular distribution for respondents with balls in the category "74 or older" such that the adjusted distribution extends into age categories not specified explicitly in Figure A1[1].

The procedure is illustrated with an example in Figure A2. The top panel illustrates one such individual response where balls have been allocated to the bin "74 or older". In the middle panel, we take all the balls in the lower bins with a minimum value a and maximum value a', and consider the triangle with height h and area γ (the fraction of balls in lower bins). The height of the triangle is $h = \frac{2\gamma}{a'-a}$, and the probability density function of the symmetric triangular distribution is:

$$h = \frac{x-a}{(c-a)^2}.$$

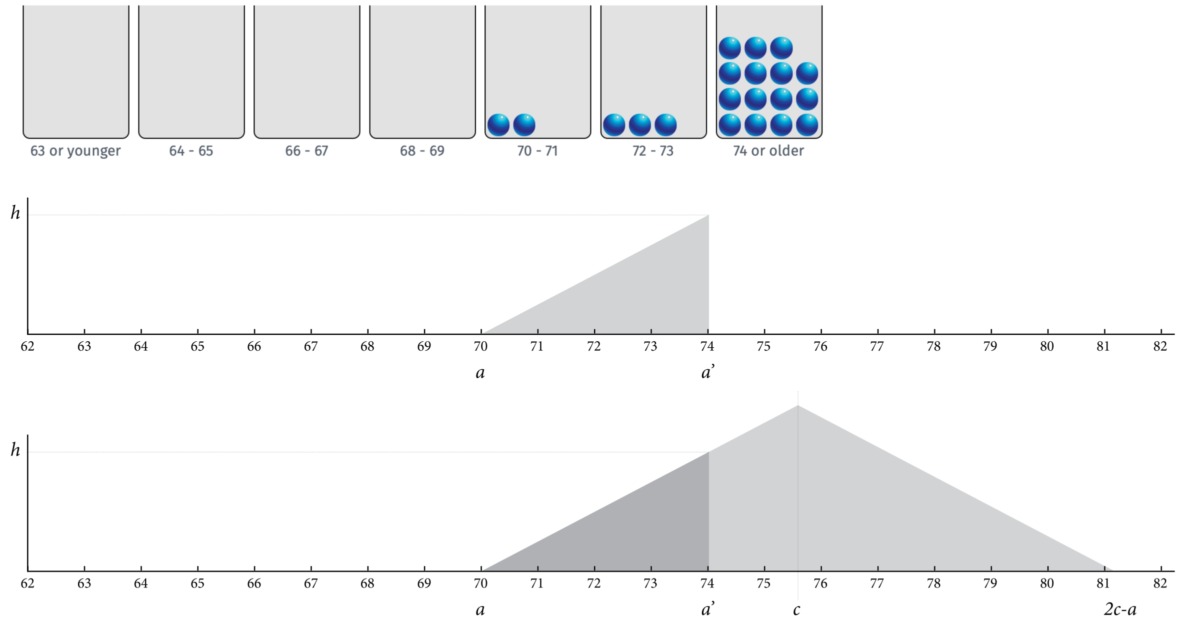
Solving for the peak value, c, gives

$$c = \sqrt{\frac{x-a}{h}} + a$$

[1]We can only impose the triangular distribution for respondents who, in addition to having balls in the category "74 or older" also have balls in at least one of the other categories, as these balls guide the parameters of the imputed triangular distribution.

which uniquely identifies the distribution. The bottom panel of Figure A2 shows how the imposed distribution looks.

Figure A2: *"Balls-in-Bins", imposing a triangular distribution*



Notes: The "balls-in-bins" instrument has a cap at "74 or older". For distributions where "74 or older" has at least one ball, we impose a triangular distribution. The top panel shows an example distribution, the middle panel shows the triangle used to infer the underlying distribution, and the bottom panel shows the underlying distribution.

