Youth Subjective Life Expectancy and Early Labor Market Choices*

Lucas Finamor†

First version: January 2024
This version: May 2024

Abstract

The literature has documented how subjective life expectancy (SLE) is strongly associated with savings and retirement outcomes for those nearing retirement. This paper assesses whether SLE matters when young individuals make consequential career decisions at the labor market entrance. Exploring survey and administrative data from Chile, I show how individuals aged 18–26 with one standard deviation higher SLE had 11.6%–14.8% higher pension wealth 15 years later. Type of employment — working formally, accounts for a large share of these pension gaps. I also show how these expectations respond to new information and how individuals revise their actions consistently with their revised expectations. In a simple theoretical framework, I show how ignoring heterogeneity in life expectancy leads to biased predictions and might induce suboptimal policies.

JEL Codes: D84, J22, J46, H55

*I would like to thank Richard Blundell, Áureo de Paula, Cormac O'Dea, David Sturrock, Eduardo Ferraz, Dalila Figueiredo, and Antonia Paredes-Haz for their comments. I thank the ESRC for support under the Centre for the Microeconomic Analysis of Public Policy (CPP) at IFS, grant number ES/T014334/1. This research uses the Chilean Social Protection Survey (*Encuesta de Protección Social*). I thank the Undersecretary of Social Protection, the intellectual owner of the survey, for the authorization to use the de-identified dataset. All the results from this research are the responsibility of the author and do not implicate the Chilean Undersecretary of Social Protection.

†Institute for Fiscal Studies. Contact: lucas.finamor@ifs.org.uk

1 Introduction

Survival probabilities are an important ingredient in any intertemporal dynamic problem, being essential to determine savings, adequate financial planning, labor market decisions, and retirement choices. Since the seminal work of Hamermesh and Hamermesh (1983) and Hamermesh (1985), the literature has explored the properties and importance of individual beliefs about their survival probabilities. Hudomiet et al. (2023) review this body of work, showing how subjective life expectancy correlates with health information (e.g., health assessment, diagnoses), lifestyle behaviors (e.g., smoking, exercising), life events (e.g., parental deaths), and, importantly, with actual mortality patterns.

Building on these findings showing how survey-elicited beliefs had desirable properties, researchers have explored the relation between survival beliefs and important economic activities. Subjective survival beliefs are associated with consumption/savings behavior, retirement, and the purchase of annuities.[1] Most of this literature uses beliefs elicited from the population over 50 years of age, close to retirement.[2] However, the most impactful decisions in terms of career progression, exposure to risk, and coverage by social insurance are taken much earlier on by workers. Even the first job is very consequential (Altonji et al., 2016, Arellano-Bover, 2020). Therefore, it is still unclear how beliefs on life span play a role more broadly over the life cycle and particularly at the onset of labor market trajectories. This paper fills this gap by exploring elicited beliefs from very young adults and how they relate to future labor market outcomes, such as labor market participation, employment type, and pension contributions.

I combine a longitudinal household survey with administrative pension data from Chile.

[1] Hurd et al. (2004), Bloom et al. (2006), O'Donnell et al. (2008), Van der Klaauw and Wolpin (2008), Salm (2010), Gan et al. (2015), Wu et al. (2015), Bissonnette et al. (2017), Heimer et al. (2019), Bresser (2021), O'Dea and Sturrock (2023)

[2] This specific age range comes from the dataset used, which is, in most cases, the Health and Retirement Survey in the United States and their equivalent in other countries. These studies only survey individuals over 50. The exception is Heimer et al. (2019), which has a sample of the US population aged 28–78.

The survey brings information on labor market status, health, and beliefs measured in several waves. The survey can be combined with precise and frequent data from the pension administrative system at the individual level, including pension wealth. This yields an extraordinary dataset with individuals' trajectories for over 15 years. I restrict the data to young individuals (18–26 years old) to assess the impact of subjective life expectations at the onset of labor market careers.

The main analysis explores how cross-sectional variation in subjective life expectancy (SLE) relates to future outcomes in the labor market. I leverage the existence of two different measures of life expectancy to use an instrumental variable approach that accounts for measurement error, including rounding.[3] I additionally explore the panel dimension to show patterns of revisions in subjective life expectancy and robustness, exploring solely within-individual variation on SLE.

First, I document great dispersion in survival chances beliefs and how they correlate with health and lifestyle behaviors. Those smoking, not exercising regularly, who have been diagnosed with diseases, who self-report bad health, or those with higher body mass index report lower subjective life expectancies. Men with higher education levels display higher life expectancy. I also show that even after controlling for various demographic, health, and life behavior variables, 80% of subjective life expectancy variance is left unexplained. This indicates that SLE has substantial content, over and above all these observed measures. These findings are in line with the literature focusing on older individuals, reviewed by Hudomiet et al. (2023). Using samples from slightly older individuals, I show how past SLE is correlated with mortality, measured in the administrative data.

In terms of subsequent labor market results, I show, using the pension administrative data, how individuals who reported higher SLE have a higher attachment to the pension

[3] The two questions were placed in different modules within the survey and elicited the same underlying beliefs with different wording.

system in all years following the initial report. At the end of my sample period in 2019, 15 years after SLE was first measured, this corresponds to a significant gap in pension contributions. Those with 10 years higher SLE have around five months more pension contributions (4% of the mean) and around 360 thousand Chilean pesos higher pension wealth (5.7% of the mean).[4] Ten years of SLE equals the 75th–25th percentile gap, or approximately 0.83 standard deviations. Instrumenting SLE with another question, which elicits the probability of being alive at age 65, yields larger (and noisier) estimates, consistent with the presence of measurement error. The 75th–25th gap in SLE translates to a 10.7% gap in pension wealth in December 2019, in the instrumental variable estimates. The Chilean pension system is based on individual capitalization accounts, where individuals make mandatory pension contributions while working. Upon retirement, their pension wealth is used to fund retirement benefits. Therefore, these pension gaps associated with SLE can be read as potential gaps in final pensions upon retirement.[5]

I investigate further what could be behind the documented pension gaps by looking at labor market information in the survey. I first document how, even though those reporting higher SLE are more likely to participate in the labor force, this is insufficient to explain the documented large gaps in pension contributions. Ten years higher SLE is associated with being 1.6pp more likely to be in the labor force (1.9% of the mean participation rate). I turn then to analyze the type of employment. Pension contributions are mandatory for those working formally in the private and public sectors, while this is not the case for the large share of individuals in informal jobs or self-employment. Those with 10 years higher SLE are 4.2pp more likely to work as formal employees (7.4% of the mean). This indicates that employment type accounts for most of the pension gap found. In short, those expecting to live longer are more likely to work and to be working as formal employees with mandatory

[4] One thousand Chilean pesos corresponded to approximately 1.6 USD in 2004, around the first wave of the survey.

[5] The gap in final pensions would still depend on the (endogenous) retirement age, minimum pension floor guaranteed by the government, and future contributions.

pension contributions.

I additionally leverage the longitudinal dimension of both datasets. I show that new disease diagnoses correlate with reductions in future SLE, with magnitudes depending on the severity of the diagnosis.[6] Moreover, I show that even conditioning on the first SLE, future beliefs of SLE are also positively associated with pension outcomes. This would be consistent with a "revision behavior" — individuals with similar initial beliefs on life expectations revise their pension contributions up or down, depending on future information revealed to them. I also use the longitudinal dimension to show how the results are robust when exploring solely the within-individual variation across surveys, showing robustness to time-invariant unobservables.

We could expect more risk-averse individuals to be more sensitive to survival beliefs as they place more value on the expected scenarios of living longer. I conduct a heterogeneity exercise interacting the SLE variable with different risk aversion measures. In all of them, the point estimates of the interaction term are positive. However, I cannot reject that they are statistically different from zero. Splitting the analysis by gender, there are two interesting facts. First, while men's reports are closer to the objective measures from the life tables, women's reports are much more pessimistic. Second, the estimated coefficients for men are always higher than for women for all outcomes of interest (total pension contributions, pension wealth, labor force participation, and working in the formal sector).

In addition to the exercises ignoring cross-sectional variation and exploring solely within individual variations, I apply several robustness exercises. They show how the results are not sensitive to the specific age range chosen, the included or excluded controls, the linear specification, and the specific SLE measure used. One concern is still if there are omitted variables that correlate with SLE report and labor market choices. To address this concern, I follow Altonji et al. (2005) and Oster (2019) and show how the estimates are very robust

[6] This is consistent with the findings by Hurd and McGarry (2002).

to the inclusion of relevant observed variables.

These novel findings might be consequential for policy-making, such as the design of social security and private savings incentives. To assess this, I develop a simple theoretical framework where individuals choose how much to invest for retirement, taking into account beliefs on life span. I show that ignoring heterogeneity in (subjective) life expectancy produces biased predictions. Models that only account for the average life expectancy across individuals tend to overestimate optimal savings for retirement and behavioral changes to social security reforms. These results highlight the importance of taking into account individual survival beliefs for dynamic decisions.

This paper offers several contributions to the literature analyzing subjective life expectancy and economic decisions.[7] First, as most of the literature focuses on individuals near retirement, I show novel evidence that early beliefs on survival are strongly associated with labor market choices and pension contributions. In particular, young individuals with higher SLE are more likely to participate in the labor market and to make pension contributions. Given our knowledge of how consequential early life decisions are for career trajectories, this is a critical finding for our understanding of labor market choices. Second, I add to both literature on SLE and informal labor markets in developing countries by providing new insights on how SLE is associated with employment type. Namely, I show that individuals with higher SLE are more likely to work formally. Employment type is extremely impactful. It is associated with exposure to risk, eligibility to several government programs, and, more generally, career progression. Lastly, the data structure combining longitudinal survey and administrative data allows me to go beyond self-reported outcomes, thus avoiding any assumptions on how reporting behavior in the outcome and subjective beliefs correlate. Additionally, it allows for the exploration of patterns of revisions on SLE. The results show

[7] Hurd et al. (2004), Bloom et al. (2006), O'Donnell et al. (2008), Van der Klaauw and Wolpin (2008), Salm (2010), Gan et al. (2015), Wu et al. (2015), Bissonnette et al. (2017), Heimer et al. (2019), Bresser (2021), O'Dea and Sturrock (2023)

how individuals who believe they will live longer choose jobs and actions consistent with a longer life span and revise their actions based on new information.

The rest of the paper proceeds as follows. The next section provides a brief description of the institutional setting in Chile and of the data. The third section discusses the life expectancy measures and their properties in the data. The following section presents the econometric strategy; the results are presented in the fifth section. In the sixth section, I present a theoretical framework. Some concluding remarks in the last section conclude the paper.

2 Institutional Setting and Data

2.1 Institutional Setting

In 1980, Chile introduced a fully funded individual capitalization pension system. Each month, employees are required to make pension contributions in their pension account at 10% of their wages, up to a cap. These pension funds are administered by private pension managers and are illiquid. Upon retirement, individuals can choose to annuitize their pension wealth entirely or partially.[8] The normal retirement age is 65 for men and 60 for women. Early retirement is an option if the resulting pension benefit is above an absolute threshold set by the government and higher than a fraction of the individual average wage. While voluntary contributions are allowed, they are very rarely observed.

For most of the sample period, self-employed individuals were not required to make pension contributions. That was introduced in a pension reform in 2008, whose implementation was postponed until 2015. Nevertheless, enforcement is a challenge. Finamor (2023) shows that 2/3 of self-employed individuals with at most high-school degrees are not registered with the tax authority, and less than one quarter had any pension contributions in

[8] For more information on options upon retirement, check Bello (2019).

12 months. Therefore, working as self-employed or informal worker is one way of avoiding making pension contributions. Even though Chile has lower rates of informal work than neighboring countries in Latin America, there is still a substantial fraction of workers in informal jobs and self-employment activities — around 1/3 of the workforce (Finamor, 2023). The government subsidizes a pension floor that was substantially increased in the pension reform in 2008. Individuals with low or no pension wealth at age 65 are likely to be eligible for this minimum pension, depending on their family income and assets.

2.2 Data

This paper uses two main datasets from Chile: a household survey (*Encuesta de Protección Social*, EPS) and administrative data from the pension system (*Historia Previsional de Afiliados*, HPA). The EPS is a longitudinal survey that has been conducted since 2002, with six waves implemented every 2–4 years. Since the second wave, in 2004, it is nationally representative, covering the adult population in Chile. It contains detailed information on demographics, labor market characteristics, family, health, income, and assets for the interviewed person in the household. Crucially for this project, EPS also has questions on subjective life expectancy.

Every individual interviewed in the EPS can be linked to the administrative data from the pension system (HPA), which contains all monthly contributions towards the pension system between 1981 and 2019. Since 2008, pension wealth in every individual account is available. It is also possible to check pension claims and payments, including those paid to family members upon the death of the primary account holder. I complement these datasets with life tables computed by the Chilean National Statistics Office.

From these datasets, I derive the main sample used in the paper as individuals between 18 and 26 years old in 2004 and observed in any wave after 2009. The age restriction is to capture young individuals before or at the beginning of their labor market careers. The age limit of 26 was chosen to yield enough sample size for the analysis, particularly when splitting the sample by gender. The restriction that individuals are observed at least once again after 2009 is to obtain their final educational level, as some of them could still be in school or at universities when they were first interviewed. I use the highest educational level reported after 2009. This last restriction is binding for 367 individuals (19% of the 1,927 individuals aged 18–26 in 2004) who were dropped from the analysis.[9]

Table 1 presents some descriptive statistics. The final sample contains 1,417 observations (710 men and 707 women). I defer the discussion on the measure of subjective life expectancy to the next section, but we can see how men, on average, believe they will live longer than women. Most individuals in the sample have high school degrees, and more than one-third of the sample have higher degrees. Around 40% of the sample live in the metropolitan region of Santiago. On parental education, a minority of parents have more than high school degrees, with almost equal shares between those with and without high school degrees. In terms of subjective life expectancy, we see positive gradients for their own and parental education.

3 Life Expectancy Measure

Surveyed individuals answered two questions on life expectancy in the EPS. The first question inquired to what age individuals believed they would live.[10] The answers were ages measured in years. A second question asked the chances of living to at least age 65.[11] Respondents were expected to give answers in percentages ranging from 0 to 100. Interestingly, these questions were placed in different modules in the survey, resulting in answers to both questions being separated by dozens of questions.

