Internet Appendix for "The Cross-Section of Household Preferences"

Laurent E. Calvet, John Y. Campbell, Francisco J. Gomes, and Paolo Sodini[1]

This version: December 2020

[1] Calvet: Department of Finance, EDHEC Business School, 393 Promenade des Anglais, BP 3116, 06202 Nice Cedex 3, France, and CEPR; e-mail: laurent.calvet@edhec.edu. Campbell: Department of Economics, Harvard University, Littauer Center, Cambridge MA 02138, USA, and NBER; e-mail: john_campbell@harvard.edu. Gomes: London Business School, Regent's Park, London NW1 4SA, UK; e-mail: fgomes@london.edu. Sodini: Department of Finance, Stockholm School of Economics, Sveavagen 65, P.O. Box 6501, SE - 113 83 Stockholm, Sweden; e-mail: Paolo.Sodini@hhs.se. We thank Todor Sapunarov for research assistance.

This Internet Appendix provides additional information on the methodology used in the paper. Section 1 presents the organization of the Swedish pension system. Section 2 discusses the empirical methodology.

1 The Swedish Pension System

This Section explains the rules of the Swedish pension system applying to households in the 1999 to 2007 sample. Section 1.1 defines key indexes used in the calculation of pensions. Section 1.2 discusses the organization of public pensions, and Section 1.3 the organization of occupational pensions. Section 1.4 explains the allocation rule of the default fund in the public DC pension system. Section 1.5 describes the imputation of private pension contributions and wealth. Section 1.6 analyzes the impact of taxation.

1.1 Key Indexes

The calculations of public and occupational pensions rely on the following three income indexes defined in the Swedish Social Insurance Code.

- PBB_t denotes the "price-related base amount" (*prisbasbelopp*). It is set by the government based on calculations produced by Statistic Sweden.
- $HBPA_t$ is the "higher price-related base amount" (*förhöjt prisbasbelopp*). It is set by the government based on calculations produced by Statistic Sweden.
- IBB_t denotes the "income-related base amount" (*inkomstbasbelopp*). It is set by the government based on calculations produced by the Swedish Pensions Agency.

The price-related base amount, PBB_t, and the higher price-related base amount, $HBPA_t$, are available from Statistics Sweden's website. The income-related base amount, IBB_t, is available from the Swedish Parliament's website. For the year 2003, the price-related base amount is 38,600 SEK, the higher-price related base amount is 39,400 SEK, and the income-related base amount is 40,900 SEK.

1.2 Public Pension System

The Swedish Pensions Agency administers both the defined benefit (DB) and defined contribution (DC) components of public pensions.[2] An individual's public pension has four

[2]See the Swedish Pensions Agency's website (https://www.pensionsmyndigheten.se) and Röstberg et al. (2004) for detailed explanations.

components:

(i) the DB income pension (*inkomstpension*), $IP_{i,t}$, which is based on the individual's lifetime income;

(ii) the DC premium pension (*premiepension*), $PPM_{i,t}$, which is contingent on the returns on the compulsory DC premium pension contributions invested by the individual;

(iii) the DB guaranteed pension (*garantipension*), $GP_{i,t}$, if the individual has a low income pension $IP_{i,t}$ or no earnings-related pensions;

(iv) the DB supplementary pension (*tilläggspension*), $ATP_{i,t}$, if the individual is born in 1953 or earlier.

Our terminology closely matches the terminology used in the annual Orange Reports of the Swedish Pensions Authority (see, e.g., 2005, 2015). Sections 1.2.1 to 1.2.4 of this Internet Appendix provide detailed definitions of the four components of public pensions.

An individual born after 1953 receives a total public pension equal to

$$\text{Public Pension}_{i,t} = IP_{i,t} + PPM_{i,t} + GP_{i,t}. \tag{1}$$

An individual born in 1953 or earlier receives

$$\text{Public Pension}_{i,t} = \begin{cases} ATP_{i,t} + GP_{i,t} & \text{if born before 1938,} \\ 4/20\ (IP_{i,t} + PPM_{i,t}) + 16/20\ ATP_{i,t} + GP_{i,t} & \text{if born in 1938,} \\ 5/20\ (IP_{i,t} + PPM_{i,t}) + 15/20\ ATP_{i,t} + GP_{i,t} & \text{if born in 1939,} \\ \dots & \\ 19/20\ (IP_{i,t} + PPM_{i,t}) + 1/20\ ATP_{i,t} + GP_{i,t} & \text{if born in 1953.} \end{cases} \tag{2}$$

All variables in (**??**) and (**??**) are pre-tax. The implications of taxes on public and occupational pensions are analyzed in Section 1.4.

1.2.1 DB Income Pension

The DB income pension is a pay-as-you-go system that applies to households in all cohorts, as defined by the 1998 Earnings Related Old Age Pension Act. The rules of the income pension system mimic the organization of individual funded accounts. During an individual's working life, contributions are credited to a notional individual account. Each year a notional interest payment is credited to the account. At retirement, the notional account becomes an annuity, whose amount is determined by the individual's notional balance and his/her cohort's life expectancy.

The precise calculation of the DB income pension proceeds in three steps.

Step 1 (Notional accumulation). The notional accumulation is computed from annual income and credits as follows.

- The relevant income of individual i in year t, $RI_{i,t}$, consists of the year's annual earnings, including sickness cash benefits, parental cash benefits, and unemployment cash benefits.
- The pension qualifying income, $PQI_{i,t}$, is defined as

$$PQI_{i,t} = RI_{i,t} - 7\% \times \min(RI_{i,t}, 8.07\ IBB_t),$$

  where IBB_t is the income-related base amount defined in Section 1.1.
- The pension qualifying amount, $PQA_{i,t}$, is a fictitious income computed for years of child care, compulsory national service, studies, and sickness or activity compensation.[3]
- The pension base (*pensionunderlag*) is the sum of the pension-qualifying income and pension-qualifying amounts, capped at 7.5 times the income-related base amount:

$$\text{Pension Base}_{i,t} = \max[PQI_{i,t} + PQA_{i,t}\ ;\ 7.5\ IBB_t]. \tag{3}$$

  Beginning in 2003, the pension base is zero if $PQI_{i,t} + PQA_{i,t}$ is lower than $42.3\% \times PBB_t$.
- The pension right accumulated by individual i in year t is

$$\text{Pension Right}_{i,t} = 16.0\% \times \text{Pension Base}_{i,t}.$$

Step 2 (Notional balance). The notional *pension balance at t* is

$$\begin{aligned} \text{Pension Balance}_{i,t} &= 99.95\% \times \text{Inheritance Gain Factor}_{i,t} \\ &\times \left[\sum_{t=S_i}^{T_i} \left(\text{Pension Right}_{i,t} \times \frac{\text{IncIndex}_{T_i}}{\text{IncIndex}_t} \right) \right], \end{aligned} \tag{4}$$

where $\text{Inheritance Gain Factor}_{i,t}$ is specific to the individual's cohort in year t, S_i is the individual's first year with a strictly positive pension-qualifying income, T_i is the year of

[3] Pension-qualifying amounts for *child care* apply for a period of up to 4 years after a child's birth; the pension-qualifying amount in year t is the maximum of the following three quantities: (i) the parent's pension-qualifying income in the year prior to the child's birthyear, (ii) 75% of year t's average pension-qualifying income of all insured persons aged between 16 and 64, and (iii) the year's income-related base amount, IBB_t. The pension-qualifying amount for *compulsory national service* is set to 50% of the average pensionable qualifying-income for all insured persons aged between 16 and 64. Pension-qualifying amounts for *study* are calculated as 138% of disbursed study grants. Pension-qualifying amounts for *sickness and activity compensation* are set to 93% of the income that the insured person would have likely received if he/she had worked according to the Swedish Social Insurance Agency (*Försäkringskassan*). See Swedish Ministry of Health and Social Affairs (2009) for further details.

retirement which is typically the 65^{th} birthday, and is at the earliest the 61^{st} birthday. IncIndex_t is the income index for year t. The 99.95% factor corresponds to administrative costs of 0.05%.

