Data Article

Data projections of the replacement ratios for the cohorts of retirees in Chile under different policy reforms

Carlos Madeira
Central Bank of Chile, Chile

ARTICLE INFO

Article history:
Received 18 May 2022
Revised 6 July 2022
Accepted 1 August 2022
Available online 5 August 2022

Keywords:
Pensions
Social security system
Income redistribution
Contributory pensions
Retirement age

ABSTRACT

This article provides five panel datasets for the projections of the mean replacement ratios of pension income relative to the worker's income. The time dimension is from year 2022 until 2055. The panel groups consider the gender, income and education of the workers. Furthermore, the variables consider different scenarios for the social security system: (1) a baseline with the current policies, (2) an increase of retirement age to 67 years, (3) an increase of the retirement age to 67 years and a 6% increase of the contribution rate, with a rate of 0–3% going to solidarity funds.

© 2022 The Author(s). Published by Elsevier Inc. This is an open access article under the CC BY license (http://creativecommons.org/licenses/by/4.0/)

Specification Table

| Subject                      | Economics, Econometrics and Finance. |
|------------------------------|--------------------------------------|
| Specific subject area        | Household Finance. Pensions. Social security. Labor Economics. |
| Type of data                 | Table (Excel format and Stata dta files) |
| How data were acquired       | Data uses publicly available raw data from the Chilean Family Expenditures Survey (EPF, 2017 wave) to calibrate the future labor income and pension contributions of each worker until retirement. |

(continued on next page)

DOI of original article: 10.1016/j.jimonfin.2022.102650
E-mail address: cmadeira@bcentral.cl
Social media: @CarlosM04459173

https://doi.org/10.1016/j.dib.2022.108516
2352-3409/© 2022 The Author(s). Published by Elsevier Inc. This is an open access article under the CC BY license (http://creativecommons.org/licenses/by/4.0/)
Hardware: data analysis was performed in a standard notebook with an Intel Core i7-4700HQ 2.40 GHz processor with 16.0 GB of RAM. Software: Stata MP-6 (version 15.1).

Data format: Analyzed

Description of data collection: Data consists of different panel groups (given by the gender, education, and income of the workers) and the replacement ratios of retiree income cohorts at a yearly frequency between 2022 and 2055. The datasets consider 6 different scenarios for the future public policy. The time horizons in which the policies are analyzed start in 2022 (because the last Chilean pension policy was implemented in January of 2022) and end in 2055 (which is when the youngest workers of the present period start retiring and the alternative pension policies will have a full effect).

Data source location: Institution: Instituto Nacional de Estadísticas (INE, in English, Bureau of Official Statistics)
City/Town/Region: Santiago
Country: Chile
https://www.ine.cl/estadisticas/sociales/ingresos-y-gastos/encuesta-de-presupuestos-familiares (EPF)

Data accessibility: With the article.
Repository name: Mendeley Data
Data identification number (DOI): 10.17632/59j9wkwd8.1.
Direct URL to data: https://data.mendeley.com/datasets/59j9wkwd8/1

Related research article: C. Madeira, The impact of the Chilean pension withdrawals during the Covid pandemic on the future savings rate, J Int Money Finance, 126 (2022) 102,650.

Value of the Data

- The data can be used to study the forecasted Chilean retirees’ pension income as a replacement ratio of their last labor earnings over the period between 2022 and 2055 (Madeira [1]). Such studies can be done by demographic groups (by gender, income or education) and over time at an annual frequency until 2055.
- The data shows how the replacement ratios are affected by different policies such as an increase in the contribution rate and delaying of the retirement age.
- Since the pension system represents a large component of national wealth and savings in most countries (OECD [5]), the data can contribute to the study of the low savings rate in Latin America (Cavallo and Serebrisky [2]) and the policies necessary to solve this problem in Chile.

1. Data

The dataset provides projections of the replacement ratios of retirement income relative to the last income of the worker (Table 1). Average values are reported for the cohorts of future retirees in Chile from 2022 until 2055. The time of the analysis starts in 2022, because the last Chilean pension policy was substantially changed due to recent policies, such as the three pension withdrawals (August of 2020, December of 2020, April of 2021) and a non-contributory public pension increase in January of 2022 (Madeira [1]). The data ends in 2055 when the different policy alternatives reach a full effect.

