

PENSION REFORMS AND THE HEALTH DISTRIBUTION OF RETIREES

INTRODUCTION AND SUMMARY

Axel Börsch-Supan and Courtney Coile

Final version 18 November 2025

ABSTRACT

The *International Social Security (ISS) project* compares the experiences of a dozen developed countries to study *Social Security Programs and Retirement Around the World*. The project was launched in the mid 1990s and was motivated by decades of decline in the labor force participation rate of older men. The first phases of the project documented that social security program provisions can create powerful incentives for retirement that are strongly correlated with the labor force behavior of older workers. Since then, countries have undertaken numerous reforms of their social security programs, disability programs, and other public benefit programs available to older workers. In a second stage of this project, we found that these reforms substantially reduced the implicit tax on work at older ages and that stronger financial incentives to work were positively correlated with labor force participation at older ages. In a third stage, we exploited time-series and cross-national variation in the timing and extent of reforms of retirement incentives and employed micro-econometric methods in order to show that the rising participation rates since the end of the 1990s have been caused by the pension reforms, in particular by the sharply increased financial incentives to work at older ages.

The pension reforms from the 1980s through 2020 may therefore be celebrated as a success story in fostering old-age labor force participation, which is important in the face of rapid demographic aging. However, there may be negative side effects as the reforms may have increased the income inequality among retirees. In the most recent stage of the project, we investigated this and found a heterogeneous picture with six out of ten countries where the reforms actually decreased income inequality. The main question to be answered by this final phase of the project is whether the reforms have increased health inequality among retirees. This is motivated by the fact that increasing the retirement age may harm workers in bad health more than those who are healthy. We use internationally comparable panel data, five different dimensions of health, and several ways to express health inequality. We do not find that health inequality has increased during the last two decades.

Axel Börsch-Supan
Munich Research Institute for the Economics of Aging
Max Planck Institute for Social Law and Social Policy
Leopoldstrasse 139
80804 Munich, Germany
and NBER
boersch-supan@mea.mpg.de

Courtney Coile
Department of Economics
Wellesley College
106 Central Street
Wellesley, MA 02481
and NBER
ccoile@wellesley.edu

Acknowledgements

This paper is part of the National Bureau of Economic Research's International Social Security (ISS) project. This phase of the ISS Project is supported by the Sloan Foundation (G-2019-12578). The findings and conclusions expressed are solely those of the authors and do not represent the views of the Sloan Foundation or NBER. The authors thank the other members of the ISS Project for important methodological contributions and for the country-specific analyses on which this chapter is based. We thank Johannes Rausch and Luca Salerno for their assistance and participants in the NBER's Summer Institute Aging Workshop for helpful comments.

1. Project Overview

Over its twenty-five year history, the International Social Security (ISS) project has used the vast differences in social security programs across countries as a natural laboratory to study the effects of retirement program provisions on the labor force participation of older persons and other questions related to the older workforce. Motivated by the decline in older men's labor force participation during the 20th century and the possibility that the spread of public pensions contributed to this development, the project's early analyses (Gruber and Wise, 1999, 2004, and 2007) documented the strong relationship across countries between social security incentives and older men's labor force participation that existed in the late 1990s, confirmed this relationship in microeconomic analysis, and estimated the labor market and fiscal implications of social security reforms. Later volumes have examined the relationship between disability insurance program provisions, health, and retirement (Wise, 2012 and 2016) and explored whether older employment affects youth unemployment (Gruber and Wise, 2010) and whether older workers are healthy enough to work longer (Wise, 2017).

Since the project's inception, there has been a remarkable reversal of employment trends, with dramatic increases in employment at older ages for both men and women in all of the ISS countries. More recent phases of the ISS project have therefore examined potential explanations for these changes in behavior. The volume edited by Coile, Milligan, and Wise (2019) examined cohort changes in health and education. The main finding was a surprisingly weak correlation between the development of old-age labor force participation and factors such as cohort changes in health and education.