The inheritance gain factor (*arvvinstfaktorer*) is computed by the National Social Insurance Board (*Riksförsäkringsverket*). It is designed to allocate the pension balances of deceased persons to survivors in the same age group. For cohorts up to the age of 60, the inheritance gain factor is equal to unity plus the ratio of the pension balances of the deceased to the pension balances of survivors. For cohorts aged more than 60, the inheritance gain factor is a survival rate, which permits the homogeneous treatment of economically active and retired individuals in the same cohort.[4]

The Income Index is computed by the National Social Insurance Board and is available from its website. By definition, the growth rate of the income index, $\text{IncIndex}_t/\text{IncIndex}_{t-1}$, is the product of 1-year inflation with the 3-year moving average of real earnings.[5]

Step 3 (Payout). The annual DB income pension is calculated by dividing the individual's pension account balance (**??**) by his/her cohort's annuitization divisor (*delningstal*):

$$IP_{i,t} = \frac{\text{Pension Balance}_{i,t}}{\text{Annuitization Divisor}_i}.$$

The annuitization divisor is computed by the National Social Insurance Board. It is common to individuals born in the same year and does not vary after age 65.

1.2.2 DC Premium Pension

The DC component of public pensions, which was introduced in 1995, is designed as a funded system. During the working life of individual i, the DC premium contribution in every year t is

$$\text{Premierate}_i \times \text{Pension Base}_{i,t},$$

where $\text{Pension Base}_{i,t}$ is given by (**??**). The rate of contribution, Premierate_i, is 2.5% for years 1999 and later, and 2% in years 1995-1998. Premium pension contributions are invested by default in the AP7 fund (see section 1.4 for details). Individual can opt out and choose funds among a large menu of mutual funds available in the system.[6]

At retirement, the insured person can either convert the funds into an annuity or keep them invested in mutual funds.

[4]The exact definitions of the Inheritance Gain Factor is provided in Appendix A of Swedish Pensions Agency (e.g., 2005).

[5]See Appendix A of Swedish Pensions Agency (e.g., 2005).

[6]In June 2003, 655 funds were available for active choice on the Premium pension platform

- If the individual chooses to receive an annuity, the pension is calculated as a guaranteed life-long annuity payable in nominal monthly instalments.
- Alternatively, the pension savings remain in the account and are invested in mutual funds chosen by the insured. The premium pension is recalculated once a year based on the value of fund shares in December.

In both cases, the premium pension, $PPM_{i,t}$, is the value of the premium pension account divided by an annuity divisor, which is based on forecasts of future life expectancy.[7]

1.2.3 DB Guaranteed Pension

The guaranteed pension provides basic retirement income for individuals who have had little or no previous earnings. Unlike other components of the public pension system, the guaranteed pension is funded by the state budget. Residents of Sweden are eligible beginning at age 65. In order to obtain a maximum guaranteed pension, one needs to have resided in Sweden for 40 years after age 25.[8] In our simulations, we assume for simplicity that all households meet this criterion.

The level of the guaranteed pension is determined by the DB income pension, $IP_{i,t}$, defined in Section 1.2.1. The guaranteed pension, $GP_{i,t}$, is given by

$$GP_{i,t} = \begin{cases} 2.13\ PBB_t - IP_{i,t} & \text{when } IP_{i,t} < 1.26\ PBB_t \\ 0.87\ PBB_t - 0.48(IP_{i,t} - 1.26\ PBB_t) & \text{when } 1.26\ PBB_t \leq IP_{i,t} \leq 3.07\ PBB_t \\ 0 & \text{when } IP_{i,t} > 3.07\ PBB_t \end{cases}$$

for unmarried individuals, and by

$$GP_{i,t} = \begin{cases} 1.90\ PBB_t - IP_{i,t} & \text{when } IP_{i,t} < 1.14\ PBB_t \\ 0.76\ PBB_t - 0.48(IP_{i,t} - 1.14\ PBB_t) & \text{when } 1.14\ PBB_t < IP_{i,t} < 2.72\ PBB_t \\ 0 & \text{when } IP_{i,t} > 2.72\ PBB_t \end{cases}$$

for married pensioners, where PBB_t is the price-related base amount defined in Section 1.1.

1.2.4 DB Supplementary Pension

Individuals born in 1953 or earlier are entitled to a public DB supplementary pension, which can be computed in three steps.

[7] See Appendix A of Swedish Pensions Agency (e.g., 2005) for further details.

[8] Years of residence in another EU or EEA country also provide guaranteed pension credits.

Step 1. We compute the individual's 15 best years of earnings, out of 30. For each of these 15 years we calculate:

$$\text{Pension Points}_{i,t} = (PQI_{i,t} - HBPA_t)/HBPA_t, \tag{5}$$

where $HBPA_t$ is the higher price-related base amount defined in Section 1.1. The average pension points for the best 15 years is denoted by AvPoints_i.

Step 2. The annual supplementary DB pension is

$$SuppDB_{i,t} = 60\% \times \text{AvPoints}_i \times PBB_t, \tag{6}$$

where PBB_t is the price-related base amount. An individual needs to work 30 years in order to get the full supplementary DB pension. The benefit (**??**) is otherwise multiplied by a factor of $N/30$, where N is the number of years for which the individual received pension rights. In our simulations, we set N equal to 30 for all individuals in the panel.

Step 3. We compute the *folkspension*:

$$Folkspension_{i,t} = \begin{cases} 96\% \times PBB_t & \text{for an unmarried person,} \\ 78.5\% \times PBB_t & \text{for a married person,} \end{cases}$$

where PBB_t is the price-related base amount.

Step 4. The total supplementary pension is

$$ATP_{i,t} = SuppDB_{i,t} + Folkspension_{i,t}.$$

1.3 Occupational Pension System

We now provide a detailed description of some of the main collective labor agreements in Sweden:

- STP and SAF-LO, which cover private-sector blue collar workers (1.3 million persons in 2002);
- ITP2, which covers private-sector white collar workers (610,000 persons in 2002);
- PA 91 and PA 03, which cover central government employees (250,000 persons in 2002);
- PFA 98 and KAP-KL, which cover local government employees (940,000 persons in 2002).

Overall, these agreements cover about 90% of the Swedish workforce. We refer the reader to Sjögren and Wadensjö (2005, ch. 6.4.1) for an excellent detailed description of these occupational plans.