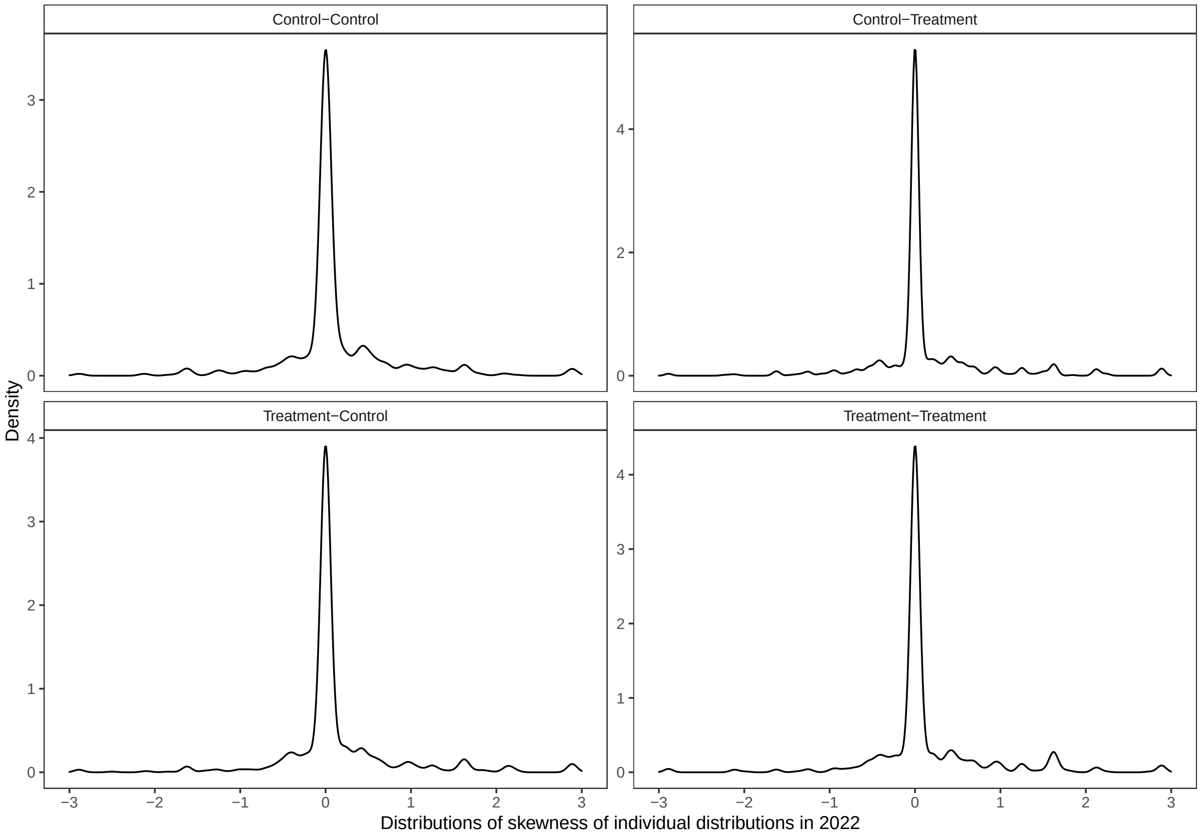
The censoring problem naturally affects younger cohorts the most. In Table A4 the fraction of responses that we have corrected is listed.

Table A4: *Fraction of observations within each cohort group on which symmetry is imposed*

Eligibility age	Fraction with imposed triangular distribution
67.00	0.02
68.00	0.04
69.00	0.08
70.00	0.17
71.00	0.28
72.00	0.38
73.00	0.44
73.50	0.51
74.00	0.52
74.50	0.53

The procedure assumes two features of the distributions: Symmetry and functional form. We test the symmetry assumption using answers from the follow-up survey, where we allowed for a wider support. Specifically, we allowed for three additional 2-year age bins such that the possible age categories span from "63 or younger" to "80 or older". In this way, the subjective distributions are in practice uncensored. In Figure A3 we plot the density of individual skewness by treatment status in both 2021 and 2022. It shows that skewness is heavily centered around zero and 87% of individual distributions have an absolute skewness below one. This show that it is reasonable to assume that subjective distributions tend to be symmetric.

Figure A3: *Distribution of Subjective Skewness by 2021-2022 Treatment Status*



Notes: Distribution of subjective skewness in 2022 for respondents by treatment status in 2021 and 2022.

C.1 Correcting for Censoring using a Gaussian Distribution

In the main analysis we impute impute uncensored subjective distributions by assuming that the underlying distributions are triangular. To examine the robustness of this, we also implement the procedure assuming that the underlying uncensored distribution is Gaussian. Following **?** we use the fraction of balls, γ, located in bins lower than the cut-off, $a = 74$ (all except ”74 or older”), and calculate the quantile function of γ (the inverse of the cumulative distribution function for a Gaussian distribution):

$$\alpha = \Phi^{-1}(\gamma)$$

We then calculate the auxiliary variable, λ, (where ϕ is the Gaussian probability density function):

$$\lambda = \frac{\phi(\alpha)}{\gamma}$$

The mean of the truncated distribution is given by (**?**):

