# Youth Subjective Life Expectancy and Early Labor Market Choices*

Lucas Finamor†

First version: January 2024
This version: January 2026

**Abstract**

The literature has documented how measures eliciting subjective life expectancy (SLE) have desirable properties, predict mortality, and are strongly associated with savings and retirement outcomes among those nearing retirement. This paper assesses whether SLE matters when young individuals make consequential career decisions at labor market entry. Using survey and administrative data from Chile, I show that the SLE elicited from a young population also exhibits desirable properties. Moreover, individuals aged 18–26 with one standard deviation higher SLE have about 5% higher pension wealth fifteen years later. In a pension system based on individual capitalization accounts, contributions made early in the career compound for longer and are therefore valuable. I employ different empirical strategies, including exploiting cross-sectional variation in SLE, longitudinal individual fixed-effects approach, and instrumental variables that exploit health and death of family members.

**Keywords:** Subjective life expectancy, early carrer decisions, labor market entry, pension contributions, longevity expectations

*I would like to thank Richard Blundell, Áureo de Paula, Cormac O'Dea, David Sturrock, Eduardo Ferraz, Dalila Figueiredo, Antonia Paredes-Haz, and anonymous referees for their comments and suggestions. I gratefully acknowledge the support of the ESRC Centre for the Microeconomic Analysis of Public Policy (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.

†Sao Paulo School of Economics — FGV and Institute for Fiscal Studies. Contact: lucas.finamor@fgv.br

# 1 Introduction

Individual subjective beliefs have proven to be a useful tool for understanding human behavior. Across different domains where expectations about the future have been elicited, a particularly fruitful area concerns beliefs about mortality and survival. It has been documented how individual survival beliefs correlate with expected health outcomes (e.g. health assessment, disease diagnoses), lifestyle behaviors (e.g. smoking, exercising regularly), and demographics (e.g. education, parental death).[1] Importantly, several studies show how individual beliefs predict individual-level mortality or are consistent with group-level mortality, with some discussion on whether they complement mortality estimates coming from life tables.[2]

Building on these findings showing how survey-elicited beliefs have desirable properties, researchers have explored the relationship between survival beliefs and individual decisions. Subjective survival beliefs are associated with consumption and savings behavior, retirement decisions, and the purchase of annuities.[3] This literature relies mostly on beliefs elicited from the population over 50 years of age, close to retirement.[4] However, the most impactful decisions in terms of career progression, risk exposure, and social insurance coverage are made much earlier in life. Therefore, it is still unclear how beliefs on life span play a broader role over the life cycle, particularly for individuals at the beginning of adulthood. This paper fills this gap by exploring elicited beliefs from very young adults and how they relate to future pension contributions, labor market participation, and employment type.

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

[1]Bloom et al. (2006), Costa-Font and Vilaplana-Prieto (2022), Delavande et al. (2017), Delavande and Rohwedder (2011), Dormont et al. (2018), Foltyn and Olsson (2024), Hamermesh (1985), Hamermesh and Hamermesh (1983), Hudomiet et al. (2023), Liu et al. (2007), Mirowsky and Ross (2000), O'Dea and Sturrock (2023), O'Donnell et al. (2008), Zhou et al. (2022).

[2]Bissonnette et al. (2017), Bloom et al. (2006), Delavande et al. (2017), Delavande and Rohwedder (2011), d'Uva et al. (2020), Elder (2013), Hurd (2009), Hurd and McGarry (2002), Kutlu-Koc and Kalwij (2017), Mirowsky (1999), O'Dea and Sturrock (2023), Perozek (2008), Post and Hanewald (2013), Van Solinge and Henkens (2010).

[3]Bissonnette et al. (2017), Bloom et al. (2006), Bresser (2021), Bucher-Koenen and Kluth (2013), Foltyn and Olsson (2024), Gan et al. (2015), Griffin et al. (2012), Heimer et al. (2019), Hurd et al. (2004), O'Dea and Sturrock (2023), O'Donnell et al. (2008), Post and Hanewald (2013), Salm (2010), Van der Klaauw and Wolpin (2008), Van Solinge and Henkens (2010), Wu et al. (2015), Zhou et al. (2022).

[4]This specific age range comes from the dataset used, which is, in most cases, the Health and Retirement Study in the United States and their equivalent in other countries. These studies, typically, only survey individuals over 50. The exception is Heimer et al. (2019) which has a sample of the US population aged 28–78, and Bucher-Koenen and Kluth (2013) who work with a population of Germans aged 26–60.

administrative system at the individual level. This yields an extraordinary dataset with individuals' trajectories for over fifteen 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. In pension systems based on individual capitalization accounts, such as the one in Chile, contributions made early in the career can compound for longer, being very consequential for retirement outcomes.

I start the analysis by assessing the properties of the elicited subjective life expectancy (SLE) measure. SLE is defined as the individual's response to the question "up to what age do you think you will live?". I show how survival beliefs correlate, in expected ways, with demographics, health, lifestyle behaviors, and personal traits. Those smoking, not exercising regularly, who have been diagnosed with diseases, self-reporting bad health, or with higher body mass index report lower SLE. Beliefs are positively correlated with own and parental education. However, all these variables explain only a quarter of the SLE variance. It is therefore important to understand whether this residual SLE has true content or is just noise. I show how SLE (and residual SLE) predict mortality fifteen years ahead. One standard deviation higher SLE is associated with 35.2% lower mortality rates for older men and 13.9% for older women (the last one not statistically significant). That is, SLE has substantial content over and above all these observed measures. The panel dimension is also useful to show that the beliefs are modestly persistent and responsive to new information, such as new disease diagnoses.

After verifying the desirable properties of the SLE measure, even when elicited from a young population, I turn the analysis to the pension contribution and labor market outcomes. I start by exploring the cross-sectional variation in SLE. Using the pension administrative data, I show how individuals who reported higher SLE have a higher attachment to the pension system. By the end of the sample period in 2019, fifteen 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 (3.8% of the mean) and around 255 thousand Chilean pesos higher pension wealth (3.8% of the mean).[5] Ten years of SLE equals the 75th–25th percentile gap, or approximately 0.83 standard deviations. 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 pension benefits after retirement.[6]

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

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

Next, I investigate what the drivers of these pension gaps by SLE are. Individuals with higher SLE might choose differently how much they participate in the labor market, or which type of employment to engage with, leading to different pension contributions. First, I find that individuals who have 10 years higher SLE are 1pp more likely to participate in the labor force (1.3% of the mean). I then turn the analysis to 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 individuals in informal jobs or self-employed. Those with 10 years higher SLE are 1.5pp more likely to work as formal employees (3.1% of the mean). Mediation analysis shows that labor participation accounts for 25.0% of the total coefficient estimated for pension contributions, while type of employment accounts for 48.8%. These results indicate that employment type accounts for most of the observed pension gap.

These results exploit the cross-sectional variation in SLE, controlling for age, educational, gender, region and parental education. Therefore, it can still be the case that there are unobservables that correlate with SLE and with the outcomes and confound the estimates. For instance, it can be the case that more patient individuals have beliefs they will live longer and are also more likely to contribute to pensions. To investigate how robust the results are to these concerns, I first explore how sensitive the results are to the inclusion of additional controls and then I turn to different empirical strategies. I show how the coefficients are robust to the inclusion of relevant variables, including those measuring health assessment and diagnosis, numeracy, risk aversion, patience, and the big five personality measures.[7] Applying the robustness procedures of Oster (2019) and Cinelli and Hazlett (2020), I discuss (and quantify) the conditions under which this result could be overturned by unobservables.

Exploiting the longitudinal dimension of the survey and the fact that SLE was elicited in several waves, I also estimate a specification with individual fixed effects, exploiting solely within-individual variation. This specification is, therefore, robust to time-invariant unobservables, for example, any relevant personal trait or preferences that are not measured directly. The fixed-effects estimates produce very similar positive results. When using the panel dimension, I show how conditioning on the first SLE, future SLE beliefs are positively associated with pension outcomes. This would be consistent with a "revision behavior" — individuals with similar initial beliefs about life expectations revise their pension contributions up or down, depending on future information revealed to them. While this approach can deal with time-invariant unobservables, it is still vulnerable to time-varying unobservables that correlate with the expectations and SLE. One particular concern is that health

[7]The "big five" is a factor model in psychology aimed at describing individual personalities into five distinct categories. It produces a score for Agreeableness, Conscientiousness, Emotional Stability, Extraversion, and Openness. For more details on the measures and questions used, please check Gosling et al. (2003).

shocks may simultaneously lower SLE and hinder individuals' ability to work. This would create an artificial positive correlation between SLE and labor market participation.

Leveraging the fact that the survey also collects information on family members, I analyze another empirical strategy that uses health of family members and parental death as instruments for SLE. These IV estimates yield larger positive coefficients than the cross-sectional analysis, however they present wide confidence intervals. This approach alleviates the concerns with the own-health shocks that are present in the panel approach, however the exclusion restrictions require strong assumptions. We need to rely on the assumption that parental deaths or the health status of family members affect labor market and pension outcomes only through their impacts on survival beliefs. For parental death, this would rule out direct effects coming from lack of family support or inheritance. For the health of family members there are similar concerns. If family members with worse health also require more assistance, and this in turn affects labor supply, then we have a direct effect that violates the exclusion restriction. For this last channel, I can at least show that the estimates are similar when restricting to family members who do not require care.

In summary, SLE exhibits desirable properties even when elicited from a very young population and, crucially, SLE, even partialling out health information, has predictive power for future mortality. Using three different empirical strategies, I show how higher SLE is associated with higher pension contributions 15 years later. Each empirical strategy has its own assumptions and shortcomings. To bridge these results with the literature that documents how individuals close to retirement exhibit different savings and retirement behaviors depending on their subjective life expectancies, I also compute these results over the life cycle. While 10 extra years of SLE is associated with 2.5–5.0% higher pension for young individuals, these numbers increase over the life cycle, reaching 9–11% for those aged 48–52 in 2004. This is consistent with concerns over longevity and how to fund consumption over retirement becoming more salient over time, and with uncertainty over lifespan decreasing as individuals get older and more information is revealed.

This paper offers several contributions to the literature studying individual subjective beliefs over lifespan. First, as most of the literature focuses on individuals near retirement, I show novel evidence that early beliefs on survival are strongly associated with pension contributions and labor market choices. 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. This in particular, complements the finding from Zhou et al. (2022) which shows that higher SLE is associated with participation in private pension insurance. Additionally, my data allows me to estimate the same empirical specification for different age groups. The results show how the estimates increase over the life cycle. Second, the richness of the data allows me to explore different empirical strategies that address some identification concerns, leveraging information collected across different waves and from family members. Third, I add to both literatures on SLE and informal labor markets in developing countries by providing new insights into 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.

The paper proceeds as follows. Section 2 describes the Chilean institutional setting and the data. Section 3 discusses the life-expectancy measures. Section 4 presents the conceptual framework and econometric strategy. Section 5 reports the results, and Section 6 concludes.

# 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 may 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 some thresholds. While voluntary contributions are allowed, they are rarely observed.

For most of the sample period, self-employed individuals were not legally required to make pension contributions. This requirement was introduced in 2008, with postponed implementation until 2015. Nevertheless, enforcement is a challenge. Finamor (2025) shows that less than one-fourth of self-employed individuals with at most high school degrees had any pension contributions over a year. Therefore, working as self-employed or informal worker is one way of avoiding making pension contributions. Although Chile has lower rates of informal work than neighboring countries in Latin America, there is still a substantial

[8]For more information on options upon retirement, check Bello (2023).

fraction of informality — around 1/3 of the workforce (Finamor, 2025). 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 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 seven waves implemented every 2–4 years.[9] It contains detailed information on demographics, labor market characteristics, family, health, income, and assets for the interviewed person in the household. Crucially for this study, EPS also includes questions on subjective life expectancy.

Every individual interviewed in the EPS can be linked to HPA, which contains all monthly contributions to 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 survivor benefits. I complement these datasets with life tables computed by the Chilean National Statistics Office.

From these datasets, I derive the main sample as individuals aged 18–26 years in 2004. The age restriction is to capture young individuals before or at the beginning of their labor market careers. The lower limit of 18 is determined by the survey, which does not interview younger individuals as primary respondents.[10] The upper limit of 26 is chosen to yield enough sample size for the analysis. As the main analysis only uses the survey data from 2004 and the administrative data from 2004–2019, all individuals are kept in the sample irrespectively of their participation in the future waves of EPS. The final sample contains 1,750 observations (894 men and 856 women). Table A.2 presents some descriptive statistics.

