PENSION REFORMS AND THE HEALTH DISTRIBUTION OF RETIREES

INTRODUCTION AND SUMMARY

Axel Börsch-Supan, Courtney Coile and Luca Salerno

Final version 02 April 2026

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 evidence 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.mpisoc.mpg.de

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

Luca Salerno
Munich Research Institute for the Economics of Aging
Max Planck Institute for Social Law and Social Policy
Leopoldstrasse 139
80804 Munich, Germany

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 and are more likely to come from socioeconomically disadvantaged parts of the population. Health inequalities and their socioeconomic determinants have been subject of a large literature (Marmot, 2005 and National Academies, 2015). The concern that economic and demographic developments may have increased health inequalities has been voiced for example by OECD (2017) and Truesdale et al. (2022). However, the explicit link between pension reforms and health inequality has not been investigated to our knowledge. It is this question that we explore in this volume with a specific interest in whether answers will differ internationally that is motivated by the international difference in socioeconomic health disadvantages (Avendano et al., 2009) and the history of the International Social Security Project.

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 goes about here ------

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 not only according to a detailed template but they also used the same international dataset for seven out of the eleven countries and applied the same common software that we developed in close consultation with the 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
France	Julie Tréguier
Germany	Axel Börsch-Supan, Luca Salerno, Frederik Fetzer 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 longitudinal harmonized health and income data available 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 nine country papers present 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. For two countries, this introduction presents their main results since capacity reasons prevented them from contributing detailed country chapters.

2. Common methodology

2.1 Data

All our work is based on microdata with detailed longitudinal information on health and socioeconomic status (SES). 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). SHARE, which covers seven of the eleven countries, provides identical measures across Continental Europe and is harmonized with ELSA in England and HRS in the US (Börsch-Supan, 2013). Nine countries have strictly comparable data with identical health and income measures. Only Canada and Japan have to use national data sets. However, since the concepts behind our health and SES measures are common and well established, results are well comparable across all eleven countries in this volume.

Since we want to study the health equality potentially affected by pension reforms, our study samples are restricted to retired individuals aged between 60 and 79. This may create a sample selectivity problem because healthier people may work longer. We therefore did key calculations also for a sample of all individuals aged between 65 and 79 and found that the results did not change in any substantive way. We therefore only present results for the retiree sample. 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 in study sample
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
Nether-lands	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 gradient", where a health measure is plotted against pension income, expressed in deciles for better comparison across countries (Figure 2). The slope of this relation is an indicator how better health is for richer individuals than for poorer individuals. In all countries we see this gradient. Our research question is whether the pension reforms since the 1980s have made the health difference between rich and poor larger or smaller. We therefore investigate whether the slope of the health-income gradient has become steeper during the period of pension reforms and thereafter, e.g. because individuals remained working longer in strenuous jobs because retirement ages have been increased.

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 Abeliansky 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.

Our main measure of socioeconomic status is income. It 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 a partner is present, we therefore combine both partner's income and divide the total by 1.5 in line with the OECD equivalence scale. Since our study sample consists of individuals aged 65 and older, additional household members such as children are rare. If they are present, we subtract their contribution to household income.

We use three outcome measures. The first is the change of the slope of the health-income gradient, shown in Figure 2 for the comprehensive health measure. For the second outcome measure, we translate the slopes and their changes 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. As a third outcome measure, we plot the cumulative share of individuals with bad health against the percentiles of income. This created a "concentration curve", which corresponds to the Gini index of health inequality (Figure 3). All three outcome measures of health inequality are purely descriptive and our analyses do not attempt to show causality.

3. Summary of results and conclusions

In spite of 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 (Figure 2), the catch-up years and the development of the concentration index over time (Figure 3).

Figure 2 shows the slope of the health-income gradients at five waves of SHARE data (2004-2022), six waves of HRS data (1996-2020), and two waves for the UK, Canada and

Japan (2001/2 and 2022). The data refers to retired men, aged 60-79. Income is measured as country-specific deciles, and health is measured by the comprehensive health index. Its average value differs by country as does the slope but in all countries, the higher income deciles have better health than the lower deciles. Figure 2 also provides 95% confidence bands. The data is rather noisy and the confidence bands wide. This also holds for the countries where we have large samples (Canada, Japan, US). No formal statistical test is necessary to see that there is no significant change in the steepness of the health-income gradient over time as a potential effect of the pension reforms. However, it is not the lack of statistical precision that let us conclude that there is so far no evidence for a negative reform effect since the gradients for different data waves cross each other and do not follow a consistent pattern such that even much narrower confidence bands would not identify a consistent steepening of the health-income gradients.