[9]Figure A.3 shows how the distributions of SLE for those with and without the final educational level being observed are very similar.

[10]In Spanish, the original wording was *"¿Hasta qué edad cree usted que va a vivir?"*

[11]In Spanish, the original wording was *"¿Cuáles son sus posibilidades de vivir hast a los 65 años?*

Table 1: Descriptive statistics

	Men			Women		
	N Obs	Prop	SLE (mean)	N Obs	Prop	SLE (mean)
Total	710	-	75.983	707	-	73.663
Educational level						
Primary	68	0.096	72.603	74	0.105	74.149
High School	382	0.538	76.031	361	0.511	73.022
Vocational	100	0.141	76.100	109	0.154	73.339
College	160	0.225	77.231	163	0.231	75.080
(subtotal)	710			707		
Region						
MR Santiago	276	0.389	75.967	280	0.396	74.289
Other	434	0.611	75.993	427	0.604	73.253
(subtotal)	710			707		
Mother education						
Less than HS	337	0.491	74.644	354	0.516	72.452
High School	290	0.423	77.234	269	0.392	74.431
More than HS	59	0.086	77.458	63	0.092	77.810
(subtotal)	686			686		
Father Education						
Less than HS	305	0.457	75.148	321	0.488	72.327
High School	292	0.437	76.712	272	0.413	74.893
More than HS	71	0.106	77.620	65	0.099	77.369
(subtotal)	668			658		

Notes: The table presents the main descriptive statistics for men (first three columns) and women (last three columns). For each variable, the table presents the number of observations, the proportion out of the total for each gender, and the mean subjective life expectancy measured in years. The first row presents the total number of observations for men and women and the average SLE. The following groups show the statistics for the level of education (primary, high-school, vocational, or college education), the region of residence (metropolitan region of Santiago or other), and parental education (less than high-school, high-school degree, or more than high school).

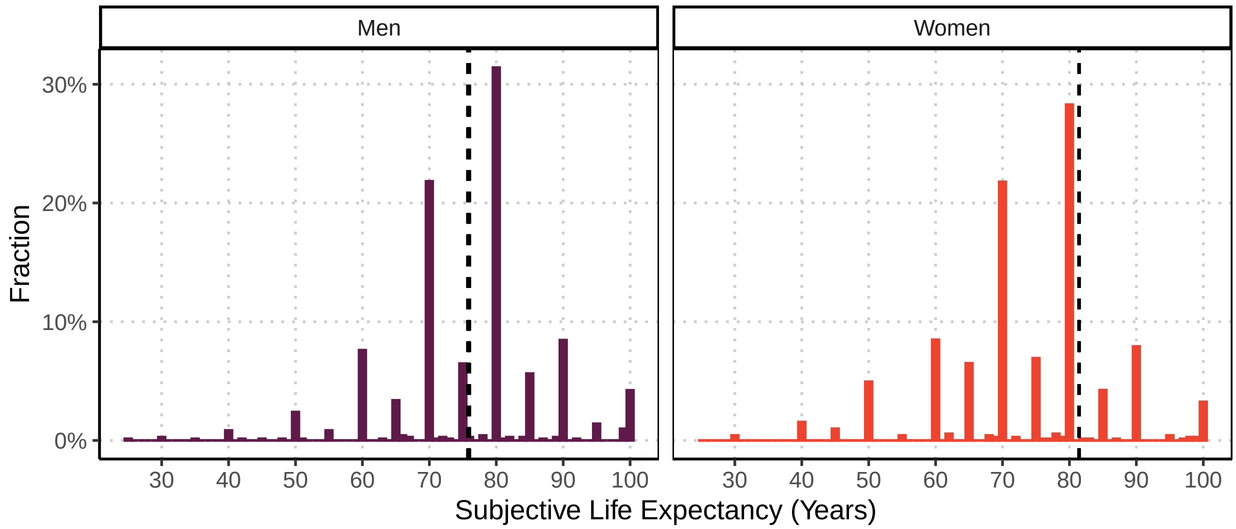
I expect the two questions to capture the same underlying subjective beliefs around mortality. Indeed, empirically, the two are highly correlated. In the analysis, I favor the first question for two main reasons. First, it is measured in years, arguably a unit of measure that all respondents are more familiar with than probabilities. Second, as I measure beliefs for a very young age range, most respondents answer 100% for the probability of being alive at age 65. While 48% of individuals answered 100% to the second question, only 3.7% answered any value above 100 years for the first question. Figure 1c plots the raw data for these two questions. We can see how there is much more dispersion in the first than in the second question. This is probably not the case for the papers analyzing beliefs for those over 50 with target ages beyond that.

The drawback of using the question measured in years is that it is unclear which statistic (mean, median, mode, or any other summary statistic) individuals are reporting. This would be more problematic if I were to recover the entire distribution of subjective life expectancy joint with any parametric restrictions. This is not the case. What I explore in the analysis is how future labor market behaviors compare with reporting high or low values of subjective life expectancy, conditional on a rich set of covariates.

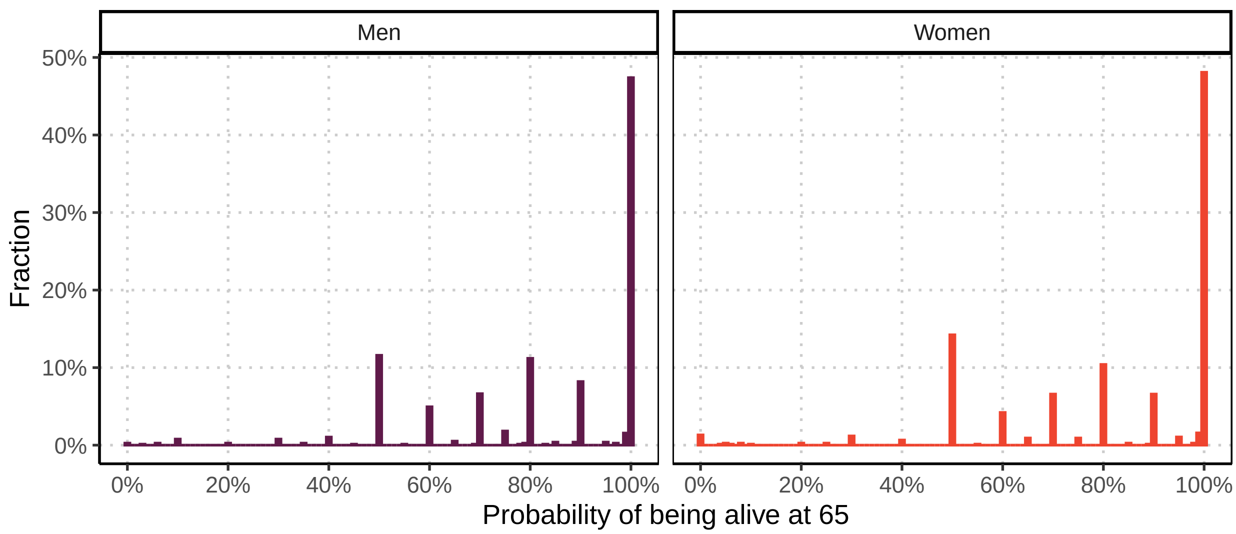
From the raw data in Figure 1c, we can see some rounding behavior. Most individuals answer in multiples of five. Men's answers are closer to the objective life expectancy from the life tables, while women's reports are significantly below. From this figure, the high dispersion of the answers is also evident. The standard deviation is around 12 years for men and 12.5 for women. Figure 2 plots the answers to the two questions for each individual, showing their positive correlation. The correlation coefficient is 0.46, which is probably affected by the ceiling at 100% in the probability of living at age 65 question. Figure A.2 plots the raw data by age and education.

I now turn to investigate what in the data is correlated with SLE. From now on, I will use the "net" SLE, which subtracts the objective life table life expectancy from each individual's answer, as the main variable. Every time I refer to SLE, I will be referring to this net variable unless explicitly written differently. The results are displayed in Table 2. The first column shows a regression, where the outcome is the net subjective life expectancy

(a) Subjective Life Expectancy (years)



(b) Probability of being alive at age 65



(c) Distribution of subjective survival belfies, by gender

Notes: The figure plots the histogram for the raw answers for the two subjective survival beliefs elicited in the survey, separately for men and women. The top panel plots the answer to the question on subjective life expectancy (Question: Up to what age do you believe you will live?) and the bottom panel to the question on probabilities of living up to age 65 (Question: What are the chances of living until 65?). The vertical dashed line shows the average life expectancy from the life tables for this sample.

(SLE) regressed on demographic cell (age-gender-education) fixed effects, region fixed effects, and parental education fixed effects. These controls explain about 16.3% of the variation in SLE. The following seven columns include additional controls related to health behaviors and diagnosis. We can see that all the coefficients have the expected sign. Smokers have,

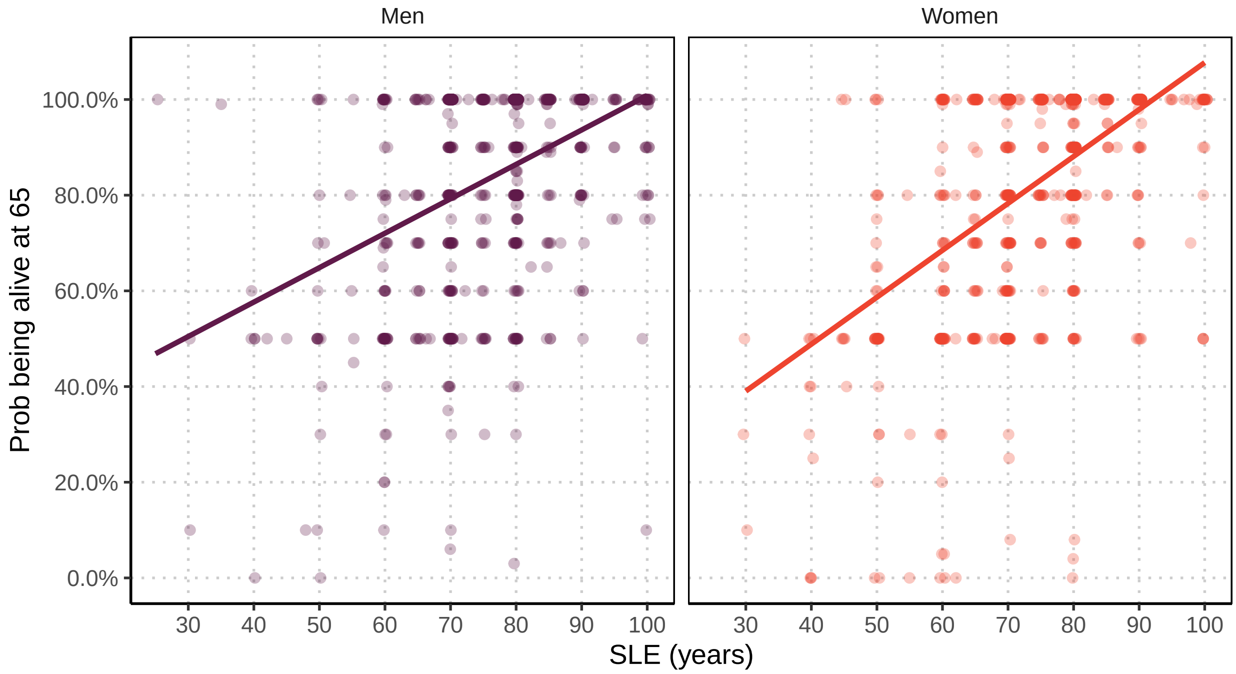


Figure 2: Subjetive life expectancy and probability of being alive at age 65

Notes: For each individual, the figure plots their answer to the two questions regarding life expectancy. On the x-axis is the SLE question measured in years, and on the y-axis is the question on the probability of being alive at age 65, measured in 0-100%. Dots are plotted with transparency, therefore, darker regions represent a higher mass of points. The solid line is the best linear fit of the data.

on average, -1.646 years of SLE. Doing physical activities regularly and assessing your own health status as good are associated with a higher SLE. While having any diagnosis in a given list of diseases, having a higher BMI, and having a deceased mother or father are associated with lower SLE.[12]

The last column includes all these variables in the same regression. Two interesting facts emerge. First, most coefficients are very stable. The exception is maternal death, which has a positive coefficient. As this is a very young sample, having a deceased mother is a rare event (less than 2.5% of the individuals in the sample have a deceased mother). More importantly, we can see that all these variables, including the rich set of baseline controls,

[12]The list of diagnoses includes asthma, pulmonary emphysema, depression, diabetes, hypertension, high blood pressure, heart problems, cancer, arthritis, osteoarthritis, renal diseases, stroke, mental illness, and HIV AIDS.

Table 2: Contemporaneously correlation of SLE

	Outcome: SLE								
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Smoking		-1.646							-1.350
		(0.669)							(0.723)
Regular Physical Activities			1.484						1.844
			(0.855)						(0.889)
Good Health				2.776					2.090
				(1.008)					(1.061)
Any Diagnosis (Diseases)					-1.560				-0.988
					(1.150)				(1.254)
BMI						-0.224			-0.264
						(0.089)			(0.095)
Deceased Mother							-1.545		0.258
							(2.595)		(2.612)
Deceased Father								-2.169	-1.959
								(1.367)	(1.397)
Controls	✓	✓	✓	✓	✓	✓	✓	✓	✓
Observations	1,417	1,417	1,416	1,417	1,414	1,416	1,412	1,239	1,231
R^2	0.163	0.167	0.165	0.168	0.164	0.168	0.163	0.179	0.196

Notes: The table presents the results from a regression of SLE on the demographic cell (age-gender-education) fixed effects, region fixed effects, and parental education fixed effects. The first columns present the baseline regression, from the second to the eighth column one control is added separately. The controls are, respectively, an indicator for smokers, an indicator for doing physical activities regularly, an indicator for reporting having good health (own assessment), having any diagnoses, body mass index, and indicators for having a deceased mother or father. The last column adds all variables simultaneously.

do not explain much of the SLE dispersion. The R-squared of this regression is 19.6%. This is not unique to this setting; Hurd and McGarry (2002), Puri and Robinson (2007), and Delavande et al. (2017) report similar findings.