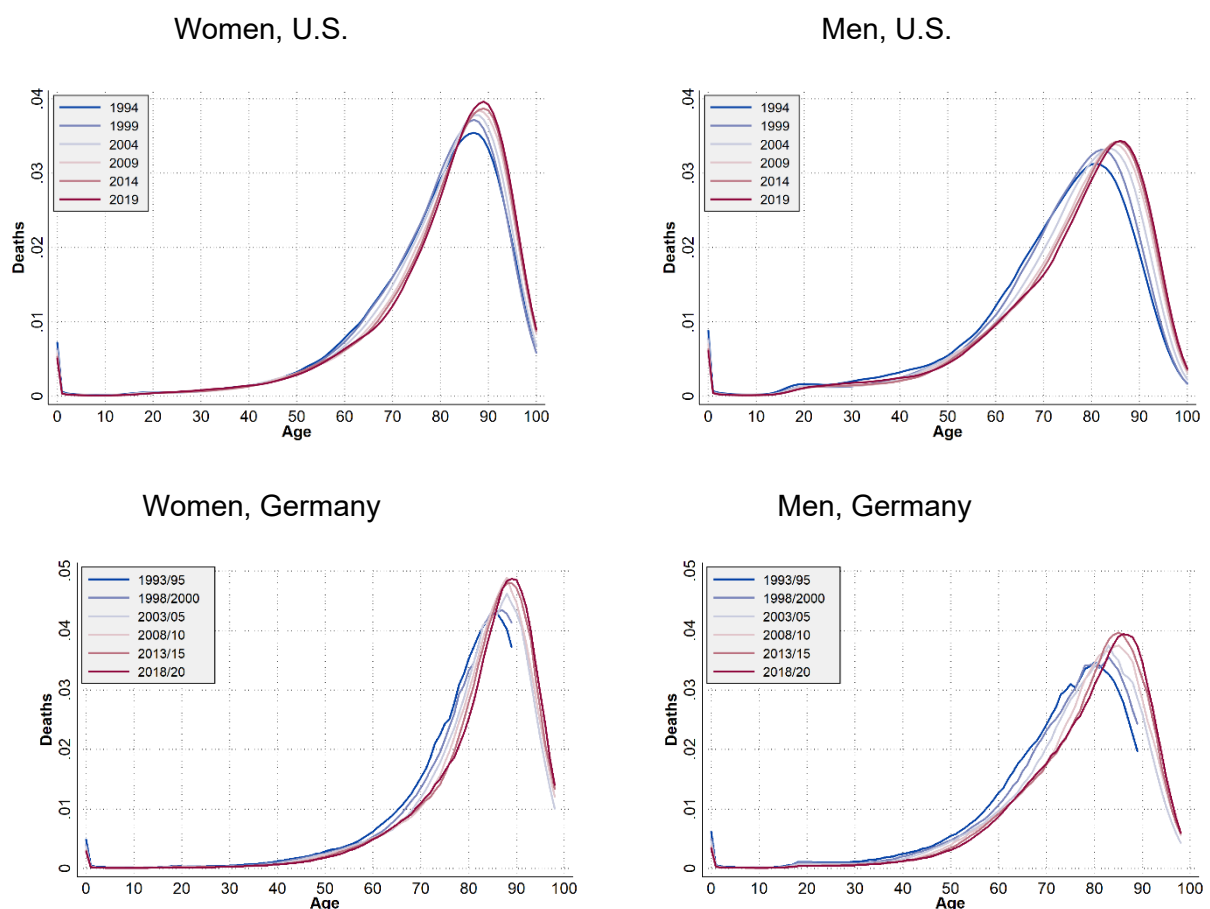
More recent volumes (Börsch-Supan and Coile, 2020 and 2025) documented the evolution of financial incentives to work at older ages from 1980 to the present in the wake of the pension reforms that most countries enacted since the late 1990s. The former volume documents that the past few decades have been a very active period of pension reform in the ISS countries, with many reforms increasing statutory retirement ages and reducing benefit generosity. The main finding in the latter volume was a strong correlation between changes in financial incentives and changes in old-age labor force participation. By exploiting the cross-country differences in social security policy across the participating countries and the inter-temporal changes in policy that have been adopted within these countries over almost four decades, we could show that the increase in labor force participation among older individuals was causally related to the stronger financial incentives to working longer.

While this may be seen as a great success of economic policy that strengthens old-age labor force participation in times of rapid demographic aging, there may be negative side

effects. The most recent volume (Börsch-Supan and Coile, forthcoming) investigated whether pension reforms have increased the income inequality among retirees, motivated by the hypothesis that workers with low earnings suffer more from benefit cuts than workers with high earnings. We found mixed evidence for this hypothesis, with the reforms actually decreasing income inequality in six out of ten countries. However, increasing the retirement age – a crucial ingredient of most pension reforms – may have worsened health inequality among retirees since working longer in strenuous jobs may come at the expense of worsening health, and this particularly for workers that suffer already from health deficits. It is this question that we explore in this volume.