------ Figure 2 goes about here ------

Figure 3 reports on the concentration index for all five health measures. The horizontal axis is time and the vertical axis represents the opposite of the Gini index such that lower values mean more inequality. If health inequality had increased as a result of pension reforms, we would expect declining slopes for all health measures. 95% confidence bands are given for the health measure that exhibits the steepest decline in the index, i.e. where an increase in health inequality is most likely. The results are more nuanced than those of Figure 2. However, the slopes are not statistically different from flat in all ten countries.

------ Figure 3 goes about here ------

A few findings stand out in the country chapters. 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.

We stress four major caveats. First, 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. Health varies much more across individuals than over time, so reform effects that have happened so far may be too small to be detected in the general health heterogeneity of a population. We therefore refrained from using the analytical apparatus that this project has developed and successfully employed in earlier studies, e.g. the study of income inequality (Börsch-Supan and Coile, forthcoming). Nevertheless, our results show that it is premature to claim that pension reforms have worsened public health is it has happened in heated European pension debates.

Second, the data may mask specific 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.

Third, 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. The negative health effects of working longer in occupations that are characterized by strenuous and/or stressful work evolve over time, and the additional months or few years due to pension reform may be marginal relative to the duration of the entire job career.

Fourth, 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. With such data a more rigorous analytical approach can be used 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.

References

Abeliansky, A. L., & Strulik, H. (2019). Long-run improvements in human health: Steady but unequal. The Journal of the Economics of Ageing, 100189.

Avendano, M., Glymour, M. M., Banks, J., & Mackenbach, J. P. (2009). Health disadvantage in US adults aged 50 to 74 years: a comparison of the health of rich and poor Americans with that of Europeans. American journal of public health, 99(3), 540-548.

Börsch-Supan, Axel and Courtney Coile (2020). Social Security Programs and Retirement Around the World: Reforms and Retirement Incentives. Chicago: University of Chicago Press.

Börsch-Supan, Axel and Courtney Coile (2025). Social Security Programs and Retirement Around the World: The Effect of Reforms on Retirement Behavior. Chicago: University of Chicago Press.

Börsch-Supan, Axel and Courtney Coile (forthcoming). Social Security Programs and Retirement Around the World: Income Inequality. Chicago: University of Chicago Press.

Börsch-Supan, Axel, Irene Ferrari, and Luca Salerno (2021). Long-run Health Trends in Europe. In: Journal of the Economics of Ageing, Volume 18.

Börsch-Supan, Axel, Martina Brandt, Christian Hunkler, Thorsten Kneip, Julie Korbmacher, Frederic Malter, Barbara Schaan, Stephanie Stuck, and Sabrina Zuber (2013): “Data Resource Profile: The Survey of Health, Ageing and Retirement in Europe (SHARE)”, In: International Journal of Epidemiology, 42(4), pp. 992-1001.

Coile, Courtney, Kevin Milligan, and David A. Wise (2019). Social Security Programs and Retirement Around the World: Working Longer. Chicago: University of Chicago Press.

Gruber, Jonathan, and David A. Wise (1999). Social Security and Retirement Around the World. Chicago: University of Chicago Press.

Gruber, Jonathan and David A. Wise (2004). Social Security Programs and Retirement Around the World: Microestimation. Chicago: University of Chicago Press.

Gruber, Jonathan and David A. Wise (2007). Social Security Programs and Retirement Around the World: Fiscal Implications of Reform. Chicago: University of Chicago Press.

Gruber, Jonathan and David A. Wise (2010). Social Security Programs and Retirement Around the World: The Relationship to Youth Employment. Chicago: University of Chicago Press.

Lawton, M. Powell and Elaine M. Brody, E. M. (1969). Assessment of older people: Self-maintaining and instrumental activities of daily living. *The Gerontologist, 9*(3, Pt 1), 179–186. https://doi.org/10.1093/geront/9.3_Part_1.179

Marmot, Michael (2005). Social determinants of health inequalities. Lancet, 365(9464):1099-104.

