Help! I'm not so self-assured!

On the design of zero-pillar compensations

Marcela Parada-Contzen*

December 10, 2018

*Departamento de Ingeniería Industrial, Facultad de Ingenería – Universidad de Concepción, Chile. Address: Edmundo Larenas 219, Concepción, Chile. Phone: (56-41)-2204151. E-mail: mparadacontzen@gmail.com.
I thank the funding granted by Fondecyt, Project No. 3180155. I also thank the Subsecretary of Social Prevision of Chile for providing and authorizing the use of the data.

Abstract

This paper provides a mechanism for the design of policies that attempt to ameliorate income disparities and poverty among the elderly in defined-contribution retirement systems. Because lifetime events impact wealth accumulation through multiple channels, this paper estimates a set of correlated equations that capture labor market behavior, portfolio and savings decisions, risk preferences, family characteristics, medical care consumption, expected duration of life, and health status. With these data, I unbiasedly quantify the effect of family and health characteristics on retirement wealth accumulation. Using data from the Chilean retirement system, I propose a bonus of between 177 and 6,252 dollars payable to women to compensate for gender inequality due to marital status and a compensation of between 1,965 and 3,912 dollars for the bottom 10% of the wealth distribution to alleviate inequality among women. For compensation due to child bearing, I propose a bonus payable up to the third quartile of the woman's wealth distribution that varies between 344 and 4,577 dollars per additional child. Finally, to compensate for health shocks, I propose a bonus that varies between 474 and 5,509 dollars.

Keywords: retirement income policy, gender inequality, elderly poverty.

JEL Classification: J26, D15, I38, C33.

1 Introduction

This paper provides a mechanism for the design of policies to ameliorate retirement income disparities and low pension incomes in defined-contribution (DC) retirement systems. I refer to this set of policies as zero-pillar schemes, following the five-pillar World Bank classification.[1] Zero-pillar compensations have gained importance as individuals reach retirement age without accumulating enough wealth for retirement. Because in DC schemes, an individual's retirement income depends on the wealth she accumulates by the end of the working life cycle, lifetime events that impact labor market outcomes and financial positions have important long-run consequences.

For example, having children may directly impact finances as well as consumption and savings behavior and may indirectly impact the time that household members can work (Hubener et al., 2015). Along this line, women are a particularly vulnerable group. Generally, this occurs because women participate less in the labor market due to child bearing and other household work, work fewer hours, and earn less (Bettio et al., 2013; Ponthieux and Meurs, 2015). Similarly, health shocks impact earnings, portfolio choices, expenses, consumption patterns, and savings behavior (Kim et al., 2017; Beshears et al., 2017; Bogan and Fertig, 2018), causing important consequences for the elderly.

The way retirement systems cope with differences in retirement income depends on the design of the system. Some systems do not address heterogeneous labor histories, while others compensate histories of unstable contributions in order to equalize retirement income (Kuivalainen et al., forthcoming). For example, some systems provide access to childcare, others provide contribution credits for periods of time during which the individual is taking care of others, and others match contributions of the poorer individuals (Jefferson, 2009). Other systems follow non-contributory programs to prevent poverty. Generally, multi-pillar pension schemes that guarantee a pension income have been recommended to account for incomplete contribution histories (Möhring, 2015).

In this context, this paper quantifies the effect of family (i.e., marital status, children and female labor opportunity cost of child bearing) and health status on retirement wealth accumulation. The objective is to provide mechanisms for the design of zero-pillar policies that attempt to ameliorate retirement income disparities and low pension

[1] Zero-pillar compensations consist of guaranteed pension income typically provided by the state. They are also called non-contributory pensions (World Bank, 2008).

incomes (e.g., bonus per illness and child bearing compensations). The zero-pillar compensations proposed here consist of bonuses payable by governments added to an individual's retirement savings accounts at the time of retirement or payable in monthly shares. These bonuses are designed so that they compensate women for wealth losses due to changes in family structure and negative health shocks so that poverty and inequality rates decrease. The paper isolates the effects of different life-cycle events to design efficient short-run programs that address poverty and inequality among retirees.

To unbiasedly quantify the effects of life-cycle events, I use maximum likelihood methods to estimate a multiple equation reduced-form life-cycle model in which individuals make decisions that directly impact wealth accumulation in a DC retirement system (e.g., employment status, occupation selection, retirement portfolio choice and optional savings). I endogenously incorporate family characteristics (such as marital status and number of children) and health status by allowing correlation with decisions through observable and unobservable characteristics. Using the model estimates, I simulate the effects of changes in family structure and health shocks to quantify their impact on retirement wealth. This paper focuses on the iconic Chilean retirement system. For the estimation, I use the first four waves (2002–2009) of the Chilean Survey of Social Protection (EPS) linked with administrative records of the Chilean Superintendence of Pensions. The EPS is a validated survey used for the evaluation of retirement policies (Behrman et al., 2011; Joubert, 2015).

With this research I contribute to the literature in several ways. First, I provide simulations for the implementation of zero-pillar schemes in DC systems. Note that many countries that implemented DC systems in the 80s and 90s are now approaching the first generation of retirees. Therefore, the policy discussion around the effects of incomplete contribution histories have increased worldwide. In the case of millennials (i.e., those born after 1980), it is possible to consider insurance mechanisms to ensure contributions during periods of inactivity.

Nevertheless, in the short run, there are generations of contributors who will not benefit from potential policy reforms. To prevent poverty and inequality among these generations of retirees, in this paper, I propose implementing fiscal pension compensations specific to the individual's position in the wealth distribution and to her characteristics. I take the experience of Chile because its model, first implemented in

1981, has served as a prototype for the design of other DC systems (Orszag and Stiglitz, 2001).[2] In this setting, this research contributes to the design of public policies for all countries that have implemented similar systems and are facing or will face similar issues.

Second, because the amounts of social pensions are generally discretionary, established after discussions between policy makers and fiscal authorities, and depend on other policy variables such as minimum wages (Bovenberg et al., 2012), in this research, I quantify the effects of family and health characteristics on life-cycle retirement wealth accumulation so that non-contributory pensions are indeed related to the dynamic effects of discontinuous contribution histories and not dependent on other variables. In addition, I identify different sources of inequalities, such as differences due to gender or heterogeneity across the wealth distribution within gender groups. With the results, I contribute to the design of such policies.[3]

Third, by jointly estimating a set of correlated equations and capturing individual behaviors and outcomes, I account for many sources of estimation biases typically present in the retirement literature, such as selection into behaviors, endogeneity due to unobservables and measurement error. Selection and endogeneity biases may occur because individual preferences over choice sets are unobserved, and if we expect these unobservable characteristics to be correlated across decisions, then the estimates are inconsistent (Gelber, 2011; Beshears and Choi, 2012).[4] An advantage of this research is that I use administrative records that allow me to recover the complete history of retirement wealth accumulation, meaning that it does not depend on self-reported amounts, which are susceptible to measurement error (Engelhardt and Kumar, 2007).

[2]Some countries in the region that have adopted some form of what is known as the Chilean model are Argentina (1994), Bolivia (1997), Colombia (1993), Costa Rica (1995), the Dominican Republic (2003), El Salvador (1998), Mexico (1997), Panama (2008), Peru (1993), and Uruguay (1996). Others in Eastern Europe are Bulgaria (2000), Hungary (1998), and Poland (1999) (Kritzer, 2008; Joubert, 2015). The system has the great advantage of alleviating the fiscal stress that pay-as-you-go systems impose. Nevertheless, some of the limitations are low pension outcomes, low coverage, and lack of gender equality (Barr and Diamond, 2016).

[3]In Chile, to assist individuals with low retirement income due to an incomplete contribution history or a low lifetime income, a non-contributory pension (solidarity pension) was introduced in 2008 for the bottom 60% of retirees (Barr and Diamond, 2016). Moreover, to alleviate retirement income disparities, a lump-sum payment to retired women with children was introduced (bonus per child), which equals a year's contribution for the minimum wage (OECD, 2007).

[4]For example, it is expected that risk-averse individuals (generally unobserved) select safer occupations and chose safer portfolios while consuming more preventative medical care, resulting in better health status and higher levels of earnings. Measurement error may come from self-reported variables measured with noise.

In addition, different from the previous literature, I use a flexible empirical approach in which few distributional assumptions about the jointly correlated unobservables are made, which allows me to incorporate the non-linearities of the individual's decision-making process without making any assumption about preferences and expectation processes, thus avoiding potential biases due to misspecification (Engelhardt and Kumar, 2007; Card and Ransom, 2011).

Fourth and finally, since I model the individual decision-making process over time, I do not rely on exogenous shocks for identification. Rather, once I have estimated the unbiased reduced-form parameters over observable and unobservable characteristics, I can simulate situations in which individuals face health shocks or changes in family structure (e.g., marriage and additional children). This avoids the issue in the pension literature that analysis generally depends on policy reforms for identification or on settings that lack external validity (Blau, 2016).

The rest of the paper is organized as follows. The next section presents the empirical model together with the institutional background of the Chilean retirement system. Section 3 presents the data and construction of the research sample, and Section 4 presents the estimation strategy. Section 5 presents the features of the estimated model and its fits, along with the simulation results. Finally, Section 6 concludes.

2 Empirical model

The model consists of a set of correlated reduced-form equations that capture individual life-cycle behavior. Upon observing a wage offer, individuals make choices with respect to employment, occupation, retirement portfolio, and optional savings. These decisions depend on observed individual characteristics and unobserved heterogeneity. Family and health characteristics are endogenized in the individual decision-making process by allowing for their future realizations to depend on current decisions, in addition to previous behaviors, and on individual unobserved heterogeneity. The empirical model also considers other characteristics and outcomes that are endogenous to decisions, such as medical care consumption, self-reported expected duration of life, and risk aversion.

All the equations are correlated through individual unobserved heterogeneity (CUH), which is correlated with equations through a permanent component (μ) and a time-varying component (ν_t). Additionally, there is uncorrelated unobserved heterogeneity (ε_t). I will later refer to the error structure in detail.

Note that I do not parametrize and solve the optimization problem but use a theoretical life-cycle model as a benchmark model to approximate the discrete choice behaviors that define wealth accumulation in the context of a DC retirement system. The model includes institutional characteristics of the Chilean pension system. These features are presented in the next section, followed by the timing of the model.

2.1 Institutional background

The Chilean DC retirement system was introduced in 1981. Since its beginning, all formal workers must contribute 10% of their income to an individual retirement account. This 10% is deducted from their before-tax paycheck and put into the accounts. Mandatory savings for retirement (from now on, retirement wealth) can only be cashed upon retirement using an annuity program and some guaranteed term scheme.

The per-period alternatives with respect to contribution rates, retirement portfolio allocation, and optional savings are defined within the Chilean retirement system framework. Retirement wealth accumulated within the retirement system is denoted A_{it}.

Since 2002, enrollees make 5 retirement portfolio choices. These choices correspond to whether to invest or not in each available retirement account. These accounts are labeled A, B, C, D, and E, and the only differences between them is that they vary in the level of financial risk. For example, the riskiest fund is Account A, which invests 40-80% in equities, while account E is the safest one, investing up to 5% in equities. The range invested in equities for Account B is 25-60%, that for Account C is 15-40%, and that for Account D is 5-20%. Individuals can chose at most 2 investment accounts. Before 2002, only one account was available (Account C). Individuals' portfolio choices are denoted by $p_{it} = \{p_{it}^A, p_{it}^B, p_{it}^C, p_{it}^D, p_{it}^E\}$, which consist of whether or not they invest in the available accounts of the system. Note that the data used in this research start in 2002, so it is not necessary to model initial portfolio conditions.

Individuals may also open a voluntary retirement account or a savings account outside of the retirement system (e.g., in the banking sector). These optional savings decisions are modeled as a dichotomous variable denoted s_{it}.

2.2 Timing and notation

The individual i starts period t with an information set Ω_{it} that includes past choices and relevant knowledge about individual and market characteristics. The information set Ω_{it} contains stock variables up to period t that summarizes the history of decisions, lagged choices, and exogenous characteristics. Stock variables include, for example, accumulated wealth for retirement (A_{it}), work experience (E_{it}), marital state (M_{it}), number of children (N_{it}), and health status (H_{it}). Lagged choices include the retirement portfolio chosen last period ($p_{i,t-1}$) and optional savings last period ($s_{i,t-1}$). I include individual exogenous characteristics (X_{it}) (e.g., gender, age, and schooling) and other exogenous market-level characteristics (Z_{it}) (e.g., prices). I use $\tilde{\Omega}_t = [p_{i,t-1}, s_{i,t-1}, A_{it}, E_{it}, M_{it}, N_{it}, H_{it}]$ to denote the set of endogenous variables influencing an individual's decision. Therefore, $\Omega_{it} = [\tilde{\Omega}_{it}, X_{it}, Z_{it}]$.

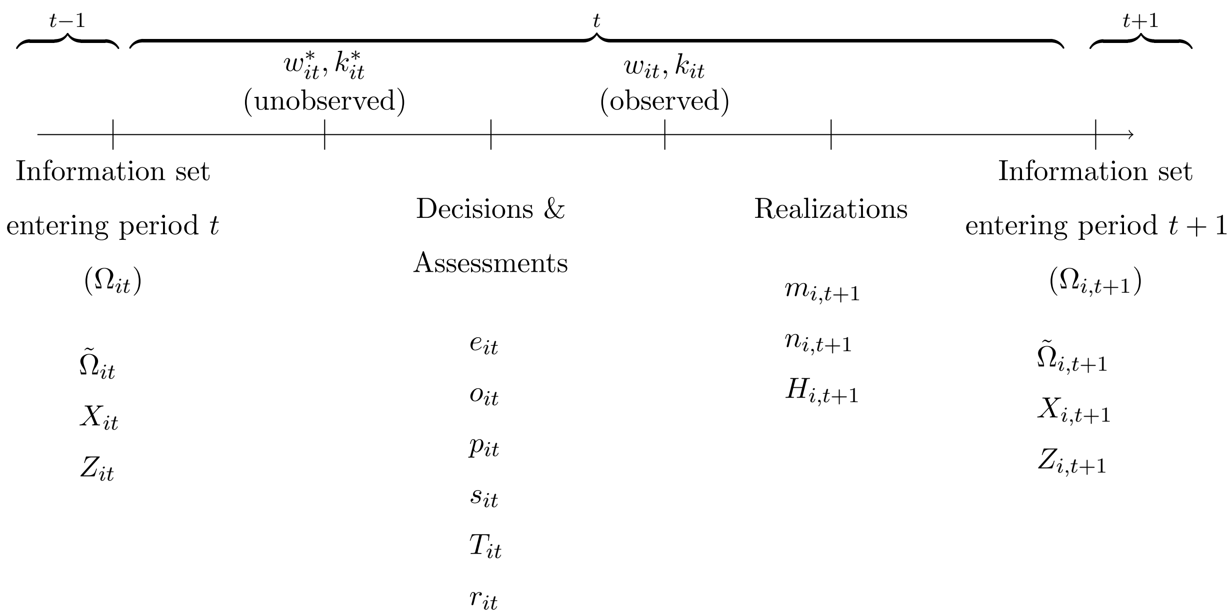
The choice probabilities for individual i at time t correspond to an approximation of non-linear demand functions that depend on observable characteristics (Ω_{it}) and individual unobserved heterogeneity (ϵ_{it}). These choice probabilities capture simultaneous decisions that impact retirement wealth accumulation, such as employment status (e_{it}), occupation selection (o_{it}), retirement portfolio choice (p_{it}), and optional savings (s_{it}). The decisions are simultaneously made once the individual has observed a wage offer (w^*_{it}) drawn from an occupation-specific wage distribution and a draw (k^*_{it}) from the medical care consumption distribution, which represents necessary consumption within the current period. Both of these draws are unobserved for the econometrician.