$$\mathbb{E}[x|x < a] = \mu + \lambda\sigma$$

and

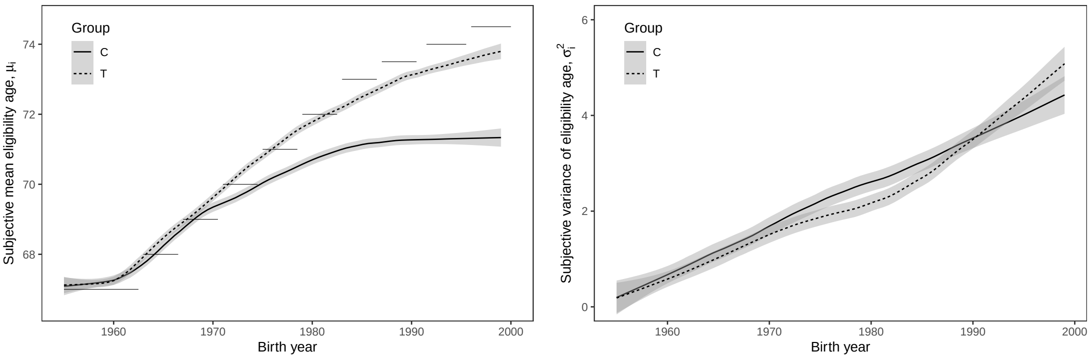
$$\alpha = \frac{a - \mu}{\sigma}$$

This can be arranged in a system of linear equations:

$$\begin{bmatrix} \mu & \lambda\sigma \\ \mu & \alpha\sigma \end{bmatrix} = \begin{bmatrix} \mathbb{E}[x|x < a] \\ a \end{bmatrix}$$

Solving these yield the parameters of the underlying distribution, μ and σ. Figure A4 shows the results. The results are practically identical to the case where we use the triangular distribution.

Figure A4: *Social Security Eligibility Beliefs in Baseline Survey. Imputed Gaussian distribution for censored 2021 answers*



(a) *Average Subjective Means* **(b)** *Average Subjective Variances*

Notes: Lines show locally weighted linear regressions for control (solid) and treatment (dotted) groups for subjective mean eligibility ages, Panel A4a, and subjective variance of eligibility ages, Panel A4b. Censored distributions are imputed assuming a Gaussian distribution. See notes to Figure **??**. In Panel A4a horizontal lines show official eligibility ages.

D Additional Survey Instruments

The questions listed below were only asked in the 2021 survey

Income 2020

What was your earned income during 2020?

Please report the most accurate value you can:

Retirement wealth 2020

Consider how much wealth you have accumulated in total in pension accounts by now.

Please report your belief about this accumulated amount.

Lowest possible amount:

Highest possible amount:

Please enter all 20 balls in the bins

Retirement I

How old do you expect to be when you retire?

Please consider the various factors that are uncertain and that may affect your retirement age (for example, health, savings, or other factors that may be important).

Please enter all 20 balls in the bins

Retirement II

Suppose that you first become eligible for social security at the age of 65.

At what age do you expect to retire?

Please enter all 20 balls in the bins

Retirement III

Suppose that you first become eligible for social security at the age of [Table age]. At what age do you expect to retire?

Please enter all 20 balls in the bins

Income from Pension Wealth

Suppose you retire at age [Table age], and suppose you stay in your current job until retirement.

How much annual income in retirement do you believe your pension would provide?

Lowest possible amount:

Highest possible amount:

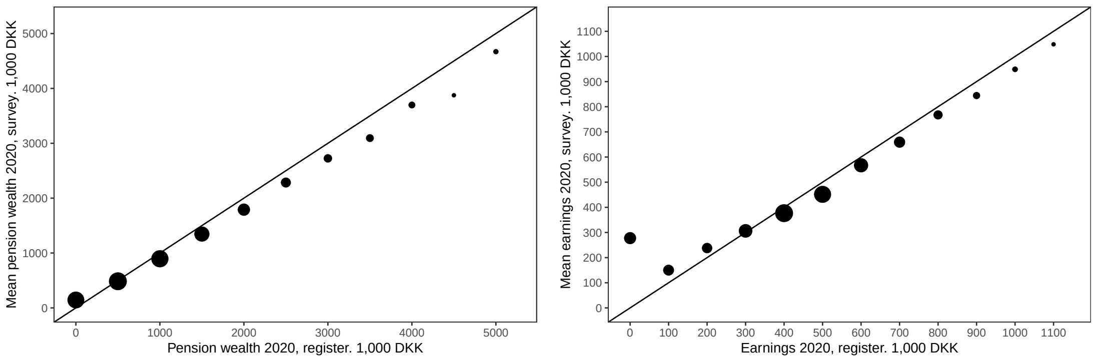
Please enter all 20 balls in the bins

E Validation

An important question is whether respondents are able and willing to respond accurately to the questions that they are asked in the survey. To assess this we asked the respondents about pension wealth and earnings in 2020 and we then compare stated pension wealth and earnings from the survey with their third-party reported counterparts from the administrative registers. Figure A5a reports average pension wealth as reported in the survey in 2020 by DKK500,000 bins of pension wealth as recorded in the administrative register for 2020. Panel A5b reports average earnings as reported in the survey in 2020 by DKK100,000 bins of earnings as it is recorded in the administrative register data for 2020. In both panels, the size of the dots indicate number of observations and the dotted line is a weighted OLS regression through the micro data with coefficients reported in the top-left of the panel.

The reported pension wealth is very close to the 45-degree line. Reported earnings is also close to the 45-degree line, except at the bottom end of the 2018 distribution. Overall, Figure A5 shows that survey responses align remarkably well with objective third-party reported measures from the administrative register data. These findings confirm that respondents are able and willing to provide meaningful answers in the survey.

Figure A5: *Validation of Survey Responses*



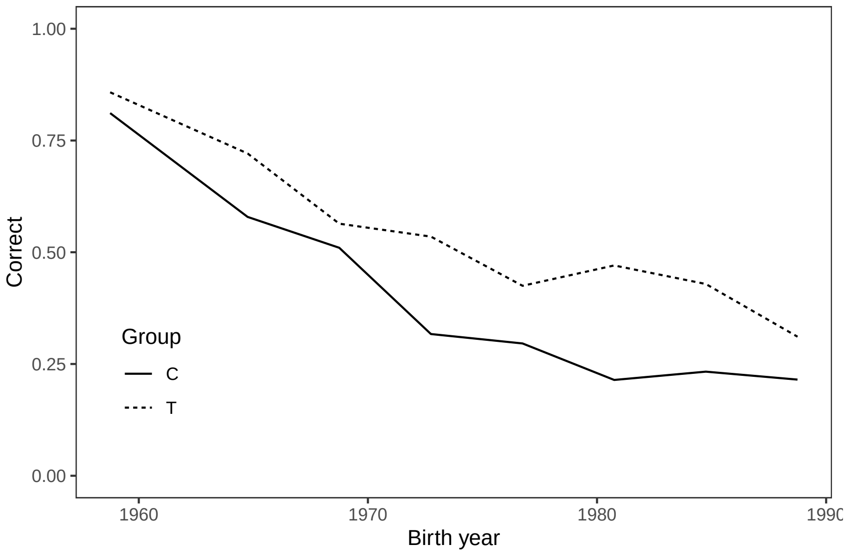
(a) *Mean reported pension savings, 2020*

(b) *Mean reported earnings, 2020*

Notes: The figure shows the relationship between reported pension wealth, Panel A5a, (earnings, Panel A5b) in 2020 against the corresponding measure that is third-party reported in the administrative register in 2020. The panels show binned scatter plots (black circles) where the bins are defined over intervals of the register measure. The size of the dots is proportional to the number of observations in the bin. The 45 line is overlaid.

F Robustness

Figure A6: *Fraction of Balls in Correct Bin 2021*



Notes: The figure shows the fraction of balls allocated into the correct bin for the information treated group and the control group in the baseline survey 2021.

G A Model of Policy Beliefs

In this Appendix we formulate a model of belief formation that conceptualizes the key forces that drive social security belief formation, and develop a quantitative version of it that is able to broadly match the average patterns of eligibility beliefs documented in section **??**.

The model takes the Danish policy environment as a starting point. The government follows either of two potential social security policies. In one policy scenario, social security eligibility age is linked to life tables. In the other policy scenario, politicians resort to a policy with a lower eligibility age because the link to the life tables would imply too drastic increases in the eligibility age for them to resist public pressure. The two policies are common for all individuals. There is inherent uncertainty about exactly how the policies will be implemented. For example, cohort-specific life expectancy will likely be updated in the future and it is uncertain exactly how the government will implement a policy with a lower eligibility age. Individuals form subjective beliefs about their social security eligibility age by weighting the probabilities of these two underlying policies. The information treatment mimics a policy maker sending out information that future social security eligibility ages will follow cohort specific life tables, and this may move the weight towards the life table policy as the signal informs the individual that politicians are likely to go through with the life table policy.