National Academies of Sciences, Engineering, and Medicine (2015). The Growing Gap in Life Expectancy by Income: Implications for Federal Programs and Policy Responses. Washington, DC: The National Academies Press.

OECD (2017). Preventing Ageing Unequally. OECD Publishing, Paris. http://dx.doi.org/10.1787/9789264279087-en

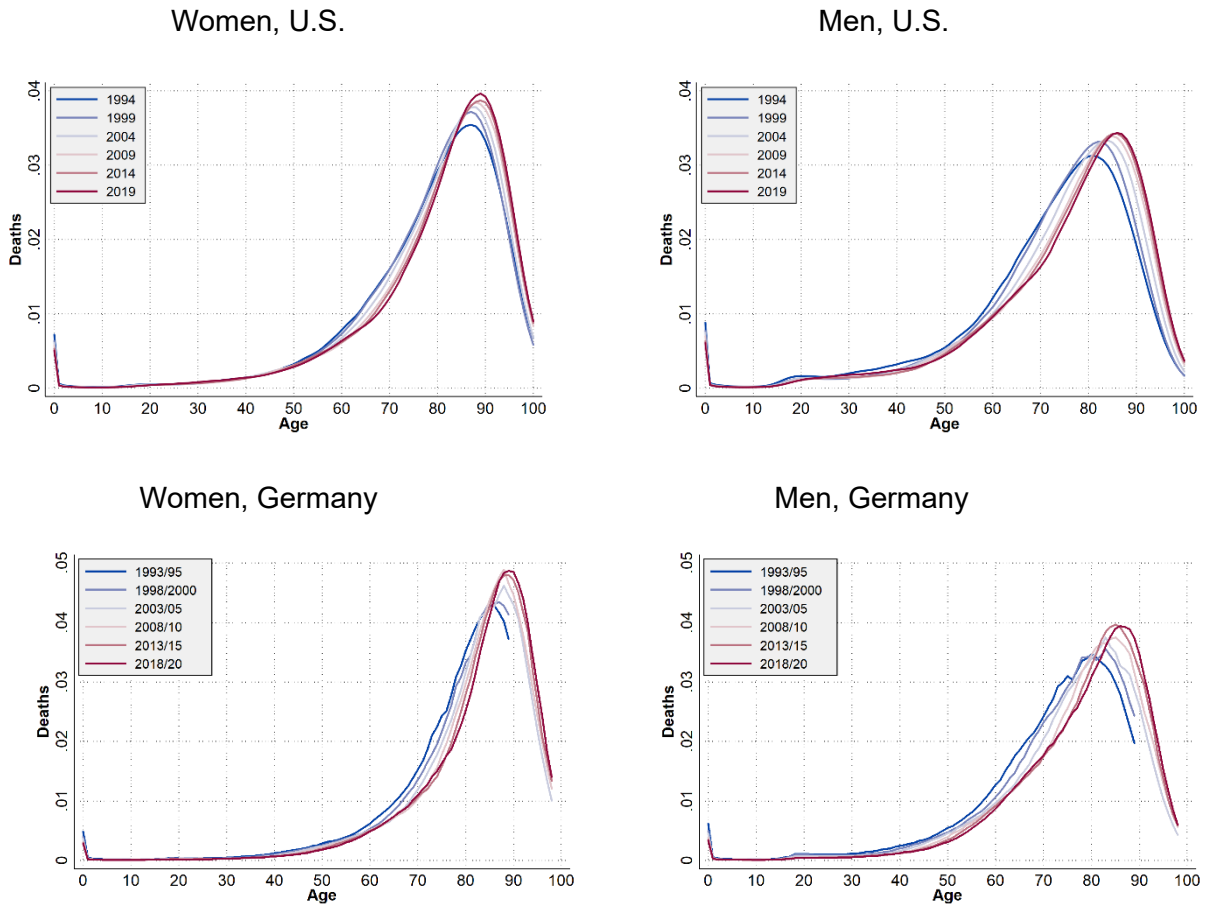
Truesdale, Beth C., Lisa F. Berkman, and Alexandra Mitukiewicz (2022). “When I’m 54: Working Longer Starts Younger Than We Think,” in Overtime: America’s Aging Workforce and the Future of Working Longer, Lisa F. Berkman and Beth C. Truesdale, eds. Oxford University Press.