The period t marital status (m_{it}), changes in family size (n_{it}), and health status (H_{it}) are observed entering period t and updated after the individual has made her choices. Note that future realizations of family and health characteristics depend on current decisions, in addition to previous behaviors and outcomes. For example, health status entering the next period ($H_{i,t+1}$) may be a function of current period employment status (e_{it}) and health consumption (k_{it}). Past marriage realizations are summarized by the marital history vector M_{it}, and past child realizations are summarized by the child history vector N_{it}.

To prevent estimation bias due to selection and endogeneity, I also incorporate into the system of equations medical care consumption (k_{it}), wages (w_{it}), self-reported expected duration of life (T_{it}) and self-reported risk aversion (r_{it}). All equations are correlated through individual unobserved heterogeneity, which represents individual characteristics that are unobserved for the econometrician but affect the individual's decisions (such as personality traits and risk preferences).

After making the period t decisions and realizing the period $t+1$ stochastic values, the individual updates her information set to $\Omega_{i,t+1}$. A summary of the timing assumptions is presented in Figure 1.

Figure 1: Timing of the model



where $\tilde{\Omega}_{it} = \{p_{i,t-1}, s_{i,t-1}, A_{it}, E_{it}, M_{it}, N_{it}, H_{it}\}$.

2.3 Estimable model

2.3.1 Wages

As in an extension of Mincer model (1974), wages are a function of exogenous individual characteristics (X_{it}), human capital accumulation such as work experience (E_{it}) or schooling, and job characteristics such as occupation category (o_{it}). Furthermore, conditional on human capital accumulation and job characteristics, individuals may vary in their level of productivity. Because productivity is unobserved for the econometrician, I proxy for it assuming that productivity depends on family characteristics such as marital status (m_{it}), number of children in the household (N_{it}), and health status (H_{it}). Wages may also depend on demand-side characteristics. Thus, I include a vector of employment demand-side shifters (Z_{it}^{E}) such as unemployment rates.

$$w_{it} = w(E_{it}, o_{it}, m_{it}, N_{it}, H_{it}, X_{it}, Z_{it}^{E}, \epsilon_{it}^{W}) \tag{1}$$

where ϵ_{it}^{W} represents unobserved wage variation, and it is decomposed into a permanent unobserved component (μ_t^W), a time-varying unobserved component (ν_{it}^W), and an idiosyncratic error term (ε_{it}^W). The terms μ_t^W and ν_{it}^W represent permanent and time-varying correlated unobserved heterogeneity.

The total unobserved heterogeneity for the wage equation is given by:

$$\epsilon_{it}^{W} = \mu_t^W + \nu_{it}^W + \varepsilon_{it}^W$$

2.3.2 Simultaneous choices

Recall that upon observing a wage offer w_{it}^*, individuals make choices with respect to employment, e_{it}, occupation, o_{it}, retirement portfolio, p_{it}, and optional savings, s_{it}.

The econometrician does not observe these wage offers, only the wage the individual receives if she decides to be employed, w_{it}. Consequently, there is potential selection bias that may arise if there are individual unobserved characteristics that simultaneously affect selection into employment, selection into occupation, and wages. To prevent the bias stemming from selection, I jointly estimate employment status, occupation category, and wages in period t by allowing correlation of the error terms that affect each equation.

Moreover, if unobserved characteristics are correlated across savings choices, retirement portfolio choices, and other behaviors, then even after correction for selection into employment and occupation, potential biases arise. It is expected that savings choices and retirement portfolios are endogenously related to earnings, as higher earnings may generate increased availability of resources to allocate to savings. Similarly, higher earnings levels generate higher contribution levels and higher levels of accumulated retirement wealth, which may impact investment choices. In addition, all decisions that impact wealth accumulation are jointly made. Thus, one can expect selection into employment to be endogenous to savings and portfolio choices.

Taking this into account, all decisions are jointly modeled. Because choices are simultaneously made with the information available at time t, all choices are a function of the same observed variables contained in the information set $\Omega_t = \{\tilde{\Omega}_{it}, X_t, Z_t\}$. Choices are also a function of permanent and time-varying correlated unobserved heterogeneity specific to each equation. As a result, in addition to the standard idiosyncratic error term, assumed to be type-I extreme value distributed (discrete outcomes), the total unobserved heterogeneity in each equation has a permanent component and a time-varying component that correlate through the equations.

Wealth accumulation choices (employment, occupation, savings and retirement portfolio) are presented in equations 2 to 5.

$$\ln\left[\frac{e_{it}=j}{e_{it}=0}\right] = e^j(\tilde{\Omega}_{it}, X_t, Z_t, \mu_i^E, \nu_{it}^E),\ j=\{1,2\} \tag{2}$$

$$\ln\left[\frac{o_{it}=j}{o_{it}=1}\right] = o^j(\tilde{\Omega}_{it}, X_t, Z_t, \mu_i^O, \nu_{it}^O),\ j=\{2,...,6\} \tag{3}$$

$$\ln\left[\frac{p_{it}^j=1}{p_{it}^j=0}\right] = p^j(\tilde{\Omega}_{it}, X_t, Z_t, \mu_i^{Pj}, \nu_{it}^{Pj}),\ j=\{A,B,C,D,E\} \tag{4}$$

$$\ln\left[\frac{s_{it}=1}{s_{it}=0}\right] = s(\tilde{\Omega}_{it}, X_t, Z_t, \mu_i^S, \nu_{it}^S) \tag{5}$$

where $\tilde{\Omega}_{it} = \{p_{i,t-1}, s_{i,t-1}, A_{it}, E_{it}, M_{it}, N_{it}, H_{it}\}$, and μ_i^z and ν_{it}^z represent the correlated permanent and time-varying individual observed heterogeneity for each equation z.

The per-period alternatives are defined according to the data availability: $e_{it} = \{0,1,2\}$ indicate unemployed, working part-time, and working full-time, respectively; $o_{it} = \{1,2,...,6\}$ indicate occupation categories (elementary occupations; legislators, senior officials and managers, professionals, technicians and associate professionals; clerical support workers; service and sales workers; skilled agricultural, forestry and fishery workers and craft and related trade workers; and plant and machine operators and assemblers); $p_{it}^j = \{0,1\}$, $j=\{A,B,C,D,E\}$ indicate no investment or investment in Account j; and $s_{it} = \{0,1\}$, indicate no optional savings or some optional savings.

Once these decisions are made, retirement wealth accumulates according to exogenous market returns $R_{i,t-1}$, which is portfolio specific, and to new contributions as follows:

$$A_{it+1} = A_{it} \cdot R_{it}(p_{it}) + 0.1 w_{it}$$

2.3.3 Simultaneous assessments

Recall that the theoretical benchmark is a life-cycle model of individual behavior. In a typical life-cycle model, rational individuals maximize their discounted expected lifetime utility, usually subject to time and budget constraints. Because the lifetime utility function depends on the static utility function in period t and future utilities, choices also depend on risk aversion.[5] Consequently, behaviors are a function of characteristics such as risk aversion. Importantly, because I am estimating reduced-form parameters, all estimated values are a function of the primitive (structural parameters) of the model.

One can expect more risk-averse individuals to select safer occupations if they decide to be employed (Bonin et al., 2007; Grazier and Sloane, 2008; Pollmann et al., 2013) or to invest in safer retirement accounts when choosing their portfolio (Bernasek and Shwiff, 2001; Arano et al., 2010; Neelakantan, 2010). Similarly, risk-averse individuals are expected to have higher levels of medical care utilization, while health shocks might affect risk aversion (Decker and Schmitz, 2016). It is also expected that risk-averse individuals are more conservative in family characteristics such as the timing of marriage (Eisenhauer and Ventura, 2003; Spivey, 2010; Light and Ahn, 2010) or number of children (Schmidt, 2008; Görlitz and Tamm, 2015).

This paper models behavior when individuals deal with occupational and financial risks and other potential sources of income variation (e.g., health shocks and variation in family structure). Thus, omitting information on risk preferences might generate inconsistent estimates. Because risk aversion depends on the curvature of the utility function, it is unobserved for the econometrician. Nevertheless, I include longitudinal survey measures of elicited risk aversion where individuals report preferences toward hypothetical lotteries of lifetime income, based on Barsky et al. (1997).

Elicited risk aversion (r_{it}) can take one of three values representing the level of risk aversion (most risk averse, intermediate, or least risk averse). This categorization follows the maximum number of categories given the number of survey questions (two) and alternative scenarios per question (two).

As before, it is expected that unobserved characteristics that affect risk aversion also affect individual choices. Hence, I allow elicited risk aversion to also depend on

[5] In the classic conceptualization of Pratt (1964) and Arrow (1965), risk aversion depends on the curvature of the static utility function.

correlated unobservables (a permanent component μ_i^R and a time-varying component ν_{it}^R). The inclusion of this correlated error also addresses potential estimation bias due to measurement error in the quantification of risk aversion. Because risk aversion depends on the inputs per period and the discounted expected future utility, I assume that elicited risk aversion is realized jointly with the modeled behaviors and is a function of the variables in the information set (Ω_{it}) and an idiosyncratic shock (ε_{it}^R) assumed to be type-I extreme value distributed.

In modern considerations, risk aversion throughout the life cycle is simultaneously determine by the curvature of the lifetime utility function and the horizon length that defines the dynamic problem (Gollier and Zeckhauser, 2002; Bommier and Rochet, 2006). For example, individuals that expect to live longer or retire later might be more aggressive in their portfolios. To capture this effect and to avoid bias due to omitted variables, I incorporate into the set of correlated equations a self-reported measure of expected life duration in years (T_{it}) that depends on observables (Ω_{it}), correlated unobservables (μ_t^T and $\nu^T it$), and an idiosyncratic shock (ε_{it}^T) assumed to be normally distributed.

Hence, elicited risk aversion and expected duration of life take the following forms:

$$\ln\left[\frac{\Pr(r_{it}=j)}{\Pr(r_{it}=1)}\right] = r^j(\tilde{\Omega}_{it}, X_t, Z_t, \mu_i^R, \nu_{it}^R),\ \ j=\{2,3\} \tag{6}$$

$$T_{it} = T(\tilde{\Omega}_{it}, X_t, Z_t, \mu_i^T, \nu_{it}^T, \varepsilon_{it}^T) \tag{7}$$

2.3.4 Family outcomes

The model assumes that family outcomes, marital status ($m_{i,t+1}$) and variation in the number of children ($n_{i,t+1}$) are realized after the individual choices are made. It it important to keep in mind that this modeling assumption does not imply that the transitioning of family outcomes is an exogenous process. Indeed, because it is expected that unobserved characteristics affect both choices and family outcomes (e.g., an individual may jointly decide her employment status and whether to have children that period), I allow correlation between family outcomes and the rest of the equations of the model through an equation-specific permanent component and a time-varying unobserved component.

The probability of transition from being single in period $t+1$ ($m_{i,t+1}=0$) to being married ($m_{i,t+1}=1$) is given in equation 8 and depends on period t choices such as employment status (e_{it}), predetermined state variables ($\tilde{\Omega}_{it}$), and exogenous individual characteristics (X_{it}). Although not modeled explicitly, I assume that there is a marriage market such that supply-side factors (Z_{it}^M) also affect marriage probability. As before, there is a permanent (μ_i^M) error term, a correlated time-varying (ν_{it}^M) error term and an idiosyncratic shock (ε_{it}^M) assumed to be type-I extreme value distributed.

$$\ln\left[\frac{m_{i,t+1}=1}{m_{i,t+1}=0}\right] = m(e_{it}, \tilde{\Omega}_{it}, X_{it}, Z_{it}^M, mu_i^M, \nu_{it}^M) \tag{8}$$

Each period, an individual's number of children in the household might increase, decrease, or remain constant. The probability of transitioning to a change in the number of children in period $t+1$ ($n_{i,t+1}=\{-1,1\}$) relative to not transitioning ($n_{i,t+1}=0$) depends on that period's employment decisions (e_{it}), predetermined endogenous choices ($\tilde{\Omega}_{it}$), individual exogenous characteristics (X_{it}), supply-side factors specific to the children market (Z_{it}^N) such as education prices, correlated permanent (μ_i^N) and time-varying (ν_{it}^N) unobserved heterogeneity, and a random shock that is type-I extreme value distributed (ε_{it}^N). The transition probability is presented in equation 9.

$$\ln\left[\frac{n_{i,t+1}=j}{n_{i,t+1}=0}\right] = n^j(e_{it}, \tilde{\Omega}_{it}, X_{it}, Z_{it}^N, mu_i^N, \nu_{it}^N),\ j=\{-1,1\} \tag{9}$$

2.3.5 Medical care utilization

Medical care utilization is measured as the number of medical visits. Medical care utilization in period $t+1$ depends on previous health status (H_{it}), individual exogenous characteristics (X_{it}) such as gender, age, and schooling, and health market supply-side factors (Z_{it}^H) such as the number of hospital beds and medical doctors available in each geographical region of an individual's residence.

It is likely that medical care demand is correlated with individual-level unobserved characteristics, such as health insurance plans or employers' health benefits, which at the same time are endogenously related to employment and occupation selection and to risk aversion and the expected duration of life. Thus, medical care utilization also depends on permanent (μ_t^K) and time-varying (ν_{it}^K) individual correlated unobserved heterogeneity. An idiosyncratic shock as usual is included (ε_{it}^K). The density function is presented in equation 10.

$$k_{i,t+1} = k(H_{it}, X_{it}, Z_{it}^{H}, \mu_{t}^{K}, \nu_{it}^{K}, \varepsilon_{it}^{K}) \quad (10)$$

2.3.6 Health outcome

The final equation in the empirical model is the health status equation. Health is an outcome that is produced at the end of each period following a specification of Grossman (1972). The transition of health into the next period is denoted by $H_{i,t+1}$ and depends on individual choices in period t and other characteristics.