Projections are given for 6 different scenarios: “Current policies,” “Increase of retirement age to 67 years,” “Increase in age plus 6% increase in personal contribution rate,” “Increase in age, 5% increase in contribution rate, 1% solidarity tax rate,” “Increase in age, 4% increase in contribution rate, 2% solidarity tax rate,” “Increase in age, 3% increase in contribution rate, 3% solidarity tax rate.” There are 5 datasets which express average panel time series for 5 different populations: (X1) All workers, (X2) by gender, (X3) by gender and worker income quartile, (X4) by gender and education, (X5) by gender, education, and income quartile. The same datasets are available in both Excel and Stata formats.
Table 1
Description of the dataset (SRates_Dib.xlsx, plus its Stata .dta version) provided in this article.

| Analyzed dataset | Description |
|-------------------|-------------|
| ReplacementRatios_X1.xlsx (ReplacementRatios_X1.dta) | Forecasts for the period 2022–2055 of the replacement ratios of each new cohort of retirees. The cohorts consider the entire population of new retirees. |
| ReplacementRatios_X2.xlsx (ReplacementRatios_X2.dta) | Forecasts for the period 2022–2055 of the replacement ratios of each new cohort of retirees, with cohorts by gender. |
| ReplacementRatios_X3.xlsx (ReplacementRatios_X3.dta) | Forecasts for the period 2022–2055 of the replacement ratios of each new cohort of retirees, with cohorts by gender and education (5 categories of worker completed education). |
| ReplacementRatios_X4.xlsx (ReplacementRatios_X4.dta) | Forecasts for the period 2022–2055 of the replacement ratios of each new cohort of retirees, with cohorts by gender and national labor income quartile (4 categories). |
| ReplacementRatios_X5.xlsx (ReplacementRatios_X5.dta) | Forecasts for the period 2022–2055 of the replacement ratios of each new cohort of retirees, with cohorts by gender, education (5 categories) and national labor income quartile (4 categories). |

This is the list of variables available in the panel datasets, with a summary description. The list is organized according to 3 categories of variables: (1) panel groups (demographic groups by gender, education, income or the whole population of new retirees), (2) time variable (year of retirement of the cohort of new retirees), and (3) the replacement ratios of retirement income relative to the last wage, according to the different public policies adopted for the pension system.

Demographic variables (gender, education, national income quartile):
These variables describe the cohort of retirees in terms of their gender and education.

1. Pop “Entire population” (this variable always takes a value of 1).
2. Sexo “Gender of the worker (1 Male, 2 Female)”.  
3. Educ_ecf “Education level of the respondent,” with values 1 “Elementary education” 2 “Secondary education” 3 “Technical or Some college” 4 “College education” 5 “Post-graduate education”.
4. Quartile “Worker income quartile: 1 lowest income workers (25% lowest earnings workers), 2 (workers between the 25 and 50% of highest labor earnings), 3 (workers between the 50 and 75% of highest labor earnings), 4 highest income (workers with the 25% highest labor earnings).”

Year of the retirement for each cohort:
This variable corresponds to the year in which each cohort of workers enters retirement, that is, the first year in which the retirees receive their pension.

1. YearR “Year of retirement for each cohort (yearly frequency between 2022 and 2055)”.

Replacement ratios of retirement income relative to worker’s last income according to the public policy scenario adopted between 2022 and 2055:
These variables correspond to the average replacement ratio of the pension income of each cohort of retirees during their first year of retirement relative to their labor income during their last working year. The alternative policies analyzed include the current policies as a baseline, an increase of the retirement age to 67 years, plus policies that increase the personal contribution (between 3 and 6%) and the solidarity tax rate (between 0 and 3%) in addition to the increase of the retirement age.

1. rr_ratioTot_B “Current policies”.
2. rr_ratioTot_S1 “Increase of retirement age to 67 years”.
3. rr_ratioTot_S2 “Increase of retirement age to 67 years plus 6% increase in personal contribution rate”.

(1) Table 1
(4) \( \text{rr ratioTot}_S3R2 \) “Increase of retirement age to 67 years, 5% increase in contribution rate, 1% solidarity tax rate”.

(5) \( \text{rr ratioTot}_S4R2 \) “Increase of retirement age to 67 years, 4% increase in contribution rate, 2% solidarity tax rate”.

(6) \( \text{rr ratioTot}_S5R2 \) “Increase of retirement age to 67 years, 3% increase in contribution rate, 3% solidarity tax rate”.

2. Experimental Design, Materials and Methods

The data consists of a simulation of the future contributory pensions plus public solidarity benefits for a sample of Chilean households (Madeira [1]), with mean forecasts presented for each new cohort of retirees between 2022 and 2055. The panel groups are given by gender, education and income.