# 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.[11] The answers are ages

[9] Since the second wave, in 2004, it is nationally representative, covering the adult population in Chile.
[10] There are 76 observations recorded with age 16 or 17, that were not targeted to be in the main sample.
[11] In Spanish, the original wording was *"¿Hasta qué edad cree usted que va a vivir?"*

measured in years. The second question asked the chances of living to at least age 65.[12] Respondents are expected to give answers in percentages ranging from 0 to 100. Interestingly, these questions are placed in different modules in the survey, separated by dozens of questions.[13] Versions of both questions are widely used in the literature.[14]

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 respondents are more familiar with than probabilities. Second, while respondents still round their answers, it does not show the issues with focal points, such as answering 0%, 50%, or 100%. This problem has been widely documented with probabilistic measures.[15] Indeed, whereas 47% of individuals answered 100% to the probabilistic question, only 4% answered any value above 100 years for the first question. Nevertheless, the two measures yield very similar results, even when predicting future mortality.

Figure 1 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. Additionally, I can use both questions to verify whether individuals provided consistent answers. Only 2.8% of individuals answered the two questions with unlikely values (answering 99-100% to the second question and less than 65 years to the first question or answering 0-1% to the second question and more than 65 for the first). I keep them in the sample, but results are robust to removing them. Figure A.1 plots the answers to the two questions for each individual, showing their positive correlation. Figure A.2 plots the raw data by age and education. Men's answers are closer to the objective life expectancy from the life tables, while women's reports are significantly below. This gender gap in subjective life expectancy is a robust finding in the literature.[16] These results extend the findings of the literature showing that younger individuals are more

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

[13] There were at least 101 questions between the two questions. The exact number depends on each respondent's answers and the exact flow of the questionnaire.

[14] Griffin et al. (2012) and Bucher-Koenen and Kluth (2013) use questions similar to the first (with answers in years), while Bell et al. (2020), Bissonnette et al. (2017), Bresser (2021), Gan et al. (2015), Heimer et al. (2019), Hurd et al. (2004), O'Dea and Sturrock (2023), O'Donnell et al. (2008), Post and Hanewald (2013), Salm (2010), Van Solinge and Henkens (2010), Wu et al. (2015), Zhou et al. (2022), and Foltyn and Olsson (2024) use probabilistic questions similar to the second one. Hamermesh and Hamermesh (1983) and Hamermesh (1985) elicited both questions.

[15] Several papers in the literature document how respondents tend to use focal points with the probabilistic questions and how some answers, for instance 50%, may reflect epistemic uncertainty rather than equal probabilities (Bissonnette et al., 2017, Dormont et al., 2018, Fischhoff and Bruine De Bruin, 1999, Fischhoff et al., 2000, Gan et al., 2015, Heimer et al., 2019, Hudomiet et al., 2023, Hurd, 2009, Kutlu-Koc and Kalwij, 2017).

[16] Bissonnette et al. (2017), De Bruin et al. (2023), Delavande et al. (2017), Dormont et al. (2018), d'Uva et al. (2020), Elder (2013), Foltyn and Olsson (2024), Hurd (2009), Kutlu-Koc and Kalwij (2017), Liu et al. (2007), Mirowsky (1999), Perozek (2008), Wu et al. (2015).

(a) Subjective life expectancy (years)


(b) Probability of being alive at age 65

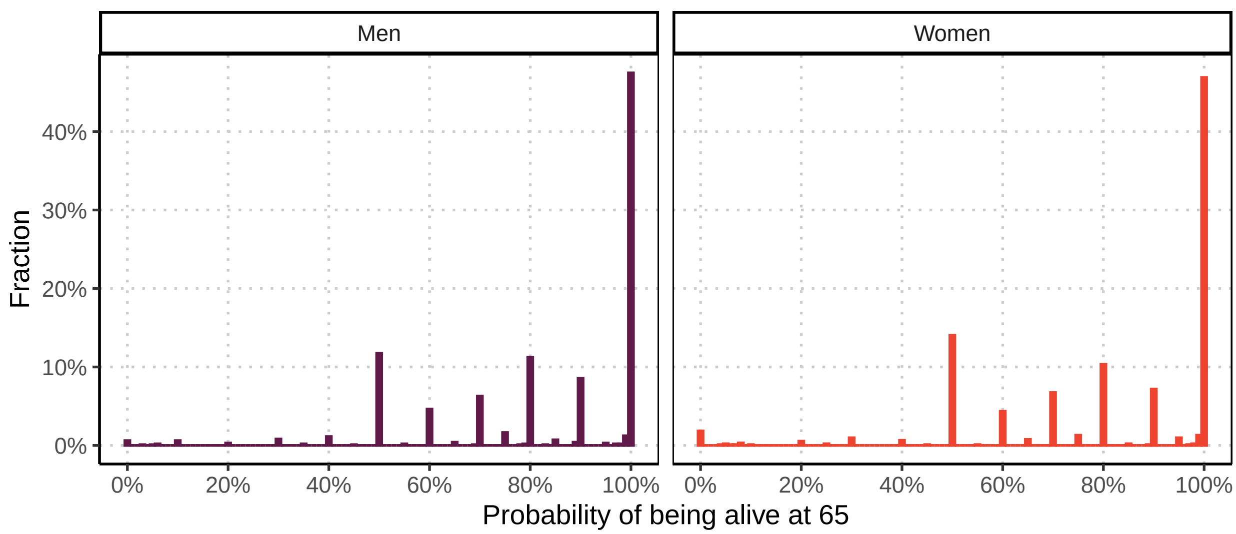



Figure 1: Distribution of subjective survival beliefs, by gender

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.

pessimistic, while older individuals are more optimistic. The evidence here indicates that such general pessimism arises at much younger ages than previously documented in the literature, particularly for women.

## 3.1 Health and mortality

I first investigate which variables correlate with SLE. The results are displayed in Table 1. We can see that gender alone explains 9.2% of the total variation of SLE. The remaining controls (age, education, region, and parental education) explain an additional 6.2 percentage points (pp) of the total variation. The variables in column 4 represent all the demographic controls that are used in the main analysis. Age is measured in years and enters discretely (9 levels, from 18–26). Education is measured in four levels (primary, high school, vocational, and college). Therefore, the gender-age-education group dummies comprise 72 different levels (2 genders $\times$ 4 education levels $\times$ 9 ages). Region considers the 13 regional divisions in Chile. Parental education is measured in three levels (primary, high school, and college) separately for mothers and fathers. That is, in total gender-age-education, region and parental education has 72+13+3+3=91 different levels. In the fifth column, I include variables related to health and lifestyle. They all have the expected sign. Smokers have, on average, -1.189 years of SLE. Doing physical activities regularly and assessing your own health status as good are associated with a higher SLE. Having any diagnosis in a given list of diseases and having a higher BMI are associated with lower SLE.[17] Together, they raise the R-squared by only 1.7pp. The next column introduces indicators for deceased mother and father, both with negative coefficients. For instance, having a deceased father is associated with a lower SLE by 2.767 years.

In the last column (7), I add variables related to preferences, skills, and personality traits. I add an indicator for the individual being risk-averse, a numeracy measure, an indicator for future-oriented preferences (patience), and measures of the "Big Five" personality traits. Appendix A shows the precise definitions and constructions of these variables.[18] The indicators for risk-aversion and patience have small positive associations with SLE, not distinguishable from zero. Surprisingly, the numeracy measure is not correlated with SLE (conditional on remaining controls). Jointly, these variables explain little of the residual variation of SLE. The R-squared rises by only 1.2pp. More importantly, we can see that all these variables, including the baseline controls, do not explain much of the SLE dispersion. The R-squared of this regression is 18.7%. This is not unique to this setting; Hurd

[17] 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 A.9 shows the same results splitting by group of diseases and separately for each disease. The results are very noisy, with no discernable patterns.

[18] One important caveat is that while all variables from columns 1–6 and the risk-aversion question were measured in the same initial wave (2004), the questions used to derive the numeracy and future-oriented preferences were only introduced in the following wave (2006) and the Big Five measures were introduced in the next round in 2009.

Table 1: Correlates of SLE

| | Outcome: SLE | | | | | | |
|---|---|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) |
| Smoking | | | | | -1.189* | -1.162* | -1.191** |
| | | | | | (0.607) | (0.605) | (0.606) |
| Physical Activities | | | | | 1.765** | 1.855** | 1.893** |
| | | | | | (0.772) | (0.767) | (0.771) |
| Good Health | | | | | 3.134*** | 3.017*** | 3.057*** |
| | | | | | (0.969) | (0.966) | (0.958) |
| BMI | | | | | -0.185** | -0.183** | -0.184** |
| | | | | | (0.083) | (0.084) | (0.084) |
| Any Diagnosis (Diseases) | | | | | -0.768 | -0.912 | -0.963 |
| | | | | | (1.089) | (1.088) | (1.080) |
| Deceased Mother | | | | | | -1.530 | -1.168 |
| | | | | | | (2.270) | (2.284) |
| Deceased Father | | | | | | -2.767** | -2.760** |
| | | | | | | (1.295) | (1.286) |
| Numeracy | | | | | | | -0.043 |
| | | | | | | | (0.234) |
| Risk-aversion | | | | | | | 0.587 |
| | | | | | | | (0.632) |
| Future-oriented Preferecnes | | | | | | | 0.094 |
| | | | | | | | (0.869) |
| Agreeableness | | | | | | | 0.155 |
| | | | | | | | (0.598) |
| Conscientiousness | | | | | | | 0.256 |
| | | | | | | | (0.494) |
| Emotional Stability | | | | | | | -0.239 |
| | | | | | | | (0.468) |
| Extraversion | | | | | | | -1.012* |
| | | | | | | | (0.553) |
| Openness | | | | | | | -0.687 |
| | | | | | | | (0.494) |
| Gender | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ |
| Gender-Age-Educ | - | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ |
| Region | - | - | ✓ | ✓ | ✓ | ✓ | ✓ |
| Parental Education | - | - | - | ✓ | ✓ | ✓ | ✓ |
| Observations | 1,750 | 1,750 | 1,750 | 1,750 | 1,750 | 1,750 | 1,750 |
| $R^2$ | 0.092 | 0.137 | 0.146 | 0.154 | 0.171 | 0.175 | 0.187 |

Notes: The table presents the results from a regression of SLE on selected variables. In each column the corresponding controls are sequentially included. Individual level clustered standard errors are in parentheses. *Signif. Codes: ***: p-value<0.01, **: p-value<0.05, *: p-value<0.10.*

and McGarry (2002), Puri and Robinson (2007), and Delavande et al. (2017) report similar findings.

A crucial question is whether these beliefs are predictive of mortality. That is, do they reveal any private information individuals might have about their survival chances. Table 2 below shows the results of an indicator for being deceased in 2019 and the reported SLE in 2004, conditional on the main controls. As individuals in my sample are only 18–26 in 2004, there is not enough time to observe their mortality. Therefore, I also include in this mortality analysis older individuals aged 27–36, 37–46, and 47–56. We can see that most of the coefficients are of the expected sign, those with higher SLE are less likely to be deceased in 2019. The coefficients are statistically significant for the older samples of men. Panel B adds all the health and lifestyle variables from Table 1 and shows that SLE is still predictive of future mortality after conditioning on the main health variables. For women, we see much weaker correlations, which are indistinguishable from zero. This result is in line with most of the literature, which finds that SLE is correlated with future mortality and with the life tables, but it is not necessarily accurate in levels.[19]

Another interesting question concerns the persistence of these beliefs and how they respond to new information. Table A.3 shows how the beliefs are modestly persistent over time. As the survey also records the diagnosis of diseases, we can see how beliefs vary after a new diagnosis. Results are presented in Table A.4. We can see that new diagnosis of more serious diseases are associated with large reductions in future SLE.

In summary, subjective survival beliefs correlate with expected demographics, health and lifestyle behaviors. Nevertheless, SLE displays substantial dispersion that cannot be explained by observed variables. Subjective beliefs are associated with future mortality and display moderate persistence, responding to health shocks.

# 4 Theoretical framework and empirical strategy

After having established the desirable properties of the elicited SLE measure, we now turn the analysis to study the relationship between SLE and future labor market decisions. I first present a simplified theoretical framework that makes explicit why we would expect SLE to be associated with pension contributions and labor market outcomes. Then, I present the empirical strategies that will be used to quantify this relationship.

[19] Bissonnette et al. (2017), Delavande et al. (2017), Delavande and Rohwedder (2011), d'Uva et al. (2020), Elder (2013), Hamermesh and Hamermesh (1983), Hurd (2009), Hurd and McGarry (1995, 2002), Kutlu-Koc and Kalwij (2017), Mirowsky (1999), O'Dea and Sturrock (2023), Perozek (2008), Van der Klaauw and Wolpin (2008), Van Solinge and Henkens (2010).