In particular, health status transitioning into the next period depends on current health status ($H_{i,t+1}$) and medical care consumption (k_{it}), which represents medical care inputs. The period t employment and occupation choices (e_{it} and o_{it}), in addition to other individual-level exogenous characteristics (X_{it}), also affect health transitions. Note that employment behavior may directly affect health or may proxy for omitted non-medical care inputs, such as nutrition and exercise. Market-level geographical characteristics (Z_{it}^{H}), such as rainfall, also impact health status. Inches of rainfall proxy for environmental and climatological conditions in the region of residence of individual i that may impact an individual's health status.

The following equation represents the probability of being in health status j in period $t+1$, denoted $H_{i,t+1} = j$, where $j = \{2, 3, 4\}$ represent the categories good, regular, and poor, respectively, relative to being in a very good health status ($H_{i,t+1} = 1$):

$$\ln\left[\frac{\Pr(H_{i,t+1} = j)}{\Pr(H_{i,t+1} = 1)}\right] = H^{j}(H_{it}, k_{it}, e_{it}, o_{it}, X_{it}, Z_{it}^{H}, \mu_{i}^{H}, \nu_{it}^{H}),\ \ j = \{2, 3, 4\} \quad (11)$$

Health status is likely correlated with permanent and time-varying unobservables that impact other health-related choices, such as medical care utilization, or assessments, such as risk aversion and expected duration of life. Therefore, the error structure is the same as in previous equations in that it allows for unobserved constant and time-varying correlated unobserved heterogeneity specific to equation 11 (μ_{i}^{H} and ν_{it}^{H}, respectively).

3 Data source and research sample

The first 4 waves of the Chilean Survey of Social Protection (EPS for *Encuesta de Protección Social*) corresponding to the years 2002, 2004, 2006, and 2009 are used. The

EPS is administered by the Ministry of Labor and Social Security in Chile jointly with the University of Chile and the Institute for Social Research from the University of Michigan. Its design, implementation, and partial funding were provided by the University of Pennsylvania. This data set is linked to administrative records on retirement wealth accumulation of the Chilean Superintendence of Pensions.

The research sample consists of a panel of 7,168 individuals observed four times (28,672 person-year observations). All individuals were aged between 25 and 59 years old (limits included) in 2002 and were observed in all four waves (no attrition or deaths) with no missing information for health status, optional savings, work experience, marital status, and region of residence. The age limit is introduced to focus on decisions of the working life cycle. The other restrictions are introduced since a dynamic panel is required for the estimation.

Except for an over-representation of the first educational category (less than high school), no major differences in demographic characteristics are observed between the full sample and the research sample, discarding selection issues in the research sample (see Table 1).

The definition of variables is presented in the online Appendix. The summary statistics for the dependent variables in the research sample are presented in Table 2. Most individuals are employed full time, invest in conservative retirement funds, do not hold voluntary savings, and are observed to be risk averse. Individuals also tend to be married, have no change in their number of children, and be in good health.

Table 1: Comparison between full and research samples (2002)

Variable	Full Sample		Research Sample	
	Mean	Std. Dev.	Mean	Std. Dev.
Age	40.633	9.461	40.715	9.275
Female	0.497	0.500	0.462	0.499
Education category				
Less Than High School	0.413	0.492	0.531	0.499
High School	0.259	0.438	0.285	0.452
Technical College	0.104	0.305	0.109	0.311
College or Post-College	0.067	0.250	0.065	0.247
Missing	0.158	0.365	0.010	0.098

Table 2: Summary statistics of dependent variables for research sample

Variable	Estimator	Mean	Std. Dev.	Min.	Max.	N
Employment (e_{it})	mlogit					21,504
Full-time employed		0.690	0.462	0	1	
Part-time employed		0.031	0.174	0	1	
Not working		0.278	0.448	0	1	
Occupation (o_{it}) (if working)	mlogit					15,327
Elementary occupations		0.219	0.414	0	1	
Legis., prof., tech., other		0.185	0.388	0	1	
Clerical support workers		0.107	0.309	0	1	
Service and sales workers		0.147	0.354	0	1	
Agricultural, craft and trade		0.057	0.231	0	1	
Operators and assemblers		0.286	0.452	0	1	
Retirement portfolio (p_{it})	logit					21,504
Account A (riskiest)		0.104	0.305	0	1	
Account B		0.231	0.422	0	1	
Account C		0.495	0.500	0	1	
Account D		0.215	0.411	0	1	
Account E (safest)		0.037	0.189	0	1	
Savings outcomes (s_{it})	logit					21,490
Any optional savings		0.263	0.441	0	1	
Expected duration of life (T_{it})	OLS	75.780	10.091	30	110	17,287
Elicited risk aversion (r_{it})	mlogit					20,557
Most risk averse		0.747	0.435	0	1	
Intermediate risk averse		0.076	0.265	0	1	
Least risk averse		0.177	0.381	0	1	
Log of wages (w_{it})	OLS	0.657	1.440	-10.219	5.255	14,705
Marital status ($m_{i,t+1}$)	logit					21,504
Married		0.571	0.495	0	1	
Variation in number of children ($n_{i,t+1}$)	mlogit					21,060
No change		0.788	0.408	0	1	
Decrease		0.184	0.387	0	1	
Increase		0.028	0.165	0	1	
Medical consumption ($k_{i,t+1}$)	OLS					21,438
Number of medical visits		6.697	12.639	0	240	
Health status ($H_{i,t+1}$)	mlogit					14,336
Very good		0.147	0.354	0	1	
Good		0.519	0.500	0	1	
Regular		0.266	0.442	0	1	
Poor		0.068	0.252	0	1	

4 Estimation

4.1 Error structure

Recall that each equation of the set of jointly estimated reduced-form equations is correlated with the other equations of the system through individual unobserved heterogeneity. The error in each equation z has the form:

$$\epsilon_{it}^{z} = \mu_{i}^{z} + \nu_{it}^{z} + \varepsilon_{it}^{z} \tag{12}$$

where $z = \{W, E, O, P^A, P^B, P^C, P^D, P^E, S, R, T, M, N, K, H\}$ represent equations 1 to 11, ϵ_{it}^{z} corresponds to the total unobserved heterogeneity, μ_{i}^{z} and ν_{it}^{z} represent the correlated permanent and time-varying unobserved heterogeneity (CUH), and ε_{it}^{z} is an idiosyncratic shock, independent and identically distributed (iid). ε_{it}^{z} is assumed to be type-1 extreme value distributed for discrete dependent variables and normally distributed for continuous variables. The correlation across equations comes from the components μ_{i}^{z} and ν_{it}^{z}.

I jointly estimate the equations using a semi-parametric full information maximum likelihood (FIML) procedure known as the Discrete Factor Method (DFM) (Guilkey and Lance, 2014). This method was first proposed in Heckman and Singer (1984) for single equations and then extended to multiple non-linear equations models in Mroz and Guilkey (1992) and Mroz (1999). Typically, FIML methods require assumptions about the distribution of μ_{i}^{z} and ν_{it}^{z} (e.g., multivariate normality) (Morales et al., 2016). Nevertheless, the DFM does not impose any distributional error assumption on the correlated terms μ_{i}^{z} and ν_{it}^{z}. Rather, the cumulative distribution function of the unobserved heterogeneity is approximated by a step function. Empirically, I determine the number of mass points, the mass point locations, and their probabilities jointly with the estimated parameters of the observed heterogeneity.

With DFM estimation, I account for biases stemming from selection, endogeneity and measurement error. The method performs as well as maximum likelihood estimation assuming normality when the true distribution of the error term is jointly normal and better in terms of both precision and bias compared to when the distribution is not normal (Mroz, 1999; Guilkey and Lance, 2014). Since the econometrician does not know the true distribution of the unobserved heterogeneity, the method accounts for estimation bias in a flexible way.

4.2 Initial conditions

Individuals are observed for the first time in 2002 when some of the state variables that explain endogenous choices are non-zero. For example, some individuals are first observed when work experience is greater than zero. Because I cannot use a dynamic specification to estimate the initially observed variation, I use a static specification to model initial conditions: initial employment status (E_{i1}), initial work experience (EX_{i1}), initial occupation category (O_{i1}), initial savings decision (S_{i1}), initial marital status (M_{i1}), initial number of children (N_{i1}), and initial health status (H_{i1}).

All initial conditions are modeled as a function of exogenous individual and market characteristics. These equations are jointly estimated with the other equations of the system by allowing for correlation through permanent unobserved heterogeneity. Thus, in each initial condition equation, $l = \{E^I, EX^I, O^I, S^I, M^I, N^I, H^I\}$, total unobserved heterogeneity takes the form $\epsilon_{it}^l = \mu_i^l + \varepsilon_{it}^l$, where μ_i^l is a permanent individual component and ε_{it}^l is an *iid* shock assumed to be type-1 extreme value distributed for discrete dependent variables and normally distributed for continuous variables.

Exogenous individual characteristics for initial employment status, initial work experience, and initial occupation category include age, gender, education, parents' years of schooling, interaction terms between gender and parents' education, and self-reported socioeconomic status of the household when growing up. Market characteristics include the vector $Z_I = (Z_I^E, Z_I^M, Z_I^N, Z_I^K, Z_I^H)$.

4.3 Likelihood function

The individual and per-period-specific contribution to the likelihood function, conditional on the values of the unobserved heterogeneity parameters μ_i and ν_{it} is:

$$\begin{aligned} L(\Theta|\mu_i, \nu_{it}) =& f_w(\epsilon_t^W|\mu_i, \nu_{it}) \cdot f_k(\epsilon_t^K|\mu_i, \nu_{it}) \\ & \cdot \prod_j^J \left\{ \Pr\left(I(d_t^j = d^j|\mu_i, \nu_{it}) \right) \cdot f_j(\epsilon_t^j|\mu_i, \nu_{it}) \right\}^{I(d_t^j = d^j)} \end{aligned} \tag{13}$$

where Θ is the entire set of estimated parameters, $f(\cdot)$ represents the density function of the error term of each equation, $\Pr(\cdot)$ is the cumulative distribution function for each choice, d_t^j represents a choice $j = \{E, O, P, S, T, R, M, N, H\}$, and $I(d_t^j = d^j)$ is an indicator function for each particular choice.

After integrating out the unobserved heterogeneity, the unconditional likelihood is:

$$L(\Theta) = \prod_{i=1}^{N} \left\{ \sum_{q}^{Q} \pi_q \left[\prod_{t=1}^{T} \sum_{g}^{G} \psi_g L(\Theta|\mu_i, \nu_{it}) \right] \right\} \tag{14}$$

where the probability weights π_q and ψ_g represent mass point-specific estimates of the joint probability of the permanent and time-varying heterogeneity, respectively, when there are Q mass points for the permanent component and G mass points for the time-varying component. The mass points are determined empirically. They are estimated together with their weights and the other parameters of the model.

4.4 Identification

Identification of the set of equations comes from several sources. First, identification comes from the non-linearity of the system of equations (Morales et al., 2016). Second, identification comes from the dynamic nature of the model (Arellano and Bond, 1991). In particular, each lagged predetermined variable (exogenous at time t) serves as an instrument for identification. With three waves for the estimation and one wave for initial conditions, there are enough instruments for the identification of the system.

Third, identification comes from the timing assumptions regarding the decision-making process, which result in theoretically justified restrictions (Morales et al., 2016; Gilleskie et al., 2017). Some of the market-level characteristics that explain choices in period t are excluded from the outcome equations in the same period, serving as exclusion restrictions. For example, conditional on the chosen behaviors in period t only characteristics of the marriage market (Z_{it}^M) and of the family market (Z_{it}^N) affect marital status and children variation, respectively.

Fourth, identification comes from the functional form assumptions regarding the distribution of the idiosyncratic *iid* component of the error term in each equation, ε_{it}^l and ε_{it}^z. Finally, identification comes from the restriction on the number of factors allowed for approximating the distribution of correlated individual unobserved heterogeneity (two, one permanent and one time varying).

5 Results

5.1 Estimation results

Recall that the multiple equation reduced-form life-cycle model consists of 22 jointly estimated equations (equations 1 to 11 plus initial conditions). The empirical specification of each of these equations is presented Table 3. Note that the table also shows sources of identification based on the theoretical exclusion restrictions and correlation across equations and over time. Determinants can be classified as predetermined endogenous variables, exogenous variables, and unobserved heterogeneity. Explanatory variables may enter linearly, as higher order moments, and as interaction variables. Predetermined variables capture the dynamic relationships between outcomes over time.

A total of 1,079 reduced-form coefficients are estimated. Note that 974 coefficients capture observed heterogeneity, while the rest capture the distribution of correlated unobserved heterogeneity. The number of mass points and their probability weights are empirically determined. I find that four permanent and four time-varying mass points are sufficient to capture the distribution of correlated unobserved heterogeneity.

Rather than focusing on the estimated coefficients and their contemporaneous marginal effects, this paper focuses on simulated long-run effects of family and health characteristics. For this, I use the estimated dynamic decision model to simulate the effects of family and health characteristics considering the sequential process by which individuals accumulate wealth over time. Estimated results are available in the online appendix.

5.2 Model fit

Since the estimated coefficients are used to simulate counterfactual scenarios, we need the model to fit the data well. Hence, I compare observed and simulated behavior using the model as a data generating process (see Table 4). Behaviors are simulated from the first period forward using each period's values to update the endogenous variables that explain subsequent behaviors. No statistical difference is found between simulated and observed data at the 1% significance level.