To formalize this, denote the life table policy as policy $q = 1$ and the alternative policy as policy $q = 2$ and assume that the eligibility age under either policy is characterized by a normal distribution capturing the underlying uncertainty: $f_{q,k} \sim N(\mu_{q,k}, \sigma^2_{q,k})$ where $q = 1, 2$, and k is an indicator for the cohort group, c.f., Table **??**. The subjective beliefs about eligibility age, f, is given by the mixture of the two normal distributions

$$f_{1,k} \sim N(\mu_{1,k}, \sigma^2_{1,k}) \tag{1}$$

$$f_{2,k} \sim N(\mu_{2,k}, \sigma^2_{2,k}) \tag{2}$$

$$f_{k,D} = p_{k,D} f_{1,k} + (1 - p_{k,D}) f_{2,k} \tag{3}$$

The parameters of $f_{q,k}$ need not be the same across cohort groups, but by randomization, they are identical across treatment and control groups. $f_{k,D}$ is the average subjective distribution at time t for individuals belonging to cohort group[2] k with treatment status D, where $D = T$ when information treated and C otherwise. $p_{k,D}$ is the average subjective weight on policy 1, the life table policy for cohort group k with treatment status D. The mean and variance of $f_{k,D}$ has the following closed form solution:

$$m_{k,D} = \mathbb{E}[f] = p_{k,D}\mu_{1,k} + (1 - p_{k,D})\mu_{2,k} \tag{4}$$

$$s^2_{k,D} = \mathbb{V}[f] = p_{k,D}\sigma^2_{1,k} + (1 - p_{k,D})\sigma^2_{2,k} + p_{k,D}(1 - p_{k,D})\left(\mu_{1,k} - \mu_{2,k}\right)^2 \tag{5}$$

The first two terms of Equation (5) are the weighted variances of the underlying distributions and the third reflects the added variance coming from the distance between the means of the underlying distributions. The behavioral parameter of interest is the subjective probability weight on the life table policy, $p_{k,D}$. We fit the parameters of the model and estimate how the information treatment works through $p_{k,D}$.

G.1 Fitting the Model

The model has six parameters for each cohort group, $p_{k,C}, p_{k,T}, \mu_{1,k}, \sigma^2_{1,k}, \mu_{2,k}, \sigma^2_{2,k}$. From the elicited distributions we use four empirical moments: average subjective means and variances for both the treatment and the control group, i.e., $\bar{m}_{k,T}$, $\bar{s}^2_{k,T}$, $\bar{m}_{k,C}$, and $\bar{s}^2_{k,C}$ for all cohort groups, k. With six parameters and four empirical moments, the model is not identified and we need to impose some additional restrictions. The restrictions we impose follow naturally from the policy environment. First, we fix $\mu_{1,k}$, the mean of the life table policy, to take the values listed

[2] We use cohort groups that correspond to the eligibility ages, c.f., Table **??**. The data used to fit the model is cohort group specific average moments.

in Table ??. Next, we assume that $\sigma^2_{1,k} \geq \sigma^2_{1,k-1}$, i.e., that uncertainty about the life table is at least as big for cohort k as it is for cohort $k-1$. This is essentially just saying that young cohorts face at least as much uncertainty as older cohorts, meaning that life expectancy is at least as hard to predict for the young as for the old because of the longer horizon. Similarly, we assume that $\sigma^2_{2,k} \geq \sigma^2_{2,k-1}$ because a long horizon leaves at least as much uncertainty about the details of a future alternative policy. Finally, we restrict $\mu_{2,k} \in \left[\overline{\min(m_{k,C})}, \mu_{1,k}\right]$ where $\overline{\min(m_{k,C})}$ is the average of the minimum possible eligibility age indicated by control group individuals in their "balls-in-bins" answer to the question about their social security eligibility age, setting a lower limit, and $\mu_{1,k}$ is the table age. We fit the six parameters by minimizing the squared distance between the empirical moments, $\bar{m}_{k,C}$, $\bar{m}_{k,T}$, $\bar{s}^2_{k,C}$, $\bar{s}^2_{k,T}$, and the corresponding model implied moments, cf., equations (4) and (5).[3]