Table 2: SLE and mortality

| | Men | | | | Women | | | |
|---|---|---|---|---|---|---|---|---|
| Age range | 18–26 | 27–36 | 37–46 | 47-56 | 18–26 | 27–36 | 37–46 | 47-56 |
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) |
| **Panel A - Main specification** | | | | | | | | |
| SLE | -0.006 | -0.004 | -0.017** | -0.031*** | -0.001 | 0.001 | -0.003 | -0.007 |
| | (0.004) | (0.004) | (0.007) | (0.010) | (0.004) | (0.002) | (0.004) | (0.007) |
| **Panel B - With health controls** | | | | | | | | |
| SLE | -0.006 | -0.003 | -0.015** | -0.020** | -0.001 | 0.002 | -0.001 | -0.004 |
| | (0.005) | (0.003) | (0.007) | (0.010) | (0.004) | (0.002) | (0.004) | (0.007) |
| Observations | 874 | 1,586 | 1,614 | 1,201 | 820 | 1,457 | 1,575 | 1,008 |
| Mean | 0.023 | 0.021 | 0.045 | 0.110 | 0.013 | 0.009 | 0.032 | 0.063 |

Notes: The table presents the results from an OLS regression of equation 1 using an indicator for being deceased by December 2019 as the outcome variable. The estimated coefficients show the coefficient associated with an increase of 10 years of SLE. The first four columns are for men, and the last four for women. Panel A is the baseline regression and Panel B presents the regression with all health variables included in Table 1. All regressions include the baseline controls (demographic cell, region, 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. *Signif. Codes: ***: p-value<0.01, **: p-value<0.05, *: p-value<0.10.*

Consider a simplified version of an intertemporal savings-consumption problem. Individual $i$ is alive in period 1 and there is a chance they survive to period 2. They believe the survival chance is $\pi$. They maximize their expected lifetime utility, given by:

$$u_1(Y(1-a)) + \rho\pi u_2(Ya),$$

where $Y$ is the income in period 1, $a$ is the fraction of income saved for period 2, $\rho$ is the discount factor, and $u_1$ and $u_2$ are the period-specific utility functions in each period. In this simplified setup, $a$ represents direct saving from period 1 to period 2. In the real setting, this can be achieved by any decision that increases resources in period 2, such as working formally and sacrificing current consumption to make pension contributions. The first-order condition for this problem is given by:

$$\frac{u_1'(Y(1-a))}{u_2'(Ya)} = \rho\pi.$$

Provided that $u$ is increasing and concave, the optimal savings $a^*$ increases in $\pi$.[20] That is, as individuals assign higher chances of surviving to period 2, they optimally decide to transfer more resources into the future.

Note, however, that the optimal decision $a^*$ also depends on other characteristics that may vary across individuals. For instance, if individuals are more patient (higher $\rho$) or put more weight in the future consumption (higher $u_2'$), they will also optimally decide to save more. Therefore, the relationship between subjective survival beliefs $\pi$ and the saving decision $a^*$ may be confounded by other characteristics correlated with both. That would be the case, for example, if more patient individuals also systematically report higher survival beliefs.

## 4.1 Empirical strategy

The empirical strategy aims to assess and quantify the relationship between SLE and labor market and pension contribution decisions. I present three different approaches. The first explores the cross-sectional variation in initial beliefs. The second approach uses the panel dimension and explores within-individual variation in SLE. The third approach uses only the initial SLE measure, but instruments initial beliefs with information on mortality and health of family members.

The first approach exploits cross-sectional variation in the initial SLE, measured in 2004, with outcomes measured in 2005–2019. The main OLS regression is:

$$Y_{it} = \beta \text{SLE}_i + \gamma X_i + \varepsilon_{it}, \tag{1}$$

where $Y_{it}$ is the outcome for individual $i$ in time $t$, regressed on the SLE reported in 2004. Therefore, $\beta$ is the coefficient of interest, measuring the association between the outcome and subjective beliefs. When running this regression, it is important to include controls $X_i$ that account for other characteristics that may be correlated with both SLE and the outcome. For instance, the theoretical framework discusses how patience may correlate with SLE and also influence the importance of future consumption. For the baseline controls, I consider $X_i = \eta_{d(i)} + \nu_{r(i)} + \varphi_{p(i)}$, which includes a demographic-cell fixed effects ($\eta_{d(i)}$), comprised of gender, age, and educational level, region fixed effects ($\nu_{r(i)}$), and parental education fixed

[20] If $\pi$ increases, the right-hand side of the first-order condition increases. Therefore, the left-hand side must also increase. As $u_1'$ is decreasing in $a$ and $u_2'$ is increasing in $a$, the only way for the left-hand side to increase is for $a$ to increase.

effects ($\varphi_{p(i)}$) which proxy for socioeconomic status.[21] However, there can be several other variables that are relevant and are not included in this baseline specification. To address this concern, I explore the stability of the $\beta$ coefficient as these baseline and several additional controls are included.

A main limitation of the cross-sectional approach is that not all relevant factors can be observed or controlled for. For instance, it could be the case the optimism of individuals correlates with both SLE and pension contributions, as argued by Puri and Robinson (2007). To alleviate some of these concerns, I move to use the full panel dimension of the data, exploiting within-individual variations in subjective beliefs. I augment the main specification to include individual fixed-effects and to consider SLE varying over time, rather than just the initial belief in 2004. I run the following regression:

$$Y_{it} = \beta^{WI}\text{SLE}_{it} + \eta_i + \varepsilon_{it} \tag{2}$$

which includes individual fixed-effects ($\eta_i$). This specification addresses the concerns on omitted variables that may correlate with SLE but are permanent characteristics of individuals. For example, important personal traits that cannot be measured, provided that they are constant over time, they will not confound this analysis. For every individual, I consider all SLE reports across the waves and use the cumulative pension contributions at the time of each report as the outcome variable. Therefore, we do not expect $\beta^{WI}$ to have the same magnitude of $\beta$, as they will be identifying correlations across different time domains. While this approach is robust to time-invariant unobservables, it is still susceptible to time-varying shocks. For example, one serious concern is with health shocks. It can be the case that individuals receive health shocks that affect both their survival beliefs and their ability to work. In that case individuals would reduce their SLE at the same time as investing less in pensions, and we would estimate a positive $\beta^{WI}$ which is not due to the beliefs, but merely coming from the health shocks and potential incapacitation effects. Thus, these results should be interpreted with caution.

The third approach takes advantage of the comprehensive data collection in the survey and use information on mortality of parents and general health of other family members in the initial wave as instruments for the initial beliefs. Let $H_i$ be this measure of whether either

[21] Each of these variables enters the model in a discrete way. There are 72 different intercepts for the demographic group, 13 for region, and 3 for each parental education (6 in total). For consistency, education is measured in the initial wave in 2004. While it is possible that some individuals will get more educated over the years, note that the specific controls are gender-age-education cells; therefore, it compares individuals with the same educational level at the same baseline age. An interesting question for future research is whether educational decisions could also be influenced by SLE. Table A.5 shows some preliminary evidence on this mechanism.

of their parents were deceased or the average health self-assessment of family members. I run the following 2SLS estimation:

$$\begin{aligned} \text{(First Stage) } \mathrm{SLE}_i &= \gamma H_i + \eta_{d(i)} + \nu_{r(i)} + \varphi_{p(i)} + \epsilon_i \\ \text{(Second Stage) } Y_{it} &= \beta^{IV}\mathrm{SLE}_i + \eta_{d(i)} + \nu_{r(i)} + \varphi_{p(i)} + \zeta_{it}. \end{aligned} \tag{3}$$

This approach alleviates the concern of other unobservables that may be correlated with the initial 2004 belief and future labor market decisions. However it requires an exclusion restriction assumption: the instrument (parental death or health shocks to family members) affects the outcome only through the subjective beliefs. This is a strong assumption, and it may fail, for instance, if individuals must care for sick family members and therefore reduce their labor supply. I explore whether this is a relevant concern in the analysis.

In summary the three approaches exploit the relationship between subjective survival beliefs and future labor market outcomes. Each approach has strengths and weaknesses. To conduct inference, 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

To facilitate interpretation, I present coefficients associated with an increase of 10 years of the "net" SLE, defined as the individual's reported life expectancy minus the objective life expectancy from the life tables, conditional on gender and age.[22] The gap in subjective life expectancy between the 25th and 75th percentiles of the SLE distribution is 10 years. The equivalent 90th–10th gap is 30 years. One standard-deviation of SLE is 12 years. Hence, all coefficients can be interpreted as changes in 10 years of subjective life expectancy, a movement from the 25th to the 75th percentile of the SLE distribution, or an almost one standard deviation increase in SLE (precisely, 0.83 standard deviations).

[22]The net measure will be given by $SLE_i^{net} = SLE_i^{subj} - LE_i^{obj}$, where $LE_i^{obj}$ is the objective life-expectancy for an individual $i$, given their gender and age. This is the (life-expectancy + age) from the 2004 life-table. For example a women who is 20 years old has a life-expectancy of 61.2 years according to the 2004 life tables. Suppose she reported she would live until age 85. Therefore, her net SLE is given by $SLE_i^{net} = 85 - (61.2 + 20) = 3.8$ years.

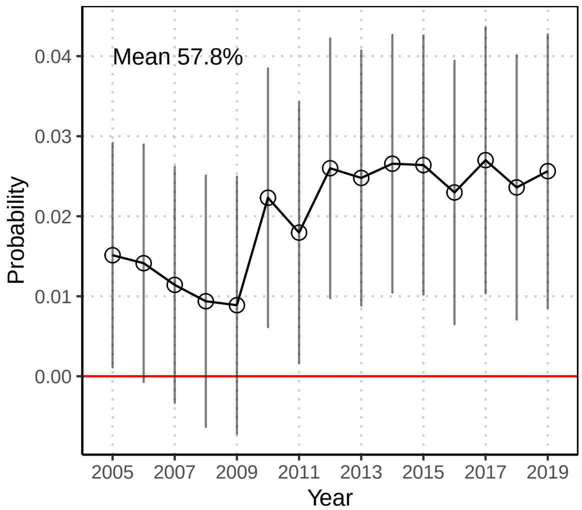
## 5.1 Cross-sectional approach

I now exploit 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 between 2005 and 2019. The results are summarized in Figure 2. In the top left plot (2a), 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 plot the $\beta$ coefficient from the baseline regression, which controls for demographic cell (gender-age-education), region, and parental education fixed-effects. For the initial years, 2004–2009, the average coefficient is around 1pp and 2.5pp for the remaining sample period (2010–2019). In terms of magnitude, the overall probability of making a pension contribution is 57.8%. That is, those with 10 years higher SLE are 2–4.5% more likely to contribute. The smaller coefficient in the 2004–2009 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.

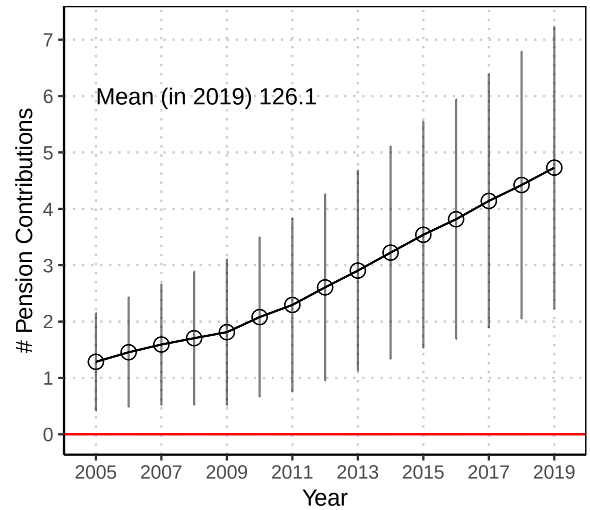
In Panel 2b, I show the results for the number of pension contributions individuals have accumulated. Given that those with higher SLE are constantly more likely to make pension contributions, it is unsurprising to see the estimates rising over time. In 2019, those with 10 years higher SLE have, on average, 4.7 months more of pension contributions. This translates to 3.8%, using the average number in December 2019 of 126.1. We can see the same in panel 2c, where the pension wealth (in thousands of Chilean pesos) is used as the outcome variable.[23] In 2019, those with 10 years higher SLE have a pension wealth 255.1 thousand Chilean pesos higher. Using the mean in December of 2019, that corresponds to 3.8%.