Table 3 shows the correlations of SLE with measures of risk aversion and financial/literacy and numeracy. All columns control for the demographic-cell, region, and parental education fixed effects. We can see no correlation between subjective life expectancy and risk aversion and numeracy measures. Risk aversion 25 is measured with a question asking if individuals prefer a job paying X with certainty or a job paying 2X and 25%X with equal probabilities. It is coded as one if the individual chooses the constant wage. Risk aversion

50 and 75 are defined similarly. The numeracy variable ranges from 0 to 6, depending on the number of corrected answers in 6 math questions.[13] Appendix Table A.4 shows how these measures correlate with gender and educational levels in the expected way.

Table 3: SLE, risk aversion, and numeracy

	Outcome: SLE			
	(1)	(2)	(3)	(4)
Risk aversion 25	0.004			
	(0.008)			
Risk aversion 50		0.006		
		(0.007)		
Risk aversion 75			0.011	
			(0.007)	
Numeracy				0.001
				(0.003)
Observations	1,411	1,410	1,410	1,397
R^2	0.164	0.165	0.166	0.164

Notes: The table presents the correlation between the SLE variable and risk aversion and numeracy variables. All the regressions include the baseline controls (demographic cell, region, and parental education). The first three columns assess the conditional correlation between SLE and dummy variables that take the value one if individuals were unwilling to accept a risky wage scheme instead of a constant one. Please refer to the text for a precise definition of the risk aversion variables — they are coded as indicators for choosing the safest option (more risk aversion). The numeracy question goes from 0 to 6, with higher values associated with higher numeracy. Heteroskedastic-robust standard errors are presented in parenthesis.

One question is whether these beliefs are predictive of mortality. As these individuals were only 18–26 years old in 2004, and I observe them until 2019, there is not enough time to observe real mortality. However, in Appendix A, I show how using an alternative sample of older individuals, we see that their life expectancy is correlated with lower mortality rates. For instance, in a sample aged 35–45 in 2004, men who reported 10 years higher SLE are 1.3 percentage points more likely to be alive 15 years later, on a base mortality of 4.2% for

[13] Unfortunately, the numeracy questions were not included in the 2004 wave. I use here responses from the first time they were asked these questions, which typically occurred in the 2006 wave.

this sample. I see similar results for different age ranges.

The proportion of non-reporting for the two variables measuring subjective survival beliefs is low. Only 8.9% and 1.9% did not answer the first and second questions, respectively. Appendix B explores the non-reporting, showing how it is not correlated with the outcomes of interest.

In summary, individuals answered two questions eliciting beliefs around survival chances that are positively correlated. The answers correlate with expected demographics and health and lifestyle behaviors. Nevertheless, it displays substantial dispersion that cannot be explained by various demographics and health controls. Risk aversion and numeracy are not correlated with their reportings after controlling for demographic characteristics. Even though I cannot test directly for my sample, subjective beliefs of older individuals are associated with future mortality.

4 Empirical Strategy

The empirical strategy contrasts future labor market behavior with initial subjective life expectancy, controlling for the main demographic variables. I run the following OLS regression:

$$Y_{it} = \beta \text{SLE}_i + \eta_{d(i)} + \nu_{r(i)} + \varphi_{p(i)} + \theta_t + \varepsilon_{it}, \tag{1}$$

where Y_{it} is the outcome for individual i in time t, regressed on the SLE reported in 2004. The regression controls for demographic-cell fixed effect ($\eta_{d(i)}$), comprised of gender, age, and educational level, for region fixed effects ($\nu_{r(i)}$), parental education fixed effects ($\varphi_{p(i)}$) separately for mothers and fathers, and time fixed effects (θ_t). β, therefore, is our coefficient of interest, measuring the correlation between the outcome and subjective beliefs, conditional on our rich set of covariates. If there is no omitted variable that correlates with SLE and Y, we could interpret β as a causal effect of subjective life expectancy.

One concern to specification 1 is measurement error. As SLE is self-reported in a survey, it potentially contains measurement error that would attenuate the estimated coefficient of interest. To address this concern, I also use an instrumental variables approach, leveraging the two questions on subjective life expectancy in the survey. Therefore, I instrument SLE using the response on the probability of being alive at age 65 (P^{65}) as an instrument, running a two-stage least squares regression:

$$\begin{aligned} \text{(First Stage) } SLE_i &= \alpha P_i^{65} + \eta_{d(i)} + \nu_{r(i)} + \varphi_{p(i)} + \epsilon_{it} \\ \text{(Second Stage) } Y_{it} &= \beta_{IV} SLE_i + \eta_{d(i)} + \nu_{r(i)} + \varphi_{p(i)} + \theta_t + \zeta_{it} \end{aligned} \tag{2}$$

To ease interpretation, I report the coefficients of changing 10 years of SLE. This is the gap in subjective life expectancy between individuals in the 25th and 75th percentiles of the SLE distribution, conditional on gender. The equivalent 90th–10th gap is 30 years. For men, one standard deviation of SLE is 12 years, and for women, 12.5. Therefore, all the coefficients can be interpreted as an increase of approximately 0.8 standard deviations. Whenever there is more than one observation per individual, standard errors are clustered at the individual level. For cross-section regressions, the standard errors are heteroskedastic robust.

5 Results

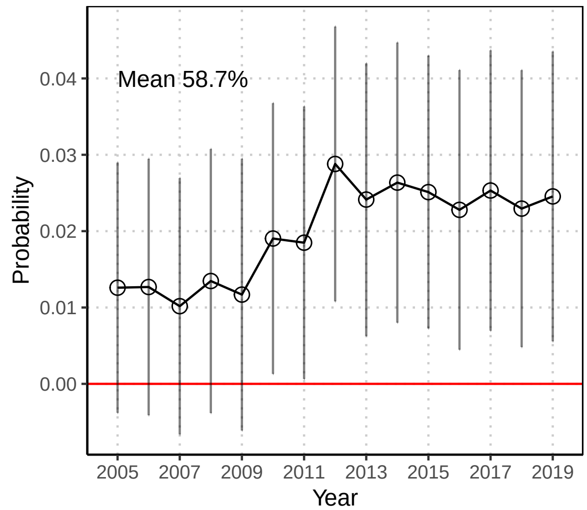
I first explore the richness of the administrative data and compute regression 1 using as outcomes the probability of making pension contributions, the number of contributions (stock), and pension wealth for all months between 2005 and 2019. The results are summarized in figure 3. In the top left plot (3a), we can see the correlation between reporting 10 years higher SLE and the probability of making a pension contribution in any given month.

The circles show the β coefficient from the baseline regression, which controls for demographic cells (gender-age-education), region, and parental education. For the first years, the average coefficient is around 1pp and 2.5pp for the remaining sample. That is, individuals who report 10 more years of SLE are more likely to make pension contributions. In terms of magnitude, the overall probability of making a pension contribution is 58.7% around this period. That is, those with 10 years higher SLE are 2–4.5% more likely to contribute. The smaller coefficient in the first years may be attributable to individuals still in schools/universities and navigating the labor market entrance. Nevertheless, I cannot statistically reject that these coefficients are all equal.

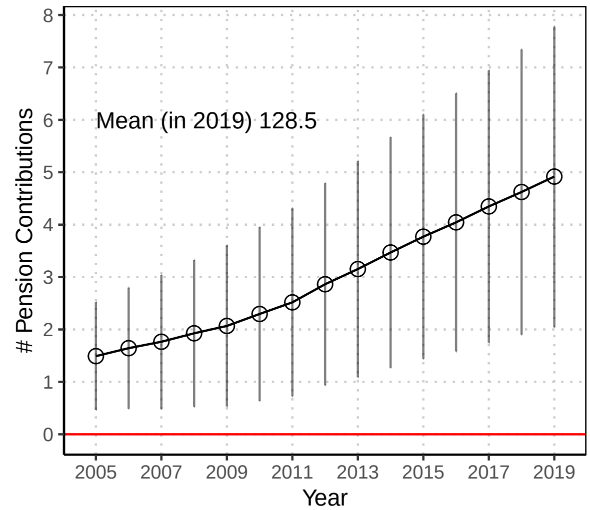
In panel 3b, I show the results for the number of pension contributions individuals have accumulated. Given that those with higher SLE were constantly more likely to make pension contributions, it is unsurprising to see the estimates rising over time. In 2019, those with 10 years more SLE have, on average, five months more of pension contributions. Which translates to 3.8%, using the average number in December 2019 of 128.5. We can see the same in panel 3c, where the pension wealth (in thousands of Chilean pesos) is used as the outcome variable.[14] In 2019, those with 10 years higher SLE have a pension wealth 360 thousand Chilean pesos higher. Using the mean in December of 2019, that corresponds to 5.7%. As an illiquid investment, pension contributions made earlier in life accumulate returns for a longer period.

Table 4, panels A and B, present these final results in 2019 for the overall sample and separately for men and women. For each sample (pooled, men, and women) and each variable (rows), the first column presents the outcome mean, and the second column presents the OLS estimation of equation 1. Looking at the pension status on the last date of the sample, in December 2019, there are two summary measures: the number of pension contributions (panel A) and the pension wealth in thousands of Chilean pesos (panel B). Both measures

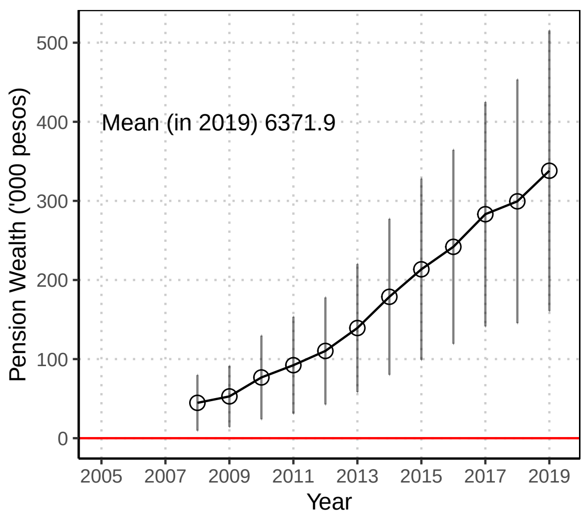
[14]The pension wealth variable is winsorized at the top 5%, separately for each month.



(a) Probability of making pension contributions



(b) # Pension contributions (stock)



(c) Pension wealth ('000 pesos)

Figure 3: SLE and future pension contributions

Notes: The figure plots the results from the OLS estimation of equation 1. Panel (a) for the binary outcome on the probability of making monthly pension contributions, panel (b) for the number of total pension contributions (stock), and panel (c) pension wealth measured in thousands of Chilean pesos. The regressions are run by pooling all monthly observations but separately for each calendar year. The circles are the estimated coefficient for β, and the solid vertical lines are the 95% confidence intervals. The text in each graph displays the mean for the outcome variable for the entire period (panel a) and for December 2019 (panels b and c). All regressions include the baseline controls (demographic cells, region, parental education, and time fixed effects). Standard errors are clustered at the individual level.

show a much higher coefficient for men than for women. Men reporting 10 extra years of SLE have, on average, 7.3 extra months of pension contributions and pension wealth 511.4

thousand pesos higher. Using the outcome averages, these correspond to 5.1% and 6.8%, respectively. For women, these numbers are smaller; 10 extra years of SLE is associated with 2.6 extra months of pension contribution (2.3%) and 202 thousand pesos higher pension wealth (3.9%).

Table 4: SLE and future labor market outcomes

	All			Men			Women		
	Mean	OLS	IV	Mean	OLS	IV	Mean	OLS	IV
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Panel A. # Pension Contributions (stock) in Dec2019									
SLE	128.478	4.917	7.860	144.239	7.306	9.581	112.778	2.649	6.583
		(1.461)	(3.136)		(2.060)	(4.920)		(2.011)	(4.036)
# Obs		1,417	1,408		710	706		707	702
# Individuals		1,417	1,408		710	706		707	702
Panel B. Pension Wealth ('000 pesos) in Dec2019									
SLE	6,371.9	362.5	685.5	7,500.8	511.4	927.0	5,247.4	202.7	488.7
		(101.2)	(224.1)		(157.8)	(389.7)		(128.1)	(266.1)
# Obs		1,417	1,408		710	706		707	702
# Individuals		1,417	1,408		710	706		707	702
Panel C. Labor Force Participation									
SLE	0.834	0.006	0.016	0.935	0.009	0.026	0.744	0.002	0.008
		(0.006)	(0.014)		(0.006)	(0.018)		(0.010)	(0.019)
# Obs		41,004	40,765		19,498	19,384		21,506	21,381
# Individuals		1,417	1,408		710	706		707	702
Panel D. Formal Sector									
SLE	0.566	0.015	0.042	0.650	0.024	0.074	0.491	0.006	0.022
		(0.008)	(0.019)		(0.012)	(0.032)		(0.011)	(0.022)
# Obs		41,004	40,765		19,498	19,384		21,506	21,381
# Individuals		1,417	1,408		710	706		707	702

Notes: The table presents the outcome mean (columns 1, 4, and 7), the OLS estimation of equation 1 (columns 2, 5, and 8), and the IV/2SLS estimation of 2 (columns 3, 6, and 9). The first 3 columns refer to the entire sample, the next 3 columns refer to the sample restricted to men and the final 3 for the sample restricted to women. There are four panels, each with one variable. Respectively, the total number of pension contributions in December 2019, pension wealth in December 2019 in thousands of Chilean pesos, indicator for labor force participation and indicator for formal employment (in a firm or the public sector). For the OLS and IV columns, the β coefficient is displayed with the estimated standard error (clustered at the individual) level in parenthesis. All regressions include the baseline controls (demographic cell, region, parental education, and time fixed effects). All panels show the number of observations in the regression and the number of individuals.