A straightforward and readily available measure of the most severe health condition is mortality. If pension reforms had increased mortality inequality, the mortality distribution by age should become more uneven, i.e. flatter, maybe even shifting to the left. However, the development of mortality distributions over the recent decades shows the opposite in all countries that participate in this study. Figure 1 takes the U.S. and Germany as examples. This is a result worth digging deeper, using detailed and internationally comparable data on health and income.

Figure 1. Mortality distributions in the U.S. and Germany



Source. Period life tables from the Social Security Administration (2024) and the German Federal Statistical Office (2015).

An important goal of the project is to present results that are as comparable as possible across countries. Thus, while model estimation and analyses were conducted for each country by analysts in that country, the papers are prepared according to a detailed template that we developed in close consultation with country participants. The country teams are:

Belgium	Giulia Klinges, Alain Jousten and Mathieu Lefèbvre
Canada	Kevin Milligan and Tammy Schirle
Denmark	Paul Bingley, Nabanita Datta Gupta, Malene Kallestrup-Lamb, and Alexander O.K. Marin
Germany	Axel Börsch-Supan, Luca Salerno, Frederik Fetzner and Johannes Rausch
Italy	Agar Brugiavini, Raluca Elena Buia, Giacomo Pasini, and Guglielmo Weber
Japan	Takashi Oshio and Satoshi Shimizutani
Netherlands	Adriaan Kalwij, and Arie Kapteyn
Spain	Cristina Bellés-Obrero, Manuel Flores, Pilar García-Gómez, Sergi Jiménez-Martín, and Judit Vall-Castelló
United Kingdom	David Sturrock, Carl Emmerson and James Banks
United States	Luca Salerno, Axel Börsch-Supan and Courtney Coile

The selection of these countries at the start of the project was guided by four main criteria. On the one hand, they should represent different pension systems that have emerged from diverse cultural-historical backgrounds and have experienced substantial reforms over the recent three decades. On the other hand, the countries should be comparable with regard to stages of the demographic transition and of economic development with its associated job composition and quality of work. Third, the countries were selected by the quality of the data that is required to precisely describe the incentives exerted by their pension systems and measure them over a relatively long time horizon, spanning the period during and after significant pension reforms. Fourth and maybe most importantly, the eleven countries have excellent research teams well experienced in this type of analyses.

The rest of this introductory chapter presents our methodology and summarizes our main results. The country papers themselves present much more detail for each country and,

in addition to the common analyses performed by all countries, often present country-specific analysis relevant to each particular country.

2. Common methodology

2.1 Data

All our work is based on microdata drawn from national surveys. Where possible, we exploit the international comparability of the global health and retirement surveys (SHARE: Survey of Health, Ageing and Retirement in Europe, ELSA: English Longitudinal Study of Ageing ELSA, and HRS: U.S. Health and Retirement Study). Table 1 presents the datasets in each country with the years covered and the number of individuals.

Table 1: Data sources

Country	Dataset names	Years covered	Individuals
Belgium	SHARE	2004-2022	1,998
Canada	Canadian Community Health Survey (CCHS).	2001-2022	17,500
Denmark	SHARE	2004-2022	2,233
France	SHARE	2004-2022	2,520
Germany	SHARE	2004-2022	3,289
Italy	SHARE	2004-2022	3,883
Japan	Comprehensive Survey of Living Conditions (CSLC)	2001-2022	150,591
Netherlands	SHARE	2004-2022	2,139
Spain	SHARE	2004-2022	1,704
United Kingdom	English Longitudinal Study of Ageing (ELSA)	2002-2019	7,847
United States	Health and Retirement Study (HRS)	1996-2022	25,696

2.2 Methodology

Our main depiction of health inequality is in terms of the “health-income gradients”, where a health measure is plotted against pension income. Health is a complex multidimensional construct. We therefore consider five different health dimensions which are defined as health capacities, such that a higher score reflects better health:

- (1) **Functional health:** Our functional health measure is based on 20 limitations in functional health self-reported by the respondents. They include mobility limitations, limitations in

activities of daily living (ADLs), and limitations in instrumental activities of daily living (IADLs) (Lawton and Brody, 1969). Mobility items cover activities like walking, sitting, climbing stairs, and other (fine) motor tasks. The ADLs refer to basic self-care tasks such as bathing, dressing, and eating, while the IADLs capture more complex activities necessary for independent living, including preparing a hot meal and shopping. Our health capacity measure is constructed by subtracting the actual limitations which the respondent reports from the maximum possible number of limitations.