Table 3: Empirical specification of the correlated set of equations

Equation	Explanatory Variables		
	Predetermined Variables	Exogenous Variables	Unobserved Heterogeneity
Employment (e_{it})	$p_{i,t-1}$, $s_{i,t-1}$, A_{it}, E_{it}, M_{it}, N_{it}, H_{it}	X_{it}, $Z_{it}^E, Z_{it}^M, Z_{it}^N, Z_{it}^K, Z_{it}^H$	μ_i^E, ν_{it}^E, ε_{it}^E
Occupation (o_{it})	$p_{i,t-1}$, $s_{i,-1}$, A_{it}, E_{it}, M_{it}, N_{it}, H_{it}	X_{it}, $Z_{it}^E, Z_{it}^M, Z_{it}^N, Z_{it}^K, Z_{it}^H$	μ_i^O, ν_{it}^O, ε_{it}^O
Savings (s_{it})	$p_{i,t-1}$, $s_{i,t-1}$,A_{it}, E_{it}, M_{it}, N_{it}, H_{it}	X_{it}, $Z_{it}^E, Z_{it}^M, Z_{it}^N, Z_{it}^K, Z_{it}^H$	μ_i^S, ν_{it}^S, ε_{it}^S
Account A (p_{it}^A)	$p_{i,t-1}$, $s_{i,t-1}$, A_t, E_{it}, M_{it}, N_{it}, H_{it}	X_{it}, $Z_{it}^E, Z_{it}^M, Z_{it}^N, Z_{it}^K, Z_{it}^H$	$\mu_i^{P^A}$, $\nu_{it}^{P^A}$, $\varepsilon_{it}^{P^A}$
Account B (p_{it}^B)	$p_{i,t-1}$, $s_{i,t-1}$, A_{it}, E_{it}, M_{it}, N_{it}, H_{it}	X_{it}, $Z_{it}^E, Z_{it}^M, Z_{it}^N, Z_{it}^K, Z_{it}^H$	$\mu_i^{P^B}$, $\nu_{it}^{P^B}$, $\varepsilon_{it}^{P^B}$
Account C (p_{it}^C)	$p_{i,t-1}$, $s_{i,t-1}$, A_{it}^r, E_{it}, M_{it}, N_{it}, H_{it}	X_{it}, $Z_{it}^E, Z_{it}^M, Z_{it}^N, Z_{it}^K, Z_{it}^H$	$\mu_i^{P^C}$, $\nu_{it}^{P^C}$, $\varepsilon_{it}^{P^C}$
Account D (p_{it}^D)	$p_{i,t-1}$, $s_{i,t-1}$, A_{it}, E_{it}, M_{it}, N_{it}, H_{it}	X_{it}, $Z_{it}^E, Z_{it}^M, Z_{it}^N, Z_{it}^K, Z_{it}^H$	$\mu_i^{P^D}$, $\nu_{it}^{P^D}$, $\varepsilon_{it}^{P^D}$
Account E (p_{it}^E)	$p_{i,t-1}$, $s_{i,t-1}$, A_{it}, E_{it}, M_{it}, N_{it}, H_{it}	X_{it}, $Z_{it}^E, Z_{it}^M, Z_{it}^N, Z_{it}^K, Z_{it}^H$	$\mu_i^{P^E}$, $\nu_{it}^{P^E}$, $\varepsilon_{it}^{P^E}$
Duration of life (T_{it})	$p_{i,t-1}$, $s_{i,t-1}$, A_{it}, E_{it}, M_{it}, N_{it}, H_{it}	X_{it}, $Z_{it}^E, Z_{it}^M, Z_{it}^N, Z_{it}^K, Z_{it}^H$	μ_i^T, ν_{it}^T, ε_{it}^T
Risk aversion (r_{it})	$p_{i,t-1}$, $s_{i,t-1}$, A_{it}, E_{it}, M_{it}, N_{it}, H_{it}	X_{it}, $Z_{it}^E, Z_{it}^M, Z_{it}^N, Z_{it}^K, Z_{it}^H$	μ_i^R, ν_{it}^R, ε_{it}^R
Log wages ($w_{it} \mid e_{it}, o_{it}$)	E_{it}, H_{it}	X_{it}, Z_{it}^E	μ_i^W, ν_{it}^W, ε_{it}^W
Medical care (k_{it})	H_{it}	X_{it}, Z_{it}^K	μ_i^K, ν_{it}^K, ε_{it}^K
Marital status ($m_{i,t+1}$)	e_{it},M_{it}, N_{it}	X_{it}, Z_{it}^M	μ_i^M, ν_{it}^M, ε_{it}^M
Children in hh ($n_{i,t+1}$)	e_{it}, M_{it}, N_{it}	X_{it}, Z_{it}^N	μ_i^N, ν_{it}^N, ε_{it}^N
Health status ($H_{i,t+1}$)	e_{it}, o_{it}, k_{it}, E_{it}, H_{it}	X_{it}, Z_{it}^H	μ_i^H, ν_{it}^H, ε_{it}^H
Initial conditions ($t = 1$)			
Employment (e_{i1})		X_{i1}, $Z_{i1}^E, Z_{i1}^M, Z_{i1}^N, Z_{i1}^K, Z_{i1}^H$	$\mu_i^{E_I}$, $\varepsilon_{i1}^{E_I}$
Work experience (E_{i1})		X_{i1}, $Z_{i1}^E, Z_{i1}^M, Z_{i1}^N, Z_{i1}^K, Z_{i1}^H$	$\mu_i^{EX_I}$, $\varepsilon_{i1}^{EX_I}$
Occupation (o_{i1})		X_1, $Z_{i1}^E, Z_{i1}^M, Z_{i1}^N, Z_{i1}^K, Z_{i1}^H$	$\mu_i^{O_I}$, $\varepsilon_{i1}^{O_I}$
Savings (s_{i1})		X_{i1}, $Z_{i1}^E, Z_{i1}^M, Z_{i1}^N, Z_{i1}^K, Z_{i1}^H$	$\mu_{i_{i1}}^{S_I}$, $\varepsilon_{i1}^{S_I}$
Marital status (m_{i1})		X_{i1}, Z_{i1}^M	$\mu_i^{M_I}$, $\varepsilon_{i1}^{M_I}$
Number of children (n_{i1})		X_{i1}, Z_{i1}^N	$\mu_i^{N_I}$, $\varepsilon_{i1}^{N_I}$
Health status (H_{i1})		X_{i1}, Z_{i1}^K, Z_{i1}^H	$\mu_i^{H_I}$, $\varepsilon_{i1}^{H_I}$

Note: (a) All equations are correlated through unobserved heterogeneity separated into μ_i and ν_{it}. Total unobserved heterogeneity includes an idiosyncratic independent and identically distributed error term ε_{it}. (b) The distributions of μ_i and ν_{it} are estimated simultaneously with the reduced-form parameters of the model.

Table 4: Model fit

Outcome	Observed Mean	Simulated Mean
Employment		
Full-time employed	0.690	0.695
Part-time employed	0.031	0.033
Not working	0.278	0.272
Occupation		
Elementary occupations	0.219	0.248
Legis., prof., tech., other	0.185	0.174
Clerical support workers	0.107	0.096
Service and sales workers	0.147	0.144
Agricultural, craft and trade	0.057	0.069
Operators and assemblers.	0.286	0.270
Retirement portfolio		
Account A (riskiest)	0.104	0.104
Account B	0.231	0.223
Account C	0.495	0.512
Account D	0.215	0.207
Account E (safest)	0.037	0.038
Optional savings	0.263	0.262
Expected duration of life	75.780	75.775
Elicited risk aversion		
Most risk averse	0.747	0.747
Intermediate risk averse	0.076	0.076
Least risk averse	0.177	0.176
Log of wages	0.657	0.534
Marital status (married)	0.571	0.575
Variation in number of children		
No change	0.788	0.784
Decrease	0.184	0.184
Increase	0.028	0.032
Medical consumption (medical visits)	6.697	6.681
Health status		
Very good	0.147	0.145
Good	0.519	0.521
Regular	0.266	0.268
Poor	0.068	0.066

Note: (a) Values simulated using observed explanatory variables, with no updating of current endogenous behaviors and with 100 replications for the types of probabilities.

5.3 Simulation results

I now simulate the effect of counterfactual scenarios to evaluate gains and losses due to family and health characteristics, together with gender differences for each case. For each scenario, I compute changes in retirement wealth at the end of seven years. These changes are computed across the wealth distribution.

Each individual is replicated 100 times, allowing draws from the unobserved heterogeneity distribution. Every individual enters the first period with their initially observed characteristics except when otherwise specified. As a baseline for comparison, unless other specified, I use simulated outcomes updated according to the estimated model. Bootstrapped standard errors are computed using 100 draws.

I use a yearly model, assuming that individuals save 10% of their annual wage and that wealth is accumulated at the annualized mean real rate of return on investment funds for the period of analysis (October of 2002 to December of 2009).

5.3.1 Accumulated wealth distribution

Table 5 presents the distribution of retirement wealth for the baseline model, and the results are expressed in dollars.[6] Note that retirement wealth is substantially different across gender. At the mean, there is a difference of more than 7,000 dollars. Note also that there is a large variation in accumulated wealth across the distribution.

There are many differences that explain these gaps. On the one hand, there is the standard gender wage gap. Indeed, this paper finds a significant gender wage gap of 19.6% that agrees with other studies for Chile for the same period of time. Using the first three waves of the EPS (2002–2006), Perticará and Bueno (2009) find a wage gap between 12.7% and 18.7%, while using administrative records on unemployment insurance between 2004 and 2009, Cruz and Rau (2017) find a wage gap of 24.5%.

Also contributing to wealth differences across genders is the role of risk preferences. Consistent with previous evidence (Grable, 2000; Grazier and Sloane, 2008; Dohmen et al., 2005, 2011; Arano et al., 2010; Le et al., 2011), I find that women are more risk averse than men and select safer retirement portfolios (all regressions are provided in the online appendix).

[6] For example, when compensations are computed for a woman's average wealth, the first row of column 2 is used.

Table 5: Distribution of retirement wealth in the last period (baseline model)

	Retirement wealth		
	All	Women	Men
	(1)	(2)	(3)
Mean	12,860	9,015	16,128
Percentile			
1%	308	183	1,027
5%	1,082	639	2,384
10%	1,849	1,118	3,295
25%	3,716	2,408	5,613
50%	7,695	4,918	10,711
75%	15,392	10,119	19,280
90%	28,669	20,580	33,372
95%	42,035	32,559	48,850
99%	81,109	60,826	93,546

Note: (a) All amounts in dollars of 2009.

5.3.2 Family characteristics

Marital status

Two scenarios are simulated (denoted I and II): I) individuals are simulated to be initially married and II) individuals are simulated to be permanently married entering the initial period with observed characteristics. For I), the percentage change in accumulated wealth is computed with respect to the baseline simulation, while II) compares the accumulated wealth of being permanently married versus being permanently single. For the design of policies, one needs to select which is the relevant counterfactual and the scenario for comparison. In the setting of marriage, it makes sense to design policies based on what individuals gain or lose due to their marital status (simulation I). However, to understand how marital status affects wealth accumulation over time, it is also important to quantify the effect of being married by itself (simulation II). Simulation results are presented in Table 6, while proposed zero-pillar amounts are presented in Table 7.

I first refer to the simulation results. Generally, on average for the entire sample, there is a positive, statistically significant effect of marriage on wealth accumulation (first row of columns 1 and 4). This result throughout the distribution is driven by men. Note that the effect varies across gender. While it is positive for men across the entire wealth distribution, the effect is negative for women in the first percentile and positive for women in the upper percentiles. No statistically significant effects are found for women in the middle of the distribution.

For men, at the mean, retirement wealth increases by 0.77% if individuals start the 7-year period married (first row of column 3) and increases by 3.37% when comparing being permanently married with being permanently single (first row of column 6). The effect of being married is substantially larger for the first deciles of the distribution. For example, at the 10 and 25% of the distribution, the effects of being permanently married on retirement wealth are 11% and 8% respectively. For women, there is no statistically significant effect at the mean (first row of columns 2 and 4), and results are substantial only for the bottom of the distribution.

While this paper does not explore the mechanisms by which marriage affects retirement wealth, the results are consistent with patterns documented in the literature.

Using a sample of women at an older age (HRS), Addo and Lichter (2013) find that married women who stay married accumulate higher wealth than those that interrupt their marriages (equivalent to scenario II). Using individual data for young individuals (aged between 22 and 35 years) in the U.S., Knoll et al. (2012) find that married individuals are more likely to participate in DC plans and that single women save less for retirement. Note from Table 6 that this gender difference only holds for the upper 10% of the distribution. Using survey data for the U.S., Honig and Dushi (2010) find that married respondents and males are more likely to enroll in retirement plans and to contribute more to their accounts.

The results show two sources of inequality. One the one hand, men accumulate higher wealth when married, generating inequality across genders. On the other hand, poor women accumulates less when married, generating inequality across the wealth distribution. Thus, I propose two policies to alleviate these issues. The first one corresponds to a lump-sum payment to compensate for gender differences computed for every percentile of the wealth distribution using both the mean and median of man's wealth (panel *i)* of Table 7). The second policy corresponds to a lump-sum payment payable to women in the bottom 10% of the wealth distribution with the objective of alleviating wealth inequality among women (panel *ii)* of Table 7). The bonuses are computed using the annualized rate of wealth differences found in Table 6 (columns 3 and 6 for gender differences and columns 2 and 4 for the extra bonus) per percentile using man's mean and median wealth for the first case and woman's mean and median wealth for the second case. Results are presented in Table 7.

To alleviate gender inequality in wealth accumulation due to marital status, I propose a bonus specific to the individual position in the wealth distribution, between 177 and 6,252 dollars, with an average of 710 dollars. For comparison purposes, I also compute the bonus at the median of men's wealth, leading to smaller compensations across the distribution. To alleviate inequality among women, I propose a bonus that varies between 1,965 and 3,912. Thus, for instance, a married women in the 10th percentile of the wealth distribution should receive a total compensation of 5,005, while a married women in the 99th percentile of the wealth distribution should receive a total compensation of 177 dollars.

One could also chose other parameters to compute the compensations (e.g., using median values) or the wealth distribution specific to married women. Policy makers could also consider a benchmark wealth level (i.e., 40 years of continued contributions) and use this wealth level as a reference for computing zero-pillar amounts. In any case, it is important to note that an advantage of this procedure is that the compensations are independent of other policy variables, such as minimum wages or adjustments that may not be consistent with inequality and poverty levels. Furthermore, compensations are based on the actual effect of the variable on life-cycle wealth accumulation.

What the policy makers need to decide are the scenarios that will be use for the computation of compensations. For this purpose, I also provide the corresponding compensations for scenario II in Table 8.