G.2 Minimum Distance

The model has six parameters for each cohort, $p_{k,C}, p_{k,T}, \mu_{1,k}, \sigma^2_{1,k}, \mu_{2,k}, \sigma^2_{2,k}$. We observe average subjective means and variances, i.e., $\bar{m}_{k,C}$, $\bar{m}_{k,T}$, $\bar{s}^2_{k,C}$, and $\bar{s}^2_{k,T}$ for all the cohort groups and for the treatment and control groups. In order to identify the model parameters, we thus need to impose some additional restrictions. First, we fix $\mu_{2,k}$, the mean of the life table policy, to take the value listed in Table ??. Next, we assume that $\sigma^2_{2,k} \geq \sigma^2_{2,k-1}$ and $\sigma^1_{2,k} \geq \sigma^1_{2,k-1}$. Finally, we restrict $\mu_{1,k} \in \left[\overline{\min(m_{k,C})}, \mu_{2,k}\right]$ where $\overline{\min(m_{k,C})}$ is the average of the minimum possible eligibility age indicated by control group individuals in their "balls-in-bins" answer to the question about their social security eligibility age, and $\mu_{2,k}$ is the table age. We find the values of the parameters, $p_{k,C}, p_{k,T}, \mu_{1,k}, \sigma^2_{1,k}, \sigma^2_{2,k}$ by minimizing the squared distance between the empirical moments, $\bar{m}_{k,C}$, $\bar{m}_{k,T}$, $\bar{s}^2_{k,C}$, and $\bar{s}^2_{k,T}$, and the corresponding model implied moments in equations (4) and (5) subject to the constraints listed above and separately for each cohort. This is summarized in equation (6):

$$\Theta = \underset{p_{k,C}, p_{k,T}, \mu_{1,k}, \sigma^2_{1,k}, \sigma^2_{2,k}}{\operatorname{argmin}} \left[(m_{k,C} - \bar{m}_{k,C})^2 + (m_{k,T} - \bar{m}_{k,T})^2 + (s^2_{k,C} - \bar{s}^2_{k,C})^2 + (s^2_{k,T} - \bar{s}^2_{k,T})^2 \right]$$

subject to

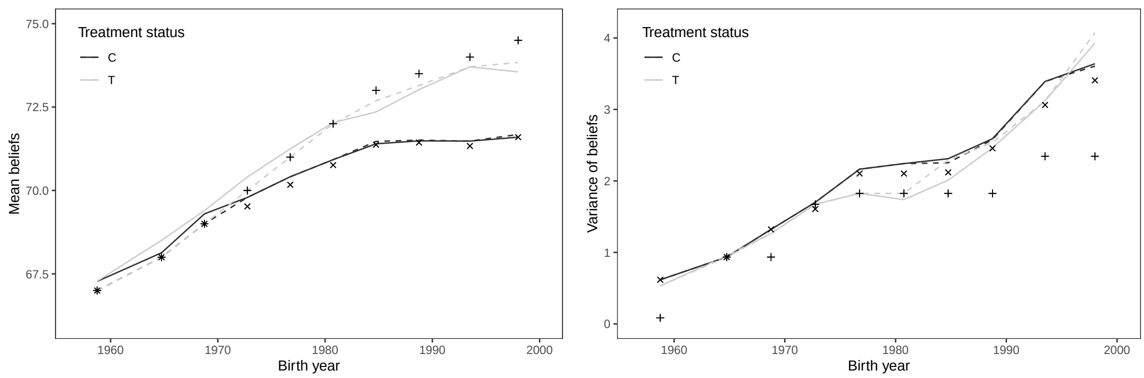
$$\begin{aligned} &\sigma^1_{2,k} \geq \sigma^1_{2,k-1} \\ &\sigma^2_{2,k} \geq \sigma^2_{2,k-1} \\ &\mu_{1,k} \in \left[\overline{\min(m_{k,C})}, \mu_{2,k}\right] \end{aligned} \tag{6}$$

[3] We refer to Appendix G.2 for details about the minimum distance procedure.

G.3 Results

In Figure A7 we report the model's ability to replicate average mean beliefs, Panel A7a, and uncertainty, Panel A7b, by information treatment status. Panel A7a shows average beliefs in the data with solid lines and model generated mean beliefs with dashed lines. The model implied mean beliefs match the data quite closely for both the treatment and the control group and for all cohorts. Moreover, the fitted model is able to replicate the effect of the information treatment. Panel A7b shows average subjective uncertainty in the data with solid lines and model implied subjective uncertainty with dashed lines. Also here there is a close correspondence between data and model implied average beliefs where average subjective uncertainty is increasing in cohort year and with no effect of the information treatment. The fact that there is no effect of treatment on the average subjective uncertainty reflects that the overall subjective uncertainty is affected by the difference in mean eligibility ages between the two policy distributions as well as by the uncertainty associated with each of the underlying policy components, cf. equation (5). Subjective uncertainty thus reflects inherent policy uncertainty associated with both policy regimes.