As discussed in Section 4, these coefficients are likely attenuated by measurement error in the self-reported life expectancy. To address this concern, I leverage the fact that the survey contains two different questions on life expectancy placed in different modules. I then use the answer to the second question, the probability of being alive at age 65, as an instrument for the SLE variable. The results are presented in the third column for each sub-panel in table 4. Across the board, we see higher coefficients in the IV/2SLS strategy than in the OLS, as would be the case if there is significant measurement error. However, the estimates are more imprecise, partially coming from the ceiling at 100% for the probability question. Table A.6 shows the first stage for these estimates. The F-statistic for the pooled sample is 341.4. For the analysis by gender, the F-statistics are 116.6 (men) and 225.5 (women). The IV estimates imply that those with (predicted) 10 years higher SLE have 6.1% higher pension contributions for the pooled sample and, respectively, 6.6% and 5.8% for men and women. In terms of pension wealth, these numbers compound to 10.8% (pooled), 12.4% (men), and 9.3% (women).

Panels C and D explore the survey data to shed some light on the mechanisms that explain these results. For this exercise, to avoid issues with recollection, I only use spells reported not later than 12 months. Therefore, there are many observations for each individual. I investigate whether individuals are contributing more because they are more likely to be employed or because they are more likely to be working in occupations with mandatory pension contributions, such as having a formal job in a firm or being a public employee. In Panel C, I run the OLS and IV regressions using an indicator for participation in the labor survey as the outcome variable. Individuals reporting 10 extra years of SLE are 0.6pp (OLS) and 1.6pp (IV) more likely to participate in the labor force. We see a much larger effect in Panel D, with an indicator for participating in the formal sector (being formally employed in a firm or the public sector). The same coefficients are 1.5pp and 4.2pp. That is, even though labor market participation explains part of the documented pension gaps, employment type — whether working formally — accounts for a larger share. Using the IV estimates, men with higher expectations are 2.8% more likely to be in the labor force and 11.4% more likely to be employed in a formal job. For women, these numbers are 1.1% and 4.5%.

One interesting aspect is whether these estimated relations depend on risk aversion. More risk-averse individuals may be particularly sensitive to beliefs on life spam. Table 5 presents results from the main specification, including an interaction of the SLE variable with the three different indicators for risk aversion. These variables are indicators for choosing the constant wage over a lottery of two times the wage and 25%/50%/75% of the wage with equal probabilities. Across all specifications, the points estimated for the interaction of risk aversion and SLE are always positive. This is consistent with more risk-averse individuals being more sensitive to their survival beliefs. However, they are not statistically different from zero at the usual significance levels.

Table 5: SLE and risk aversion

	Outcome: # Pension Contributions (stock) in Dec2019					
	OLS			IV		
	Risk25 (1)	Risk50 (2)	Risk75 (3)	Risk25 (4)	Risk50 (5)	Risk75 (6)
SLE	2.081	2.612	3.184	2.332	1.663	5.836
	(2.930)	(2.455)	(2.155)	(6.637)	(5.499)	(4.642)
SLE x Risk Aversion	3.417	3.218	2.590	6.807	8.603	2.849
	(3.298)	(2.953)	(2.789)	(7.421)	(6.513)	(6.091)
Observations	1,411	1,410	1,410	1,402	1,401	1,401

Notes: The table presents the results of a regression where the outcome variable is the total number of pension contributions in December 2019 on the SLE variable and an interaction of SLE and a dummy for risk aversion. Refer to the text for the construction of the risk aversion variable — one indicates more risk aversion (choosing a constant wage over a lottery). The first three columns are from the OLS estimation, and the last three are from the IV/2SLS estimation. The instrument is the probability of being alive at age 65 and the interaction of this variable with the risk aversion dummy. Each column uses a different variable measuring risk aversion. All regressions include the baseline controls (demographic cell, region, and parental education). Heteroskedastic robust standard errors are presented in parentheses.

5.1 Revisions

One question that naturally arises in this setting is how persistent these expectations are and if individuals revise their actions when revising their expectations. I leverage the longitudinal dimension of the survey and the fact that this question was repeatedly asked until the 2012 wave.[15] Table 6 shows the correlation between SLE in 2004 and subsequent waves, controlling for the same set of controls as in equation 1. We can see that 1 extra year of SLE in 2004 corresponds to 0.4 extra years in their subjective expectations in 2006, almost 0.3 in 2009, and around 0.15 in 2012, 8 years after it was initially measured. That is, expectations show some stability, even after controlling for a rich set of variables. However, the long-term correlation is far from one. We would expect that, particularly in two cases, (i) if there is measurement error when eliciting these expectations and (ii) individuals indeed revise their expectations given that in these long spans, they are subjected to several shocks and discover new information. The second panel in Table 6 investigates the role of measurement error, using the second elicited question on life expectancy as an instrument. The estimated coefficients are larger for 2006 and 2012 and slightly smaller for 2009, being, however, estimated with much lower precision.

To investigate the role of new shocks and information, I explore the role of new diagnoses. In each survey, individuals were asked whether they had been diagnosed with a given list of health diseases, which I divided into three groups. The less severe diseases (group 1) include asthma, pulmonary emphysema, diabetes, arthritis, and osteoarthritis. The more severe diagnoses (group 2) include hypertension, high blood pressure, heart problems, cancer, renal diseases, stroke, and HIV AIDS. And lastly, group 3 with any mental illness and depression. Table 7 shows the results of a regression contrasting SLE in two consecutive surveys and a dummy variable for whether the individual received a new diagnosis of diseases in groups 1–3 between the two surveys. We can see that the overall correlation between the two

[15]The question is absent from the 2015 and 2019 waves.

Table 6: SLE in 2004 and subsequent waves

	Outcome: SLE in		
	2006	2009	2012
	(1)	(2)	(3)
Panel A. OLS			
SLEin 2004	0.399	0.283	0.148
	(0.037)	(0.036)	(0.034)
Observations	1,088	1,008	828
Panel B. IV			
SLEin 2004	0.447	0.271	0.236
	(0.073)	(0.071)	(0.076)
Observations	1,080	1,002	822

Notes: The table shows the results from a regression of SLE measured in future surveys (2006, 2009, and 2012, respectively) on the initial SLE reported in 2004. All regressions included the baseline controls (demographic cell, region, and parental education). The results in panel A are from the OLS estimation. The results in panel B use an IV/2SLS strategy, using the probability of being alive at age 65 as an instrument for SLE in 2004. Standard errors clustered at the individual level are presented in parentheses.

measures of SLE is 0.33, in line with the results presented in Table 6. A diagnosis of a group 1 disease is associated with a reduction of 0.9 years in SLE, statistically indistinguishable from zero. At the same time, a diagnosis of a group 2 disease is associated with a reduction of SLE in 5.2 years and a reduction of 2.9 years for group 3 diseases, both statistically significant at the 5% level. It is reassuring to observe that individuals revise their subjective life expectancy in response to new health shocks.

The main results showed that individuals with higher subjective life expectations in 2004 had more pension contributions, higher labor force attachment, and higher chances of being employed in formal occupations. One question is whether future subjective life expec-

Table 7: SLE in subsequent surveys and new diagnosis

	Outcome: SLE $_{t+1}$		
	(1)	(2)	(3)
SLE $_t$	0.331	0.330	0.330
	(0.026)	(0.026)	(0.026)
New Diagnosis Group 1	-0.909		
	(1.710)		
New Diagnosis Group 2		-5.226	
		(1.768)	
New Diagnosis Group 3			-2.908
			(1.372)
Observations	2,007	2,010	2,010

Notes: The table shows the results from an OLS regression comparing two SLE reports in two consecutive surveys. The outcome variable is the SLE in the second survey for all three columns. The first role shows the coefficient associated with the initial SLE. Each of the three columns adds a variable, indicating whether the individual received a new diagnosis between the two surveys. Group 1 includes new diagnoses for asthma, pulmonary emphysema, diabetes, arthritis, and osteoarthritis. Group 2 includes hypertension, high blood pressure, heart problems, cancer, renal diseases, stroke, and HIV AIDS. Lastly, group 3 includes mental illness and depression. All regressions include the baseline controls (demographic cell, region, and parental education). Standard errors clustered at the individual level are displayed in parentheses.

tations have additional effects, controlling for the initial measure in SLE. Table 8 shows the results of an exercise controlling for SLE in different waves, with the baseline set of controls. The first column reproduces the main effect presented in Table 4 — 10 extra years of SLE in 2004 are associated with around five extra months of pension contribution in December 2019. Column 2 adds the observed SLE in 2006 as a control. The estimated coefficient shows that controlling for the initial SLE in 2004, an extra 10 years of subjective expectations in 2006 are associated with further increases of 2.8 months in pension contribution. The estimates are not statistically significant, though, as the two variables are highly correlated. Similar results are displayed for SLE in 2009 and 2012, with the last one presenting a coefficient marginally significant. The last three columns show the same exercise as the IV approach, using the reported probability of being alive up to 65 as an instrument. Column 5 only reproduces the IV results from Table 4. Columns 6 and 7 use as instruments the P^{65} variable in 2004 and 2006/2009. Unfortunately, this second question was not recorded in 2012. As we saw before, the estimates are overall higher when using the instrument to correct for measurement error. Therefore, this evidence is consistent with individuals acting in ways aligned with revisions of their subjective expectations; those revising their SLE upwardly exhibit more pension contributions.

Table 8: SLE and future revisions

	# Pension Contributions (stock) in Dec2019						
	OLS				IV		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
SLE in 2004	4.917	3.212	3.463	4.450	7.860	3.476	2.251
	(1.461)	(1.858)	(1.854)	(1.871)	(3.136)	(4.853)	(4.415)
SLE in 2006		2.762				4.594	
		(1.805)				(5.606)	
SLE in 2009			2.152				12.868
			(2.089)				(6.143)
SLE in 2012				4.011			
				(2.191)			
Observations	1,417	1,088	1,008	828	1,408	1,080	1,001

Notes: The table shows the results from regressions where the outcome variable is the total number of pension contributions in December 2019. The first four columns are the results of an OLS estimation. The first column only includes the baseline regressor, SLE in 2004. The next three columns include SLE measured in 2006, 2009, and 2012. The last three columns are from an IV/2SLS estimation using the reported probability of living at age 65 in the two surveys as instruments. For instance, in column 6, the instruments are P^{65}_{2004} and P^{65}_{2009}. This question was not included in the 2012 wave. All regressions include the baseline controls (demographic cell, region, and parental education). Hetoreskedastic-Robust standard errors are presented in parentheses.

Another exercise exploring the panel dimension of the data is presented in Table 9. In this exercise, I explore variation in SLE solely coming from time variation by augmenting the main specification to include individual fixed effects. This specification addresses the concerns on omitted variables that may correlate with SLE but are permanent characteristics of individuals. For every individual, I consider all SLE reports between 2004 and 2009 and use the cumulative pension contributions at the time of each report as the outcome variable.[16] The first column shows that 10 extra years of within-individuals SLE are associated with 0.22 extra months of cumulative pension contributions. In the 2SLS/IV specification in the second column, I instrument each SLE with the contemporaneous report on the probability of being alive at age 65. We can see that 10 extra years of (predicted) within-individual SLE are associated with an increase of 2.12 months of pension contribution. The distance across two reports is, on average, 27 months. These results are similar to the one obtained using the cross-sectional variation. The OLS result is smaller, while the IV estimate is larger.

Table 9: SLE and Pension Contributions — Within-person variation

Outcome: # Pension Contributions		
	OLS	IV
	(1)	(2)
SLE	0.224	2.121
	(0.387)	(1.317)
Observations	2,271	2,270

Notes: The table shows the results of a regression where the outcome variable, the stock of pension contributions, is regressed on the SLE variable, individual fixed effects, and time since the last survey. The second column presents a 2SLS/IV estimation, using the probability of being alive at age 65 as an instrument for SLE. Standard errors are clustered at the individual level.

5.2 Robustness

In this section, I present a series of exercises to assess the robustness of the results presented so far. In the main specification, I control for a rich set of covariates: demographic

[16] I left the observations from 2012 out since only one of the questions eliciting survival beliefs was asked in this wave.

cell (gender-age-education) fixed effects with 2x8x4=72 levels, region fixed effects (13 levels), and maternal (3 levels) and paternal education (3 levels). In Table A.9, I show how the coefficient is stable when we include each of these controls sequentially, using the number of pension contributions in December 2019 as the outcome variable. The R-square goes from 0.2-1.2% in the regression with no controls to 20.5-21.6%. That is, even including controls that explain a great degree of the variation of the outcome variable, the estimated coefficient barely changes, actually increasing in magnitude. This lessens the concern that the estimated coefficient merely reflects unobserved variables that correlate with subjective life expectations and the main outcomes analyzed here.[17]

One concern is whether the positive results on pension contribution and labor force attachment could stem just from the fact that individuals with higher subjective life expectations also have better health status, as documented in Table 2. In Figure A.5, I reproduce the results from Figure 3, controlling for health status (own assessment) in 2004. The results barely change. That is, even when controlling for the initial health status, the association between SLE and future pension outcomes is essentially the same. I opt to leave health out of the main specification as it is likely that it is part of subjective life expectancy, which is the focus of the analysis. Nevertheless, I showed how SLE presents a substantial variation on top of all health and lifestyle behaviors observed in the data and how the associations presented are not exclusively driven by health.

Another concern is with the initial status in the labor market. In Figure 3, we can see that, using the baseline specification, individuals reporting higher SLE were also more likely to contribute to pensions and had higher stock already in 2005. Table A.5 also shows how individuals with formal jobs are more likely to report higher values for SLE. This is expected, as it has been explored throughout this paper, individuals who expect to live longer have behaviors in the labor market consistent with valuing stable jobs with pension contributions.

[17] As the estimated coefficients increase in magnitude, it is not feasible to compute the ratio of unobserved/observed impact on the primary outcome to explain the total effect as proposed by Oster (2019).

However, one can be concerned that this initial status may dominate future behavior and is not connected to life expectancy per se. In Figure A.5 and Table A.7, I reproduce the main results controlling for initial labor market status. The results are smaller but have the same overall pattern. Even conditioning on the same initial status in the labor market, those with higher SLE are more likely to contribute to pensions, accumulate more pension wealth, participate in the labor force, and work in formal jobs. I do not include initial status as a control in the main specification, as the contemporaneous choices in the labor market could already be chosen in response to each individual's beliefs.