1.3.1 Private-Sector Blue Collar Workers

In 1973, the Swedish Employers Association (SAF) and the central confederation of blue-collar workers (LO) agreed to provide a special defined benefit pension plan, the *särskild tilläggspension* (STP). The STP introduced a defined contribution component in 1991, which came into effect in 1992. Since 1996, the main pension scheme for blue-collar workers is the SAF-LO defined contribution agreement. However, workers previously enrolled in the STP defined benefit scheme still receive pension payments between 1999 and 2007.

Defined Benefits. Yearly DB pension payouts are given by

$$10\% \times (\text{relevant pension points} + 1) \times PBB_t,$$

where the relevant pension points are defined as the average of Pension Points$_{i,t}$ (defined in (**??**)) over the best 3 years between the ages of 55 and 59. Following the introduction of SAF-LO, the DB pension payouts are adjusted for birthyear using the same ratios as the ones applied to the public DC premium pension discussed in Section 1.2.2.

Defined Contributions. The average yearly premium is

$$\begin{array}{ll} 0.53\% \times PQI_{i,t} & \text{in 1992,} \\ 0.54\% \times PQI_{i,t} & \text{in 1993 and 1994,} \\ 3.90\% \times PQI_{i,t} & \text{in 1996,} \\ 4.08\% \times PQI_{i,t} & \text{in 1997,} \\ 3.50\% \times PQI_{i,t} & \text{from 1998 to 2007.} \end{array} \tag{7}$$

There was no contribution in 1995 as the pension system transitioned to a pure defined contribution scheme.

The premium in (**??**) is the sum of the contributions paid by both the worker and the employer. Specifically, the worker and the employer respectively contributed $2\% \times PQI_{i,t}$ and $1.9\% \times PQI_{i,t}$ to the pension plan in 1996, $2\% \times PQI_{i,t}$ and $2.08\% \times PQI_{i,t}$ in 1997, and $2\% \times PQI_{i,t}$ and $1.5\% \times PQI_{i,t}$ in 1998 and 1999. Starting in 2000, employees pay a contribution of $3.5\% \times PQI_{i,t}$ (no ceiling), and there are no contributions from employers.

The age at which defined contributions start accruing is 28 until 1999, 22 in 2000 and 2001, and 21 in 2002 onward. DC premia continue to be paid if the worker is on sick leave, pregnancy leave, or parental leave.[9] Blue collar workers covered by SAF-LO can allocate their contributions to traditional insurance (characterized by conservative investments and guaranteed payments) and unit-linked insurance (riskier investments).

[9] In these cases, the premia are paid by an insurance tool, called *premiebefrielseförsäkring*, instead of being deducted from the salary. The rationale is that long-term illness, for instance, should not result in lower occupational pension on retirement.

1.3.2 Private-Sector White Collar Workers

The *industrins och handelns tilläggspension* (ITP) covers white-collar workers from the private sector. It consists of two schemes: ITP2 for members born before 1979 and ITP1 for members born therafter. ITP1 is not relevant for our paper and is therefore not discussed here. In addition, the ITPK agreement was introduced in 1977 as a compulsory DC scheme for private-sector white collar workers (Hagen 2013).

Defined Benefits. ITP2 provides a specific share of the final salary after thirty years of service. Let $SIYBR_i$ denote the salary in the year before retirement. The ITP2 DB payout is given by the following formula:

$$\begin{aligned}&(10\% \times \text{portion of } SIYBR_i \text{ below } 7.5\ IBB)\\&+(65\% \times \text{portion of } SIYBR_i \text{ between } 7.5\ IBB \text{ and } 20\ IBB)\\&+(32.5\% \times \text{portion of } SIYBR_i \text{ between } 20\ IBB \text{ and } 30\ IBB).\end{aligned}$$

Levels of $SIYBR_i$ above 30 IBB do not provide an ITP2 payout. Before 2003, the formula was based on the increased price base amount, $HBPA$, instead of the income base amount, IBB.

Defined Contributions. The yearly contribution to the ITPK defined-contribution plan is

$$2\% \times \min(PQI_{i,t}; 30\ IBB_t)$$

over the entire period, where $PQI_{i,t}$ is the pension qualifying income that applies to enrolled workers. Workers start contributing to the ITPK plan at age 28. DC premia continue to be paid if the worker is on sick leave, pregnancy leave, or parental leave.

Before 1990, the ITPK plan was handled as if each payment gave rise to a benefit and ITPK payouts were part of the collective refund (i.e. with base year and pension allowance). Pre-1990 ITPK can thererefore be seen as a hybrid between a defined contribution and a defined benefit plan. In 1990 this changed and it became possible to invest ITPK premia in garanteed or unit-linked products provided by companies such as Collectum or Valcentralen. For these reasons, we include pre-1990 ITPK in the DB payout defined for regular ITP2.

High Earners. Workers earning more than 7.5 IBB_t in a given year, who are known as *tiotaggare* (high earners), are entitled to leave their standard ITP arrangement and join a defined contribution scheme, in agreement with their employer. According to Collectum (2015), the contributions to these alternative schemes, are calculated using either the *frilagd premie* (free premium) or the *premietrappa* (premium ladder) for the portion of income between 7.5 IBB_t and 30 IBB_t.

Our dataset does not provide the type of premium chosen by high earners and their employers. However, a feature of the system allow us to make a simplifying assumption. The "free premium" matches the costs the employers would incur within the DB component of ITP 2; it tends to vary substantially, which makes it difficult for employers to budget and control their pension costs. By contrast, the "premium ladder" is stable and easy to budget, and the variation of pension premia with age and salary levels can easily be explained to employees. The premium ladder is therefore the more popular option, so we assume that high earners who leave the ITP 2 system always choose the "premium ladder."

The dataset does not allow us to determine if a high earner is enrolled in the ITP 2 system or in an alternative scheme. As a compromise, we assume that a high earner invests half of the premium in the ITP 2 DB scheme and the other half in the alternative DC scheme.

1.3.3 Central Government Employees

Central government employees are covered by two agreements: PA 91 from 1992 to 2002, and PA 03 since 2003.

Defined Benefits. The pension basis is the average of the 5 last salaries before retirement. The PA 91 defined benefit payout is

$$\begin{aligned}&(10\% \times \text{portion of pension basis below } 7.5\ HBPA_t)\\&+(65\% \times \text{portion of pension basis between } 7.5\ HBPA_t \text{ and } 20\ HBPA_t)\\&+(32.5\% \times \text{portion of pension basis between } 20\ HBPA_t \text{ and } 30\ HBPA_t).\end{aligned}$$

The PA 03 defined benefit payout is

$$\begin{aligned}&(0\% \times \text{portion of pension basis below } 7.5\ HBPA_t)\\&+(60\% \times \text{portion of pension basis between } 7.5\ HBPA_t \text{ and } 20\ HBPA_t)\\&+(30\% \times \text{portion of pension basis between } 20\ HBPA_t \text{ and } 30\ HBPA_t).\end{aligned}$$

Payments of PA 91 continued after the introduction of PA 03 for persons born before 1943.