Figure A7: *Estimated model parameters*

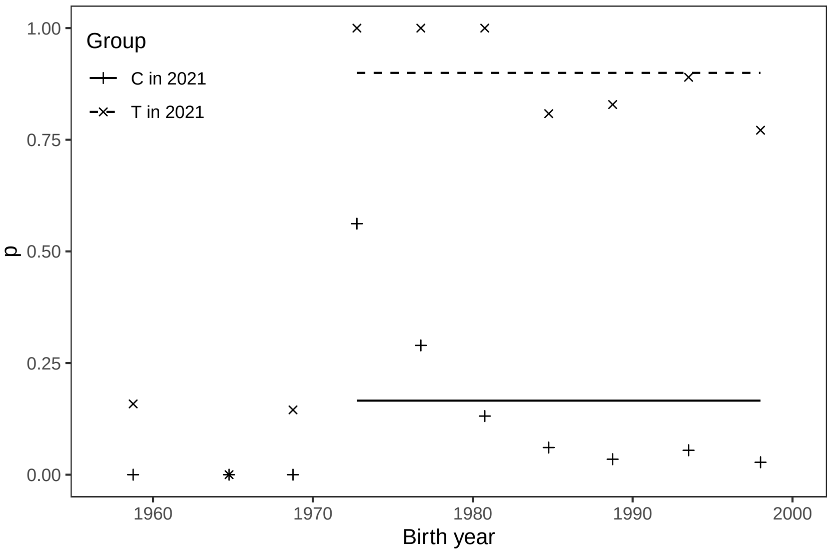


(a) *Mean fit* **(b)** *Variance fit*

Notes: Panels A7a and A7b show the results from fitting Equations (4) and (5) to average subjective means, Panel A7a, and variances, Panel A7b, separately for all cohort groups and for treated and untreated individuals. Solid lines are data moments and dashed lines are fitted values. In Panel A7a, '+'s indicate $\mu_{1,k}$, which is fixed to match official eligibility ages listed in Table **??**, and '×'s are fitted values of $\mu_{2,k}$. In Panel A7b, '+'s are fitted values of $\sigma^2_{1,k}$ and '×'s are fitted values of $\sigma^2_{2,k}$.

In the model described by Equations (4) and (5) the effect of the information treatment operates through shifting the subjective weight on the life table policy, $p_{k,D}$. In Figure A8 we plot with '+'s and '×'s the estimates of $p_{k,D}$ for $D = (C, T)$, i.e., for the control and treatment groups for all cohort groups. For the three oldest cohort groups, $p_{k,D}$ is not identified as there is no discernible difference between the beliefs of the treatment and control groups. The weights are approximately constant across cohorts within the treatment and control groups as indicated by the horizontal lines. For the control group, the average weight put on the life table policy is 0.17, compared to 0.90 for the treatment group. This means that the treatment induces a large increase in the weight assigned to the life table policy, $p_{k,D}$. In other words, the information treatment is extremely successful in shifting the average subjective weight from the alternative policy to the life table policy, such that people who have been information treated predominantly form their beliefs based on the life table policy scenario.

Figure A8: *Fitted model weights*

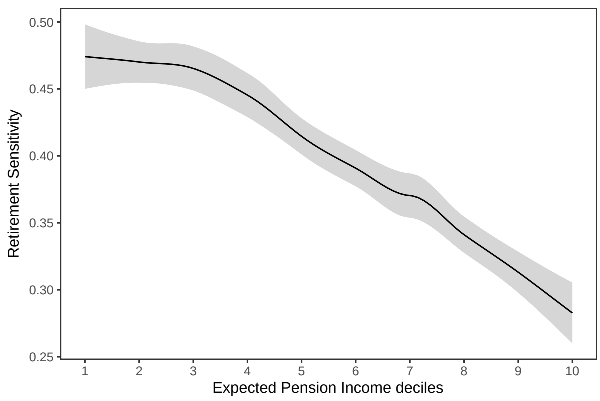


Notes: Fitted values of the weight on the life table policy, $p_{k,D}$, for $D = (C, T)$, i.e., for the control and treatment groups, for all cohort groups, k. Mean values for all but the three oldest cohort groups are overlaid. For the oldest cohort groups, $p_{k,D}$ is not identified as there is no discernible difference between the beliefs of the treatment and control groups. Figure A7 shows the fit of the four moments of the common underlying distributions, $f_{q,k}$.