To put these numbers in context, Table A.10 shows several statistics for these two variables. We can see for instance, for pension wealth, the increase of 10 years of subjective beliefs is associated to an increase of 3.3% of the interquartile range, 10.7% of the gender-gap, and 7.1% of the gap between individuals with college education and those with only high school degree. The coefficient equals around 52.7% of the 12-month increase in pension wealth and 68.6% of the 12-month increase in pension contributions. That is, fifteen years later, the 25th-75th gap in SLE translates to about half to two-thirds of a typical 1-year increase in pension contributions. Finally, in terms of elasticity, one standard deviation increase in SLE is associated with 0.048 standard deviations increase in pension wealth and 0.080 standard deviations increase in the number of pension contributions.

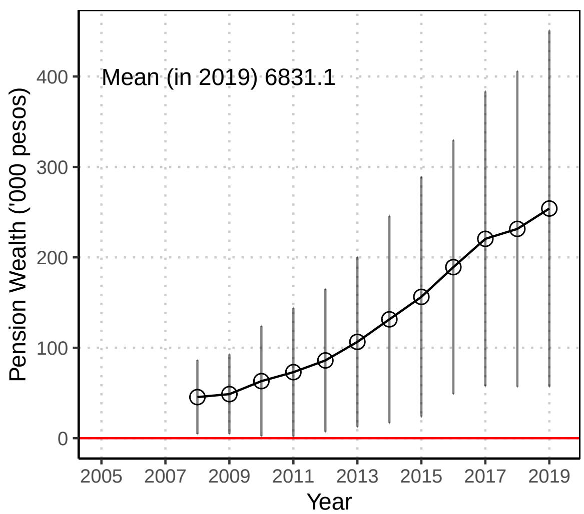
[23]The pension wealth variable is Winsorized at the top 5%, separately for each month.




(a) Probability of making pension contributions




(b) # Pension contributions (stock)




(c) Pension wealth ('000 pesos)

Figure 2: SLE and future pension contributions

Notes: The figure plots the results from the OLS estimation of equation 1. The estimated coefficients show the coefficient associated with an increase of 10 years of SLE. 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 $\beta$, 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.

Table 3 presents these final results in 2019 for the overall sample and separately for men and women. For each sample and variable, the table presents the OLS estimation of

Table 3: SLE and future labor market outcomes

| Outcome: | # Pension Contributions | Pension Wealth | Labor Force Participation | Formal Sector Participation |
|---|---|---|---|---|
| | (1) | (2) | (3) | (4) |
| **Panel A. All** | | | | |
| SLE | 4.730*** | 255.078** | 0.010** | 0.015** |
| | (1.278) | (111.470) | (0.005) | (0.006) |
| N Obs | 1.750 | 1.750 | 1.750 | 1.750 |
| Mean | 126.099 | 6831.148 | 0.752 | 0.480 |
| **Panel B. Men** | | | | |
| SLE | 7.018*** | 445.138*** | 0.009* | 0.021** |
| | (1.776) | (167.329) | (0.005) | (0.009) |
| N Obs | 894 | 894 | 894 | 894 |
| Mean | 140.358 | 7999.776 | 0.833 | 0.544 |
| **Panel C. Women** | | | | |
| SLE | 2.256 | 47.415 | 0.010 | 0.008 |
| | (1.799) | (147.385) | (0.008) | (0.008) |
| N Obs | 856 | 856 | 856 | 856 |
| Mean | 111.208 | 5610.641 | 0.666 | 0.413 |

Notes: The table presents the OLS estimation of equation 1 for four different outcomes: the total number of pension contributions in December 2019 (column 1), pension wealth in December 2019 in thousands of Chilean pesos (column 2), share of time periods in the labor force (column 3), and share of time periods in formal employment (column 4). The $\beta$ coefficient is displayed with the estimated standard error (clustered at the individual) level in parenthesis. The estimated coefficients show the coefficient associated with an increase of 10 years of SLE/predicted SLE. All regressions include the baseline controls (demographic cell, region, and parental education fixed effects). Panel A presents the results for the entire sample, and panels B and C, separately for men and women. *Signif. Codes: ***: p-value<0.01, **: p-value<0.05, *: p-value<0.10.*

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 (first column) and the pension wealth in thousands of Chilean pesos (second column). Those reporting 10 extra years of SLE have, on average, 4.7 extra months of pension contributions and pension wealth 255.1 thousand pesos higher.

Splitting by gender, 10 extra years of SLE is associated with 7.0 more contributions for men and 2.2 for women. In percentage terms, the estimates imply an increase of 5.0%

for men and 2.0% for women. For pension wealth, there is a similar pattern. Those with 10 extra years of SLE have pension wealth 5.6% higher for men and 0.8% higher for women. These gender differences are large and difficult to rationalize. Women have higher (objective) life expectancy and are on average more pessimistic. These results show that women may be less sensitive to their subjective beliefs when compared to men. One potential explanation is if given their relative lower labor market attachment, women expect overall lower pension contributions and therefore lower returns to investing in pensions, particularly given the existence of minimum pension guarantees. With a larger sample one could investigate if this is a relevant channel by testing whether those with (predicted) higher labor market attachment are more sensitive to their subjective beliefs.

These results show how those who report higher SLE have more pension contributions. One potential explanation is that individuals who have higher survival beliefs are more likely to be participating in the labor force and therefore making pension contributions. Another potential explanation is that those with higher survival beliefs, independently of their labor market participation, are more likely to work formally and therefore making pension contributions. The third and fourth columns of Table 3 explore the survey data on employment to shed some light on these mechanisms.[24] For each individual, I compute the share of months that they participated in the labor force and worked formally. We can see that individuals spend on average 75.2% of months in the labor force (83.3% for men and 66.6% for women). Individuals reporting 10 extra years of SLE have 1.0pp higher labor force attachment. This is statistically significant, but relatively small, representing an increase of 1.3% of the mean participation. In the last column we test whether the type of employment, whether working formally, also differs by SLE. We can see that the average share of months working formally is 48%. If the effect was entirely driven by labor force participation we should see coefficients for these regression that were (0.480/0.752)*0.010 = 0.006pp. That is, if there is no difference in this type of contract we should only see the additional effect coming from the expected fraction of those additional participation (1pp) that work formally (48% of the time). That is not what we see. Those with 10 extra years of SLE are 1.5pp more likely to be working formally. This is statistically significant and represents an increase of 3.1% of the baseline formal share of employment. Implementing a mediation analysis, we see that labor force participation accounts for 25.0% of the total coefficient on pension contributions, while type of employment accounts for 48.8%. That is, employment type — whether working formally — accounts for a larger share of the documented pension gaps.

Using questions and empirical strategies very close to this approach, several papers

[24] To avoid issues with recollection, I only use employment information reported not later than 12 months.

have found that older individuals make economic decisions that correlate with subjective life expectancies. In particular, those with higher SLE are more likely to save more and delay retirement. Using my dataset, I can assess how SLE and pension contributions are correlated over the life cycle, taking advantage of the high-quality administrative pension data. Figure 3 plots the results for estimates of equation 1 using two distinct pension outcomes: the stock of pension contribution and total pension wealth in 2019 for different age groups. To enhance precision, I group individuals by 5 years. To ease comparison across outcomes and age ranges, I plot the coefficients in percentage terms. We can see how estimates get larger for older individuals. The percentage effects increase linearly over the life-cycle, and the estimates for the younger population are about one-third of those in their 50s, which is at the lower range of the ages used in the HRS and their worldwide counterparts. This would be consistent with pension decisions and SLE becoming more salient as individuals age. Moreover, uncertainty decreasing over the lifespan would also produce similar results.

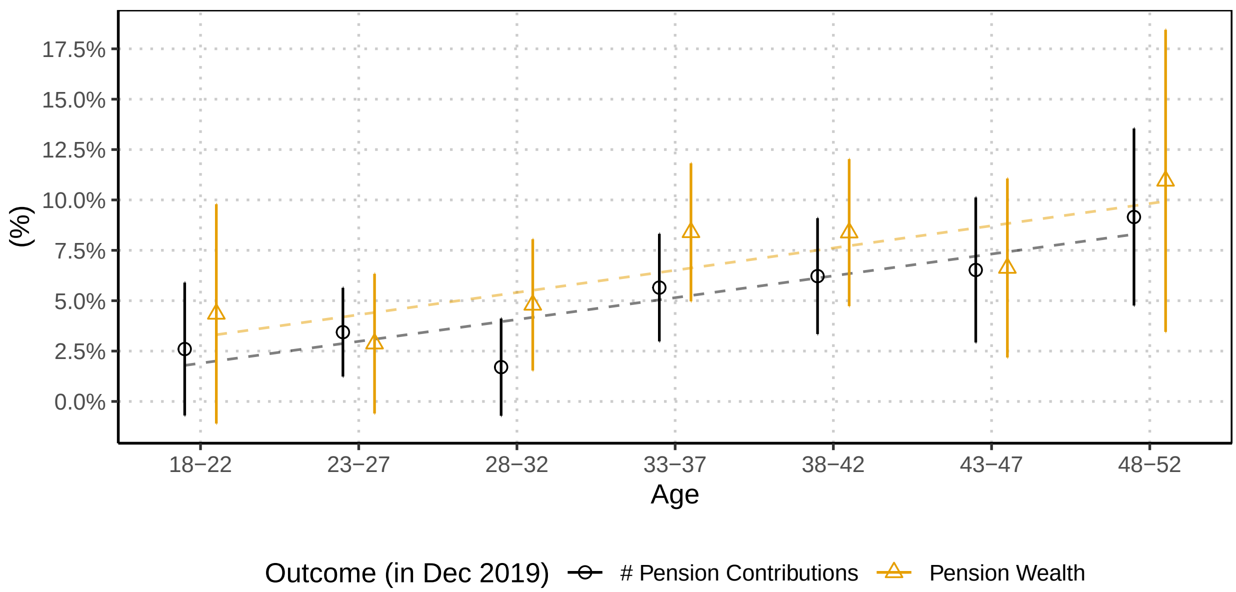



Figure 3: Estimates by age group

Notes: The figure presents the OLS estimation of equation 1 for two different outcomes: the total number of pension contributions in December 2019 (in black circles), pension wealth in December 2019 in thousands of Chilean pesos (in yellow triangles) for different age groups. I plot the estimated $\beta$ coefficient divided by the outcome mean, to display the coefficient in percentage terms. The straight lines show the 95% confidence intervals. The estimated coefficients show the coefficient associated with an increase of 10 years of SLE. All regressions include the baseline controls (demographic cell, region, and parental education fixed effects). The dashed lines show the best-fit line connecting all estimated coefficients.

The $\beta$ coefficient from equation 1 may not be only capturing the association between SLE and the labor market outcomes, but also additional variables that are correlated with the beliefs and the labor market outcomes. We have many candidates of unobservables that may correlate with both: productivity, human capital, patience, optimism, among many others. We now turn to investigate what happens with this coefficient as we augment our specification with additional controls. Table A.6 shows the coefficient of the OLS estimation of equation 1 with different set of controls. The only significant impact is when adding the gender variable. The coefficient goes from 6.959 to 3.909. This is not surprising, as we saw stark gender difference on beliefs and on labor market participation. The rest of the columns shows how remarkably stable the coefficient is, ranging from 3.909 to 4.730 when a variety of different variables are introduced, including several relevant health, lifestyle and personal traits measures.

Importantly these variables are relevant for the outcome measure. Together they raise the R-squared from 1.6% to 25.6%. With these results we can apply the procedure presented by Oster (2019) to assess how important the unobservables must be to drive these effects. Comparing the point estimates and R-squared from the first and last column, we conclude that the unobservables need to be 72% as important as all the observed variables to drive the results to zero.[25] Another approach to assess the plausibility of unobservables driving the results is the one proposed by Cinelli and Hazlett (2020). Using their approach, unobservables need to explain at least 8.6% of residual variance of both the outcome variable and the SLE to drive the results to zero. Note how gender, the variables with the most impact in our analysis would not pass this criteria as it only explains 3.1% of the outcome variable. No other observable variable comes close to this threshold. However, I cannot rule out that a combination of different omitted variables would be strong enough to drive this coefficient toward zero.

## 5.2 Longitudinal approach

In this section, I exploit the longitudinal dimension of the survey to explore alternative identification strategies. I first use jointly the initial and beliefs elicited in subsequent surveys. Results are presented in Table 4. The first column reproduces the main effect presented in Table 3. 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 2.5 months in pension contribution. The estimates are statistically significant at the 10% level. Similar results are displayed for

[25]This considers an $R^2$ maximum of 0.700 for any model explaining the stock of pension contributions. Figure A.3 presents these value for several values of the maximum $R^2$. With my set of covariates the maximum reached is 0.256.