- (2) **Diagnosed health:** This measure counts the number of eleven health conditions that a doctor has diagnosed since a respondent has been interviewed in the panel data. To construct a number that increases with better health, we subtract the actual number of conditions from the total number of conditions.
- (3) **Comprehensive health:** This most comprehensive measure of health is based on the health-deficiency index as specified in Börsch-Supan et al. (2021), similar to Abeliatsky and Strulik (2019). It summarizes the individual health status by aggregating a set of self-reported health deficits, including chronic conditions, functional limitations, and difficulties with daily activities. Each deficit is coded as present or absent, and the index is calculated as the proportion of observed deficits relative to the total number of non-missing health indicators. To construct a number that increases with better health, we subtract the actual number of deficiencies from the maximum number of possible health deficiencies.
- (4) **Mental health:** We use standardized measures of depressive symptoms (Euro-D for the SHARE countries and CES-D for the other countries). They are targeted at clinical depression and are based on binary items related to mood and behavior, such as sadness, pessimism, sleep disturbance, and lack of interest. Each affirmative response scores one point, with higher scores indicating more depressive symptoms. To construct a number that increases with better health, we subtract the sum of affirmative responses from the total number of items.
- (5) **Cognitive health:** We compute a cognition score based on three cognitive function tests: immediate word recall (from a list of ten words, the test counts how many words a respondent can recall), delayed word recall (after about five minutes, the respondent is asked again to recall these words), and the serial sevens subtraction task (subtract seven from 100, and then four times keep subtracting seven from the result). The total score ranges from 0 to 30, with higher scores indicating better cognitive functioning.

Income is measured in deciles of the respondents' income from all sources. While individual respondents are our analytical unit, we recognize that health is likely influenced by the household context. This is particularly relevant for income. Where partner information is available, we therefore combine both partner's income and divide the total by 1.5 in line with

the OECD equivalence scale. We then investigate whether the slope of this gradient becomes steeper during the period of pension reforms.

We also translate these results into a metric that is comparable across health measures and has a real world meaning. We call it “catch-up time”: the number of years that a poorer individual would need to catch up to the same level of health as a richer individual. Additionally, we plot the cumulative share of individuals with bad health against the percentiles of income (“concentration curve”), which corresponds to the Gini index of health inequality. We acknowledge that our analysis of health inequality is purely descriptive and do not claim causality.

3. Summary of results and conclusions

In spite of all the differences in the breadth and depth of pension reforms across the countries involved in this study, we find the same result in all countries: Taking all measures together, health inequality has neither significantly nor materially increased during the time period in which benefits decreased and retirement ages went up. This holds for the steepness of the health-income gradient, the catch-up years and may best be seen in the development of the concentration index over time. However, the data is rather noisy, even in countries where we have a large number of observations. We find irregular patterns in the data, such as negative health-income gradients and frequent ups and downs in our inequality measures.

A few findings stand out. In the U.S., we find that functional and comprehensive health display a steady rise in inequality over time, but not the other three measures, and some health-income gradients became steeper for women, but not for men. In the UK, the gaps in terms of catch-up years tend to have grown more at the bottom of the income distribution. In Germany, the concentration index for diagnosed health increased, but insignificantly so, and there was no such effect for the other health measures. The Danish findings highlight that, while certain physical health measures, such as functional and comprehensive health, have seen narrowing disparities owing to improvements among lower-income groups, mental health remains the domain with the most pronounced and enduring inequalities. However, the overall structure of health inequality has been stable between 2004 and 2022.

There are three main caveats. First, the data may mask heterogeneities. There may well be occupation groups which have suffered from the pension reforms. Examples are blue collar jobs such as construction or steel workers. The sample size in the SHARE countries is sufficiently large to separate healthy and unhealthy individuals in each income decile, but not large enough to further stratify by occupation. Arduous jobs make up only a small share of all occupations in Europe. Hence, while reducing pension benefits and increasing the retirement age may be a policy option without negative health effects for most, policy makers may want to combine them with special programs if they want to protect workers in arduous occupations.

Second, we should stress that the period under investigation may be too short to pick up the longer-term effects, particularly in countries in which the main pension reforms took place only recently.

Third, our analyses are purely descriptive and therefore have inherent limitations. In particular, we are unable to directly link specific pension reforms to changes in health inequalities. Furthermore, various other factors may exert an influence on health inequalities over that period as well. Arguably, larger samples, such as administrative data which is often used to analyze the income-mortality gradient, are needed to investigate the evolution of health inequalities before drawing firm conclusions, and a more rigorous analytical approach will be needed in future work to disentangle the underlying mechanisms and to identify which design features of the pension system most strongly influence health outcomes. Such evidence would enable policymakers to better anticipate and mitigate unintended negative effects on vulnerable groups, and to align social and health programs more effectively with pension policy.