The data uses the pooled cross-section sample of households from the Chilean Family Expenditures Survey (in Spanish, Encuesta de Presupuestos Familiares, hence on EPF) wave of 2017. The calibration estimates the worker’s average yearly labor income using their job income while working and their average time in unemployment, with unemployment risk given according to 538 workers’ types which are obtained from the multivariate vector of the workers’ sex, age, education, industry and region (Madeira [3]). The life expectancy and population weights of workers of each gender (male, female) with different ages are then adjusted over time using population forecasts from the United Nations (Madeira [1]).

The methodology starts by first computing the permanent income of each worker \( k \) at time \( t \) (\( P_{k,t} \)), according to its average time employed (\( 1 - u_{k,t} \)) versus unemployed (\( u_{k,t} \)):

\[
P_{k,t} = G_k W_{k,t-1} \left( 1 - u_{k,t} - u_{k,t} RR_k \right),
\]

with \( G_k, W_{k,t}, u_{k,t}, RR_k \) denoting, respectively, the yearly growth rate of labor income in the occupation of worker \( k \), the wage of worker \( k \) in its occupation in the previous year, the unemployment probability of \( k \)’s worker type, and the replacement ratio of labor income during unemployment. These components of the model are obtained for 538 workers’ types which are obtained from the multivariate vector of the workers’ sex, age, education, industry and region (Madeira [3]).

I then obtain the worker’s contributions (\( PWI_{k,t} \)) to the social security retirement system until age \( S(t,k) \):

\[
PW_{k,t} = \sum_{h=25}^{S(t,k)-1} cr_h r_h \min \left( \frac{mc_h}{P_{k,h}}, P_{k,h} \right) pc_{k,h} - r_{2021} pw_{k}^{1+2+3}
\]

with \( cr_h, r_h, mc_h, P_{k,h}, pc_{k,h} \) denoting, respectively, the contribution rate, the average rate of return between time \( h \) and time \( t \), the top contributory income value, the permanent income, and the probability of contribution (given by the probability of being both in the labor force and in forma work) at time \( h \). \( pw_{k}^{1+2+3} \) denotes the wealth withdrawn by each worker from its pension account during the pandemic.

When worker \( k \) at age \( R_k \) in year \( t \), its accumulated pension turns into a monthly annuity for their life, \( p_{a_{k,t}} = \frac{rp_{k,t}(1/\beta)^{r_{k,t}}}{1-(1/\beta)^{r_{k,t}-S(t,k)}} \), with \( T_{k,t}, \beta, r = (1/\beta) - 1 \) denoting, respectively the worker’s life expectancy, the discount factor and the risk neutral interest rate.

Each member of a family within the nine lowest deciles of income (\( SB_d \)) is also entitled to solidarity benefits according to the law of 2022 (\( B_{k,t} \)) and to additional solidarity benefits funded from the “solidarity tax rate” that may be implemented in the future.

The new solidarity benefits of each retiree \( k \) can therefore be expressed as

\[
B_{k,t} = b_1 1(p_{a_{k,t}} < b_2) + b_1 \left(1 - \frac{p_{a_{k,t}} - b_2}{b_3 - b_2}\right) 1(b_2 < p_{a_{k,t}} < b_3),
\]
with $b_1 = 185,000$, $b_2 = 630,000$, and $b_3 = 1,000,000$. Therefore, retirees receive 185,000 pesos as solidarity benefits if their contributory pension income is below 630,000 pesos and then receive gradually fewer solidarity benefits until their contributory pension income is already one million Chilean pesos.

The total pension $TP_{k,t}$ is therefore given by:

$$TP_{k,t} = pa_{k,t} + SB_k(B_{k,t} + SP_{k,t}).$$  \hspace{1cm} (4)

with $SP_{k,t} = \frac{rSPT_{k,t}(1/\beta)^{B_{k,t} - S_{k,t} - S_{k,t}^T}}{1-(1/\beta)^{-2(A_{k,t} - S_{k,t}^T - S_{k,t})}}$ and $SPT_{k,t} = \sum_{h=0}^{T_{k,t} - R_k} \beta^h(MP_t - (pr_k - 1))$ denoting, respectively, the new solidarity benefits received by worker $k$ retiring in year $t$ and the total discounted value of all the new solidarity benefits that this worker will receive during its life in retirement. In this formula, $MP_t = \frac{Sc}{NR_t} + \frac{mp-1}{2}$, with $Sc_t$, $NR_t$ denoting, respectively, the total funds collected through the new solidarity tax in each year and the total number of retirees receiving solidarity benefits. In this expression $mp = 10,001$ is the total number of groups among whom the solidarity funds are distributed and $pr_k$ is the ranking of the group of retiree $k$ among the total number of retirees that can benefit from the solidarity system. This mechanism ensures that more solidarity funds are distributed towards those of lowest income.