SLE in 2009 (not statistically significant) and 2012 (significant at the 10% level). Therefore, this evidence is consistent with individuals acting in ways aligned with revisions of their subjective expectations; those revising their SLE up exhibit more pension contributions. An alternative explanation would be that SLE is measured with error, and, therefore, SLE measured in different surveys would display positive coefficients because they are measures of the same underlying belief in addition to measurement error.

Table 4: Future SLE elicitations

| Outcome: | # Pension contributions in Dec2019 | | | | # Pension contributions in time $t$ |
|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) |
| SLE in 2004 | 4.730*** | 3.948*** | 4.422*** | 4.301*** | |
| | (1.278) | (1.360) | (1.306) | (1.284) | |
| SLE in 2006 | | 2.535* | | | |
| | | (1.513) | | | |
| SLE in 2009 | | | 1.815 | | |
| | | | (1.938) | | |
| SLE in 2012 | | | | 3.515* | |
| | | | | (2.061) | |
| SLE in $t$ | | | | | 1.446*** |
| | | | | | (0.457) |
| Observations | 1750 | 1750 | 1750 | 1750 | 5095 |
| Individual Fixed-Effects | - | - | - | - | ✓ |

Notes: The first four columns show results from estimating equation 1 augmented with SLE measures in 2006 (column 2), 2009 (column 3), and 2012 (column 4). The outcome variable is the total number of pension contributions in December 2019. The estimated coefficients show the coefficient associated with an increase of 10 years of SLE. All regressions include the baseline controls (demographic cell, region, and parental education). The last column display the results from the panel estimation, equation 2, exploiting within-individual variation, as this specification include individual fixed-effects. For this analysis, the outcome is the stock of pension contributions at time $t$. Individual-level clustered standard errors are presented in parentheses. *Signif. Codes: ***: p-value<0.01, **: p-value<0.05, *: p-value<0.10.*

Another exercise exploring the panel dimension of the data is presented in the last column of Table 4. In this exercise, I exploit variation in SLE solely coming from time variation, running the equation 2, which includes individual fixed-effects. This specification addresses the concerns on omitted variables that may correlate with SLE but are permanent characteristics of individuals. For instance, if patience or optimism are permanent, even if they correlate with SLE and with the outcomes, their effects will be purged. For every individual, I consider all SLE reports across the waves and use the cumulative pension contributions at the time of each report as the outcome variable. The result shows that 10 extra years of within-individual SLE are associated with 1.4 extra months of cumulative pension contributions (statistically significant). These results are similar to the one obtained using the cross-sectional variation. Note that these estimates explore shorter time frame, as it only captures changes in outcomes in between survey, instead of the levels in the entire 15-year period as the other approach. The distance across two reports is, on average, 39.5 months. To make it comparable to the cross-sectional estimate we would need to compare the same time-frame. The average time between the initial wave and the last observation in December 2019 is 178 months. Therefore, this coefficient implies $178/39.5 \times 1.446 = 6.606$ for the entire period, compared to 4.730 with the cross-sectional approach.

The results in this section explore how beliefs evolve over time and are robust to time-invariant unobservables that are correlated with the beliefs and pension contributions. Examples of concerns that are alleviated with this approach are fixed preferences, skills, and personal traits that are not captured with the survey questions. However, the coefficients would still be biased if there are time-varying unobservables that correlate with changes in SLE and pension contributions. One potential channel is, for instance, the incidence of health shocks that may drive the changes in beliefs but also impede individuals to work. This would create a mechanical link of (negative) SLE changes and (negative) pension contributions.

## 5.3 Instrumental variables

In this section, I take advantage of the comprehensive data collection in the survey and use information on mortality of parents and general health of other family members in the initial wave as instruments for the initial beliefs. I first use parental deaths as an instrument for the initial SLE. In the first column of Table 5 we can see the first stage. Having deceased parents decreases the initial SLE by 3.05 years.[26] The F-statistic is 8.11, slightly below the typical threshold of 10. The second column shows the IV result for the instrumented SLE on the stock of pension contributions in December 2019. An extra 10-years of (predicted) SLE corresponds to 35 extra monthly pension contributions. This estimate is very noisy, and this coefficient is only statistically significant at the 10% level. In order for this instrument to be valid, the only channel through which parental deaths

[26] The SLE variable is measured in 10 years.

could affect pension contributions is through life expectancy. This assumption would be violated if for instance, through parental deaths, individuals change their behavior in the labor market due to lower family support, lower family income, or due to receiving a large inheritance. With this data, it is difficult to assess the plausibility of these channels. It is worth mentioning though, that some of them would act in the opposite direction of the current findings. For instance, if parental death leads to higher labor market participation due to lower family income and support, that would generate a negative link between SLE and our measured outcomes. The inheritance channel would likely act in the other direction biasing the coefficient up.

Table 5: Instrumental variables: mortality and health status of family members

| | Parental Death | | Health family members | | Health family members | |
|---|---|---|---|---|---|---|
| Stage | First | Second | First | Second | First | Second |
| | (1) | (2) | (3) | (4) | (5) | (6) |
| SLE | | 35.891* | | 25.428 | | 22.208 |
| | | (21.311) | | (17.087) | | (18.699) |
| Deceased parents | -0.305*** | | | | | |
| | (0.118) | | | | | |
| Health family members | | | 0.149*** | | 0.144*** | |
| | | | (0.049) | | (0.053) | |
| Observations | 1,528 | 1,528 | 1,303 | 1,303 | 1,196 | 1,196 |
| F-stat | 8.110 | | 10.500 | | 9.060 | |

Notes: The table shows the results from regressions where the outcome variable is the total number of pension contributions in December 2019. The estimated coefficients show the coefficient associated with an increase of 10 years of SLE. The first two columns show, respectively the first and second stage when using parental death as the instrument variable. The third and fourth colum show the results when using as instrument the average health assessment of family members (spouse or parents). The fifth and sixth columns show the same, but removing any cases where any family member reported to be invalid for work. Heteroskedastic-robust standard errors are presented in parentheses. *Signif. Codes: ***: p-value<0.01, **: p-value<0.05, *: p-value<0.10.*

The next results use as instruments the average health status of spouses and parents that were also interviewed.[27] I use their self-reported health status, in the 1–6 scale, as the instrument. The third column shows that 1 extra point in the average health status is associated with an increase of 1.49 years of SLE. The F-statistic is 10.5. The second stage result in the fourth column shows a coefficient of 25.42 (non-statistically significant). One potential violation of the exclusion restriction is if those family members with worse health status require more care from individuals, which might influence their labor market status. To alleviate this concern, in the fifth and sixth columns I repeat this exercise excluding

[27] All family members living in the same place were interviewed.

individuals who live with family members that reported to have any disability. The results are very similar, although we cannot reject that these coefficients differ from zero.

The coefficients estimated with these three instrumental variables approach are 3–6 times larger than the baseline results with the cross-section OLS approach. However, they are considerably less precise, with F-statistics for the first stage just around the typical thresholds considered. For the three estimates, the 95% confidence interval would contain the cross-sectional and fixed-effect approaches, as well as the zero. In terms of validity, the parental death requires stronger assumptions. The estimates in the last column are less demanding as they are identifying SLE reporting being induced from family members that self-report worse health who however do not have any disability and do not require assistance.

### 5.4 Robustness

In this section, I present several robustness checks using as benchmark the baseline estimate using the cross-sectional OLS estimation. The first robustness check assesses the role of the initial status in the labor market. In Figure 2, we can see that individuals reporting higher SLE were also more likely to contribute to pensions and had higher stock already in 2005. This is expected, as it has been explored throughout the paper, individuals who expect to live longer have behaviors in the labor market consistent with valuing stable jobs with pension contributions. However, one can be concerned that this initial status may dominate future behavior and is not connected to life expectancy per se. Figure A.5 reproduces the main results controlling additionally for the 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 and could be argued as part of the main effect.

Table A.7 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 25th-75th percentile gap. The overall pattern is very similar across the two distinct measures. 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. Results from the linear specification, represented by the solid black curve, are within the 95% confidence intervals. Lastly, Appendix B investigates the role of non-reporting in the SLE variable, not finding any consistent pattern between non-reported SLE and the observed outcomes.

## 5.5 Additional outcomes

The administrative data is great because of its coverage and richness, displaying precise monthly contributions and pension wealth for all individuals. However, we would ultimately like to test also what happens with total savings and not only pension wealth. In the survey, individuals also self-reported other measures of wealth. I use a measure of total assets summing all available wealth data as outcome in the main specification. Results are presented in Table A.5. Given these answers are infrequent and self-reported, the estimates are very imprecise. The point estimates are all positive, which is reassuring, as large negative results could indicate that total savings are not higher for those with higher SLE.

In the survey, individuals also respond to questions related to their knowledge of the pension system. Individuals were inquired about what were the normal retirement ages for men and women, the value for the mandatory contribution rates and how the pension benefits are computed. I use indicators for correctly answering these questions as outcome variables, using the main specification. Results are presented in Table A.5. Positive coefficients imply that those with higher SLE are more likely to give correct answers for each assessed information. For instance the first column shows that those with 10 extra years of SLE are 1.3pp more likely to know the correct age for retirement for men. The estimates show positive but statistically insignificant coefficients for correctly answering the age-threshold for retirement, the contribution rates, and the way to compute pension benefits.

One other channel it would be plausible to see correlations of SLE is with educational decisions. One margin of adjustment could indeed be whether those with higher beliefs are also more likely to acquire more education. Table A.2 already shows that those with higher levels of education in 2004 also have higher SLE. To further investigate this channel, Table A.5 shows two additional results. The first uses as outcome whether individuals reported they were studying in the initial wave. We can see that 10 extra years of SLE is associated with individuals being 1.1pp more likely to be studying. In the second column, I classify individuals whether they acquired more education. This is an indicator if individuals reported higher educational levels in any wave after the initial one. We also see that 10 years higher SLE is associated with individuals being 1.2pp more likely to have acquired more education, in a base of 30.6%. However, none of these results are statistically significant.

This channel is interesting and deserves future research with larger sample and better data on education progression.

# 6 Discussion and 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.

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 fifteen years of the initial report, the 25th-75th gap in SLE translates to individuals having about 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. Comparing the estimates with the ones obtained with older individuals, we can see that the percent effects increase with age.

The positive link between SLE and labor market outcomes was found exploring the cross-sectional variation and the within-individual variation in the longitudinal dimension. Results for both approaches were in a similar range. However, both approaches could be susceptible to the existence of unobservables (only time-varying unobservables in the latter approach) that correlate both with SLE and the outcomes. 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 and the within-individual variation results. I additionally presented results using an instrumental variable approach, which also presented positive associations between SLE and labor market outcomes, much larger and more imprecise than the other two approaches.

From a measurement perspective alone, it is interesting that SLE measured at young ages correlates so well with future pension contributions and labor market choices. On its own, this fact already advocates for the broader use of these measures. 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


Bell, D., Comerford, D. and Douglas, E. (2020). How do subjective life expectancies compare with mortality tables? similarities and differences in three national samples, *The Journal of the Economics of Ageing* p. 100241.

Bello, P. (2023). Gender-based price discrimination in the annuity market: Evidence from chile, *European Economic Review* **151**: 104356.

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.

Bucher-Koenen, T. and Kluth, S. (2013). Subjective life expectancy and private pensions.

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

Cinelli, C. and Hazlett, C. (2020). Making sense of sensitivity: Extending omitted variable bias, *Journal of the Royal Statistical Society Series B: Statistical Methodology* **82**(1): 39–67.

Costa-Font, J. and Vilaplana-Prieto, C. (2022). Biased survival expectations and behaviours: Does domain specific information matter?, *Journal of Risk and Uncertainty* **65**(3): 285–317.

De Bruin, W. B., Chin, A., Dominitz, J. and Van der Klaauw, W. (2023). Household surveys and probabilistic questions, *Handbook of Economic Expectations*, Elsevier, pp. 3–31.

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.

Delavande, A. and Rohwedder, S. (2011). Differential survival in europe and the united states: Estimates based on subjective probabilities of survival, *Demography* **48**(4): 1377–1400.

Dormont, B., Samson, A.-L., Fleurbaey, M., Luchini, S. and Schokkaert, E. (2018). Individual uncertainty about longevity, *Demography* **55**(5): 1829–1854.

d'Uva, T. B., O'Donnell, O. and Van Doorslaer, E. (2020). Who can predict their own demise? heterogeneity in the accuracy and value of longevity expectations, *The Journal of the Economics of Ageing* **17**: 100135.