Defined Contributions. Yearly PA 91 pension contributions are made under the Kåpan scheme and are defined as follows:

$$\begin{aligned}&1.5\% \times PQI_{i,t} \text{ in years 1992 to 1994,}\\&1.7\% \times PQI_{i,t} \text{ in years 1995 to 2002,}\end{aligned}$$

for workers who are at least 28 years of age. DC premia continue to be paid if the worker is on sick leave, pregnancy leave, or parental leave.

Yearly PA 03 pension contributions are made up of two parts: the *individuell ålderspension* premium and the Kåpan premium.[10] From the age of 23, central government employees pay the *individuell ålderspension* premium:

$$2.3\% \times PQI_{i,t} \text{ in } 2003,$$
$$2\% \times PQI_{i,t} \text{ thereafter.}$$

From the age of 28, employees also pay the Kåpan premium:

$$1.9\% \times PQI_{i,t} \text{ in } 2003,$$
$$2.3\% \times PQI_{i,t} \text{ thereafter.}$$

Central government employees younger than 23 pay no DC contributions.

1.3.4 Local Government Employees

Local government workers are covered by two occupational pensions schemes during the period: PFA 98 between 1998 and 2006, and KAP-KL thereafter.

Defined Benefits. For the 1999 to 2006 period, the PFA 98 pension payout is

$$\begin{aligned}&(0\% \times \text{portion of pension basis below } 7.5\ IBB_t)\\&+(62.5\% \times \text{portion of pension basis between } 7.5\ IBB_t \text{ and } 20\ IBB_t)\\&+(31.25\% \times \text{portion of pension basis between } 20\ IBB_t \text{ and } 30\ IBB_t),\end{aligned}$$

where the pension payout basis is the annual average salary in the 5 best years out of the last 7 years before retirement. For the year 2007, the pension payout specified by KAP-KL, whose definition is provided in Table IA.IV of this Internet Appendix.

Defined Contributions. The DC premia are not the same whether the employes is a union member of (i) OFR-förbunden and Akademikeralliansen or (ii) Kommunal. Unfortunately, the data set does not provide the union membership of individual employees. Since the OFR-förbunden and Akademikeralliansen have more members overall, we apply their pension rules to all local government employees.

For the period 1998-2006, the PFA 98 premium is

$$\begin{aligned}&3.5\% \times (\text{portion of } PQI_{i,t} \text{ up to } 7.5\ HBPA_t)\\&+1.1\% \times (\text{portion of } PQI_{i,t} \text{ above } 7.5\ HBPA_t),\end{aligned}$$

[10]Source: Sjögren & Wadensjö (2005:160,162,180) and to page 6 of the pension brochure of Arbetsgivarverket, “the Swedish Agency for Government Employers”,

where $PQI_{i,t}$ is the pension-qualifying income. For 2007, the KAP-KL yearly premium is[11]:

$$4\% \times (\text{portion of } PQI_{i,t} \text{ up to } 30\ IBB).$$

Between 1998 and 2004, DC premia are paid by local government employees who are at least 28 years old. In 2004, that age was reduced to 21 for many local government employees. DC premia continue to be paid if the worker is on sick leave, pregnancy leave, or parental leave.

1.4 Allocation rule of the PPM default fund.

To calculate retirement wealth held in the PPM system, we assume that it is invested fully in the AP7 fund, which is the default option in the DC public pension system (premiepension system or PPM). According to the AP7 fund policy during our sample period, each worker's PPM wealth is invested in only two funds: the AP7 Fixed Income Fund (*AP7 Räntefond*) and the AP7 Equity Fund (*AP7 Aktiefond*). The fixed income fund is effectively invested in the risk free rate. The equity fund is 130% levered in the unhedged MSCI World Index. The allocation depends on the age of the worker in the following way.

Between 0-55 years of age, the proportion is at 100. At 56 years old it is 97%. Between 57 and 60 years of age, it decreases by 3.5 percentage points per year to 83%. Between 61 and 65 years of age, it decreases by 3.2 percentage points per year to 67%. Between 66 and 70 years of age, it decreases by 3.4 percentage points per year until it is 50% at 70 years old.

To convert this proportion into the equity share invested in the unlevered MSCI World Index, we multiply the allocation in the equity fund by 130%.

1.5 Private pension contributions and wealth

This section discusses the organization of the private pension system in Sweden and the imputation of individual private pension contribution and wealth. The Swedish wealth register contains exact information on private pension contributions from 1994, and only reports a capped version of the variable from 1991. We impute full contributions from 1991 to 1993 using information in the years during which we observe both variables (i.e. 1994 onwards). More specifically, for each of the three education groups, we separately estimate the following individual fixed effect regression model over the years 1994 to 2007:

$$\log(T_{i,t}) = a_i + b_1 RI^1_{i,t} + ... + b_{20} RI^{20}_{i,t} + c \log(C_{i,t}) + c_1 C^1_{i,t} + ... + c_{20} C^{20}_{i,t} + c_{21} age^{21}_{i.t} + ... + c_{64} age^{64}_{i.t} + e_{i,t}$$

where T is the true contribution, C is the capped contribution, RI is the relevant income defined in Section 1.2.1 and variables with superscripts are dummy variables. Superscripts

[11]Source: OFR (2012, p. 21).

of RI and C dummies stand for the equally spaced percentiles (5%, 10%, ..., 90%, 95%) and superscripts of age dummies refer to the corresponding age of individuals. We then use the estimated model to predict T over the years 1991-1993.[12]

Since we do not have information on private pension contribution before 1991, we redistribute the aggregate stock of private pension wealth held by the working population in proportion to the individual share of private pension contributions in 1991. To calculate the stock of private pension wealth held by the working population we proceed in two steps. First, in 1991, the national accounts do not distinguish between private pensions and capital insurance (kapitalförsäkring), and only report the aggregate stock of both. We take the closest year national accounts have information on the split, and use the 80/130 fraction reported in 1989. Second, as noted in the main text, we follow Bach, Calvet and Sodini (2020) and further assume that 58% of private pension wealth in 1991 belongs to workers.

1.6 Taxation of DC Contributions and DB Payouts

The tax rate for withdrawing your pension before the age of 65 is higher than the tax rate for withdrawing the pension after the age of 65.

First, we calculate the tax rate on the pension income of already retired people, splitting them into three education groups (basic or missing education, high school, post-high school).

Second, we apply the obtained tax rate (specific for each education group) from the retired people and tax with it the DB payouts of the working/retiring people having the same education level. On average, the tax rate is almost invariably about 32% irrespective of education level. Additionally, the standard deviations of the tax rate by education level are negligibly small. We then tax the DC accumulations with a flat rate of 32%.

2 Empirical Methodology

2.1 Parameter Estimation

As the main text explains, we consider the labor income specification of Cocco, Gomes, and Maenhout (2005):

$$\log(Y_{h,t}) = a_c + b' x_{h,t} + \nu_{h,t} + \varepsilon_{h,t}, \tag{8}$$

where $Y_{h,t}$ denotes real income for household h in year t, a_c is a cohort fixed effect, $x_{h,t}$ is a vector of characteristics, $\nu_{h,t}$ is a permanent random component of income, and $\varepsilon_{h,t}$ is a transitory component.

[12] We exclude observations for which the true variable or the capped variable are equal to zero.