Table 6: Effect of marital status on retirement wealth (percentage change at the end of seven years)

	I Initially married			II Permanently married		
	All	Women	Men	All	Women	Men
	(1)	(2)	(3)	(4)	(5)	(6)
Mean	0.65***	0.28	0.77***	2.41***	0.71	3.37***
	(0.15)	(0.38)	(0.14)	(0.53)	(0.99)	(0.61)
Percentile						
1%	-5.29**	-7.35**	6.98***	-14.73**	-31.14***	17.95***
	(2.35)	(3.38)	(1.57)	(7.08)	(9.36)	(4.69)
5%	-2.31	-4.97**	3.99***	-5.93	-18.15***	12.46***
	(1.54)	(2.26)	(1.00)	(4.42)	(6.30)	(2.94)
10%	-0.15	-3.75**	3.35***	-0.36	-12.61**	10.59***
	(1.11)	(1.86)	(0.74)	(3.24)	(5.14)	(2.24)
25%	1.47***	-1.27	2.59***	4.55***	-4.13	7.95***
	(0.52)	(1.32)	(0.42)	(1.55)	(3.67)	(1.39)
50%	1.44***	0.17	1.30***	4.31***	0.52	4.88***
	(0.30)	(0.80)	(0.24)	(0.85)	(2.09)	(0.87)
75%	0.84***	0.64	0.59***	3.03***	1.69	2.99***
	(0.16)	(0.53)	(0.13)	(0.62)	(1.19)	(0.65)
90%	0.34***	0.61**	0.31***	1.75***	1.64**	2.05***
	(0.08)	(0.30)	(0.08)	(0.43)	(0.75)	(0.47)
95%	0.27***	0.44**	0.20***	1.50***	1.14**	1.66***
	(0.08)	(0.19)	(0.05)	(0.36)	(0.51)	(0.36)
99%	0.10	0.29	0.19***	1.25***	1.06**	1.48***
	(0.07)	(0.19)	(0.06)	(0.30)	0.48	(0.30)

Note: (a) Columns 1-3, percentage change in retirement wealth with respect to the baseline simulation. Columns 4-6, percentage change in retirement wealth when permanently married versus permanently single. (b) Permanently married starting at year 2. (c) Bootstrapped standard errors are given in parentheses, obtained using 100 draws.

* Significant at the 10 percent level; ** 5 percent level; *** 1 percent level.

Table 7: Simulated bonus to women to compensate for gender and distributional differences due to marital status (dollars)

	I Initially married				
	Annualized	Yearly bonus		40-years bonus	
	rate (%)	Mean	Median	Mean	Median
i) Gender differences					
Mean	0.11	18	12	710	472
Percentile					
1%	0.97	156	104	6,252	4,152
5%	0.56	90	60	3,612	2,399
10%	0.47	76	50	3,040	2,019
25%	0.37	59	39	2,358	1,566
50%	0.19	30	20	1,195	794
75%	0.08	14	9	542	360
90%	0.04	7	5	283	188
95%	0.03	5	3	188	125
99%	0.03	4	3	177	118
ii) Inequality among women (extra bonus for bottom 10%)					
1%	1.08	98	53	3,912	2,134
5%	0.73	65	36	2,615	1,427
10%	0.55	49	27	1,965	1,072

Note: (a) Annualized rate of retirement wealth growth computed based on Table 6 column 3 for i) and column 2 for ii). (b) The yearly bonus is the product of the annualized rate of retirement wealth and the mean (or median) wealth for men in i) and for women in ii). (c) The 40-years bonus is 40× the yearly bonus. (d) All amounts in dollars of 2009.

Table 8: Simulated bonus to women to compensate for gender and distributional differences due to marital status (dollars)

	II Permanently married				
	Annualized rate (%)	Yearly bonus		40-years bonus	
		Mean	Median	Mean	Median
i) Gender differences					
Mean	0.49	79	53	3,170	2,105
Percentile					
1%	2.87	462	307	18,492	12,280
5%	1.92	309	206	12,379	8,220
10%	1.61	260	173	10,400	6,906
25%	1.19	192	128	7,683	5,102
50%	0.72	116	77	4,630	3,075
75%	0.43	70	47	2,806	1,864
90%	0.30	48	32	1,914	1,271
95%	0.24	39	26	1,542	1,024
99%	0.21	34	23	1,378	915
ii) Inequality among women (extra bonus for bottom 10%)					
1%	3.80	342	187	13,696	7,472
5%	2.35	212	116	8,492	4,633
10%	1.68	152	83	6,066	3,309

Note: (a) Annualized rate of retirement wealth growth computed based on Table 6 column 3 for i) and column 2 for ii). (b) The yearly bonus is the product of the annualized rate of retirement wealth and the mean (or median) wealth for men in i) and for women in ii). (c) The 40-years bonus is 40× the yearly bonus. (d) All amounts in dollars of 2009.

Child effect

To evaluate the effect of parenthood on retirement wealth, I shock individuals with I) an initial extra child in the first period compared with the observed number of children and II) an initial extra child compared to no child in the initial period. The objective of comparing these two scenarios is to evaluate whether there are differences when we consider "a marginal child" or "the first child." There are no major differences between both counterfactuals, so zero-pillar compensations are calculated using only I). All simulation results are presented in Table 9 and child compensations in Table 10.

The effect of an additional child on retirement wealth is substantially different for women than for men. Almost no significant effects were found for men, except for the bottom 5% of the wealth distribution, finding a positive effect of an additional child on wealth. On the contrary, there is a significantly negative effect of an additional child for women in the bottom 75% of the distribution for scenario I and in the bottom 50% of the distribution for scenario II. The effect of a marginal child decreases (in absolute value) as one moves up in the distribution, starting at -8.55% for the bottom 1% and decreasing to -0.67% for the third quantile. The pattern is the same for the effect of the first child. Thus, in this case, there is no gender gap in the effect of a child (as no effect is found for men), and the only source of inequality comes from differences among women across the wealth distribution.

Accordingly, I compute a zero-pillar compensation using the annualized rate of wealth growth based on column 2 from Table 9. Two reference wealth levels are considered: women's mean and median wealth (see column 2 of Table 5). This compensation is only computed up to the 3rd quartile of the distribution, as no significant effects were found above this in the distribution. Computed at the mean, the zero-pillar compensation varies between 4,577 and 344 dollars per additional child.

Currently in Chile, there is a per-child compensation of 843 dollars that corresponds to a year's contribution for the minimum wage, payable to all women (i.e., regardless of her specific position on the wealth distribution).[7] With the compensations proposed in this paper, poorer women will be better off, while women above the mean will be worse off. Nevertheless, the policy proposed here helps alleviate inequality among elderly women.

[7] Amount obtained from the Chilean Superintendence of Pensions website.

Table 9: Child effect on retirement wealth (percentage change at the end of seven years)

	I One initial extra child versus observed (marginal child)			II One initial child versus no initial child (first child)		
	All (1)	Women (2)	Men (3)	All (4)	Women (5)	Men (6)
Mean	-0.12	-0.61*	0.08	-0.12	-0.63*	0.10
	(0.16)	(0.32)	(0.18)	(0.18)	(0.36)	(0.20)
Percentile						
1%	-5.43**	-8.55***	4.67***	-6.64**	-8.45**	4.61***
	(2.31)	(3.04)	(1.54)	(2.66)	(3.52)	(1.63)
5%	-3.41**	-5.99***	1.78*	-3.64**	-6.73***	2.04*
	(1.33)	(2.08)	(1.05)	(1.45)	(2.28)	(1.11)
10%	-2.10**	-4.75***	1.10	-2.12**	-5.05***	1.23
	(0.93)	(1.56)	(0.79)	(1.03)	(1.79)	(0.85)
25%	-0.60	-2.81**	0.44	-0.58	-3.03**	0.46
	(0.49)	(1.12)	(0.52)	(0.53)	(1.26)	(0.56)
50%	-0.18	-1.39**	0.13	-0.21	-1.35*	0.19
	(0.31)	(0.68)	(0.31)	(0.33)	(0.77)	(0.34)
75%	-0.04	-0.67*	-0.01	-0.06	-0.61	0.01
	0.18	0.40	0.17	0.20	0.44	0.21
90%	-0.02	-0.16	-0.02	-0.07	-0.07	-0.02
	(0.10)	(0.23)	(0.12)	(0.11)	(0.24)	(0.14)
95%	0.02	-0.06	0.05	-0.04	-0.02	-0.01
	(0.08)	(0.14)	(0.09)	(0.09)	(0.16)	(0.10)
99%	0.01	0.00	0.11	0.07	0.06	-0.07
	(0.08)	(0.13)	(0.11)	(0.08)	(0.13)	(0.11)

Note: (a) Percentage change in accumulated assets with respect to indicated case. (b) Bootstrapped standard errors are given in parentheses, obtained using 100 draws.

* Significant at the 10 percent level; ** 5 percent level; *** 1 percent level.

Table 10: Simulated bonus to women to compensate for differences in wealth due to child bearing (dollars)

	I One initial extra child versus observed (marginal child)				
	Annualized rate (%)	Yearly bonus		40-years bonus	
		Mean	Median	Mean	Median
Mean	0.09	8	4	314	171
Percentile					
1%	1.27	114	62	4,577	2,497
5%	0.88	79	43	3,169	1,729
10%	0.69	62	34	2,498	1,363
25%	0.41	37	20	1,466	800
50%	0.20	18	10	719	392
75%	0.10	9	5	344	188
90%	n/a	0	0	0	0
95%	n/a	0	0	0	0
99%	n/a	0	0	0	0

Note: (a) Annualized rate of retirement wealth growth computed based on Table 9 columns 2 and 4 when the effect in such table is statistically significant. (b) n/a = wealth difference is not statistically different from zero. (c) The yearly bonus is the product of the annualized rate of retirement wealth and the mean (or median) wealth for women. (d) The 40-years bonus is 40× the yearly bonus. (e) All amounts in dollars of 2009.

5.3.3 Health characteristics

Health status

To evaluate the effect of health status throughout the life cycle on retirement wealth, I simulate three scenarios: I) all individuals are hit with a transitory shock at the initial period, meaning that they start the seven year simulations with poor health, and are compared with the observed baseline scenario, II) all individuals start the seven year period with poor health and are compared with the reference where all individuals start with good health status, and III) individuals start the seven year period with regular health and are compared as if all individuals start with good health. All simulation results are presented in Table 11.

The general pattern of results is the same across genders, with only a few point estimates statistically different for men and women (these are indicated in Table 11). Therefore, these results show no inequality across genders (except when the difference is statistically significant) and only evidence of inequality across the wealth distribution. For simplicity, I refer to the aggregated results (columns 1, 4, and 7) and compute compensations at an aggregate level. On average, a transitory shock has a total negative effect of 2.09%. The effect decreases (in absolute value) as the position in the wealth distribution increases, ranging between 7.26% and 0.64%.

When considering the difference between individuals starting in poor and regular health versus starting in good health (simulations II and III), there is an average effect for better health of 2.35% that ranges between 12.11% and 0.65% across the wealth distribution for II) and of 1.19% that ranges between 5.92% and 0.42% for III).

Table 12 presents the aggregated compensation to individuals who suffered a health shock. Compensations are computed for the three simulations. As stated earlier in this paper, it is up to the policy makers to decide which reference scenario should be considered when calculating the compensations. I refer to scenario I), although all compensations are presented. The most straightforward computation consists of that for I, which, on average, compensates individuals for the shock considering their observed conditions. The other two scenarios are a conservative case (scenario III), which does not compensate individuals for the worst-case health shock, and a benchmark for health policies (scenario II), which compensates individuals compared to being in good health.

All compensations are computed at the mean aggregated wealth.

The compensation for suffering a health shock is payable to individuals who indeed suffer a shock. The way to compute the bonus depends on when the shock occurs. Generally, the 40 years bonus payable to individuals who face the shock at age 25 varies between 5,509 and 474 dollars, with an average of 1,548. In the benchmark scenario, the bonus varies between 8,471 and 475 dollars. If an individual suffers the health shock later in her life cycle, then the way to compute the compensations is to take the yearly bonus times the number of years between retirement age and the health shock.

Table 11: Effect of transitory health shocks (percentage change at the end of seven years)

	I Initial health shock versus observed			II Initial good health versus initial poor health			III Initial good health versus initial regular health		
	All	Women	Men	All	Women	Men	All	Women	Men
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Mean	-2.09***	-2.62***	-1.85***	2.35***	3.06***	2.05***	1.19***	1.50***	1.06***
	(0.45)	(0.51)	(0.45)	(0.51)	(0.60)	(0.50)	(0.21)	(0.25)	(0.21)
Percentile									
1%	-7.26***	-8.81***	-7.83***	12.11***	13.19***	12.96***	5.92***	5.74***	6.39***
	(1.72)	(2.28)	(1.64)	(2.64)	(3.56)	(2.63)	(1.18)	(1.62)	(1.17)
5%	-6.79***	-7.33***	-6.68***	10.30***	11.33***	9.41***	4.77***	5.38***	4.11***
	(1.15)	(1.35)	(1.43)	(1.67)	(1.99)	(2.01)	(0.74)	(0.93)	(0.84)
10%	-6.41***	-6.59***	-5.64***	9.17***	9.86***	7.40***	4.31***	4.63***	3.55***
	(1.02)	(1.10)	(1.32)	(1.41)	(1.56)	(1.76)	(0.61)	(0.71)	(0.66)
25%	-5.17***	-5.79***	-4.47***	6.66***	7.90***	5.19***	3.18***	3.76***	2.56***
	(0.94)	(0.90)	(1.04)	(1.21)	(1.22)	(1.25)	(0.48)	(0.54)	(0.46)^c
50%	-3.56***	-4.43***	-2.85***	4.05***	5.53***	3.07***	2.01***	2.63***	1.53***
	(0.72)	(0.74)	(0.65)	(0.84)	(0.93)	(0.73)^b	(0.33)	(0.39)	(0.29)^b
75%	-2.33***	-3.24***	-1.65***	2.48***	3.65***	1.84***	1.23***	1.78***	0.98***
	(0.54)	(0.64)	(0.47)^b	(0.59)	(0.74)	(0.52)^b	(0.25)	(0.31)	(0.22)^b
90%	-1.37***	-2.10***	-1.24***	1.45***	2.19***	1.26***	0.72***	1.07***	0.71***
	(0.37)	(0.50)	(0.36)	(0.40)	(0.54)	(0.39)	(0.18)	(0.25)	(0.18)
95%	-1.01***	-1.24***	-0.79***	1.08***	1.32***	0.88***	0.56***	0.75***	0.54***
	(0.33)	(0.36)	(0.29)	(0.35)	(0.39)	(0.31)	(0.16)	(0.18)	(0.14)
99%	-0.64**	-0.96***	-0.79**	0.65**	0.98***	0.88***	0.42***	0.49***	0.59***
	(0.27)	(0.32)	(0.31)	(0.29)	(0.36)	(0.34)	(0.14)	(0.18)	(0.15)

Note: (a) Percentage change in accumulated wealth with respect to the baseline in columns 1-3, to poor health in columns 4-6, and to regular health in columns 7-10. (b) Bootstrapped standard errors are given in parentheses, obtained using 100 draws.

* Significant at the 10 percent level; ** 5 percent level; *** 1 percent level.

^{a}, ^{b}, ^{c} indicate a statistically significant difference between men and women at the 1, 5 and 10 percent levels.