Table A.8 also shows how the results do not depend on the chosen age range. The estimated coefficients are remarkably close for a variety of different age ranges. Table A.10 shows how we obtain similar results when using the probability of being alive at age 65 as the main variable. To facilitate the comparison, for the probability at age 65, I compute the coefficient associated with increases of 30pp in this probability, as this corresponds to the 75th-25th percentile gap for both men and women. The overall pattern is very similar across the two distinct measures. Two advantages of the probabilities measure are the lower proportion of non-response: 1.9% versus 8.9% for the question measuring in years and the correspondence with a well-defined object.

Another concern could be with the chosen specification, where SLE enters linearly. I use the non-parametric bins-regression proposed by Cattaneo et al. (2024) to assess this hypothesis. The results are presented in Figure A.4. We can see that the results from the linear specification, represented by the solid black curve, are within the 95% confidence intervals for men and women. Lastly, Appendix B investigates the role of non-reporting in the SLE variable, not finding any consistent pattern between missing and the observed outcomes.

5.3 Discussion

The above results showed that individuals who reported beliefs of longer life spans are also more likely to be employed in formal jobs in the private and public sectors, with mandatory pension contributions. After 15 years of the initial report, the 75th-25th gap in SLE translates to individuals having 9.3%–12.4% higher pension wealth. These results are consistent with individuals making labor market choices that depend on how beneficial they are in the future. Individuals who believe they will live longer value making tax-advantaged pension contributions more. The flip side is that individuals with shorter expectations will value more jobs that do not require those pension contributions.

In the analysis, I show how these elicited beliefs have good properties. They correlate with expected health and life behavior variables. They are internally consistent, modestly persistent over time, respond to new information, and predict future mortality. In order to interpret the estimated coefficients as causal links between subjective expectations and life choices, one needs to assume no relevant omitted variable. What the above exercises could show us is that these associations seem to be very stable, even when we include a rich set of covariates that are very relevant to the outcomes being measured, in the spirit of Altonji et al. (2005) and Oster (2019). Additionally, the results exploring solely within-individual variation in SLE showed how fixed unobserved terms are not a concern.

One challenge is, for instance, if reporting beliefs of longer life spans is also correlated with other beliefs that are important for determining behavior. For example, Puri and Robinson (2007) show how SLE can be correlated with general optimism. While I do not have any optimism variable to investigate this in my sample, all the results presented here seem consistent with the large number of papers reviewed by Hudomiet et al. (2023), where these beliefs indeed predict mortality. So, it seems these variables capture agents' private information over their survival chances. Additionally, the general optimism hypothesis would be difficult to reconcile with some of my results, for instance, the large drops in SLE following a new diagnosis, the interaction with risk aversion, and the within-individual variation results.

6 Theoretical Framework

In this section, I develop a simple theoretical framework that shows how ignoring heterogeneity in (subjective) life expectancy can lead to biased conclusions in life-cycle models. In this model, individuals live in a younger and an older stage of life. The instantaneous utility in both periods is u, a concave, increasing, and differentiable function. In the younger stage, an individual receives an exogenous income Y and sets the amount of savings a she can consume in the older stage. The interest rate and the policy environment are such that, to get a stream of income a in the older stage, the agent needs to invest τa in the younger stage. This can encompass private investment for retirement, including working and making social security contributions. There are two states of the world in the older stage. The good state, where individuals live "longer", is realized with probability π. To capture a longer life span in a simple two-period model, I assume that the good state generates H times more utility than the bad state, where $H > 1$.[18] Finally, the relative value of the older stage of life with respect to the younger stage is β. Therefore, an individual faces the following problem:

$$\max_{a \in [0, Y/\tau]} \left\{ u(Y - \tau a) + \beta\big(\pi H u(a) + (1 - \pi) u(a)\big) \right\} \tag{3}$$

Since the optimal amount of savings depends on how likely an agent believes she will live longer, we can define the $a^*(\pi)$ as

$$a^*(\pi) = \arg\max_{a} \left\{ u(Y - \tau a) + \beta\big(\pi H u(a) + (1 - \pi) u(a)\big) \right\}. \tag{4}$$

[18]Note, for instance, that if we assume that individuals live for two periods in the good state, that is equivalent of assuming $H = 1 + \beta$, if there are no savings between these two periods.

That is, this function records, for any π, the optimal level of investments across the two periods.

Figure 4 left plot shows the $a^*(\pi)$ function for a given choice of parameters.[19] We can see that individuals with higher π choose higher values of a. This is consistent with the evidence presented in this paper — young individuals with higher SLE are more likely to make pension contributions.

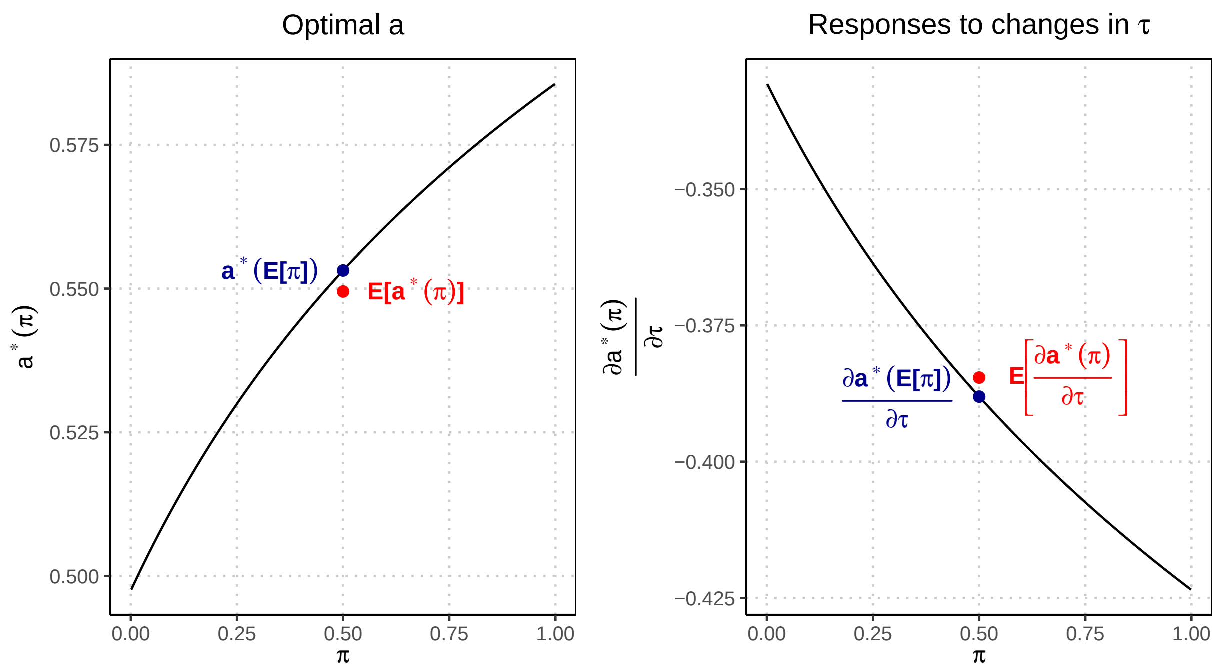


Figure 4: Theoretical Framework

Notes: The plot on the left shows the optimal choices of a as a function of the probabilities π (solid line). The blue point shows the optimal choice for the average value of π. The red point shows the average optimal choice across all values of π. The plot on the right has the same structure, plotting the changes in optimal response to changes in τ.

We often do not observe individual π and solve the model for the average $\mathbb{E}[\pi]$ in this economy. For example, in the case of life expectancy, it is common to obtain $\mathbb{E}[\pi]$ directly from life tables for some sub-groups, for example, men and women. The rationale of using $\mathbb{E}[\pi]$ is to assume that individuals have rational expectations and *do not have private*

[19] For this example, $Y = 1$, $\tau = 1$, $\beta = 0.97$, $H = 1 + \beta + \beta^2$, $u(c) = c^{(1-\sigma)}/(1-\sigma)$, $\sigma = 3$, and $\pi \sim U[0.001, 0.999]$.

information over own survival chances. The blue circle shows the prediction of a correctly specified model that uses only $\mathbb{E}[\pi]$, ignoring heterogeneity over π. That is, it just computes the optimal behavior function $a^*(.)$ at the average π. In contrast, the red circle shows the true average optimal choice over the distribution of π. Given the concavity of the $a^*(\pi)$ function and Jensen's inequality, we have that $a^*(\mathbb{E}[\pi]) \geq \mathbb{E}[a^*(\pi)]$, where expectations are taken over the distribution of π.[20] This exercise shows that models that ignore heterogeneity in subjective life expectancy and assume a representative agent with average life expectancy always overestimate the optimal average savings for retirement. This result reiterates the importance of the empirical findings, showing how individuals with different beliefs have different labor market and social security choices.

This theoretical framework can also be used to assess policy changes, particularly around social security. We can interpret τ as a reduced-form parameter relating investments for retirement to pension benefits. Social security reforms that decrease benefits or strengthen the requirements to access public pensions can be seen as increases in τ. Using the structure above, we can assess the changes in the optimal choice of investment from changes in τ: $\frac{\partial a^*(\pi)}{\partial \tau}$. The right plot from Figure 4 plots the values of this derivative for each value of π. When τ increases, all individuals reduce their investments for period 2, as transferring resources across the two periods is more costly. Individuals with higher π exhibit larger changes, reducing a^* to a greater extent. Similarly to the abovementioned arguments, we can contrast the predictions of a model that ignores heterogeneity in π. A model considering the response only on the average $\mathbb{E}[\pi]$ would overestimate the behavioral changes to policies that enhance τ. This could affect policy-making, as several policies could be decided upon the prediction of models that have biased predictions of individuals' responses to the policy changes.

In summary, this simple two-period model shows how individuals with different life

[20] Concavity of $a^*(\pi)$ does not depend on the specific parameters chosen. It inherits the concavity of $u(c)$. Appendix C shows how this holds for a large class of utility functions.

expectancy beliefs adopt different strategies relating to retirement. This is consistent with this paper's findings and, more broadly, with the literature analyzing workers closer to retirement. Moreover, this model highlights how ignoring heterogeneity in life-span beliefs can be consequential. "Naive" models, even if correctly specified and with the true parameters, would produce biased results. Such models would overestimate the optimal savings rates and behavioral responses to policy changes. This may lead to the adoption of suboptimal policies, as they wrongly predict the optimal response of individuals to the policy environment.

7 Conclusion

While most of the literature focuses on subjective life expectations of individuals near retirement, in this paper, I explore longitudinal survey data and administrative data from Chile to investigate SLE of young individuals. I show how even if these beliefs are elicited from a young population, they still present good properties: they are internally consistent, correlate with expected behavior, are modestly persistent, respond to new information, and predict future mortality.

Looking at early labor market choices, SLE is correlated with future labor market participation, with the likelihood of working in the formal sector and making pension contributions. I observe relatively large pension gaps associated with SLE, 15 years after it was measured. There is some evidence that men and more risk-averse individuals exhibit correlations that are more responsive to SLE. Lastly, exploring the longitudinal dimension for the SLE measures, I show results consistent with individuals revising their labor market choices according to their revised beliefs. My results also illustrate the importance of correcting for measurement errors in these variables, leveraging the existence of different questions eliciting the same underlying beliefs.

From a pure measurement perspective, it is interesting that SLE measured at young ages correlates so well with future labor market choices. On its own, this fact already advocates for the broader use of these measures in economics and policy. Given how consequential early labor market choices are for career progression and overall welfare, these results ask for new research to further explore how these beliefs are formed, their accuracy, and their consequences.

REFERENCES

Altonji, J. G., Elder, T. E. and Taber, C. R. (2005). Selection on observed and unobserved variables: Assessing the effectiveness of catholic schools, *Journal of political economy* **113**(1): 151–184.

Altonji, J. G., Kahn, L. B. and Speer, J. D. (2016). Cashier or consultant? entry labor market conditions, field of study, and career success, *Journal of Labor Economics* **34**(S1): S361–S401.

Arellano-Bover, J. (2020). The Effect of Labor Market Conditions at Entry on Workers' Long-Term Skills, *The Review of Economics and Statistics* pp. 1–45.
URL: *https://doi.org/10.1162/rest_a_01008*

Bello, P. (2019). Gender price discrimination in the annuity market: Evidence from chile, *Technical report*.

Bissonnette, L., Hurd, M. D. and Michaud, P.-C. (2017). Individual survival curves comparing subjective and observed mortality risks, *Health economics* **26**(12): e285–e303.

Bloom, D. E., Canning, D., Moore, M. and Song, Y. (2006). The effect of subjective survival probabilities on retirement and wealth in the united states, *Technical report*, National Bureau of Economic Research.

Bresser, J. d. (2021). Evaluating the accuracy of counterfactuals the role of heterogeneous expectations in life cycle models.

Cattaneo, M. D., Crump, R. K., Farrell, M. H. and Feng, Y. (2024). On binscatter, *American Economic Review* **114**(5): 1488–1514.

Delavande, A., Lee, J. and Menon, S. (2017). Eliciting survival expectations of the elderly in low-income countries: Evidence from india, *Demography* **54**(2): 673–699.

Finamor, L. (2023). Labor market informality, risk, and public insurance.

Gan, L., Gong, G., Hurd, M. and McFadden, D. (2015). Subjective mortality risk and bequests, *Journal of Econometrics* **188**(2): 514–525.

Hamermesh, D. S. (1985). Expectations, life expectancy, and economic behavior, *The Quarterly Journal of Economics* **100**(2): 389–408.

Hamermesh, D. S. and Hamermesh, F. W. (1983). Does perception of life expectancy reflect health knowledge?, *American Journal of Public Health* **73**(8): 911–914.

Heimer, R. Z., Myrseth, K. O. R. and Schoenle, R. S. (2019). Yolo: Mortality beliefs and household finance puzzles, *The Journal of Finance* **74**(6): 2957–2996.

Hudomiet, P., Hurd, M. D. and Rohwedder, S. (2023). Mortality and health expectations, *Handbook of Economic Expectations* pp. 225–259.

Hurd, M. D. and McGarry, K. (2002). The predictive validity of subjective probabilities of survival, *The Economic Journal* **112**(482): 966–985.