We enrich the Cocco, Gomes, and Maenhout model by distinguishing between shocks that are common to all households in a group and shocks that are specific to each household in the group. To simplify notation, we neglect the group index g in the rest of this section. We assume that the permanent component of income, $\nu_{h,t}$, is the sum of a group-level component, ξ_t, and an idiosyncratic component, $z_{h,t}$:

$$\nu_{h,t} = \xi_t + z_{h,t}. \tag{9}$$

The components ξ_t and $z_{h,t}$ follow independent random walks:

$$\xi_t = \xi_{t-1} + u_t, \tag{10}$$

$$z_{h,t} = z_{h,t-1} + w_{h,t}. \tag{11}$$

We assume that the three income shocks are i.i.d. Gaussian:

$$(u_t, w_{h,t}, \varepsilon_{h,t})' \sim \mathcal{N}(0, \Omega) \tag{12}$$

where Ω is the diagonal matrix with diagonal elements σ_u^2, σ_w^2, and σ_e^2.

We estimate the labor income process as follows. Consider

$$y_{h,t} = \log(Y_{h,t}) - a_c - b' x_{h,t}.$$

We infer from equation (**??**) that

$$y_{h,t} = \nu_{h,t} + \varepsilon_{h,t} = \xi_t + z_{h,t} + \varepsilon_{h,t}.$$

The sample mean of $y_{h,t}$, that is $\bar{y}_t = \sum_h y_{h,t}/N$ allows us to estimate ξ_t. The sample variance of $\bar{y}_t - \bar{y}_{t-1}$ $(t = 2, \ldots, T)$, provides an estimate of σ_u^2.

Let $y_{h,t}^* = y_{h,t} - \bar{y}_t$. We estimate σ_u^2 and σ_e^2 as in Carroll and Samwick.

2.2 Hypothesis Testing

We consider the null hypothesis

$$\mathbf{H_0} : R(\theta_g^*) = 0, \tag{13}$$

where $R(\theta)$ is a matrix of dimension $r \times \dim(\theta)$. We now present the Wald test and likelihood ratio-type test of this restriction.

Wald Test Under the null hypothesis, the unrestricted estimator satisfies

$$N_g R(\hat{\theta}_g^U)' \left[\frac{\partial R}{\partial \theta}(\hat{\theta}_g^U) \hat{V}_g \frac{\partial R}{\partial \theta'}(\hat{\theta}_g^U) \right]^{-1} R(\hat{\theta}_g^U) \xrightarrow{d} \chi^2(r), \tag{14}$$

which can be easily implemented.

We are interested in testing restrictions on preference parameters across groups. Examples include:

- rate of time preference homogeneous across groups,
- EIS homogeneous across groups,
- RRA homogeneous across groups.

We denote by $\Theta = (\theta_1', ..., \theta_G')'$ the vector of parameters in each group. We develop test of the null hypothesis

$$\mathbf{H_0} : R(\Theta) = 0, \tag{15}$$

where R is a matrix of dimension $r \times \dim(\Theta)$.

2.2.1 Wald Test

The Wald test only requires the estimation of *unconstrained* parameters. That is, we estimate preference parameters separately for each group (as explained in Section 1 of these notes). Let $N = N_1 + ... + N_g$ denote the total number of observations, and let

$$k_g = N_g/N$$

denote the fraction of observations in group g. Under the null,

$$NR(\hat{\Theta}^U)' \left[\frac{\partial R}{\partial \Theta}(\hat{\Theta}^U)\hat{V}^{-1}\frac{\partial R}{\partial \Theta'}(\hat{\Theta}^U)\right]^{-1} R(\hat{\Theta}^U) \xrightarrow{d} \chi^2(r), \tag{16}$$

where

$$\hat{V} = \begin{bmatrix} \hat{V}_1/k_1 & & \\ & \ddots & \\ & & \hat{V}_G/k_G \end{bmatrix}, \tag{17}$$

and $\hat{V}_g$ is given by (**??**). The result holds for fixed proportions k_g and for a total number of observations N going to infinity. We refer the reader to the Appendix for the derivation of this result.

3 Indirect Inference

The auxiliary estimator $\hat{\mu}_N$ is asymptotically normal:

$$\sqrt{N}\,[\hat{\mu}_N - \mu(\theta)] \to N(0, V^*).$$

We estimate θ by minimizing the criterion

$$[\hat{\mu}_N - \mu(\theta)]' \, \Omega \, [\hat{\mu}_N - \mu(\theta)] \, .$$

The system is exactly identified.

The indirect inference estimator $\hat{\theta}_N$ is asymptotically normal:

$$\sqrt{N}(\hat{\theta}_N - \theta^*) \to N(0, \Sigma),$$

where

$$\Sigma = \left(1 + \frac{1}{S}\right) \left[\frac{\partial \mu}{\partial \theta'}(\theta^*)\right]^{-1} V^* \left[\frac{\partial \mu}{\partial \theta}(\theta^*)\right]^{-1} .$$

In practice, $\partial \mu / \partial \theta(\theta^*)$ is estimated by numerical differentiation of μ around $\hat{\theta}$. The asymptotic variance covariance matrix of the auxiliary estimator, V^*, is estimated by the jackknife estimator:

$$\hat{V} = \frac{N-1}{N} \sum_{i=1}^{N} (\hat{\mu}^J_{[i]} - \overline{\mu^J})(\hat{\mu}^J_{[i]} - \overline{\mu^J})',$$

where $\hat{\mu}^J_{[i]}$ is the auxiliary estimator obtained by excluding the i^{th} observation, and $\overline{\mu^J} = N^{-1} \sum \hat{\mu}^J_{[i]}$.

3.1 Overidentified Case

We now consider a vector of auxiliary statistics $\hat{\mu}_N \in \mathbb{R}^q$. The vector of structural parameters is an element of $\mathbb{R}^p$, and we assume that $q > p$. In the current setting, $p = 3$ and $q = 16$. The efficient indirect inference estimator and the overidentification test are computed as follows.

3.1.1 Method 1

First, the asymptotic variance covariance matrix of the auxiliary estimator, $\hat{\mu}_N$, is estimated by the jackknife estimator

$$\frac{\hat{V}}{N} = \frac{N-1}{N} \sum_{i=1}^{N} (\hat{\mu}^J_{[i]} - \overline{\mu^J})(\hat{\mu}^J_{[i]} - \overline{\mu^J})',$$

where $\hat{\mu}^J_{[i]}$ is the auxiliary estimator obtained by excluding the i^{th} observation, and $\overline{\mu^J} = N^{-1} \sum \hat{\mu}^J_{[i]}$.

Second, we consider the weighting matrix

$$\hat{\Omega} = \hat{V}^{-1}$$

and compute the efficient indirect inference estimator

$$\hat{\theta}_N = \arg\min_{\theta} \left[\tilde{\mu}(\theta) - \hat{\mu}_N\right]' \hat{\Omega} \left[\tilde{\mu}(\theta) - \hat{\mu}_N\right].$$

The asymptotic variance-covariance matrix of the indirect inference estimator $\hat{\theta}_N$ is

$$\left(1 + \frac{1}{S}\right) \left[\frac{\partial \mu'}{\partial \theta}(\hat{\theta}_N)\hat{\Omega}\frac{\partial \mu}{\partial \theta'}(\hat{\theta}_N)\right]^{-1}.$$

We know from Gourieroux and Monfort (1993) that $\hat{\theta}_N$ is asymptotically efficient.