Table 12: Simulated bonus to all individuals to compensate for differences in wealth due to health shocks (dollars)

	I Initial health shock versus observed			II Initial good health versus initial poor health			III Initial good health versus initial regular health		
	Annual. rate	Yearly bonus	40-years bonus	Annual. rate	Yearly bonus	40-years bonus	Annual. rate	Yearly bonus	40-years bonus
Mean	0.30	39	1,548	0.33	43	1,710	0.17	22	872
Percentile									
1%	1.07	138	5,509	1.65	212	8,471	0.83	106	4,246
5%	1.00	129	5,140	1.41	181	7,257	0.67	86	3,439
10%	0.94	121	4,845	1.26	162	6,487	0.60	78	3,107
25%	0.76	97	3,885	0.93	119	4,762	0.45	58	2,308
50%	0.52	66	2,658	0.57	73	2,927	0.28	37	1,465
75%	0.34	43	1,729	0.35	45	1,800	0.17	22	897
90%	0.20	25	1,016	0.21	26	1,057	0.10	13	526
95%	0.15	19	749	0.15	20	790	0.08	10	414
99%	0.09	12	474	0.09	12	475	0.06	8	306

Note: (a) Annualized rate of retirement wealth growth computed based on Table 11 columns 1, 4, and 7. (b) All values computed at the average wealth for the entire sample. (c) The yearly bonus is the product of the annualized rate of retirement wealth and mean wealth for the entire sample. (d) The 40-years bonus is 40× the yearly bonus. (e) All amounts in dollars of 2009.

6 Conclusion

Life-time events that affect labor outcomes may negatively affect retirement resources. Using a non-standard econometric approach, this paper quantifies the effect of exogenous changes in family structure and health shocks on retirement wealth accumulation. The results show that women are most affected by shocks.

Based on this, I propose zero-pillar compensations to alleviate poverty and inequality among gender and among the distribution. These compensations consist of bonuses that are added into the individual's retirement savings accounts at the time of retirement. The mechanisms used for computing wealth gains and losses allow for calculating compensations based on individual characteristics independent of other policy variables (e.g., minimum wages). A woman located in the middle of the wealth distribution, married, with 2 children, and who suffered a health shock in the last 10 years of her working life cycle is eligible to receive a total compensation of 3,293 dollars.

These results are relevant not only for Chile but also for all countries that have implemented similar DC systems. Note that these results correct for estimation biases coming from correlated unobserved heterogeneity and do not impose any restrictions on preferences or expectation processes.

References

Addo, Fenaba R and Daniel T Lichter (2013) "Marriage, marital history, and black–white wealth differentials among older women," *Journal of Marriage and Family*, Vol. 75, No. 2, pp. 342–362.

Arano, Kathleen, Carl Parker, and Rory Terry (2010) "Gender-based risk aversion and retirement asset allocation," *Economic Inquiry*, Vol. 48, No. 1, pp. 147–155.

Arellano, Manuel and Stephen Bond (1991) "Some tests of specification for panel data: Monte Carlo evidence and an application to employment equations," *The review of economic studies*, Vol. 58, No. 2, pp. 277–297.

Arrow, Kenneth Joseph (1965) *Aspects of the theory of risk-bearing*: Yrjö Jahnssonin Säätiö.

Barr, Nicholas and Peter Diamond (2016) "Reforming pensions in Chile," *Polityka Społeczna*, Vol. 1, No. 12, pp. 4–8.

Barsky, Robert B, F Thomas Juster, Miles S Kimball, and Matthew D Shapiro (1997) "Preference parameters and behavioral heterogeneity: An experimental approach in the health and retirement study," *The Quarterly Journal of Economics*, Vol. 112, No. 2, pp. 537–579.

Behrman, Jere, Maria Cecilia Calderon, Olivia Mitchell, Javiera Vasquez, and David Bravo (2011) "First-Round Impacts of the 2008 Chilean Pension System Reform."

Bernasek, Alexandra and Stephanie Shwiff (2001) "Gender, risk, and retirement," *Journal of Economic Issues*, pp. 345–356.

Beshears, John and James J Choi (2012) "Default Stickiness among Low-Income Individuals."

Beshears, John, James J Choi, Joshua Hurwitz, David Laibson, and Brigitte C Madrian (2017) "Liquidity in Retirement Savings Systems," *Insights in the Economics of Aging*, p. 45.

Bettio, Francesca, Platon Tinios, and Gianni Betti (2013) "The gender gap in pensions in the EU," *Publications Office of the European Union Luxembourg*.

Blau, David M (2016) "Pensions, household saving, and welfare: A dynamic analysis of crowd out," *Quantitative Economics*, Vol. 7, No. 1, pp. 193–224.

Bogan, Vicki L and Angela R Fertig (2018) "Mental health and retirement savings: Confounding issues with compounding interest," *Health economics*, Vol. 27, No. 2, pp. 404–425.

Bommier, Antoine and Jean-Charles Rochet (2006) "Risk aversion and planning horizons," *Journal of the European Economic Association*, Vol. 4, No. 4, pp. 708–734.

Bonin, Holger, Thomas Dohmen, Armin Falk, David Huffman, and Uwe Sunde (2007) "Cross-sectional earnings risk and occupational sorting: The role of risk attitudes," *Labour Economics*, Vol. 14, No. 6, pp. 926–937.

Bovenberg, Lans, Casper Van Ewijk, and Ed Westerhout (2012) *The future of multi-pillar pensions*: Cambridge University Press.

Card, David and Michael Ransom (2011) "Pension plan characteristics and framing effects in employee savings behavior," *The Review of Economics and Statistics*, Vol. 93, No. 1, pp. 228–243.

Cruz, Gabriel and Tomás Rau (2017) "The Effects of Firms' Pay Policies and Equal Pay Laws on the Gender Wage Gap in Chile," *Working paper*.

Decker, Simon and Hendrik Schmitz (2016) "Health shocks and risk aversion," *Journal of health economics*, Vol. 50, pp. 156–170.

Dohmen, Thomas, Armin Falk, David Huffman, Uwe Sunde, Jürgen Schupp, and Gert G Wagner (2011) "Individual risk attitudes: Measurement, determinants, and behavioral consequences," *Journal of the European Economic Association*, Vol. 9, No. 3, pp. 522–550.

Dohmen, Thomas J, Armin Falk, David Huffman, Uwe Sunde, Jürgen Schupp, and Gert G Wagner (2005) "Individual risk attitudes: New evidence from a large, representative, experimentally-validated survey."

Eisenhauer, Joseph G and Luigi Ventura (2003) "Survey measures of risk aversion and prudence," *Applied Economics*, Vol. 35, No. 13, pp. 1477–1484.

Engelhardt, Gary V and Anil Kumar (2007) "Employer matching and 401 (k) saving: Evidence from the health and retirement study," *Journal of Public Economics*, Vol. 91, No. 10, pp. 1920–1943.

Gelber, Alexander M (2011) "How do 401 (k) s Affect Saving? Evidence from Changes in 401 (k) Eligibility," *American Economic Journal: Economic Policy*, Vol. 3, No. 4, pp. 103–22.

Gilleskie, Donna B, Euna Han, and Edward C Norton (2017) "Disentangling the contemporaneous and dynamic effects of human and health capital on wages over the life cycle," *Review of Economic Dynamics*, Vol. 25, pp. 350–383.

Gollier, Christian and Richard J Zeckhauser (2002) "Horizon length and portfolio risk," *Journal of Risk and Uncertainty*, Vol. 24, No. 3, pp. 195–212.

Görlitz, Katja and Marcus Tamm (2015) "Parenthood and risk preferences."

Grable, John E (2000) "Financial risk tolerance and additional factors that affect risk taking in everyday money matters," *Journal of Business and Psychology*, Vol. 14, No. 4, pp. 625–630.

Grazier, Suzanne and Peter Sloane (2008) "Accident risk, gender, family status and occupational choice in the UK," *Labour Economics*, Vol. 15, No. 5, pp. 938–957.

Grossman, Michael (1972) "On the concept of health capital and the demand for health," *Journal of Political economy*, Vol. 80, No. 2, pp. 223–255.

Guilkey, David K and Peter M Lance (2014) "Program impact estimation with binary outcome variables: Monte Carlo results for alternative estimators and empirical examples," in *Festschrift in Honor of Peter Schmidt*: Springer, pp. 5–46.

Heckman, James and Burton Singer (1984) "A method for minimizing the impact of distributional assumptions in econometric models for duration data," *Econometrica: Journal of the Econometric Society*, pp. 271–320.

Honig, Marjorie and Irena Dushi (2010) "Participation and contributions in tax-deferred retirement accounts: Evidence from Social Security records," *Working paper*.

Hubener, Andreas, Raimond Maurer, and Olivia S Mitchell (2015) "How family status and social security claiming options shape optimal life cycle portfolios," *The review of financial studies*, Vol. 29, No. 4, pp. 937–978.

Jefferson, Therese (2009) "Women and retirement pensions: A research review," *Feminist Economics*, Vol. 15, No. 4, pp. 115–145.

Joubert, Clement (2015) "Pension design with a large informal labor market: Evidence from Chile," *International Economic Review*, Vol. 56, No. 2, pp. 673–694.

Kahneman, Daniel and Amos Tversky (1979) "Prospect theory: An analysis of decision under risk," *Econometrica: Journal of the econometric society*, pp. 263–291.

Kim, Hyungsoo, Serah Shin, Claudia J Heath, Qun Zhang, and E Tory Higgins (2017) "Regulatory mode and willingness to increase retirement savings contributions," *Journal of Applied Social Psychology*, Vol. 47, No. 8, pp. 436–445.

Knoll, Melissa AZ, Christopher R Tamborini, and Kevin Whitman (2012) "I do... want to save: Marriage and retirement savings in young households," *Journal of Marriage and Family*, Vol. 74, No. 1, pp. 86–100.

Kritzer, Barbara E (2008) "Chile's next generation pension reform," *Soc. Sec. Bull.*, Vol. 68, p. 69.

Kuivalainen, Susan, Satu Nivalainen, Noora Järnefelt, and Kati Kuitto (forthcoming) "Length of working life and pension income: empirical evidence on gender and socioeconomic differences from Finland," *Journal of Pension Economics & Finance*, pp. 1–21.

Le, Anh T, Paul W Miller, Wendy S Slutske, and Nicholas G Martin (2011) "Attitudes towards economic risk and the gender pay gap," *Labour economics*, Vol. 18, No. 4, pp. 555–561.

Light, Audrey and Taehyun Ahn (2010) "Divorce as risky behavior," *Demography*, Vol. 47, No. 4, pp. 895–921.

Mincer model (1974) "Schooling and earnings," in *Schooling, experience, and earnings*: NBER, pp. 41–63.

Möhring, Katja (2015) "Employment histories and pension incomes in Europe: a multilevel analysis of the role of institutional factors," *European Societies*, Vol. 17, No. 1, pp. 3–26.

Morales, Leonardo Fabio, Penny Gordon-Larsen, and David Guilkey (2016) "Obesity and health-related decisions: An empirical model of the determinants of weight status across the transition from adolescence to young adulthood," *Economics & Human Biology*, Vol. 23, pp. 46–62.

Mroz, Thomas A (1999) "Discrete factor approximations in simultaneous equation models: Estimating the impact of a dummy endogenous variable on a continuous outcome," *Journal of Econometrics*, Vol. 92, No. 2, pp. 233–274.

Mroz, Thomas Alvin and David K Guilkey (1992) "Discrete factor approximation for use in simultaneous equation models with both continuous and discrete endogenous variables.," *Working paper – University of North Carolina Carolina Population Center*.

Neelakantan, Urvi (2010) "Estimation and impact of gender differences in risk tolerance," *Economic Inquiry*, Vol. 48, No. 1, pp. 228–233.

OECD (2007) *Economic Surveys: Chile 2007*: OECD Publishing.

Orszag, Peter R and Joseph E Stiglitz (2001) "Rethinking pension reform: Ten myths about social security systems," *New ideas about old age security*, pp. 17–56.

Perticará, Marcela and Ivonne Bueno (2009) "A new approach to gender wage gaps in Chile," *CEPAL Review*, No. 99.

Pollmann, Daniel, Thomas Dohmen, and Franz Palm (2013) "Robust estimation of wage dispersion with censored data: An application to occupational earnings risk and risk attitudes."

Ponthieux, Sophie and Dominique Meurs (2015) "Gender inequality," in *Handbook of income distribution*, Vol. 2: Elsevier, pp. 981–1146.

Pratt, John W (1964) "Risk aversion in the small and in the large," *Econometrica: Journal of the Econometric Society*, pp. 122–136.

Schmidt, Lucie (2008) "Risk preferences and the timing of marriage and childbearing," *Demography*, Vol. 45, No. 2, pp. 439–460.

Spivey, Christy (2010) "Desperation or desire? The role of risk aversion in marriage," *Economic Inquiry*, Vol. 48, No. 2, pp. 499–516.

World Bank (2008) "The World Bank Pension Conceptual Framework."

Online Appendix

A Data Appendix

A.1 Definition of variables

Employment category (e_{it}): 0 = unemployed, 1 = working part-time, and 2 = working full-time. Categories depend on the reported weekly hours typically worked in period t. More than 20 hours a week is considered full time.

Occupation category (o_{it}): $\{1, 2, ..., 6\}$ based on a regrouping of the 1-digit ISCO classification in period t. 1 = Elementary occupations. 2 = Legislators, senior officials and managers, professionals, technicians and associate professionals. 3 = Clerical support workers. 4 = Service and sales workers. 5 = Skilled agricultural, forestry and fishery workers and craft and related trade workers. 6 = Plant and machine operators and assemblers.

Retirement portfolio ($p_{it} = \{p_{it}^A, p_{it}^B, p_{it}^C, p_{it}^D, p_{it}^E\}$): This is a set of five funds for investing introduced in August of 2002. Each of these variables takes 1 of 2 values, $\{0, 1\}$, where 0 represents no investment in account $\{A, B, C, D, E\}$ and 1 represents investment in this account. Account A is the riskier account. An individual may chose to invest in one or in two accounts. When the individual did not report a fund, the legal default account, according to the individual's gender and age, was assigned.

Optional savings (s_{it}): Dichotomous variable that takes the value of 1 if an individual reports having any optional savings in period t and 0 otherwise.

Accumulated required assets (A_{it}): Assets accumulated in the retirement system. Monthly wages computed from administrative records using the account choice reported in EPS from 2002 onwards and Account C before. When two accounts are available, investments are accumulated using the mean return of the two accounts. In thousand of dollars of 2009.

Work experience (E_{it}): Years of labor experience since 1980.

Wage (w_{it}): Hourly wage, measured by the reported after taxes (and legal deductions) monthly wage divided by 4 times the reported weekly hours typically worked. In 2009 dollars.