Hurd, M. D., Smith, J. P. and Zissimopoulos, J. M. (2004). The effects of subjective survival on retirement and social security claiming, *Journal of Applied Econometrics* **19**(6): 761–775.

Ministerio del Trabajo y Prevision Social (2002-2020). Encuesta de Proteccion Social.

O'Dea, C. and Sturrock, D. (2023). Survival pessimism and the demand for annuities, *The Review of Economics and Statistics* **105**(2): 442–457.

O'Donnell, O., Teppa, F., Van Doorslaer, E. et al. (2008). Can subjective survival expectations explain retirement behaviour, *DNB WP188* .

Oster, E. (2019). Unobservable selection and coefficient stability: Theory and evidence, *Journal of Business & Economic Statistics* **37**(2): 187–204.

Puri, M. and Robinson, D. T. (2007). Optimism and economic choice, *Journal of financial economics* **86**(1): 71–99.

Salm, M. (2010). Subjective mortality expectations and consumption and saving behaviours among the elderly, *Canadian Journal of Economics/Revue canadienne d'économique* **43**(3): 1040–1057.

Superintendencia de Pensiones, Chile (1980–2019). Historia Previsional de Afiliados Activos, Pensionados Y Fallecidos.

Van der Klaauw, W. and Wolpin, K. I. (2008). Social security and the retirement and savings behavior of low-income households, *Journal of econometrics* **145**(1-2): 21–42.

Wu, S., Stevens, R. and Thorp, S. (2015). Cohort and target age effects on subjective survival probabilities: Implications for models of the retirement phase, *Journal of Economic Dynamics and Control* **55**: 39–56.

Online Appendices

A Subjective Life Expectancy and Mortality

The main sample is too young to allow me to draw any inference on the true mortality patterns. Individuals were 18–26 years old in 2004 and, therefore, can be observed with the administrative data when they were, at most, 32–40. To shed some light on the predictive power of SLE in predicting survival, I analyze future mortality and initial SLE responses for older cohorts. Specifically those that were 35–45 and 45–55 in 2004.

I use mortality as measured in the administrative data of the pension system. Therefore, I do not have information for those not belonging to the system. That could be because they entered the labor force before 1980 and chose to remain in the older system or because they never had any contribution to either pension system. The results are presented in Table A.1 below.

In columns 1 and 3, I show the results for estimating the main equation using an indicator for whether the individual was ever registered in the pension system as of December 2019 as the outcome variable. Not surprisingly, we see a positive association in the two samples. That means that using the data from the pension system to observe mortality may underestimate the true associations between SLE and mortality, as those with lower expectations are less likely to be observed and more likely to die. Nevertheless, in columns 2 and 4, I show the result using as the outcome an indicator for being deceased by December 2019. All coefficients are negative and statistically significant except for women in the sample 35–45. Not only that, but the magnitudes are relevant. For men in the 45–55 sample, 10-year higher SLE is associated with a reduction in mortality by 3.1pp or 30.7%, using the mean of 10.1%. Once again, this is potentially underestimated because of measurement error on SLE and the non-random missing in the mortality variable. All regressions include

Table A.1: SLE and mortality

	Men		Women	
Outcome:	Registered	Deceased in 2019	Registered	Deceased in 2019
Model:	(1)	(2)	(3)	(4)
Panel A - Aged 35–45				
SLE	0.005	-0.013	0.009	0.000
	(0.003)	(0.006)	(0.005)	(0.003)
Observations	1,876	1,823	1,931	1,749
Mean	0.970	0.043	0.903	0.028
Panel B - Aged 45–55				
SLE	0.015	-0.031	0.013	-0.009
	(0.006)	(0.009)	(0.011)	(0.005)
Observations	1,531	1,383	1,574	1,198
Mean	0.894	0.101	0.757	0.061

Notes: The table presents the results from an OLS regression of equation 1 using an indicator for being registered in the pension system in December 2019 (columns 1 and 3) and being deceased by December 2019 (columns 2 and 4) as the outcome variables. The first two columns are for men, and the last two are for women. Panel A is for the sample of individuals aged 35–45 in 2004, and Panel B is for those aged 45–55. All regressions include the baseline controls (demographic cell, age, and parental education). Heteroskedastic-robust standard errors are included in parenthesis. The mean for each outcome and sample is also shown below the number of observations.

the baseline controls, fixed effects for demographic cells (gender-age-education), region, and parental education.

B Missing Subjective Life Expectancy

While the proportion of missing values for the subjective life expectancy variable is small (9%), in this section, I investigate who are the individuals not answering this question. Table A.2 presents the proportion of missing overall (first column) and by gender and education (column 2). There is no gender difference and a borderline significant gap in education. Those with more education are more likely to respond.

Table A.2: Proportion of missing values for SLE

	(1)	(2)
Constant	0.090	0.114
	(0.007)	(0.027)
Women		0.007
		(0.014)
High School		-0.022
		(0.027)
Vocational		-0.007
		(0.033)
College		-0.071
		(0.028)
Observations	1,557	1,557

Notes: The table shows the results of an OLS regression using an indicator for not reporting the SLE as the outcome variable. In the first column, only a constant is included. The second column adds an indicator for women and each educational level. In this column, men without high school degrees are the omitted category. Heteroskedastic-robust standard errors are presented in parenthesis.

Table A.3 investigates the correlation between not answering the SLE question and the main outcomes investigated. In no cases are the differences statistically significant. For men, the point estimates would suggest that those not reporting have a lower density of pension contributions and lower attachment to the formal sector. We would expect this if they were among those with lower subjective life expectancy. For women, we see the opposite. Women who do not report seem to have higher labor market attachment. However, none of these comparisons are statistically significant.

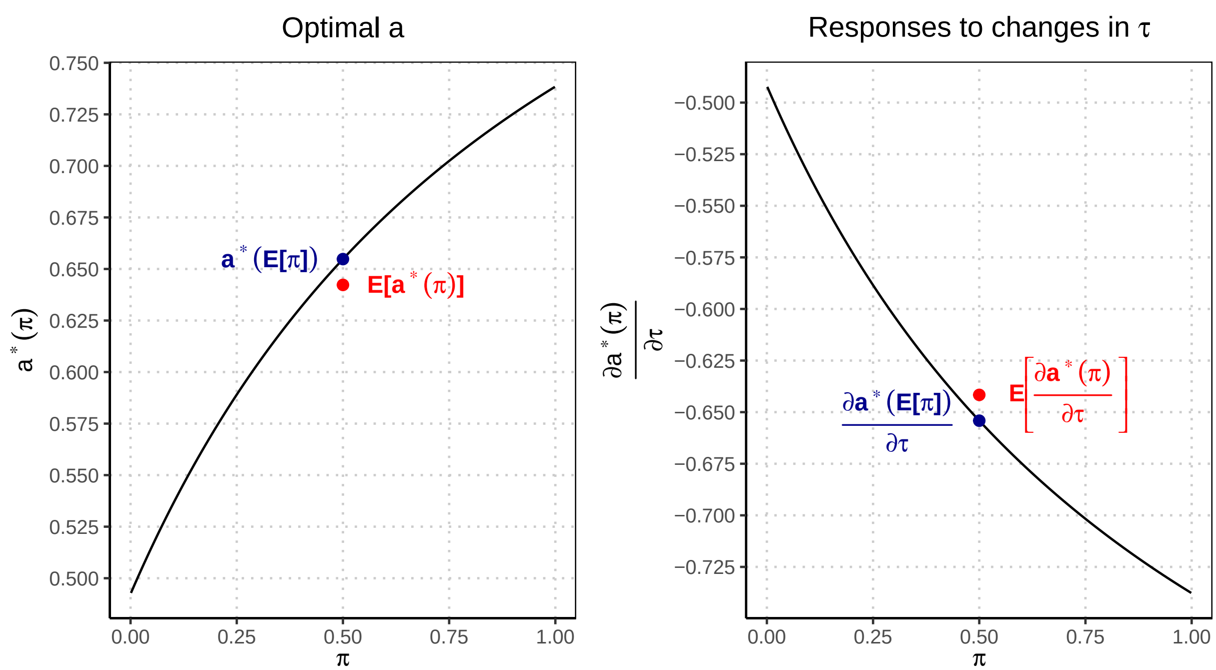
Table A.3: Missing SLE and main outcomes

Outcomes	# Pension Contrib	Pension Wealth	LFP	Formal
	(1)	(2)	(3)	(4)
Panel A. Men				
Missing SLE	-3.496	-593.673	0.006	-0.079
	(9.017)	(611.025)	(0.023)	(0.059)
Observations	777	777	21,149	21,149
Panel B. Women				
Missing SLE	3.665	327.677	0.007	0.026
	(8.188)	(494.260)	(0.042)	(0.046)
Observations	780	780	23,598	23,598

Notes: The table displays the correlation between not reporting the SLE (missing SLE) with four different outcomes: total number of pension contributions in December 2019 (column 1), pension wealth in December 2019 (column 2), indicator for labor force participation (column 3), and indicator for formal employment (column 4). Results are presented separately for men (panel A) and women (panel B). All regressions include the baseline controls (demographic cell, region, parental education, and time fixed effects). Standard errors are clustered at the individual level.

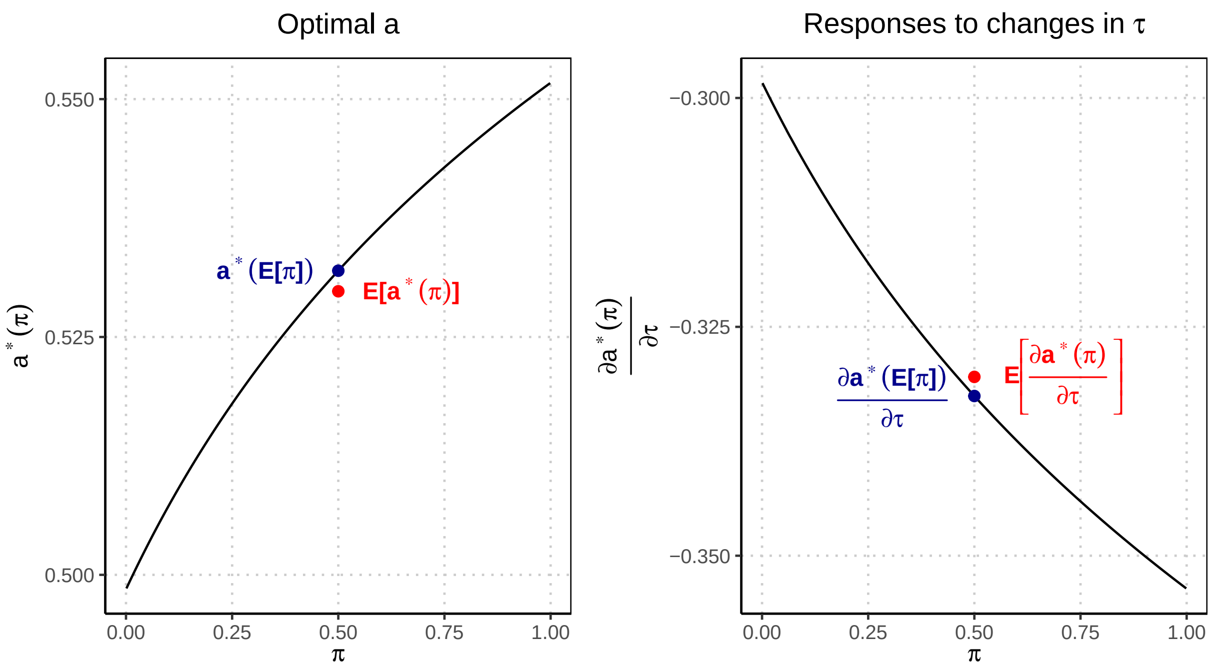
C Theoretical Framework — Other utility functions

The figure below shows the same results from section 6 with different utility functions.

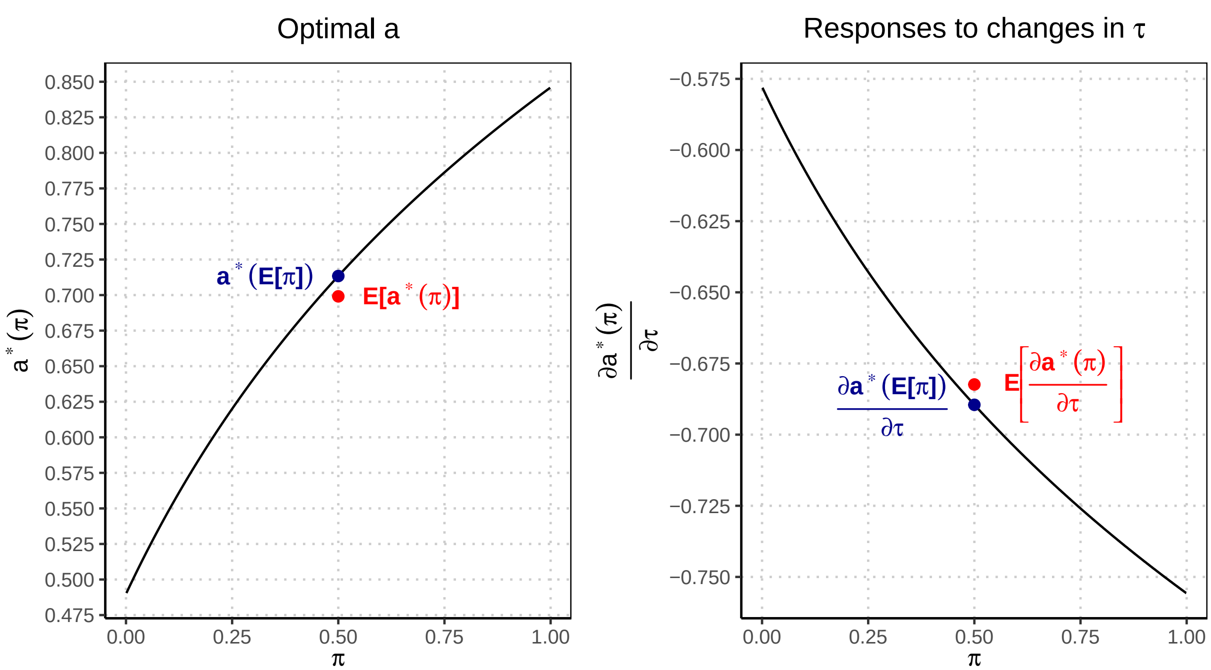


(a) Log-utility, $u(c) = log(c)$

Figure A.1: Theoretical framework with different utility functions



(b) CRRA, $u(c) = c^{(}1-\sigma)/(1-\sigma), \quad \sigma = 5$



(c) CARA, $u(c) = 1 - \exp(-\alpha c), \quad \alpha = 1.5$

Figure A.1: Theoretical framework with different utility functions (cont.)