Elder, T. E. (2013). The predictive validity of subjective mortality expectations: Evidence from the health and retirement study, *Demography* **50**(2): 569–589.

Finamor, L. (2025). Labor market informality, risk, and public insurance.

Fischhoff, B. and Bruine De Bruin, W. (1999). Fifty–fifty= 50%?, *Journal of Behavioral Decision Making* **12**(2): 149–163.

Fischhoff, B., Parker, A. M., De Bruin, W. B., Downs, J., Palmgren, C., Dawes, R. and Manski, C. F. (2000). Teen expectations for significant life events, *The Public Opinion Quarterly* **64**(2): 189–205.

Foltyn, R. and Olsson, J. (2024). Subjective life expectancies, time preference heterogeneity, and wealth inequality, *Quantitative Economics* **15**(3): 699–736.

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

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

Griffin, B., Hesketh, B. and Loh, V. (2012). The influence of subjective life expectancy on retirement transition and planning: A longitudinal study, *Journal of Vocational Behavior* **81**(2): 129–137.

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. (2009). Subjective probabilities in household surveys, *Annu. Rev. Econ.* **1**(1): 543–562.

Hurd, M. D. and McGarry, K. (1995). Evaluation of the subjective probabilities of survival in the health and retirement study, *Journal of Human resources* pp. S268–S292.

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.

Kutlu-Koc, V. and Kalwij, A. (2017). Individual survival expectations and actual mortality: Evidence from dutch survey and administrative data, *European Journal of Population* **33**(4): 509–532.

Liu, J.-T., Tsou, M.-W. and Hammitt, J. K. (2007). Health information and subjective survival probability: Evidence from taiwan, *Journal of Risk Research* **10**(2): 149–175.

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

Mirowsky, J. (1999). Subjective life expectancy in the us: correspondence to actuarial estimates by age, sex and race, *Social science & medicine* **49**(7): 967–979.

Mirowsky, J. and Ross, C. E. (2000). Socioeconomic status and subjective life expectancy, *Social Psychology Quarterly* pp. 133–151.

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.

Perozek, M. (2008). Using subjective expectations to forecast longevity: Do survey respondents know something we don't know?, *Demography* **45**(1): 95–113.

Post, T. and Hanewald, K. (2013). Longevity risk, subjective survival expectations, and individual saving behavior, *Journal of Economic Behavior & Organization* **86**: 200–220.

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.

Van Solinge, H. and Henkens, K. (2010). Living longer, working longer? the impact of subjective life expectancy on retirement intentions and behaviour, *European journal of public health* **20**(1): 47–51.

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.

Zhou, M., Wang, Y., Liang, Y., Shi, R. and Zhao, S. (2022). The effect of subjective life expectancy on the participation in commercial pension insurance of chinese elderly, *Frontiers in Psychology* **13**: 969719.

# Youth Subjective Life Expectancy and Early Labor Market Choices

Lucas Finamor

## Online Appendices

## A Definition of variables

**Risk-aversion** — The indicator for risk-aversion is 1 if individuals answered that they preferred a fixed wage $X$ over a lottery that pays $2X$ with probability 0.5 and $0.75X$ with probability 0.5.

**Numeracy** — The numeracy variable is constructed summing whether individuals responded correctly to 6 questions on basic math and financial concepts. The range of the variable goes from 0 (no correct answer) to 6 (all correct answers). There are only integer values with no partial correct answer. These questions were first introduced in the 2006 wave.

**Future-oriented preferences** — This is an indicator equal to 0 if individuals answered that they only considered the next months or equal to 1 if they considered any period longer than that (next year, next years, next 5-10 years, more than 10 years) to the question of how far in the future they consider when planning their investments and expenses. This question was introduced in the 2009 wave.

**Big-Five measures** — The big-five measures are constructed using the 10 questions on the big-five personality traits using the *ten item personality measure (TIPI)* proposed by Gosling et al. (2003). The dimensions are agreeableness, conscientiousness, emotional stability, extraversion, and openness. They each range from 1–7.

# 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. Columns 1 and 2 of Table A.1 presents the proportion of missing overall (first column) and by gender and education (column 2). There is no gender difference and a negative gradient in education. Those with more education are more likely to respond this question. Columns 3 and 4 show that answering or not this questions has no consistent relationship with our main outcome variables: the total number of pension contributions (column 3) and pension wealth (column 4) in December of 2019. The coefficients are not statistically significant and go in opposite directions.

Table A.1: Missing values for SLE

| Outcome: | Missing SLE (1) | Missing SLE (2) | # Pension contrib (3) | Pension wealth (4) |
|---|---|---|---|---|
| Constant | 0.091*** | 0.132*** | | |
| | (0.007) | (0.024) | | |
| Women | | 0.000 | | |
| | | (0.013) | | |
| High-School | | -0.032 | | |
| | | (0.025) | | |
| Vocational | | -0.041 | | |
| | | (0.029) | | |
| College | | -0.089*** | | |
| | | (0.025) | | |
| Missing SLE | | | 1.191 | -247.557 |
| | | | (5.675) | (395.823) |
| Main Controls | - | - | ✓ | ✓ |
| Observations | 1,925 | 1,925 | 1,925 | 1,925 |

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 parentheses. *Signif. Codes: ***: p-value<0.01, **: p-value<0.05, *: p-value<0.10.*

# C Additional figures and tables

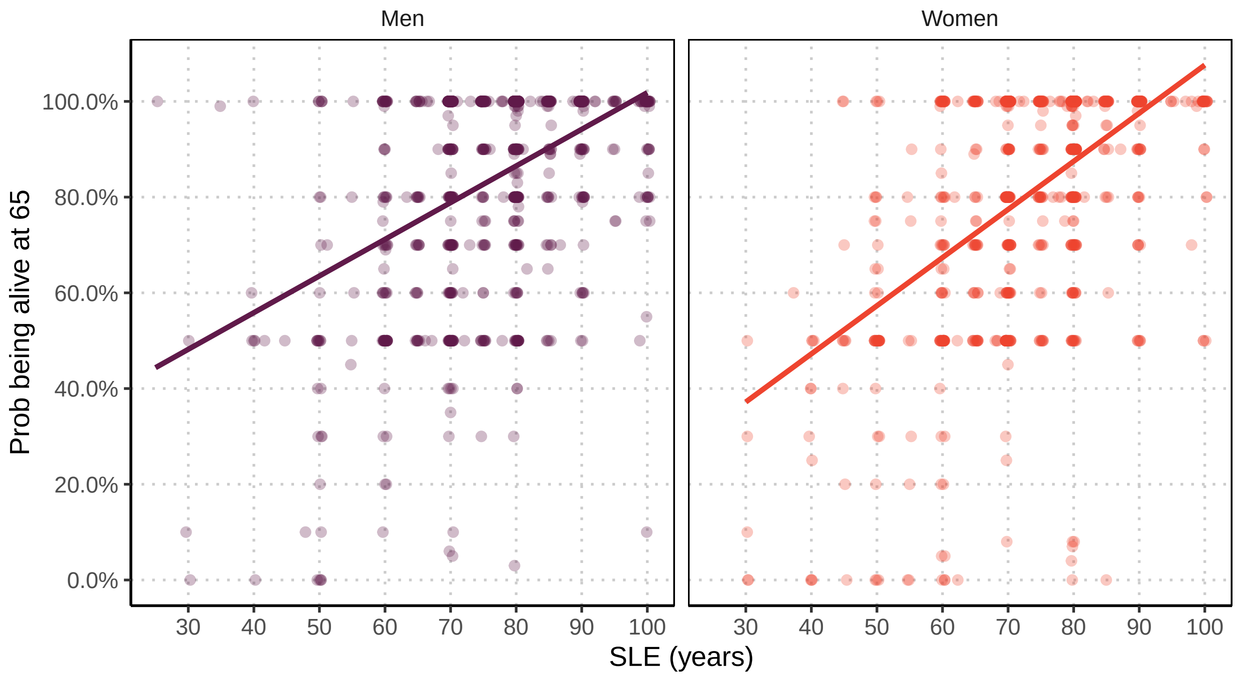
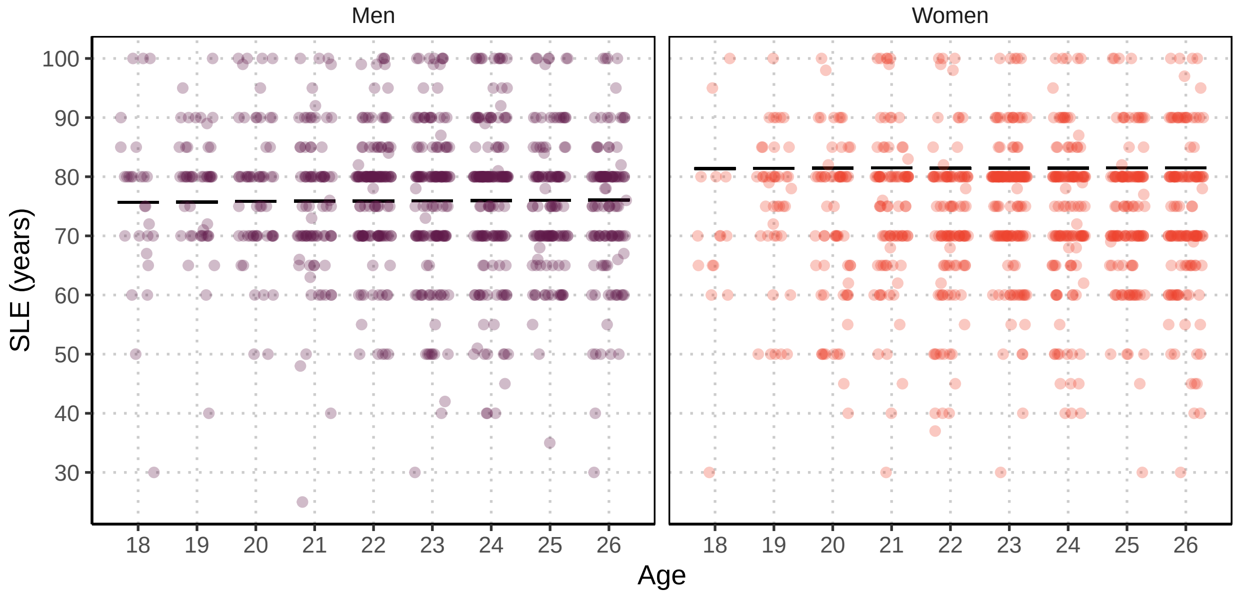



Figure A.1: Subjective life expectancy and probability of being alive at age 65

Notes: For each individual, the figure plots their answer to the two questions regarding life expectancy. On the x-axis is the SLE 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.

(a) SLE by age




(b) SLE by education

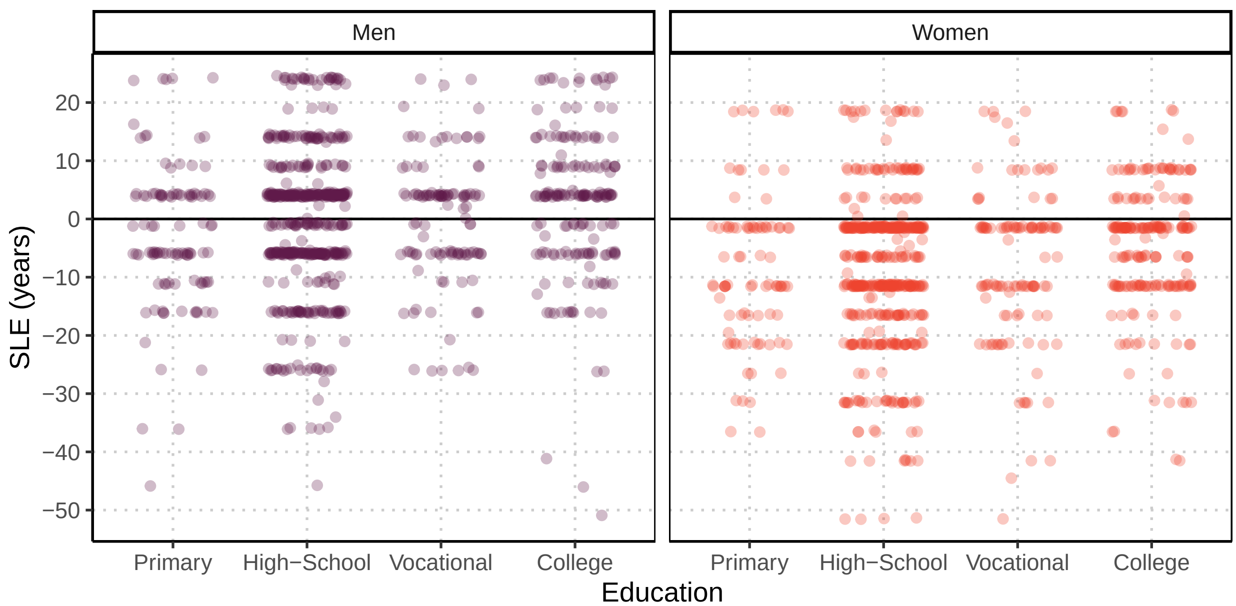



Figure A.2: SLE by age and education

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 lines on panel (a) 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.