Marital status (m_{it}): Takes a value of 1 if the individual reports to be married in period t and 0 otherwise.

Marital history (M_{it}): May include lagged marital state, number of marriages and cohabitations, and duration of most recent marriage state.

Changes in number of children (n_t): Takes 1 of 3 values, which represent changes in the total number of children of 18 years of age or younger in period t (total number refers to children in and outside the household). 0 = no change in the number of children. -1 = decrease in the number of children. 1 = increase in the number of children.

Children history (N_{it}): Includes birth in the last period, total number of children and ages of each child.

Number of medical visits (k_{it}): Reported number of medical visits in period t.

Health status (H_{it}): Takes 1 of 4 values, $\{1, ..., 4\}$, where 1 = very good, 2 = good, 3 = fair, and 4 = poor.

Expected duration of life (T_{it}): Reported expected duration of life in years (expected age of death) at the beginning of period t.

Elicited risk aversion (r_{it}): Takes 1 of 3 values based on the answers to hypothetical gambles, 1 being the most risk averse category and 3 the least risk averse category. Individuals are classified into a category of elicited risk aversion based on their answers to three hypothetical gambles (2 questions, 2 scenarios each). At the beginning of period t.

Other characteristics (X_{it}): Age: Age from administrative data. **Gender:** Gender from administrative data. **Education:** Education category. It takes four categories: less than high school, high school, technical college, and college and some post-college. **Region of residence:** Set of dummy variables based on the reported region of residence. Uses the old Chilean administrative division, which labels regions from 1 to 13, for 2002, 2004, and 2006. Uses the new Chilean administrative division, which labels regions from 1 to 15, for 2009. Used for geographical classification for exclusion restrictions. When region of residence is missing, region of place of work if working is used.

Other variables: Time trend: 0 in 2002, 2 in 2004, 4 in 2006, and 7 in 2009. **Market characteristics (Z_{it}):** Z_{it}^E includes unemployment rate by region of residence. Z_{it}^M includes number of marriages in a year per 1,000 people by region of residence and mean college tuition in 2009 dollars by region of residence. Z_{it}^N includes number of marriages in a year per 1,000 people by region of residence and mean college tuition in 2009 dollars by region of residence. Z_{it}^K includes number of beds available per 1,000 people by region of residence and number of medical doctors available per 1,000 people by region of residence. Z_{it}^{H} includes inches of rainfall in a year by region of residence.

A.2 Construction of Elicited Risk Aversion

The questions asked in waves 2, 3, and 4 of the EPS are as follow.[8] The first question asks: "*Suppose that you are the only income earner in the household. You need to choose between two jobs. Which option do you prefer? (Option A) a job with a lifetime-stable and certain salary or (Option B) a job where you have the same chances of doubling your lifetime income or earning only* 1/4 *of your lifetime income.*" If the answer to the question is "option A", the interviewer continues. "*Now what do you prefer? (Option A) a job with a lifetime-stable and certain salary or (Option B) a job where you have the same chances of doubling your lifetime income or earning only half of your lifetime income.*"

The least risk averse categories comes directly from question 1. Elicited risk aversion equals 3 for individuals who selected "option B" in the first question. If the individual chooses "option A" in the first question, then the index of risk aversion is constructed using the second question. Individuals who choose "option B" in the second question belong to the second category (elicited risk aversion of 2), and individuals who choose "option A" in the second question belong to the most risk averse category, as individuals assigned to this category exhibit that they are not willing to accept any gamble (elicited risk aversion equals 1).

[8] Wave 1 uses "decreasing up to 75%". Since the first wave is used to set the initial conditions, observed risk aversion from wave 1 does not enter the model. There is only one specification where initial elicited risk aversion is modeled. Although the questions are mathematically equivalent, to prevent loss aversion biases through changes in the wording (Kahneman and Tversky, 1979), the specification accounts for, among other potential sources of bias, measurement error.

B Estimation results

Table B.1: Wage equation

Variable	Wage (log)	
	Coeff.	St.Er.
Work Experience	0.006	0.003*
Experience Squared	0.000	0.000
Legislators	0.561	0.022***
Clerical	0.339	0.022***
Service and Sales	0.118	0.023***
Agricultural	-0.079	0.023***
Plant Operators	-0.042	0.021**
Health: Very Good	0.060	0.013***
Health: Fair	-0.107	0.013***
Health: Poor	-0.196	0.026***
Number of Children	0.003	0.007
Lagged Marital Status	0.092	0.011***
Age	0.001	0.001
Female	-0.196	0.013***
High School	0.257	0.012***
Technical College	0.686	0.021***
College	0.875	0.040***
Missing: Occupation	0.139	0.044***
Unemployment Rate	-0.003	0.003
Missing: Education	0.365	0.059***
Missing: Number of Children	0.000	0.031
Constant	0.572	0.039***
Permanent Unob. Het.	-0.263	0.028***
Permanent Unob. Het.	-0.411	0.024***
Permanent Unob. Het.	-0.314	0.029***
Time-varying Unob. Het.	0.039	0.014***
Time-varying Unob. Het.	-10.294	0.039***
Time-varying Unob. Het.	0.180	0.019***

*, **, *** Significant at the 10%, 5%, and 1% levels.

Table B.2: Multinomial logit on employment status (relative to working full-time)

Variable	Part-Time		Not Working	
	Coeff.	St.Er.	Coeff.	St.Er.
Work Experience	-0.065	0.021***	-0.078	0.011***
Experience Squared	0.001	0.001	-0.001	0.000***
Lagged Investment in A	-0.164	0.340	-0.077	0.098
Lagged Investment in B	-0.089	0.293	-0.100	0.081
Lagged Investment in C	-0.093	0.311	-0.100	0.079
Lagged Investment in D	-0.043	0.325	0.051	0.094
Lagged Investment in E	0.265	0.483	-0.047	0.139
Lagged Assets	-0.042	0.006***	-0.002	0.002
Lagged Optional Savings	-0.148	0.097	-0.143	0.049***
Lagged Marital Status	-0.399	0.138***	-0.249	0.069***
Number of Children	-0.052	0.075	-0.078	0.035**
Interaction Female-Married	0.519	0.174***	0.698	0.092***
Interaction Female-Children	0.140	0.085*	0.233	0.043***
Health: Very Good	-0.007	0.126	0.003	0.066
Health: Fair	0.083	0.099	0.328	0.050***
Health: Poor	0.455	0.172***	1.005	0.088***
Age	0.126	0.064**	0.162	0.029***
Age Squared	-0.044	0.033	-0.072	0.015***
Age Cubed	0.006	0.005	0.014	0.002***
Female	0.619	0.147***	0.602	0.077***
High School	-0.276	0.107***	-0.486	0.052***
Technical College	-0.221	0.168	-1.031	0.093***
College	-0.106	0.849	-1.581	0.347***
Unemployment Rate	-0.017	0.025	0.033	0.012***
Number of Hospital Beds	0.201	0.201	-0.087	0.092
Number of Doctors	1.174	0.512**	0.191	0.213
Number of Marriages	0.166	0.212	0.272	0.082***
Inches of Rainfall	0.010	0.004**	0.006	0.002***
College Tuition	0.093	0.091	-0.063	0.045
Missing: Number of Children	0.189	0.871	-0.317	0.194
Missing: Education	-0.261	0.785	-0.176	0.317
Time Trend	0.086	0.066	0.065	0.019***
Constant	-6.321	0.916***	-2.654	0.406***
Permanent Unob. Het.	-0.543	0.258**	-1.229	0.124***
Permanent Unob. Het.	0.395	0.154**	0.883	0.091***
Permanent Unob. Het.	-0.499	0.176***	-1.399	0.120***
Time-varying Unob. Het.	0.297	0.140**	0.028	0.064
Time-varying Unob. Het.	0.678	0.310**	1.637	0.409***
Time-varying Unob. Het.	0.312	0.177*	-0.146	0.095

*, **, *** Significant at the 10%, 5%, and 1% levels.

Table B.3: Multinomial logit on occupation category (relative to elementary occupation)

Variable	Prof and Tech		Clerical Support		Service and Sales		Agricul and Craft		Plant and Machine	
	Coeff.	St.Er.	Coeff.	St.Er.	Coeff.	St.Er.	Coeff.	St.Er.	Coeff.	St.Er.
Work Experience	-0.072	0.029**	-0.013	0.031	-0.058	0.024**	-0.003	0.029	-0.014	0.029
Experience Squared	0.001	0.001	0.000	0.001	0.001	0.001*	0.002	0.001**	0.000	0.001
Lagged Investment in A	-0.108	0.205	-0.078	0.200	0.000	0.209	-0.134	0.251	-0.161	0.191
Lagged Investment in B	-0.118	0.157	0.174	0.155	0.334	0.156**	-0.001	0.204	-0.083	0.150
Lagged Investment in C	-0.401	0.160**	-0.016	0.157	0.063	0.158	-0.347	0.206*	-0.245	0.147*
Lagged Investment in D	-0.241	0.220	-0.124	0.218	-0.149	0.215	-0.026	0.232	-0.196	0.196
Lagged Investment in E	-0.568	0.373	-0.310	0.386	0.348	0.316	-0.404	0.379	-0.237	0.275
Lagged Assets	0.051	0.004***	0.054	0.004***	0.041	0.004***	-0.004	0.006	0.028	0.004***
Lagged Optional Savings	0.317	0.085***	0.135	0.089	-0.007	0.089	0.118	0.104	-0.191	0.085**
Lagged Marital Status	0.401	0.174**	0.591	0.171***	0.137	0.177	0.068	0.120	0.192	0.133
Number of Children	-0.077	0.063	-0.162	0.068**	0.012	0.065	-0.111	0.053**	0.104	0.049**
Interaction Female-Married	0.134	0.259	-0.390	0.245	0.098	0.258	0.471	0.249*	-0.210	0.243
Interaction Female-Children	-0.030	0.086	0.083	0.089	-0.133	0.088	0.226	0.102**	-0.239	0.089***
Health: Very Good	0.236	0.115**	0.042	0.119	0.121	0.119	0.077	0.135	-0.092	0.111
Health: Fair	-0.265	0.115**	-0.072	0.114	-0.078	0.110	0.079	0.102	-0.067	0.098
Health: Poor	-0.151	0.446	0.042	0.406	0.121	0.372	0.076	0.227	-0.171	0.305
Age	0.010	0.025	-0.080	0.026***	-0.054	0.023**	-0.046	0.026*	0.013	0.025
Age Squared	0.002	0.006	0.010	0.006	0.011	0.005**	0.006	0.006	-0.003	0.006
Female	-0.227	0.180	0.324	0.174*	0.752	0.184***	-1.039	0.202***	-2.275	0.175***
High School	2.656	0.115***	2.778	0.118***	1.558	0.109***	-0.503	0.121***	1.075	0.105***
Technical College	6.471	0.275***	4.494	0.291***	2.771	0.269***	-0.271	0.477	1.523	0.285***
College	8.027	0.602***	5.560	0.710***	3.578	0.732***	1.209	1.048	1.302	0.867
Unemployment Rate	0.027	0.025	0.021	0.027	-0.025	0.024	0.018	0.026	0.061	0.025**
Number of Hospital Beds	-0.136	0.206	-0.199	0.223	-0.275	0.202	-0.217	0.212	-0.058	0.195
Number of Doctors	0.743	0.392*	0.446	0.467	1.156	0.412***	-1.558	0.536***	0.418	0.417
Number of Marriages	0.063	0.173	0.423	0.183**	0.323	0.164**	-0.620	0.195***	0.253	0.161
Inches of Rainfall	0.001	0.005	0.004	0.005	-0.003	0.004	0.027	0.005***	0.005	0.004
College Tuition	0.204	0.089**	0.439	0.094***	0.006	0.089	-0.765	0.102***	0.164	0.090*
Missing: Number of Children	-0.084	0.308	-0.235	0.356	-0.586	0.356*	0.081	0.523	0.181	0.332
Missing: Education	4.686	0.538***	3.454	0.633***	2.151	0.613***	-11.427	1.186***	1.692	0.642***
Time Trend	0.023	0.038	-0.041	0.038	-0.026	0.037	-0.006	0.048	0.038	0.035
Constant	-2.650	0.696***	-5.162	0.770***	-2.541	0.696***	4.692	0.814***	0.597	0.660
Permanent Unob. Het.	-1.440	0.193***	1.370	0.232***	-1.106	0.222***	-1.406	0.339***	-4.461	0.168***
Permanent Unob. Het.	-3.777	0.259***	-1.824	0.269***	-0.729	0.209***	0.755	0.248***	-4.240	0.144***
Permanent Unob. Het.	1.585	0.228***	1.103	0.307***	3.710	0.217***	-1.595	0.547***	-3.281	0.253***
Time-varying Unob. Het.	0.005	0.117	-0.037	0.118	-0.034	0.117	-0.004	0.131	-0.163	0.109
Time-varying Unob. Het.	1.171	0.330***	0.358	0.361	0.887	0.344***	0.453	0.501	0.311	0.327
Time-varying Unob. Het.	0.709	0.176***	0.477	0.180***	0.211	0.183	-0.294	0.227	-0.068	0.175

*, **, *** Significant at the 10%, 5%, and 1% levels.