Third, we know that the objective function at the optimum satisfies

$$\frac{NS}{S+1}\left[\tilde{\mu}(\hat{\theta}_N) - \hat{\mu}_N\right]' \hat{\Omega} \left[\tilde{\mu}(\hat{\theta}_N) - \hat{\mu}_N\right] \to \chi^2(q-p),$$

which provides a convenient overidentification test.

3.1.2 Method 2

The previous method might not work well if the jackknife estimator is very noisy due the small size of the group. In this case, we could use a more standard multistep approach.

First, we compute

$$\hat{\theta}_N^{(1)} = \arg\min_{\theta} \left[\tilde{\mu}(\theta) - \hat{\mu}_N\right]' \hat{\Omega}^{(0)} \left[\tilde{\mu}(\theta) - \hat{\mu}_N\right],$$

where $\hat{\Omega}^{(0)}$ is an arbitrary weighting matrix.

Second, we compute a jackknife estimate of $\tilde{\mu}(\hat{\theta}_N^{(1)})$:

$$\frac{\hat{V}}{N} = \frac{NS-1}{NS}\sum_{i=1}^{NS}(\tilde{\mu}_{[i]}^J - \overline{\mu^J})(\tilde{\mu}_{[i]}^J - \overline{\mu^J})'$$

based on *simulations obtained under* $\hat{\theta}_N^{(1)}$. If S is large, we can obtain an accurate estimate of $\hat{V}$.

Third, we set $\hat{\Omega}^{(1)} = \hat{V}^{-1}$ and compute the second-stage indirect inference estimator.

$$\hat{\theta}_N^{(2)} = \arg\min_{\theta} \left[\tilde{\mu}(\theta) - \hat{\mu}_N\right]' \hat{\Omega}^{(1)} \left[\tilde{\mu}(\theta) - \hat{\mu}_N\right].$$

This estimator is asymptotically efficient.

Fourth, we compute

$$\frac{NS}{S+1}\left[\tilde{\mu}(\hat{\theta}_N^{(2)}) - \hat{\mu}_N\right]' \hat{\Omega} \left[\tilde{\mu}(\hat{\theta}_N^{(2)}) - \hat{\mu}_N\right] \to \chi^2(q-p),$$

which provides a convenient overidentification test.

3.2 Interpolation

For each group, we simulate the lifecycle model on a grid of preference parameters. The grid is defined by 12 values of the RRA ranging from 2 to 10, 11 values of the TPR ranging from -0.05 to 0.22, and 14 values of the EIS ranging from 0.1 and 2.5. The grid values of the RRA are 2, 3, 4, 4.5, 5, 5.5, 6, 6.5, 7, 8, 9, and 10. To contruct the grid values of the TPR, we assume that the patience parameter $\delta \in \{0.80, 0.89, 0.92, 0.94, 0.96, 0.97, 0.98, 0.99, 1.00, 1.05\}$, so that $TPR = -\ln(\delta)$ is contained in $\{-0.05, 0, 0.01, 0.02, 0.03, 0.04, 0.06, 0.08, 0.11, 0.16, 0.22\}$. The grid values of the EIS are 0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8, 1.2, 1.4, 1.6, 1.8, 2, and 2.5. Overall, the grid contains 1,848 ($= 12 \times 11 \times 14$) parameter vectors θ. For each vector θ on the grid, we calculate the value of the objective function:

$$[\tilde{\mu}_S^g(\theta) - \hat{\mu}^g]' \, \Omega \, [\tilde{\mu}_S^g(\theta) - \hat{\mu}^g] \, ,$$

where Ω is the weighting matrix defined in the main text.

We next evaluate the objective function on a finer grid defined as follows. The RRA grid has a grid step of 0.01 and contains 81 equally-spaced grid points ranging from 2 to 10. The EIS grid has grid step of 0.01 and contains 241 values of the EIS ranging from 0.1 to 2.5. We consider an evenly spaced grid of the patience parameter δ ranging from 0.8 to 1.05 with a grid step of 0.001, which generates a TPR grid containing 251 values of the TPR ranging from -0.05 to 0.22. The finer grid therefore contain 4,899,771 ($= 81 \times 241 \times 251$) preference parameters. We evaluate the objective function on the finer grid by interpolating the values computed on the original grid. We use modified Akima cubic Hermite interpolation, which is known to reduce interpolation overshoots and oscillations compared to standard spline methods.

For each group, we obtain the indirect inference estimator by determining the parameter vector on the finer grid that minimizes the objective function. This value may occasionally be slightly negative due to interpolation error.

For inference purposes, such as calculations of the root mean squared error or the Jacobian matrix, we compute each of the 16 auxiliary statistics by a separate interpolation. These interpolations are based on a cubic spline using not-a-knot end conditions.

3.3 Monte Carlo Simulations

Let $\hat{\theta}^g$ denote the parameter vector estimated for group g, and let

$$J_g = \left[\tilde{\mu}_S^g(\hat{\theta}^g) - \hat{\mu}^g\right]' \Omega \left[\tilde{\mu}_S^g(\hat{\theta}^g) - \hat{\mu}^g\right]$$

denote the corresponding objective function. We obtain a p value for J_g as follows.

- We simulate K histories of the yearly wealth-to-income ratio and risky share, $m_k^g \in \mathbb{R}^{16}$, $k = 1, ..., K$, under the data-generating process $\theta_0 = \hat{\theta}^g$.
- For each $k \in \{1, \dots, K\}$, we estimate the model by indirect inference, treating the simulated vector m_k^g as if it were the vector of empirical observations. That is, we compute
$$V_k^g = \min_\theta \left[\tilde{\mu}_S^g(\theta) - m_k^g\right]' \Omega \left[\tilde{\mu}_S^g(\theta) - m_k^g\right], \tag{18}$$
and we denote by θ_k^g the corresponding solution.
- The p value of J_g is the fraction of elements in the set of simulated objective functions, $\{V_1^g, \dots, V_K^g\}$, that are larger than the empirical objective function J_g.

We conduct the simulations for each $k \in \{1, ..., K\}$ as follows.

Simulation of the pseudo-history. We simulate $m_k^g \in \mathbb{R}^{16}$ by sampling idiosyncratic permanent and transitory income shocks of each household in the group. We apply the policy function to obtain each household's consumption and risky share each year. Finally, we aggregate up households to obtain the group's simulated wealth-to-income ratio and wealth-weighted average risky share at the end of every year t. This produces a vector m_k. Hence our methodology will account for the variability of the estimates due to the randomness of idiosyncratic labor income shocks.