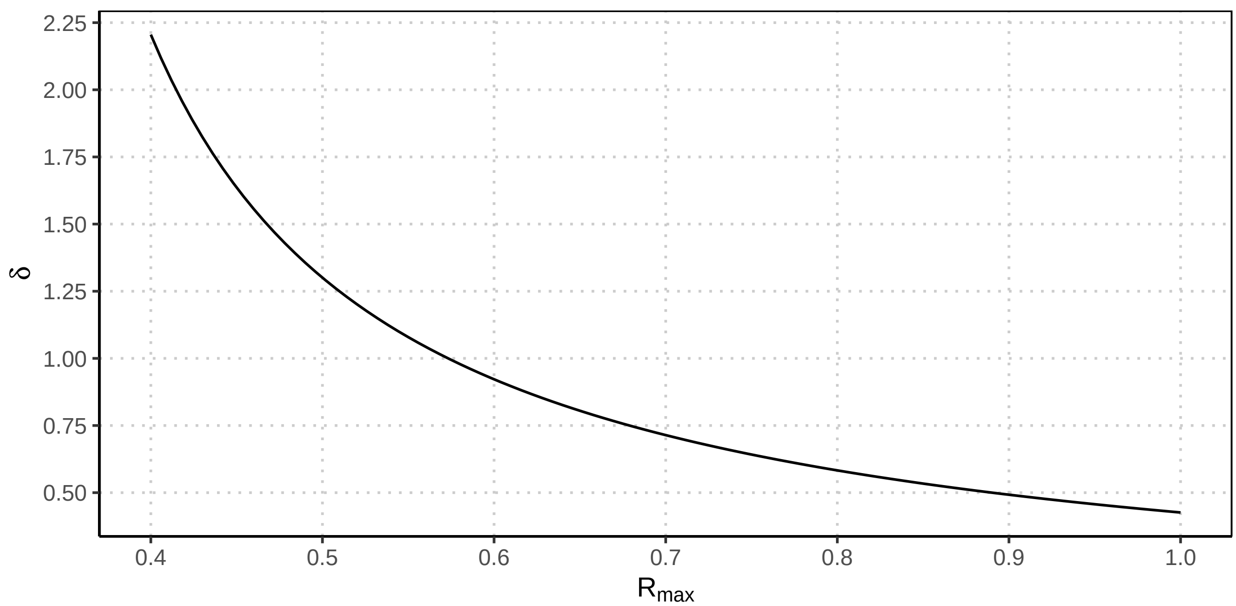
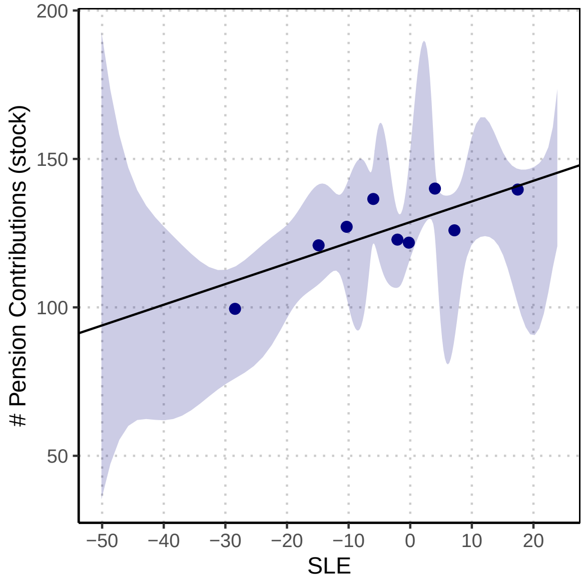


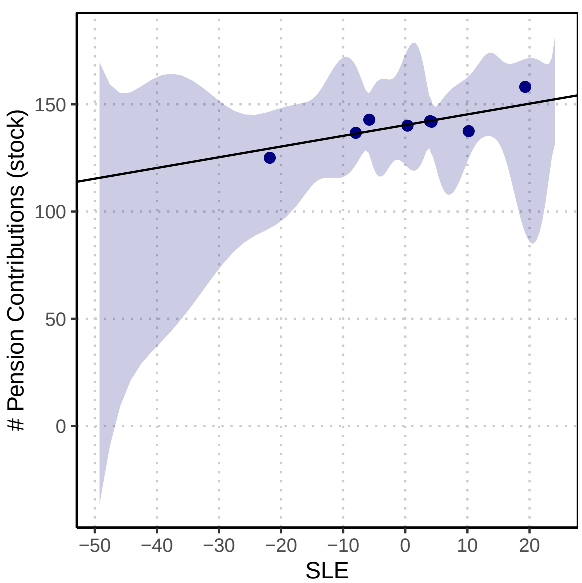

Figure A.3: Conditional independence assumption — Oster (2019)

Notes: The figure shows the results for the procedure proposed by Oster (2019) to assess how strong the unobservables must be (as a ratio of the explanatory power of the observables) to drive the results of the OLS estimation of equation 1. The curve plots this ratio ($\delta$ in the y-axis) as a function of the maximum R-squared that can be obtained. These results used as outcome variable the stock of pension contributions at December of 2019.

(a) All




(b) Men




(c) Women

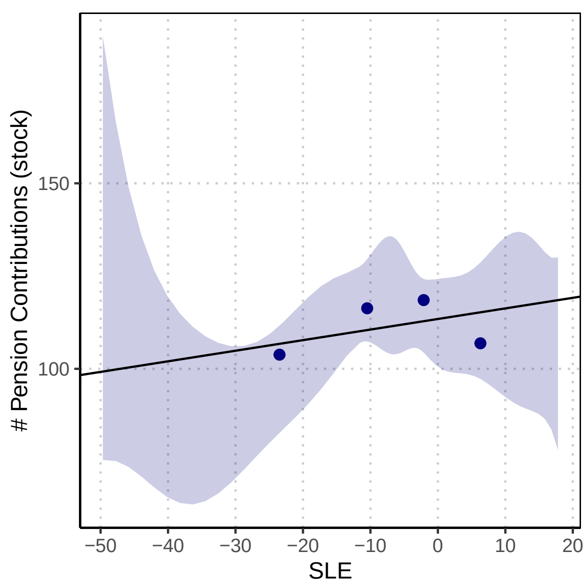
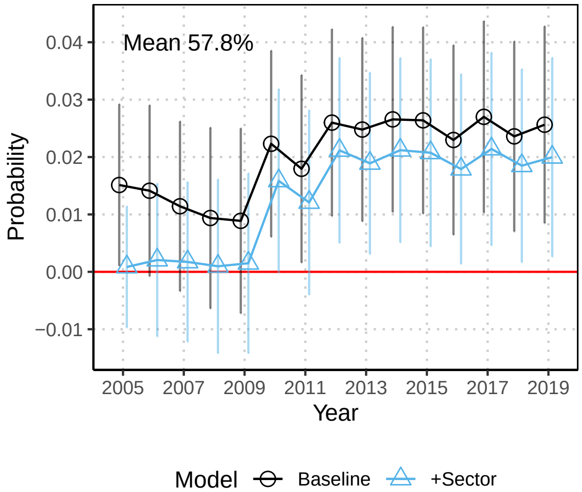
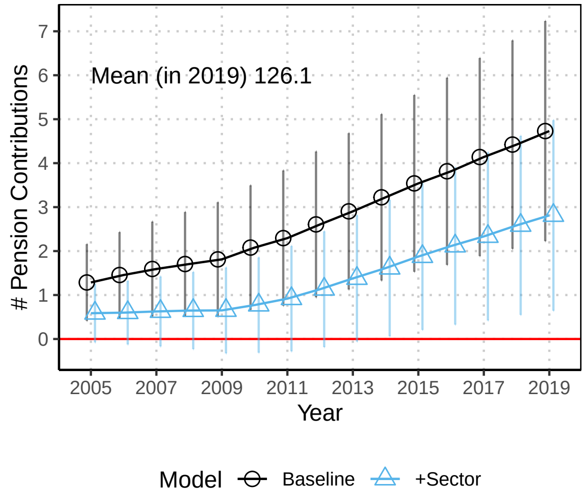



Figure A.4: Robustness — non-linear

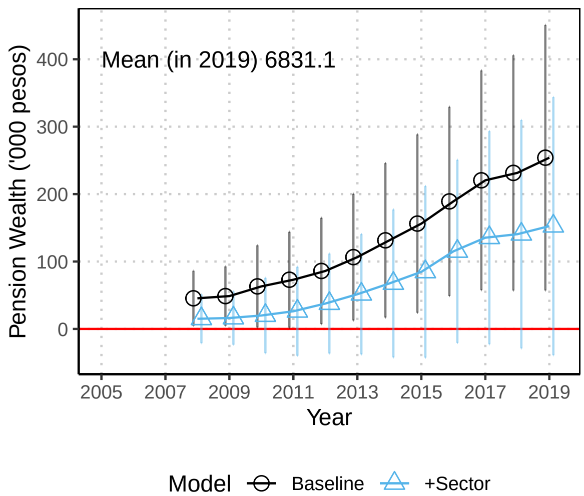
Notes: The figures show the results of applying a non-parametric binscatter regression proposed by Cattaneo et al. (2024). The dots are the resulting bins when using their algorithm to compute the optimal number and position of bins. The shaded area shows the 95% confidence interval for the relationship 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.




(a) Probability of making pension contributions




(b) # Pension contributions (stock)




(c) Pension wealth ('000 pesos)

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

Notes: The figure plots the results from the OLS estimation of equation 1. The estimated coefficients show the coefficient associated with an increase of 10 years of SLE. 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/triangles are the estimated coefficient for $\beta$, 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 which controls are included. The baseline controls (black circles) include demographic cells, region, parental education, and time fixed effects. The blue triangles include, additionally, fixed effects for the initial status in the labor market in 2004 (out of the labor force, unemployment, formal, informal, self-employment, public employment, or any other employment). Standard errors are clustered at the individual level.

Table A.2: Descriptive statistics

| Objective Life Expectancy (years) | Men [75.924] | | | Women [81.464] | | |
|---|---|---|---|---|---|---|
| | N Obs | Prop | SLE (mean) | N Obs | Prop | SLE (mean) |
| Total | 894 | - | 76.002 | 856 | - | 73.659 |
| **Educational level** | | | | | | |
| Primary | 105 | 0.117 | 73.210 | 86 | 0.100 | 72.849 |
| High-School | 497 | 0.556 | 75.765 | 455 | 0.532 | 72.473 |
| Vocational | 116 | 0.130 | 75.759 | 113 | 0.132 | 73.531 |
| College | 176 | 0.197 | 78.500 | 202 | 0.236 | 76.748 |
| *(subtotal)* | 894 | | | 856 | | |
| **Region** | | | | | | |
| MR Santiago | 370 | 0.414 | 76.146 | 357 | 0.417 | 74.364 |
| Other | 524 | 0.586 | 75.901 | 499 | 0.583 | 73.154 |
| *(subtotal)* | 894 | | | 856 | | |
| **Mother education** | | | | | | |
| Less than HS | 406 | 0.470 | 74.776 | 411 | 0.495 | 72.221 |
| HS | 374 | 0.433 | 77.067 | 328 | 0.395 | 74.534 |
| More than HS | 83 | 0.096 | 77.880 | 92 | 0.111 | 77.652 |
| *(subtotal)* | 863 | | | 831 | | |
| **Father education** | | | | | | |
| Less than HS | 369 | 0.439 | 75.171 | 369 | 0.461 | 72.033 |
| HS | 375 | 0.446 | 76.509 | 337 | 0.421 | 74.964 |
| More than HS | 96 | 0.114 | 78.125 | 95 | 0.119 | 77.168 |
| *(subtotal)* | 840 | | | 801 | | |

Notes: The table presents the main descriptive statistics for men (first three columns) and women (last three columns). For each variable and gender, the table presents the number of observations, the proportion within the group, and the mean subjective life expectancy measured in years. The first row presents the total number of observations 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). In the table header, the numbers in brackets show the average objective life expectancy, according to the 2005 life tables.