H Incentives

Figure A9 shows how retirement sensitivity varies with the importance of social security in retirement income. It plots retirement sensitivity, RS, as defined in equation (**??**), against deciles of expected retirement income. At high levels of expected retirement income, social security is relatively less important as a source of income in retirement, i.e., the financial incentive provided by social security is smaller the higher is expected retirement income. The figure shows that retirement sensitivity is negatively correlated with the expected retirement income level. This confirms that the financial incentive matters for the retirement sensitivity to the social security eligibility age.

Figure A9: *Pension income, deciles*

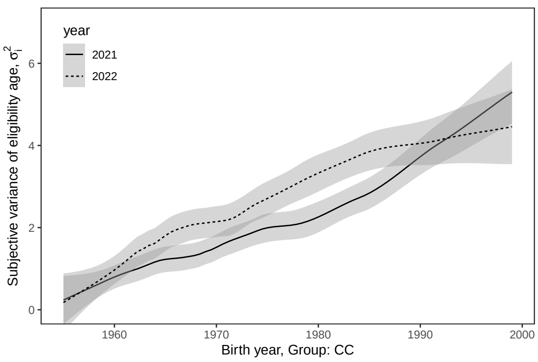


Notes: The figure shows a locally weighted linear regression for the relationship between retirement sensitivity (c.f., equation (**??**)) and deciles of expected pension income. See notes to Figure **??**.

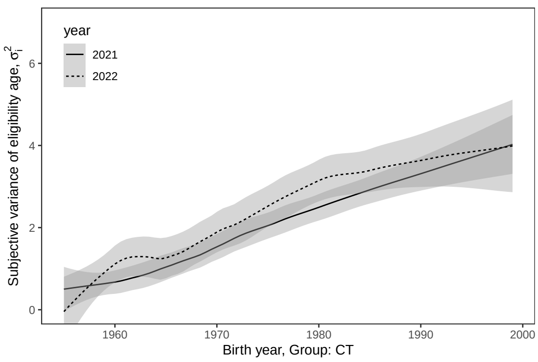
I Further Evidence from the Follow-Up Survey

In Figure A10 we show the average subjective variances of social security eligibility ages from the follow-up survey, where the panels are organized in the same way as in Figure **??**. The samples underlying the panels in Figure A10 are much smaller than in the baseline survey in 2021 and the relationships are therefore less precisely estimated. In all panels the average of subjective uncertainty is increasing in distance to eligibility and thus display the same behavior as in Figure **??**. Generally there appears to be no effect of the information treatment, albeit in Figure A10d, showing the variances for the group that was information treated in 2021, but not in 2022, the variance appears to increase for younger cohorts going from 2021 to 2022.

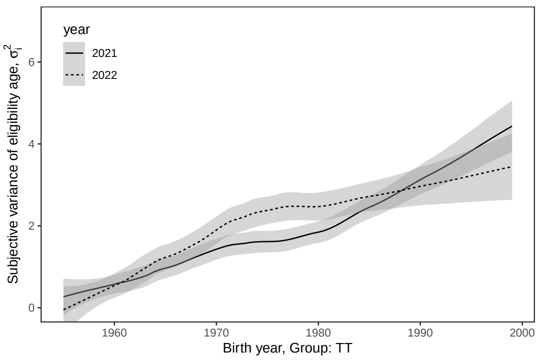
Figure A10: *Follow-Up Survey by Treatment Status in 2021/2022, Variance*



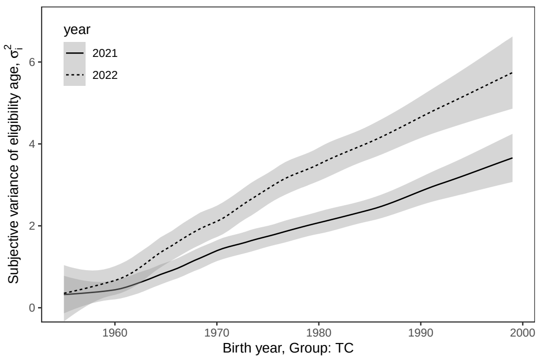
(a) *Control-Control*



(b) *Control-Treatment*



(c) *Treatment-Treatment*



(d) *Treatment-Control*

Notes: Lines show locally weighted linear regressions for subjective variances of eligibility ages for 2021 survey (solid) and 2022 survey (dotted). The panels show each combination of control and treatment in the 2021 and the 2022 survey. Results are only for the 3,540 respondents who participated in both surveys.See notes to Figure **??** for details.