For each group g, we simulate $K = 400$ paths of the wealth-income and risky share over $T = 8$ years and estimate on each path the vector of preference parameters. In every simulation and for every period t, we take as given the empirical wealth-income ratio of the group at the end of year $t - 1$ and the group-level income shocks and returns (common to all groups) in year t.[13] Each simulation of group g proceeds as follows for each period t.

i. We simulate the idiosyncratic permanent and transitory components of income of each household in the group, which produces the simulated income $\tilde{Y}_{i,t}$ and simulated permanent income $\tilde{Y}_{i,t}^P$ for every $i \in \{1, \dots, N^g\}$.

ii. Using the lifecycle model's policy functions $\alpha_t^*(\,\cdot\,)$ and $C_t^*(\,\cdot\,)$, we compute the risky share and consumption

$$\begin{aligned} \tilde{\alpha}_{i,t-1} &= \alpha_t^*(\tilde{Y}_{i,t}, W_{i,t-1}, \tilde{Y}_{i,t}^P; \theta), \\ \tilde{C}_{i,t} &= C_t^*(\tilde{Y}_{i,t}, W_{i,t-1}, \tilde{Y}_{i,t}^P; \theta), \end{aligned}$$

of each simulated unit during year t. Consistent with the model, the simulated unit sets both quantities at the end of year $t - 1$ and keeps them constant throughout year t.

[13] We assume that wages are announced at the beginning of each year.

iii. We compute the predicted wealth of each simulated unit i at the end of year t :

$$\hat{W}_{i,t} = (R_f + \tilde{\alpha}_{i,t-1} R^e_{p,t})(W_{i,t-1} + \tilde{Y}_{i,t} - \tilde{C}_{i,t}). \tag{19}$$

iv. At the end of each year t, this method produces the group's predicted wealth-income ratio:

$$\tilde{\hat{\mu}}^g_{1,t} = \frac{\sum_{i=1}^{N^g} \hat{W}_{i,t}}{\sum_{i=1}^{N^g} \tilde{Y}_{i,t}} \tag{20}$$

and predicted risky share

$$\tilde{\hat{\mu}}^g_{2,t} = \frac{\sum_{i=1}^{N^g} \tilde{\alpha}_{i,t} \tilde{W}_{i,t}}{\sum_{i=1}^{N^g} \tilde{W}_{i,t}}, \tag{21}$$

at the end of year t. We stack these values into the column vector m^g_k.

Estimation. We apply the methodology of Section 4 of the main text. Hence the estimation will be based on one-step forecasts of the wealth to income ratio and risky share.

This analysis will quantify how the variability of the risky asset returns and income shocks impact the variability of the estimated parameters and coefficients. We can use the pseudo estimates θ^g_k to plot histograms/box plots assessing the accuracy of our estimation methodology.

3.4 Measuring Parameter Heterogeneity

We denote by $\mathbb{E}$ the expectation operator computed across random realization of income shocks and by E^* the cross-sectional expectation operator. The mean preference parameter vector in the population is $\mu_\theta = E^*(\theta) = \sum_{g=1}^{G} w_g \theta^g$, where w_g is the population share of group g and θ^g is the preference vector of households in the group. Similarly, $V_\theta = E^*[(\theta - \mu_\theta)(\theta - \mu_\theta)']$ is the variance-covariance matrix of the preference vector.

Our objective is to estimate V_θ from the group-level indirect inference estimators $\hat{\theta}^g$ $(g = 1, \ldots, G)$ defined in the main text. For every group g, the indirect inference estimator can be written as

$$\hat{\theta}^g = \theta^g + b^g(\theta^g) + \epsilon^g,$$

where

$$b^g(\theta^g) = \mathbb{E}_{\theta^g}(\hat{\theta}^g - \theta^g)$$

is the bias of the estimator for a group of size N_g, and ϵ^g is the estimation noise: $\mathbb{E}(\epsilon^g) = 0$. The estimation of V_θ requires us to control for both the bias and the noise in $\hat{\theta}^g$.

Bias Adjustment. Since the parameter space Θ is bounded, the bias and the error are bounded as well. For every group g, we know that $\sqrt{N^g}(\hat{\theta}^g - \theta^g)$ converges in distribution to $\mathcal{N}(0, V_g)$. Since b^g and ϵ^g are bounded, convergence in distribution implies convergence in expectation,[14] so that $\sqrt{N^g}\, b^g(\theta^g)$ converges to 0 as N^g goes to infinity. We therefore write that $b^g(\theta^g) = o(1/\sqrt{N^g})$.

We adjust the empirical estimator $\hat{\theta}_g$ to correct the estimation bias. This step is based on the Monte Carlo estimates $\theta_1^g, \ldots, \theta_K^g$ of simulated datasets obtained under $\hat{\theta}_g$. We let

$$\bar{\theta}^g = \frac{1}{K}\sum_{k=1}^{K} \theta_k^g.$$

We estimate of the bias $b^g(\theta^g)$ in a sample of size N^g by

$$\hat{b}_g = \bar{\theta}^g - \hat{\theta}^g = \frac{1}{K}\sum_{k=1}^{K}(\theta_k^g - \hat{\theta}^g).$$

We define the bias-adjusted estimator by

$$\hat{\hat{\theta}}^g = \hat{\theta}^g - \hat{b}_g = 2\hat{\theta}^g - \bar{\theta}^g,$$

for every g. The estimator $\hat{\hat{\theta}}^g$ is approximately unbiased.

Lemma 1 *The adjusted estimator satisfies*

$$\hat{\hat{\theta}}^g = \hat{\theta}^g - b(\theta^g) + o_p\left(\frac{1}{N^g}\right)$$

as N^g and K/N^g go to infinity.

Proof. There is a small bias due to the fact that we simulate the data under $\hat{\theta}^g$, not under θ^g. Specifically, we observe that

$$\sqrt{N^g\, K}[\hat{b}^g - b(\hat{\theta}^g)] = \sqrt{N^g}\sqrt{K}\left[\frac{1}{K}\sum_{k=1}^{K}\theta_k^g - \mathbb{E}(\theta_1^g)\right]$$

converges in distribution. We infer that

$$\hat{b}^g \quad = \quad b(\hat{\theta}^g) + O_p\left(\frac{1}{\sqrt{K\, N^g}}\right)$$

[14]See, e.g., Billingsley (1999) for further details.

The condition $K/N^g \to \infty$ guarantees that simulation noise does not create a larger bias than the one it is intended to correct. Hence:

$$\hat{b}^g = b(\hat{\theta}^g) + o_p\left(\frac{1}{N^g}\right) \tag{22}$$

We next expand the bias function around θ^g:

$$b(\hat{\theta}^g) = b(\theta^g) + \frac{\partial b}{\partial \theta'}(\theta^g)\,(\hat{\theta}^g - \theta^g).$$

Since $\partial b/\partial \theta'^g) = o(1/\sqrt{N^g})$ and $\hat{\theta}^g - \theta^g = O_p(1/\sqrt{N^g})$ we infer that

$$b(\hat{\theta}^g) - b(\theta^g) = o_p\left(\frac{1}{N^g}\right). \tag{23}$$

Combining (**??**) and (**??**), we obtain:

$$\hat{b}^g - b(\theta^g) = o_p\left(\frac{1}{N^g}\right).$$

Hence $\hat{\hat{\theta}}^g = \hat{\theta}^g - \hat{b}^g = \hat{\theta}^g - b(\theta^g) + o_p(1/N^g)$ and we conclude that the proposition holds. ■