Table A.3: SLE in 2004 and subsequent waves

| Outcome: SLE in | 2006 | 2009 | 2012 |
|---|---|---|---|
| | (1) | (2) | (3) |
| SLE in 2004 | 0.415*** | 0.307*** | 0.189*** |
| | (0.055) | (0.052) | (0.045) |
| Observations | 458 | 458 | 458 |

Notes: The table shows the results from a regression of SLE measured in future surveys (2006, 2009, and 2012) on the initial SLE reported in 2004. All regressions include the baseline controls (demographic cell, region, and parental education). The results in the first three columns are from the OLS estimation. The results in columns 4–6 use an IV/2SLS strategy, using the probability of being alive at age 65 as an instrument for SLE in 2004. Heteroskedastic-robust standard errors in parentheses. *Signif. Codes: ***: p-value<0.01, **: p-value<0.05, *: p-value<0.10.*

Table A.4: SLE in subsequent surveys and new diagnosis

| | Outcome: SLE $_{t+1}$ | |
|---|---|---|
| | (1) | (2) |
| New diagnosis group 1 | 0.559 | 1.015 |
| | (1.692) | (1.567) |
| New diagnosis group 2 | -3.984$^{**}$ | -3.828$^{**}$ |
| | (1.716) | (1.709) |
| New diagnosis group 3 | -2.345$^{*}$ | -2.146$^{*}$ |
| | (1.354) | (1.268) |
| SLE$_t$ | | 0.326$^{***}$ |
| | | (0.025) |
| Observations | 2,507 | 2,507 |
| $R^2$ | 0.072 | 0.177 |

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. The explanatory variables are indicators for a new diagnosis of diseases in each of three groups. 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. The second column includes *SLE* in the first survey as a control. All regressions include the baseline controls (demographic cell, region, and parental education). Standard errors clustered at the individual level are displayed in parentheses. *Signif. Codes: ***: p-value<0.01, **: p-value<0.05, *: p-value<0.10.*

Table A.5: Additional outcomes

| Outcome: | Knowledge of the pension system | | | | Wealth | Education | |
|---|---|---|---|---|---|---|---|
| | Ret Age Men | Ret Age Women | Contrib Rate | Pension Formula | Non-Pension Wealth | Studying in 2004 | Acquired More Education |
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) |
| SLE | 0.013 | 0.011 | 0.005 | 0.011* | 29.885 | 0.011 | 0.012 |
| | (0.010) | (0.009) | (0.010) | (0.006) | (173.960) | (0.008) | (0.010) |
| Obs | 1,750 | 1,750 | 1,750 | 1,750 | 1,319 | 1,750 | 1,272 |
| Constant | 0.709 | 0.686 | 0.359 | 0.078 | 5,002.9 | 0.290 | 0.306 |

Notes: The table presents the results of the OLS regression of equation 1 for additional outcomes. The estimated coefficients show the coefficients associated with an increase of 10 years of SLE. In the first four columns, the dependent variable are indicators for correct answers to the questions about the pension system. The questions are: the retirement age for men (column 1), for women (column 2), the mandatory contribution rates (column 3), and the pension formula (column 4). In the fifth column the oucome is the non-pension wealth (self-reported in the survey), winsorized at the 10% and 90% level, measured in thousands of Chilean pesos. Lastly, in the sixth and seventh columns, the outcomes are related to educational decisions. Column (6) uses as outcome a binary indicator whether the individual was studying when they answered the main survey in 2004. Column (7) uses as outcome whether the individual acquired more education after the initial survey. For this it was coded as individuals getting further education if they showed a higher educational level (measured in the ordered scale: Primary, High-School, Vocational, College) in any survey after 2004. The number of observation is smaller because we lose individuals who were already with College education in 2004 or were not observed in any other wave of the survey. All regressions include the baseline controls (demographic cell, region, and parental education fixed effects). For the fifth column, I also include age interacted with the wave in which the wealth measure was recorded. The regression in the sixth column removes education as one of the controls, while the regression in the seventh column adds a further control for what was the last wave observed for each individual. Heteroskedastic-robust standard errors are in parentheses. *Signif. Codes: ***: p-value<0.01, **: p-value<0.05, *: p-value<0.10.*

Table A.6: Additional control variables

| | No controls | +Gender | **Baseline** | +Health | +Parental death | +Risk-aversion | +Numeracy & Pref | +Big-five |
|---|---|---|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) |
| SLE | 6.959*** | 3.909*** | 4.730*** | 4.577*** | 4.296*** | 4.234*** | 3.954*** | 3.960*** |
| | (1.280) | (1.338) | (1.278) | (1.293) | (1.289) | (1.292) | (1.287) | (1.283) |
| Obs | 1750 | 1750 | 1750 | 1750 | 1750 | 1750 | 1750 | 1750 |
| $R^2$ | 0.016 | 0.047 | 0.211 | 0.222 | 0.237 | 0.239 | 0.246 | 0.256 |

Notes: The table presents estimates of the main equation with different sets of control variables. The estimated coefficients show the coefficient associated with an increase of 10 years of SLE. The first column shows the coefficient of SLE, when no additional control is included. Second columns includes gender and the third column all the baseline controls (gender-age-education, region, and parental education fixed effects). The next follows the pattern of Table 1 including sequentially, all health & life-style variables (column 4), parental death (column 5), risk-aversion (column 6), numeracy and future-oriented preferences (column 7), and the Big-Five personality traits (column 8). Standard errors are clustered at the individual level. *Signif. Codes: ***: p-value<0.01, **: p-value<0.05, *: p-value<0.10.*

Table A.7: Future labor market outcomes — SLE in years and Prob at age 65

| Outcome: | # Pension contrib | | Pension Wealth | |
|---|---|---|---|---|
| Variable | SLE | P65 | SLE | P65 |
| | (1) | (2) | (3) | (4) |
| Estimate | 4.730*** | 3.862* | 255.078** | 350.157** |
| (s.e.) | (1.278) | (1.981) | (111.470) | (164.274) |
| Num.Obs. | 1750 | 1890 | 1750 | 1890 |

Notes: The table presents the results of the OLS regression of equation 1 using the SLE variable (columns 1 and 3) or the $P^{65}$ variable (columns 2 and 4) as the regressor. For the SLE results, I show the coefficient associated with an increase of 10 years of SLE. 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. In the first two columns the outcome variable is the total number of pension contributions in December 2019, and in the last two columns pension wealth in December 2019 in thousands of Chilean pesos. All regressions include baseline controls (demographic cell, region, parental education, and time fixed effects). heteroskedastic-robust standard errors are in parentheses. *Signif. Codes: ***: p-value<0.01, **: p-value<0.05, *: p-value<0.10.*

Table A.8: SLE and mortality — Probabilistic question

| | Men | | | | Women | | | |
|---|---|---|---|---|---|---|---|---|
| Age range | 18–26 | 27–36 | 37–46 | 47-56 | 18–26 | 27–36 | 37–46 | 47-56 |
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) |
| **Panel A - Main specification** | | | | | | | | |
| SLE | -0.006 | -0.004 | -0.017** | -0.031*** | -0.001 | 0.001 | -0.003 | -0.007 |
| | (0.004) | (0.004) | (0.007) | (0.010) | (0.004) | (0.002) | (0.004) | (0.007) |
| $R^2$ | 0.066 | 0.047 | 0.059 | 0.090 | 0.060 | 0.050 | 0.090 | 0.075 |
| **Panel B - With health controls** | | | | | | | | |
| SLE | -0.006 | -0.003 | -0.015** | -0.020** | -0.001 | 0.002 | -0.001 | -0.004 |
| | (0.005) | (0.003) | (0.007) | (0.010) | (0.004) | (0.002) | (0.004) | (0.007) |
| $R^2$ | 0.093 | 0.061 | 0.070 | 0.113 | 0.066 | 0.057 | 0.100 | 0.079 |
| **Panel C - Main specification** | | | | | | | | |
| P65 | -0.003 | -0.007 | -0.020*** | -0.036*** | -0.001 | -0.005 | 0.006 | -0.011 |
| | (0.008) | (0.005) | (0.006) | (0.013) | (0.004) | (0.005) | (0.004) | (0.009) |
| $R^2$ | 0.075 | 0.036 | 0.063 | 0.077 | 0.056 | 0.050 | 0.086 | 0.086 |
| **Panel D - With health controls** | | | | | | | | |
| P65 | -0.001 | -0.005 | -0.017** | -0.025** | 0.000 | -0.004 | 0.008* | -0.009 |
| | (0.008) | (0.005) | (0.007) | (0.012) | (0.004) | (0.004) | (0.004) | (0.009) |
| $R^2$ | 0.096 | 0.051 | 0.070 | 0.096 | 0.062 | 0.056 | 0.095 | 0.090 |

Notes: The table presents the results from an OLS regression of equation 1 using an indicator for being deceased by December 2019 as the outcome variable. The estimated coefficients show the coefficients associated with an increase of 10 years of SLE (panels A and B) and of 30 pp of P65 (panels C and D). Both of them represent approximately the 25th-75th gap in its respective distribution. The first four columns are for men, and the last four for women. Panels A and C are the baseline regressions. Panel B and D present the regression with all health variables included in Table 1. All regressions include the baseline controls (demographic cell, region, and parental education). Heteroskedastic-robust standard errors are included in parenthesis. *Signif. Codes: ***: p-value<0.01, **: p-value<0.05, *: p-value<0.10.*

Table A.9: Correlates of SLE and diseases diagnosis

| | Outcome: SLE | | |
|---|---|---|---|
| | (1) | (2) | (3) |
| Any Diagnosis | -0.955 | | |
| | (1.080) | | |
| Diagnosis Group 1 | | -2.613 | |
| | | (1.706) | |
| Diagnosis Group 2 | | 0.885 | |
| | | (1.558) | |
| Diagnosis Group 3 | | -1.101 | |
| | | (1.902) | |
| Diagnosis Asthma and Pulmonary Emphysema | | | -3.087 |
| | | | (1.990) |
| Diagnosis Diabetes | | | 2.914 |
| | | | (3.694) |
| Diagnosis Arthritis and Osteoarthritis | | | -5.856 |
| | | | (4.761) |
| Diagnosis Hypertension and High Blood Pressure | | | 0.568 |
| | | | (2.035) |
| Diagnosis Heart Problems | | | -0.745 |
| | | | (2.677) |
| Diagnosis Cancer | | | 0.741 |
| | | | (3.226) |
| Diagnosis Renal Diseases | | | 3.695 |
| | | | (3.537) |
| Diagnosis Depression | | | -0.912 |
| | | | (1.957) |
| Diagnosis Mental Illness | | | -3.625 |
| | | | (6.052) |

Notes: The table presents the results from a regression of SLE on selected variables. All regressions include as controls the baseline controls (demographic cell, region, and parental education) as well as all variables from Table 1. 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. *Signif. Codes: ***: p-value<0.01, **: p-value<0.05, *: p-value<0.10.*

Table A.10: Moments of outcome variables

| Sample | All | | Men | | Women | |
|---|---|---|---|---|---|---|
| Variable | Pension Wealth<br>(1) | Stock Pension Contrib<br>(2) | Pension Wealth<br>(3) | Stock Pension Contrib<br>(4) | Pension Wealth<br>(5) | Stock Pension Contrib<br>(6) |
| Mean | 6831.1 | 126.1 | 7999.8 | 140.4 | 5610.6 | 111.2 |
| Median | 5070.5 | 133.0 | 6477.6 | 151.5 | 3990.4 | 112.0 |
| Avg change in 1 year | 483.9 | 6.9 | 563.8 | 7.6 | 400.4 | 6.1 |
| Standard Deviation | 6324.5 | 70.7 | 6655.8 | 68.8 | 5712.7 | 69.6 |
| 75th-25th gap | 7764.7 | 109.5 | 8027.3 | 99.8 | 7049.2 | 113.0 |
| Gender gap (Men-Women) | 2389.1 | 29.1 | - | - | - | - |
| Education gap (College-HighSchool) | 3585.5 | -15.2 | 2377.0 | -37.5 | 4973.8 | 8.3 |

Notes: The table show moments of the outcome variables used in the analysis (pension wealth and number of monthly pension contributions, both measured in Dec2019). The moments are the mean, the median, the average change in 12 months (for the entire 2004–2019 period), the standard deviation, the interquartile range (75th-25th gap), the gender gap (average value for men minus average value for women), and the education gap (average value for college educated minus average value for high school educated). The first two columns are for the entire sample, the third and fourth for men, and the last two for women.