D Additional Figures and Tables

Table A.4: Risk aversion and numeracy by gender and education

Outcome Variables:	Risk25	Risk50	Risk75	Numeracy
	(1)	(2)	(3)	(4)
Constant	0.766	0.660	0.567	1.591
	(0.033)	(0.038)	(0.041)	(0.124)
Women	0.103	0.112	0.109	-0.348
	(0.021)	(0.023)	(0.025)	(0.071)
High School	-0.027	-0.013	-0.001	0.843
	(0.034)	(0.039)	(0.042)	(0.126)
Vocational	-0.038	-0.006	-0.031	1.587
	(0.041)	(0.046)	(0.050)	(0.147)
College	-0.049	-0.084	-0.117	1.858
	(0.038)	(0.044)	(0.047)	(0.139)
Observations	1,550	1,548	1,548	1,529
R^2	0.017	0.019	0.021	0.153

Notes: The first column shows a regression of a binary variable for risk aversion, where one indicates that the individual would choose a constant wage over a lottery. The regressors are a dummy for women and dummies for high school, vocational education, and college. The omitted category is men with primary education. The second and third columns are similar, with different risk measures. The last column presents the same regression, using a numeracy variable, defined in the 0–6 range, as the outcome. Standard errors clustered at the individual level are presented in parentheses.

Table A.5: SLE and contemporary labor market choices

	Outcome: SLE					
	Men			Women		
	(1)	(2)	(3)	(4)	(5)	(6)
# Pension Contrib	0.077			0.018		
	(0.023)			(0.024)		
Labor Force Participation		-1.117			-0.364	
		(1.386)			(1.056)	
Formal			3.902			0.646
			(1.498)			(1.687)
Observations	710	708	465	707	707	325
R^2	0.083	0.069	0.124	0.099	0.099	0.150

Notes: The table presents the results of a regression where the outcome variable is the SLE variable on number of total pension contributions on the date of the survey (columns 1 and 4), indicator for labor force participation at the survey (columns 2 and 5), and indicator for formal employment, in a firm or the public sector (columns 3 and 6). The first three columns are for men, and the last three are for women. All regressions control for the baseline controls (demographic cell, region, and parental education).

Table A.6: Instrumental variables — first stage

	Outcome: SLE		
	All	Men	Women
	(1)	(2)	(3)
Prob Living 65	0.245	0.220	0.266
	(0.016)	(0.023)	(0.021)
Observations	1,408	706	702
F-stat	341.4	116.6	225.5

Notes: The table presents the results of the first stage IV/2SLS regression from equation 2. The SLE variable is being instrumented by the P^{65} variable. The F-statistic is displayed in the last row. The first column is for the entire sample, and the next two are for men and women. Standard errors are clustered at the individual level.

Table A.7: SLE and future labor market outcomes with additional controls

	All			Men			Women		
	Mean	OLS	IV	Mean	OLS	IV	Mean	OLS	IV
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Panel A. # Pension Contributions (stock) in Dec2019									
SLE	128.478	2.857	4.570	144.239	4.390	4.230	112.778	1.590	5.037
		(1.281)	(2.775)		(1.868)	(4.656)		(1.770)	(3.411)
Obs		1,415	1,406		708	704		707	702
Panel B. Pension Wealth ('000 pesos) in Dec2019									
SLE	6,371.9	257.3	539.6	7,500.8	335.8	691.8	5,247.4	156.6	414.4
		(96.9)	(216.7)		(153.7)	(392.8)		(124.2)	(249.0)
Obs		1,415	1,406		708	704		707	702
Panel C. Labor Force Participation									
SLE	0.834	0.003	0.011	0.935	0.006	0.014	0.744	0.001	0.008
		(0.006)	(0.013)		(0.006)	(0.018)		(0.010)	(0.018)
Obs		40,954	40,715		19,448	19,334		21,506	21,381
Panel D. Formal Sector									
SLE	0.566	0.008	0.033	0.650	0.016	0.058	0.491	0.001	0.018
		(0.008)	(0.018)		(0.012)	(0.033)		(0.011)	(0.021)
Obs		40,954	40,715		19,448	19,334		21,506	21,381

Notes: The table presents the outcome mean (columns 1, 4, and 7), the OLS estimation of equation 1 (columns 2, 5, and 8), and the IV/2SLS estimation of 2 (columns 3, 6, and 9). The first 3 columns refer to the entire sample, the next 3 columns refer to the sample restricted to men and the final 3 for the sample restricted to women. There are four panels, each with one variable. Respectively, the total number of pension contributions in December 2019, pension wealth in December 2019 in thousands of Chilean pesos, indicator for labor force participation and indicator for formal employment (in a firm or the public sector). For the OLS and IV columns, the β coefficient is displayed with the estimated standard error (clustered at the individual) level in parenthesis. All regressions include the baseline controls (demographic cell, region, parental education, and time fixed effects), health self-assessment in 2004, and status in the labor market in 2004.

Table A.8: Robustness — age range

	Outcome: # Pension Contributions in December 2019 (stock)						
Age range	18–26	18–24	18–22	18–28	18–30	20–26	22–26
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Panel A. Men							
SLE	7.306	7.520	7.056	6.512	5.246	7.818	6.250
	(2.060)	(2.405)	(3.424)	(1.877)	(1.802)	(2.158)	(2.431)
Observations	710	517	268	947	1,207	653	543
Panel B. Women							
SLE	2.649	2.903	2.857	1.820	0.340	2.825	4.221
	(2.011)	(2.285)	(3.108)	(1.935)	(1.841)	(2.073)	(2.352)
Observations	707	486	266	926	1,191	664	533

Notes: The table shows the results of the OLS estimation of equation 1 using the total number of pension contributions (stock) in December 2019 as the outcome variable. Each column uses a sample with a different age range, indicated in the top header. Panel A shows the results for men and panel B for women. All regressions include the baseline controls (demographic cells, region, and parental education). Standard errors are heteroskedastic robust.

Table A.9: Robustness — inclusion of controls

	Outcome: # Pension Contributions in December 2019 (stock)					
	(1)	(2)	(3)	(4)	(5)	(6)
Panel A. Men						
SLE	6.241	7.275	7.293	7.346	7.412	7.306
	(2.201)	(2.117)	(2.064)	(2.060)	(2.064)	(2.060)
Observations	710	710	710	710	710	710
R^2	0.012	0.111	0.179	0.199	0.213	0.216
Controls	None	+Age	+Age-Educ	+Region	+ Mother Educ	+Father Educ
Panel B. Women						
SLE	2.621	2.486	2.268	2.324	2.842	2.649
	(2.123)	(2.136)	(2.080)	(2.032)	(2.026)	(2.011)
Observations	707	707	707	707	707	707
R^2	0.002	0.028	0.178	0.199	0.203	0.205
Controls	None	+Age	+Age-Educ	+Region	+ Mother Educ	+Father Educ

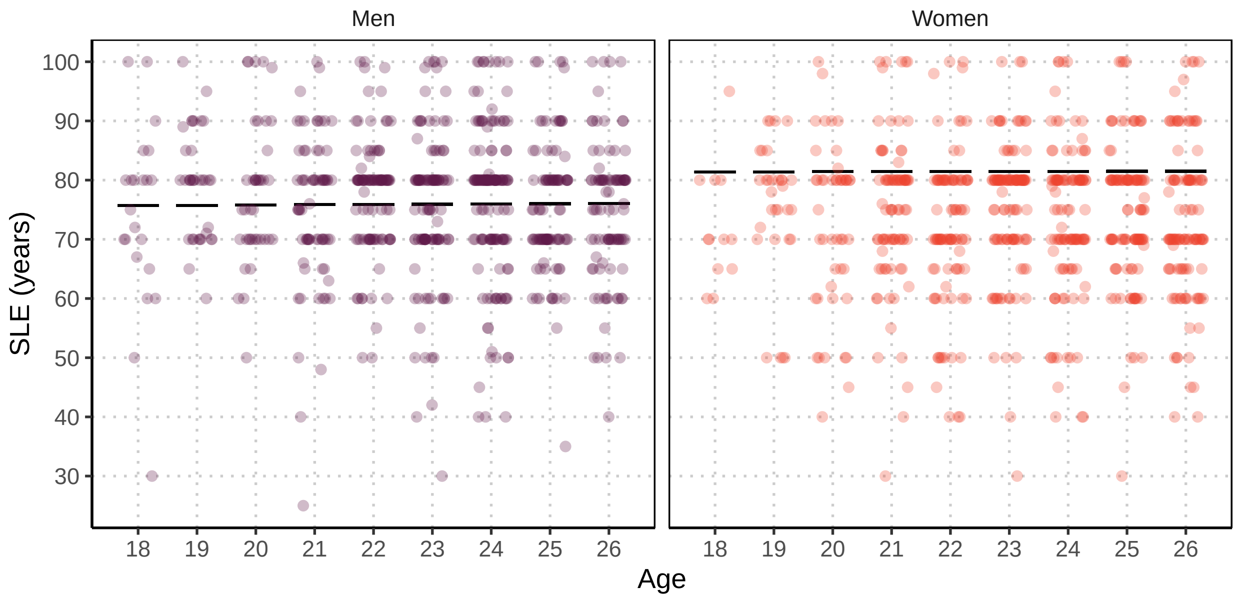
Notes: The table shows the results of the OLS estimation of equation 1 using the total number of pension contributions in December 2019 as the outcome variable. Each column progressively adds a control. The first column does not have any additional covariate. Age, age-education level, region, and mother and father's education are progressively added. Panel A shows the results for men and panel B for women. Standard errors are heteroskedastic robust.

Table A.10: Future labor market outcomes — SLE in years and Prob at age 65

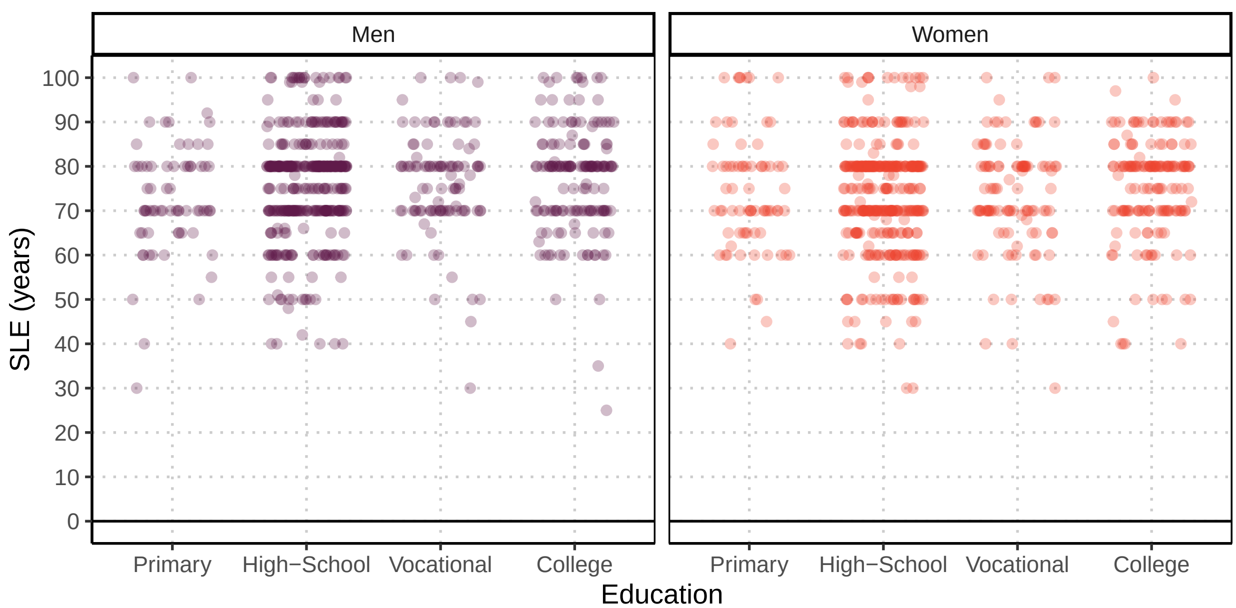
	SLE			Prob Age 65		
	All	Men	Women	All	Men	Women
	(1)	(2)	(3)	(4)	(5)	(6)
Panel A. # Pension Contributions (stock) in Dec2019						
SLE	4.917	7.306	2.649	4.965	5.594	4.279
	(1.461)	(2.060)	(2.011)	(2.188)	(3.097)	(3.054)
Obs	1,417	710	707	1,527	768	759
Panel B. Pension Wealth ('000 pesos) in Dec2019						
SLE	362.508	511.362	202.683	493.554	606.195	376.252
	(101.156)	(157.831)	(128.131)	(154.542)	(238.283)	(201.593)
Obs	1,417	710	707	1,527	768	759
Panel C. Labor Force Participation						
SLE	0.006	0.009	0.002	0.010	0.015	0.007
	(0.006)	(0.006)	(0.010)	(0.009)	(0.010)	(0.015)
Obs	41,004	19,498	21,506	43,906	20,861	23,045
Panel D. Formal Sector						
SLE	0.015	0.024	0.006	0.029	0.039	0.021
	(0.008)	(0.012)	(0.011)	(0.013)	(0.019)	(0.017)
Obs	41,004	19,498	21,506	43,906	20,861	23,045

Notes: The table presents the results of the OLS regression of equation 1 using the SLE variable (first three columns) or the P^{65} variable (last three columns) as the regressor. To make the two sets of results more comparable, I multiply the P^{65} coefficient by 30 to represent the 75th-25th gap. Therefore, it can be read as the coefficient associated with increasing P^{65} by 30 percentage points. Columns 1 and 4 are for the entire sample, columns 2 and 5 are for men, and columns 4 and 6 are for women. Each panel has a different outcome variable. Respectively, the total number of pension contributions in December 2019, pension wealth in December 2019 in thousands of Chilean pesos, indicator for labor force participation, and indicator for formal employment (in a firm or the public sector). All regressions include baseline controls (demographic cell, region, parental education, and time fixed effects). Standard errors are clustered at the individual level.