The calibration of the permanent income requires estimating the yearly growth rate of labor income in the occupation of worker $k$, the wage of worker $k$ in its occupation in the previous year, the unemployment probability of $k$’s worker type, and the replacement ratio of labor income during unemployment. These heterogeneous parameters were obtained from the rotating panel samples of the Chilean Employment Survey (Madeira [3]) from the labor earnings of each individual worker $i$ of the sample in year $t$ ($y_{i,t}$), the characteristics of each worker ($x_i = x_k$) and their unemployment status ($U_{i,t} = 1$):

$$G_k = \frac{E[y_{i,t} | t = h, x_i = x_k]}{E[y_{i,t} | t = h - 1, x_i = x_k]} - 1;$$  \hspace{1cm} (5)

$$W_{k,t} = \max(y_{i,t}, E[y_{i,t} | t = h, x_i = x_k]);$$  \hspace{1cm} (6)

$$u_{k,t} = Pr(U_{i,t} = 1 | t, x_i = x_k);$$  \hspace{1cm} (7)

$$RR_k = \frac{E[y_{i,t} | t = h, U_{i,t} = 1, x_i = x_k]}{E[y_{i,t} | t = h, U_{i,t} = 0, x_i = x_k].}$$  \hspace{1cm} (8)

These parameters are obtained for each year of the Chilean Employment Sample since 1990 until now for the workers’ types given by the multivariate vector $x_k$ of the workers’ sex, age, education, industry and region (Madeira [3]). Finally, the probability of contribution (given by the probability of being both in the labor force and in forma work) at time $t$ is also obtained from the Chilean Employment Survey. The calibration of the worker’s career paths for the future years is kept similar to the year 2017 (which had the lowest unemployment rate and is therefore close to full employment), but the parameters take into account the aging of the workers.

Finally, to calculate the savings and pensions for each cohort retiring between 2018 and 2055, each year uses the life expectancy estimates after age 60 for each gender in Chile to obtain $T_{k,t}$ and adjusts the population weights as follows: $w_{k,t} = w^{EPF}_k Pop_t(s_k, age_k)/Pop_{2017}(s_k, age_k)$, with $w^{EPF}_k$ denoting the original EPF weights in 2017 and $Pop_t(s_k, age_k)$ being the number of people in each sex-age bracket. Life expectancy $T_{k,t}$ and population by sex-age $Pop_t(s_k, age_k)$ for each year $t$ are obtained from United Nations estimates.

Users can download (at no cost) the raw data of all the EPF surveys, including the 2017 wave, from the website of the Chilean Institute of National Statistics.

EPF: [https://www.ine.cl/estadisticas/sociales/ingresos-y-gastos/encuesta-de-presupuestos-familiares](https://www.ine.cl/estadisticas/sociales/ingresos-y-gastos/encuesta-de-presupuestos-familiares).

The datasets are available at Mendeley Data (Madeira [4]): [data.mendeley.com/datasets/59j9wkvd8w/1](data.mendeley.com/datasets/59j9wkvd8w/1).
Declaration of Competing Interest

The author declares that he has no known competing financial interests or personal relationships which have or could be perceived to have influenced the work reported in this article. I received no funding from any institution besides my employer which is the Central Bank of Chile. Furthermore, there are no patents or impediments to publication, including the timing of publication, with respect to the intellectual property of the article or the associated dataset.

Data Availability

Projections of the replacement ratios for the cohorts of retirees in Chile for the period 2022 until 2055 (Original data) (Mendeley Data).

CRediT Author Statement

Carlos Madeira: Conceptualization, Methodology, Software, Validation, Formal analysis, Investigation, Resources, Data curation, Writing – original draft, Writing – review & editing, Visualization, Supervision, Project administration, Funding acquisition.

Acknowledgments

The author thanks comments and suggestions from Solange Berstein, Rodrigo Alfaro and Patricio Toro. All errors are my own.

References

[1] C. Madeira, The impact of the Chilean pension withdrawals during the Covid pandemic on the future savings rate, J. Int. Money Financ. 126 (2022) 102650, doi:10.1016/j.jimonfin.2022.102650.
[2] E. Cavallo, T. Serebrisky, Saving For Development: How Latin America and the Caribbean Can Save More and Better, Inter-American Development Bank, DC., USA, 2016.
[3] C. Madeira, Identification of earnings dynamics using rotating samples over short periods: the case of Chile, Central Bank of Chile Work Paper 754 (2015).
[4] C. Madeira, Projections of the replacement ratios for the cohorts of retirees in Chile for the period 2022 until 2055, Mendeley Data V1 (2022), doi:10.17632/59j9wkwv8d.1.
[5] OECD, Pensions at a Glance, OECD, Paris, 2021.