Table B.4: Logit on investment decisions (relative to not investing in that account)

Variable	Account A		Account B		Account C		Account D		Account E	
	Coeff.	St.Er.	Coeff.	St.Er.	Coeff.	St.Er.	Coeff.	St.Er.	Coeff.	St.Er.
Work Experience	0.053	0.024**	0.001	0.012	-0.115	0.017***	0.068	0.014***	0.033	0.020*
Experience Squared	-0.002	0.001***	0.000	0.000	0.004	0.001***	-0.003	0.000***	-0.001	0.001
Lagged Investment in A	2.507	0.177***	0.020	0.104	-0.424	0.230**	-0.058	0.141	-0.465	0.202**
Lagged Investment in B	0.246	0.176	1.325	0.088***	-1.123	0.190***	0.103	0.117	0.013	0.150
Lagged Investment in C	0.559	0.194***	0.344	0.088***	0.889	0.175***	-0.168	0.111	0.039	0.137
Lagged Investment in D	-0.067	0.342	0.369	0.114***	-1.739	0.264***	1.210	0.130***	-0.100	0.162
Lagged Investment in E	0.672	0.481	0.570	0.158***	-0.110	0.354	-0.305	0.201	1.255	0.179***
Lagged Assets	0.015	0.003***	0.013	0.002***	0.002	0.002	-0.009	0.002***	0.007	0.003***
Lagged Optional Savings	0.365	0.112***	0.021	0.053	-0.106	0.092	-0.119	0.069*	-0.047	0.087
Lagged Marital Status	0.338	0.189*	0.068	0.074	-0.173	0.122	0.016	0.097	0.226	0.120*
Number of Children	-0.049	0.067	0.013	0.034	-0.275	0.054***	0.210	0.046***	-0.102	0.058*
Interaction Female-Married	-0.289	0.288	-0.139	0.102	0.165	0.174	0.040	0.130	-0.146	0.164
Interaction Female-Children	0.048	0.103	-0.002	0.046	0.655	0.078***	-0.426	0.063***	0.134	0.075*
Health: Very Good	0.146	0.138	-0.168	0.066**	0.228	0.115*	-0.022	0.092	0.121	0.109
Health: Fair	-0.116	0.155	-0.074	0.062	-0.004	0.101	0.046	0.073	0.060	0.093
Health: Poor	0.239	0.312	-0.394	0.131***	-0.030	0.184	0.221	0.125*	-0.106	0.171
Age	0.311	0.033***	-0.347	0.013***	1.207	0.041***	-0.208	0.018	0.035	0.021*
Age Squared	-0.095	0.007***	0.053	0.003***	-0.318	0.010***	0.104	0.004	-0.006	0.005
Female	-0.314	0.238	0.062	0.084	-1.257	0.152***	1.085	0.110	0.167	0.139
High School	0.705	0.155***	0.261	0.057***	-0.239	0.096**	-0.104	0.072	0.139	0.094
Technical College	1.391	0.199***	0.562	0.086***	-0.798	0.170***	-0.402	0.116	-0.032	0.149
College	1.911	0.427***	0.717	0.210***	-1.079	0.642*	-0.647	0.362	-0.076	0.719
Unemployment Rate	0.079	0.027***	0.015	0.014	0.023	0.023	0.013	0.017	0.006	0.022
Number of Hospital Beds	-0.135	0.233	-0.004	0.110	0.278	0.197	-0.001	0.137	0.380	0.166**
Number of Doctors	-0.052	0.496	-0.442	0.326	0.223	0.510	-0.562	0.347	-1.065	0.382***
Number of Marriages	0.433	0.196**	-0.108	0.134	-0.427	0.162***	-0.111	0.133	-0.107	0.151
Inches of Rainfall	0.015	0.004***	0.005	0.003*	-0.009	0.004***	-0.005	0.003	-0.002	0.004
College Tuition	-0.517	0.107***	-0.107	0.054**	0.257	0.081***	-0.121	0.064	-0.421	0.078***
Missing: Number of Children	0.424	0.357	-0.132	0.155	-0.490	0.378	-0.053	0.261	-0.278	0.385
Missing: Education	0.892	0.815	-0.369	0.650	0.018	0.715	-0.872	0.538	0.354	0.711
Time Trend	-0.268	0.041***	-0.217	0.023***	-0.024	0.038	-0.056	0.026	-0.116	0.034***
Constant	-7.674	0.924***	1.951	0.769**	-2.374	0.724***	-4.353	0.690	-4.342	0.703***
Permanent Unob. Het.	0.176	0.191	0.195	0.086**	-0.425	0.163***	-0.027	0.117	0.048	0.150
Permanent Unob. Het.	-0.500	0.198**	-0.258	0.081***	0.096	0.143	0.157	0.103	0.399	0.124***
Permanent Unob. Het.	0.111	0.181	-0.146	0.090	-0.198	0.162	0.126	0.114	0.068	0.151
Time-varying Unob. Het.	2.051	0.210***	2.987	0.087***	-7.760	0.290***	3.405	0.124	2.794	0.267***
Time-varying Unob. Het.	2.438	0.308***	1.872	0.175***	-5.302	0.357***	2.142	0.213	1.399	0.521***
Time-varying Unob. Het.	9.663	0.337***	1.132	0.198***	-21.710	27.740	-4.549	0.462	2.718	0.310***

*, **, *** Significant at the 10%, 5%, and 1% levels.

Table B.5: Logit on savings decision (relative to not holding optional savings)

Variable	Optional Savings	
	Coeff.	St.Er.
Work Experience	-0.007	0.009
Experience Squared	0.000	0.000
Lagged Investment in A	0.203	0.068***
Lagged Investment in B	0.116	0.056**
Lagged Investment in C	0.180	0.055***
Lagged Investment in D	0.138	0.071*
Lagged Investment in E	0.121	0.103
Lagged Assets	0.006	0.001***
Lagged Optional Savings	0.825	0.034***
Lagged Marital Status	0.074	0.050
Number of Children	0.001	0.023
Interaction Female-Married	-0.061	0.070
Interaction Female-Children	-0.042	0.032
Health: Very Good	0.001	0.046
Health: Fair	-0.077	0.041*
Health: Poor	-0.201	0.081**
Age	-0.061	0.009***
Age Squared	0.008	0.002***
Female	0.131	0.057**
High School	0.294	0.038***
Technical College	0.529	0.057***
College	0.893	0.135***
Unemployment Rate	-0.016	0.010*
Number of Hospital Beds	0.049	0.069
Number of Doctors	0.152	0.150
Number of Marriages	0.037	0.062
Inches of Rainfall	0.005	0.002***
College Tuition	0.043	0.034
Missing: Number of Children	-0.071	0.112
Missing: Education	0.665	0.197***
Time Trend	0.008	0.014
Constant	-1.297	0.283***
Permanent Unob. Het.	-0.057	0.059
Permanent Unob. Het.	-0.356	0.052***
Permanent Unob. Het.	-0.091	0.056
Time-varying Unob. Het.	0.042	0.047
Time-varying Unob. Het.	-0.142	0.113
Time-varying Unob. Het.	0.172	0.066***

*, **, *** Significant at the 10%, 5%, and 1% levels.

Table B.6: Multinomial logit and OLS on subjective assessments

Variable	Elicited Risk Aversion (Relative to Most)				Expected Duration of Life	
	Intermediate		Least			
	Coeff.	St.Er.	Coeff.	St.Er.	Coeff.	St.Er.
Work Experience	0.024	0.015*	-0.019	0.011*	0.018	0.013
Experience Squared	-0.001	0.001	0.001	0.000		
Lagged Investment in A	0.098	0.117	0.123	0.084	-0.158	0.503
Lagged Investment in B	-0.021	0.098	0.002	0.071	0.402	0.357
Lagged Investment in C	0.072	0.094	-0.028	0.072	0.338	0.361
Lagged Investment in D	0.052	0.117	-0.044	0.089	0.435	0.539
Lagged Investment in E	-0.390	0.200*	-0.026	0.124	0.284	0.685
Lagged Assets	0.001	0.002	-0.001	0.002	0.005	0.006
Lagged Optional Savings	0.029	0.060	-0.002	0.042	0.526	0.172***
Lagged Marital Status	0.058	0.080	-0.006	0.055	0.785	0.310**
Number of Children	0.001	0.038	0.016	0.026	0.145	0.108
Interaction Female-Married	0.046	0.114	0.062	0.081	-0.766	0.467
Interaction Female-Children	-0.036	0.054	-0.032	0.037	-0.195	0.151
Health: Very Good	0.115	0.076	0.186	0.052***	1.253	0.220***
Health: Fair	0.024	0.067	0.046	0.048	-2.485	0.192***
Health: Poor	-0.176	0.129	-0.075	0.091	-5.987	0.402***
Age	-0.006	0.005	-0.010	0.003***	-0.120	0.042***
Age Squared					0.055	0.010***
Female	-0.090	0.096	-0.370	0.067***	-0.657	0.350*
High School	0.011	0.068	0.109	0.045**	0.513	0.191***
Technical College	0.175	0.102*	0.283	0.067***	1.662	0.353***
College	-0.111	0.613	0.267	0.185	1.735	0.693**
Unemployment Rate	-0.022	0.016	-0.020	0.011*	-0.182	0.047***
Number of Hospital Beds	0.346	0.118***	0.202	0.091**	0.164	0.431
Number of Doctors	0.519	0.295*	0.078	0.302	0.942	0.677
Number of Marriages	-0.281	0.123**	-0.212	0.126*	-0.867	0.335***
Inches of Rainfall	-0.015	0.003***	-0.007	0.002***	-0.030	0.009***
College Tuition	0.077	0.057	0.104	0.045**	0.328	0.157**
Missing: Number of Children	-0.295	0.214	0.106	0.125	0.968	0.556*
Missing: Education	0.720	0.595	0.423	0.334	0.102	1.000
Time Trend	0.045	0.024*	0.012	0.020	0.084	0.082
Constant	-2.805	0.626***	-0.895	0.741	52.038	1.014***
Permanent Unob. Het.	-0.152	0.100	-0.200	0.070***	1.337	0.437***
Permanent Unob. Het.	-0.059	0.085	-0.275	0.062***	0.060	0.367
Permanent Unob. Het.	0.081	0.096	0.154	0.064**	0.222	0.421
Time-varying Unob. Het.	0.135	0.079*	-0.049	0.054	0.135	0.223
Time-varying Unob. Het.	0.000	0.170	0.065	0.115	1.281	0.788
Time-varying Unob. Het.	0.321	0.108***	0.169	0.074**	0.276	0.474

*, **, *** Significant at the 10%, 5%, and 1% levels.

Table B.7: Marital status and variation in the number of children

Variable	Marital Status (Relative to Married)		Children Variation (Relative to No Change)			
			Decrease		Increase	
	Coeff.	St.Er.	Coeff.	St.Er.	Coeff.	St.Er.
Duration of Marriage	-0.025	0.004***	0.066	0.004***	-0.098	0.014***
Lagged Marital Status	-4.382	0.106***	-1.133	0.115***	0.798	0.195***
Number of Children	-0.258	0.035***	1.161	0.032***	0.691	0.065***
Interaction Female-Married	-0.097	0.106	-0.316	0.095***	-0.076	0.213
Interaction Female-Children	0.100	0.048**	0.177	0.041***	-0.035	0.098
Full-Time Employed	-0.047	0.071	0.297	0.060***	0.554	0.194***
Part-Time Employed	-0.029	0.153	0.254	0.127**	0.148	0.463
Age	0.063	0.028**	0.515	0.017***	-0.153	0.025***
Age Squared	-0.037	0.017**	-0.113	0.004***	0.006	0.009
Age Cubed	0.006	0.003**				
Female	0.357	0.090***	0.263	0.098***	0.005	0.211
High School	0.016	0.060	-0.078	0.049	0.202	0.118*
Technical College	-0.079	0.092	-0.131	0.080*	0.068	0.187
College	-0.452	0.159***	-0.075	0.127	0.037	0.583
Number of Marriages	-0.317	0.085***				
College Tuition			-0.001	0.039	-0.217	0.087***
Missing: Marriage Duration	-0.082	0.441	1.595	0.443***	-0.026	0.988
Missing: Number of Children	-0.641	0.158***				
Missing: Education	-0.374	0.553	0.114	0.426	0.941	0.893
Constant	3.257	0.388***	-8.618	0.261***	-2.371	0.463***
Permanent Unob. Het.	0.184	0.093**	-0.107	0.079	-0.053	0.200
Permanent Unob. Het.	0.016	0.078	0.041	0.064	-0.112	0.206
Permanent Unob. Het.	0.045	0.093	-0.099	0.076	-0.183	0.198
Time-varying Unob. Het.	0.015	0.089	-0.011	0.079	-0.199	0.212
Time-varying Unob. Het.	-1.795	0.352***	0.866	0.319***	3.972	0.439***
Time-varying Unob. Het.	-0.043	0.130	0.254	0.105**	-0.072	0.271

*, **, *** Significant at the 10%, 5%, and 1% levels.

Table B.8: Health status and medical care consumption

Variable	Health Status (Relative to Very Good)						Medical Consumption	
	Good		Regular		Poor			
	Coeff.	St.Er.	Coeff.	St.Er.	Coeff.	St.Er.	Coeff.	St.Er.
Health: Very Good	-0.528	0.060***	-0.789	0.084***	-0.889	0.203***	-1.047	0.246***
Health: Fair	0.289	0.081***	1.526	0.084***	1.845	0.122***	4.887	0.207***
Health: Poor	0.678	0.329**	2.353	0.322***	4.108	0.333***	15.679	0.424***
Number of Medical Visits	0.010	0.003***	0.022	0.004***	0.027	0.004***		
Work Experience	0.003	0.005	-0.004	0.006	-0.005	0.008		
Legislators	-0.296	0.142**	-0.442	0.175**	-0.288	0.330		
Clerical	-0.025	0.143	0.007	0.172	0.282	0.352		
Service and Sales	0.011	0.156	-0.090	0.187	0.084	0.322		
Agricultural	-0.165	0.178	-0.244	0.204	-0.191	0.342		
Plant Operators	0.062	0.141	-0.018	0.163	0.208	0.264		
Age	0.034	0.014**	0.084	0.017***	0.163	0.032***	-0.048	0.040
Age Squared	-0.004	0.003	-0.009	0.004**	-0.021	0.007***	0.019	0.009**
Female	0.170	0.064***	0.379	0.075***	0.618	0.115***	4.149	0.177***
High School	-0.098	0.066	-0.537	0.077***	-0.693	0.121***	1.370	0.198***
Technical College	-0.214	0.105**	-0.924	0.139***	-1.301	0.274***	2.881	0.378***
College	-0.489	0.253*	-1.445	0.520***	-1.873	0.826**	3.974	0.943***
Inches of Rainfall	0.001	0.002	0.006	0.002**	0.003	0.004		
Number of Hospital Beds							-0.038	0.299
Number of Doctors							0.550	0.671
Missing: Occupation	-0.096	0.327	-0.341	0.438	-0.405	0.691		
Missing: Education	-0.201	0.492	-0.657	0.712	-0.766	0.922	2.248	1.000**
Unemployed	0.123	0.333	0.254	0.448	0.713	0.686		
Constant	0.869	0.200***	-0.946	0.244***	-4.435	0.508***	1.537	0.882*
Permanent Unob. Het.	-0.079	0.139	-0.130	0.168	-0.220	0.294	-0.302	0.413
Permanent Unob. Het.	0.072	0.118	0.409	0.136***	0.749	0.206***	-0.201	0.480
Permanent Unob. Het.	0.075	0.137	0.093	0.169	0.296	0.288	-0.657	0.434
Time-varying Unob. Het.	-0.068	0.075	-0.055	0.090	0.009	0.150	0.215	0.340
Time-varying Unob. Het.	1.084	1.442	1.105	1.442	1.624	1.670	-1.633	0.699**
Time-varying Unob. Het.	-0.095	0.103	-0.273	0.126**	-0.325	0.210	0.947	0.598

*, **, *** Significant at the 10%, 5%, and 1% levels.