Noise Adjustment. The estimation of V_θ requires us to take account of estimation noise. We let

$$\begin{aligned}(\hat{\hat{\theta}}^g - \mu_\theta)(\hat{\hat{\theta}}^g - \mu_\theta)' &= (\hat{\hat{\theta}}^g - \theta^g)(\hat{\hat{\theta}}^g - \theta^g)' + (\theta^g - \mu_\theta)(\theta^g - \mu_\theta)' \\ &\quad + (\hat{\hat{\theta}}^g - \theta^g)(\theta^g - \mu_\theta)' + (\theta^g - \mu_\theta)(\hat{\hat{\theta}}^g - \theta^g)'.\end{aligned}$$

We take the expectations across realization of income shocks:

$$\begin{aligned}\mathbb{E}\left[(\hat{\hat{\theta}}^g - \mu_\theta)(\hat{\hat{\theta}}^g - \mu_\theta)'\right] &= \mathbb{E}\left[(\hat{\hat{\theta}}^g - \theta^g)(\hat{\hat{\theta}}^g - \theta^g)'\right] + (\theta^g - \mu_\theta)(\theta^g - \mu_\theta)' \\ &\quad + \mathbb{E}(\hat{\hat{\theta}}^g - \theta^g)\;\;(\theta^g - \mu_\theta)' + (\theta^g - \mu_\theta)\;\;\mathbb{E}(\hat{\hat{\theta}}^g - \theta^g)'.\end{aligned}$$

We apply the cross-sectional expectation operator and obtain:

$$\Psi^{(1)} = \Psi^{(2)} + V_\theta + \Psi^{(3)} + [\Psi^{(3)}]',$$

where

$$\begin{aligned}\Psi^{(1)} &= E^*\mathbb{E}\left[(\hat{\hat{\theta}}^g - \mu_\theta)(\hat{\hat{\theta}}^g - \mu_\theta)'\right], \\ \Psi^{(2)} &= E^*\mathbb{E}\left[(\hat{\hat{\theta}}^g - \theta^g)(\hat{\hat{\theta}}^g - \theta^g)'\right], \\ \Psi^{(3)} &= E^*\left[\mathbb{E}(\hat{\hat{\theta}}^g - \theta^g)\;\;(\theta^g - \mu_\theta)'\right].\end{aligned}$$

The cross-sectional variance-covariance matrix of the preference parameter vector satisties

$$V_\theta = \Psi^{(1)} - \Psi^{(2)} - \Psi^{(3)} - \left[\Psi^{(3)}\right]'.$$

When the estimators are unbiased, this equation is equivalent to

$$E^*\mathbb{E}\left[(\hat{\hat{\theta}}^g - \mu_\theta)(\hat{\hat{\theta}}^g - \mu_\theta)'\right] = E^*\left[Var(\hat{\hat{\theta}}^g)\right] + V_\theta,$$

as the law of total variance implies. By the lemma, this formula can be applied since $\Psi^{(3)} = E^*(1/N^g)o_p(1)$. In finite samples, we estimate V_θ by as follows.

First, we estimate μ_θ by the size-weighted mean of group estimates:

$$\bar{\theta} = \sum_g w_g \hat{\hat{\theta}}^g,$$

where $w_g = N^g/(\sum_k N^k)$ is the share of group g in the population. The estimator $\bar{\theta}$ is a consistent estimator of μ_θ as the group sizes $N^1, \ldots, N^G$ go to infinity.

Second, we estimate $E^*\,\mathbb{E}\left[(\hat{\hat{\theta}}^g - \mu_\theta)(\hat{\hat{\theta}}^g - \mu_\theta)'\right]$ by the size-weighted variance of group estimates:

$$\hat{\Phi}^{(1)} = \sum_{g=1}^{G} w_g(\hat{\hat{\theta}}^g - \bar{\theta})(\hat{\hat{\theta}}^g - \bar{\theta})'.$$

We estimate $E^*\left[Var(\hat{\hat{\theta}}^g)\right]$ by the average variance-covariance matrix of $\hat{\theta}^g$:

$$\hat{\Phi}^{(2)} = \sum_{g=1}^{G} w_g \frac{\hat{V}^g}{N^g},$$

where $\hat{V}^g$ is the asymptotic variance-covariance matrix defined in the main text. A loose explanation is that $\hat{\hat{\theta}}^g$ and $\hat{\hat{\theta}}^g$ differ by a constant, but this needs to be tidied up further. We therefore estimate the variance-covariance of θ by

$$\hat{V}_\theta = \hat{\Phi}^{(1)} - \hat{\Phi}^{(2)}.$$

PROOF

Wald test

By (**??**), the unconstrained estimator for each group g satisfies

$$\sqrt{N}(\hat{\theta}_g^U - \theta_g) \to N(0, V_g/k_g).$$

The unconstrained SMM estimator, $\hat{\Theta}^U = (\hat{\theta}_1^{U\prime}, ..., \hat{\theta}_G^{U\prime})$, is therefore asymptotically normal:

$$\sqrt{N}(\hat{\Theta}^U - \Theta) \xrightarrow{d} \mathcal{N}(0, V), \tag{24}$$

where the asymptotic variance-covariance matrix is block-diagonal:

$$V = \begin{bmatrix} V_1/k_1 & & \\ & \ddots & \\ & & V_G/k_G \end{bmatrix}. \tag{25}$$

Note that the variance-covariance matrix is conditional on the realization of the common shocks. In practice, we estimate V by the block diagonal matrix (**??**).

References

[1] Anderson, Karen M., 2015, Occupational pensions in Sweden, Working paper, Friedrich Ebert Stiftung.

[2] Barr, Nicholas, 2013, The pension system in Sweden, *Report to the Expert Group on Public Economics* 2013:7, Stockholm: Ministry of Finance.

[3] Granqvist, Lena, and Ann-Charlotte Ståhlberg, 2002, Ökad Jämställdhet- men fortfarande sämre pension för kvinnor, Working paper, Pensions Forum.

[4] Hagen, Johannes, 2015, The determinants of annuitization: evidence from Sweden, *International Tax and Public Finance* 22, 549-578.

[5] Rudolp, Heinz P., Roberto Rocha, and Dimitri Vittas, 2010, The payout phase of pension systems. A comparison of 5 countries, Working paper, The World Bank.

[6] Röstberg, Anna, Björn Andersson, and Thomas Lindh, 2004, Simulating the future pension wealth and retirement saving in Sweden, Working paper, Institute for Futures Studies.

[7] Sjögren Lindqvist, Gabriella, and Eskil Wadensjö, 2005, Inte bara socialförsäkringar, Working paper, Stockholm University.

[8] Swedish Ministry of Health and Social Affairs, 2009, The Swedish pension agreement and pension reform. Available at http://www.government.se

[9] Swedish Pensions Agency, *Orange Report: Annual Report of the Swedish Pension System*, Years 1999 to 2015, Stockholm. Available at https://www.pensionsmyndigheten.se

Table 1: Sizes of Education and Income Risk Categories

	No High School	High School	Post-High School	All
Low	17,980	43,613	66,006	127,599
Medium	10,531	51,810	28,530	90,871
High	16,212	28,590	28,216	73,018
All	44,723	124,013	122,752	291,488

This table reports number of households in groups with 3 levels of education and working in sectors with 3 levels of income volatility given in Table 2 of the main text, and for aggregates of these groups.