Figure A.2: SLE by age and education



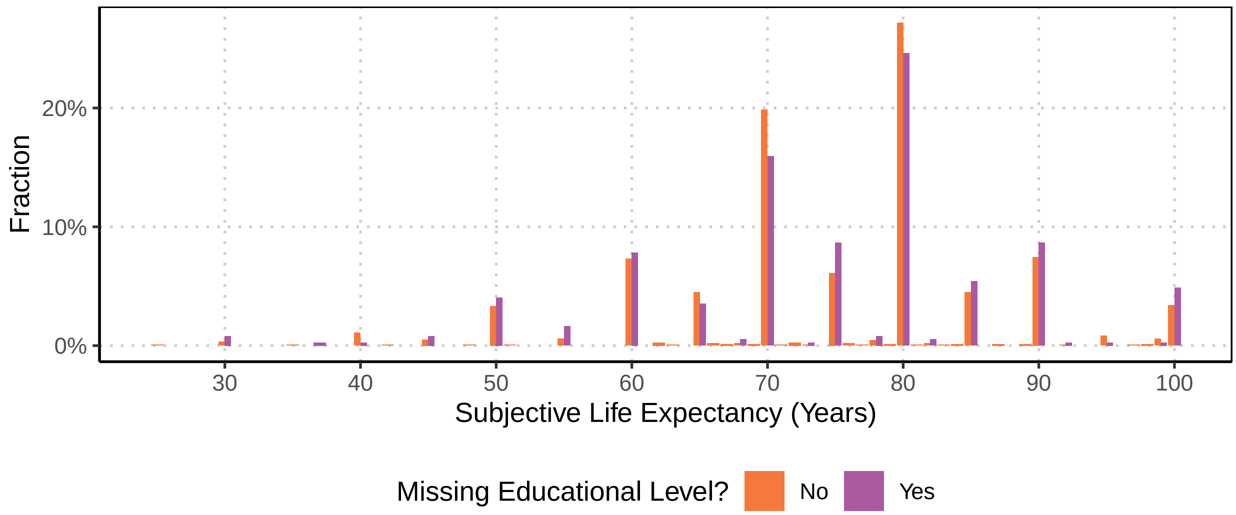
(a) SLE by age



(b) SLE by education

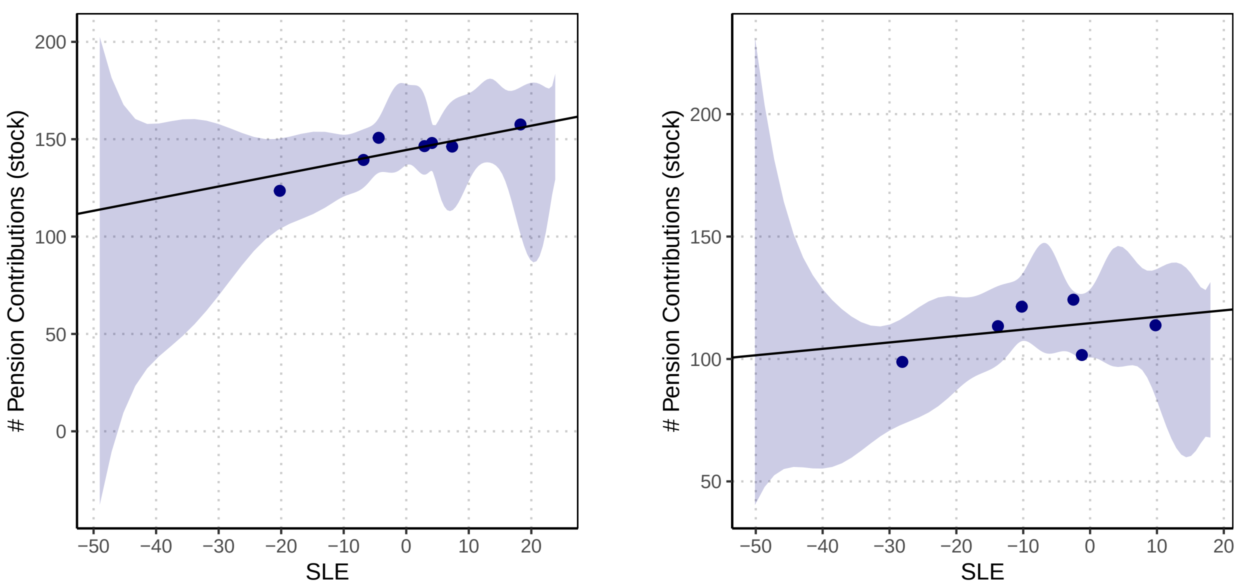
Notes: The figure plots the subjective life expectation by age (in the top plot) and by educational level (in the bottom plot), separately for men and women. Each dot displays each respondent's (raw) answer to the question "Up to what age do you believe you will live?". The solid black on panel (a) lines show the average expected mortality from the life table. To improve on the visualization, dots are spread (jittered) over the x-axis, but not over the y-axis. Additionally, they are plotted with transparency, therefore darker regions represent more mass of points.

Figure A.3: SLE by missing on educational level



Notes: The figure plots the histogram for the raw answers for the subjective life expectancy by missing status on the final educational level. The orange bars show the histogram for those observed at least once after 2009, and the purple bars for those not observed after 2009.

Figure A.4: Robustness — non-linear

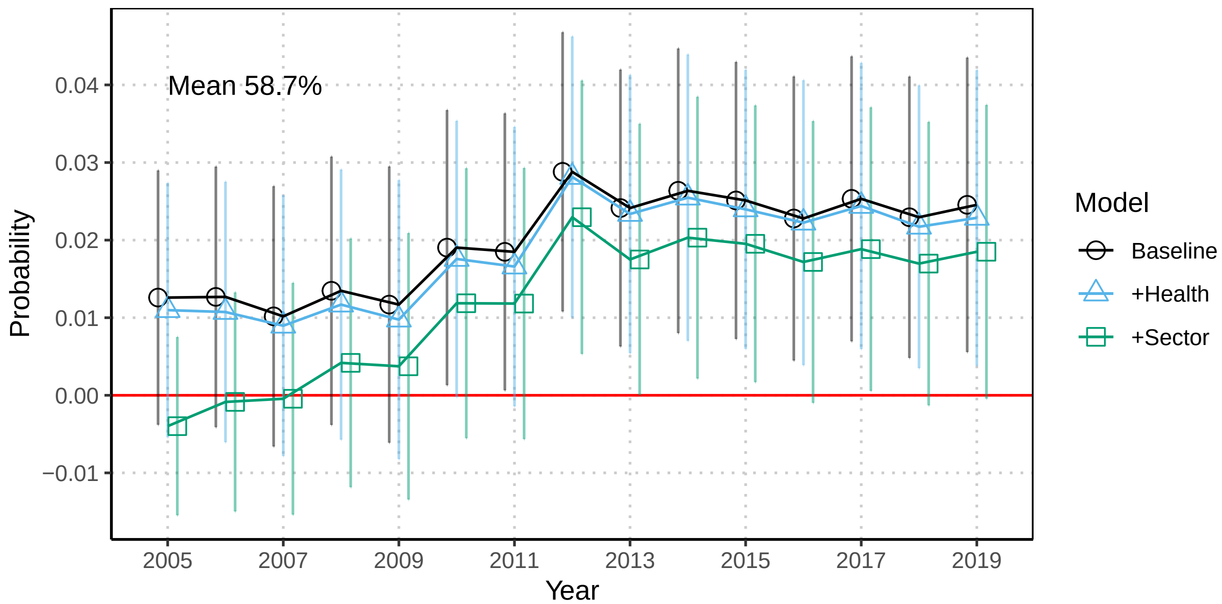


(a) Men

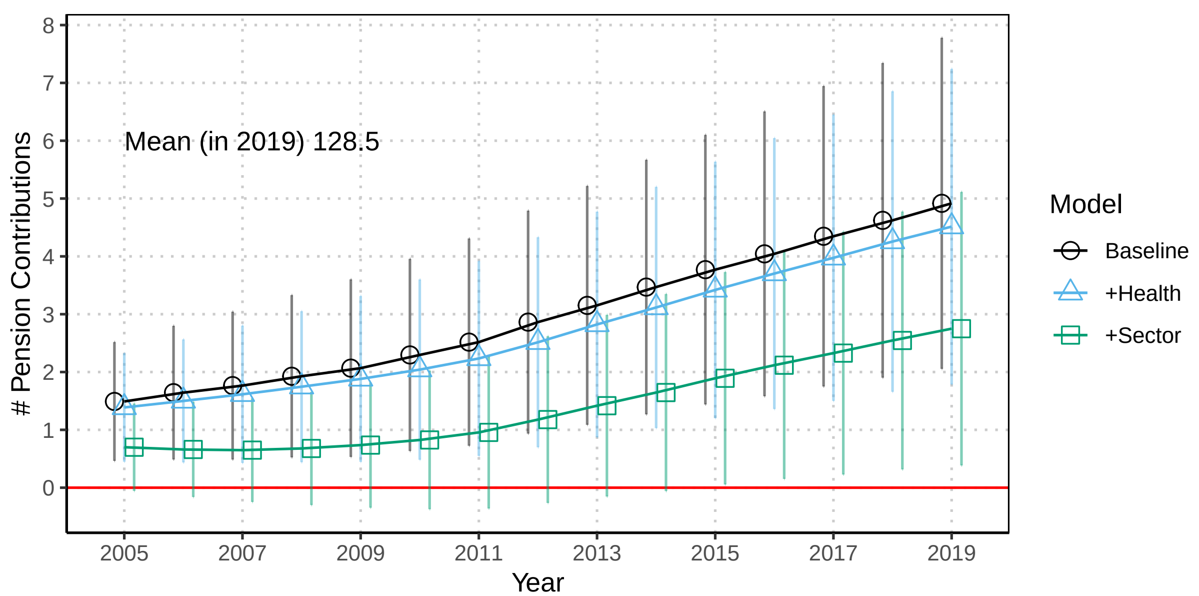
(b) Women

Notes: The figures show the results of applying a non-parametric binscatter regression proposed by Cattaneo et al. (2024). The dots are the resulting bins when using their algorithm to compute the optimal number and position of bins. The shaded area shows the 95% confidence interval for the relation between SLE(in the x-axis) and the total number of pension contributions in December 2019 (in the y-axis). I included the baseline controls (demographic cell, region, and parental education). The solid black curve shows the results from the linear specification from equation 1.

Figure A.5: SLE and future pension contributions with additional controls

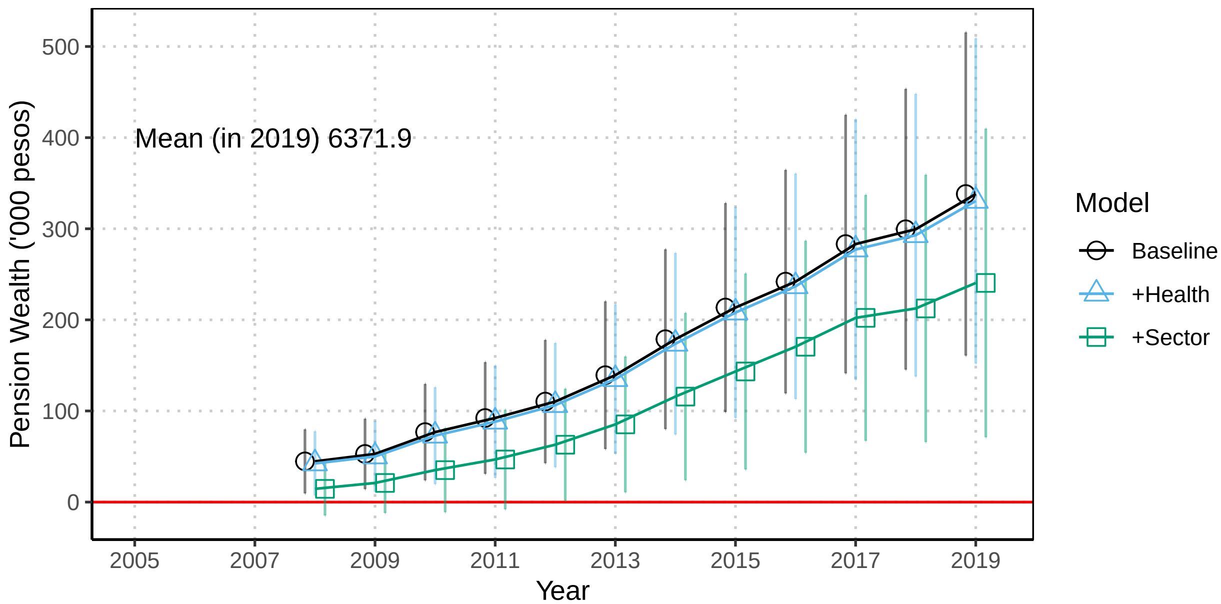


(a) Probability of making pension contributions



(b) # Pension contributions (stock)

Figure A.5: SLE and future pension contributions with additional controls (cont.)



(c) Pension wealth ('000 pesos)

Notes: The figure plots the results from the OLS estimation of equation 1. Panel (a) for the binary outcome on the probability of making monthly pension contributions, panel (b) for the number of total pension contributions (stock), and panel (c) pension wealth measured in thousands of Chilean pesos. The regressions are run by pooling all monthly observations but separately for each calendar year. The circles/dots/triangles are the estimated coefficient for β, and the solid vertical lines are the 95% confidence intervals. The text in each graph displays the mean for the outcome variable, for the entire period (panel a), and for December 2019 (panels b and c). The color and shape indicate the controls included. The baseline controls (black circles) include demographic cells, region, parental education, and time fixed effects. The blue triangles include, additionally, health self-assessment in 2004. The green squares introduce status in the labor market in 2004. Standard errors are clustered at the individual level.