Wise, David A. (2012). Social Security Programs and Retirement Around the World: Historical Trends in Mortality and Health, Employment, and Disability Insurance Participation and Reform. Chicago: University of Chicago Press.

Wise, David A. (2016). Social Security Programs and Retirement Around the World: Disability Insurance Programs and Retirement. Chicago: University of Chicago Press.

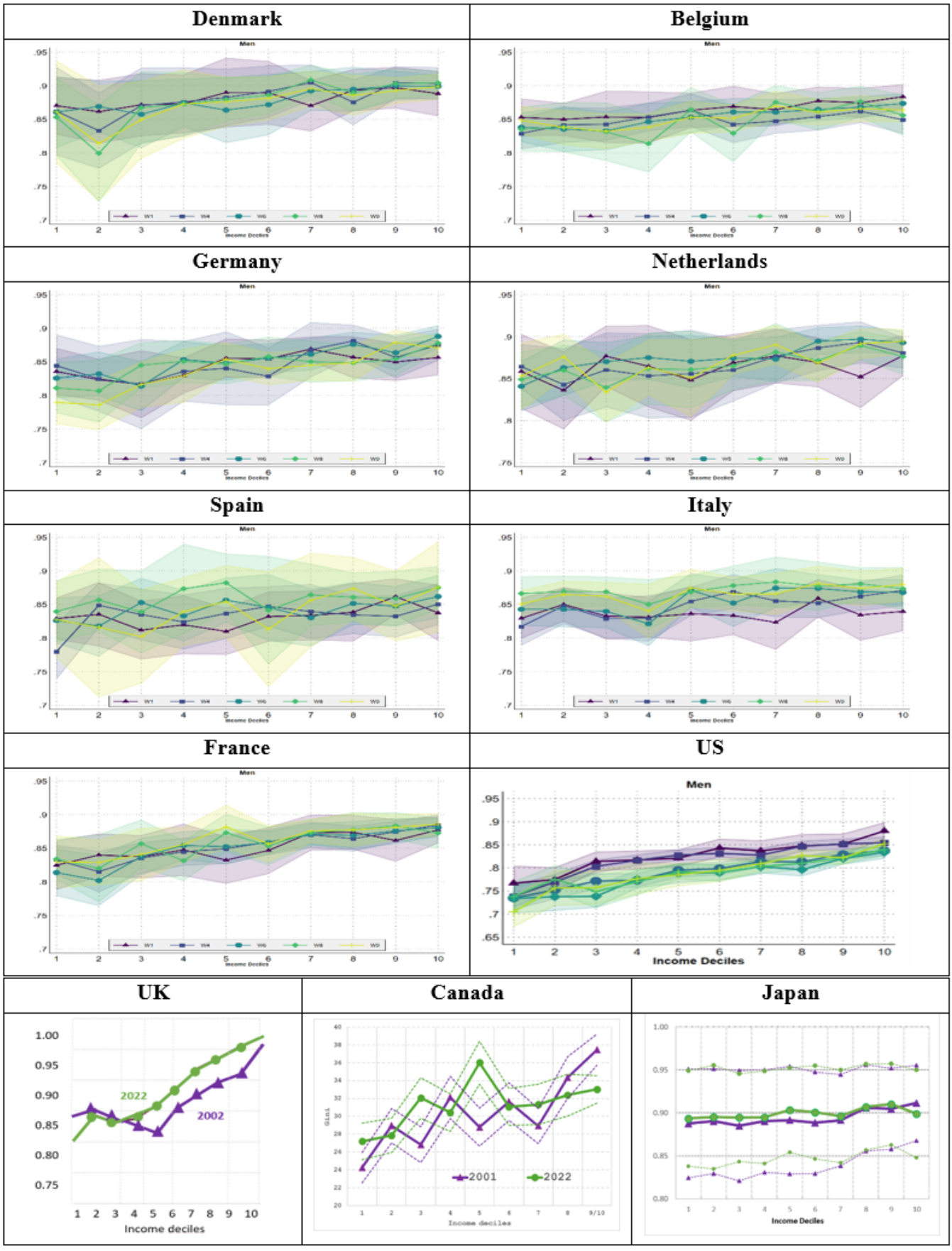
Wise, David A. (2017). Social Security Programs and Retirement Around the World: The Capacity to Work at Older Ages. Chicago: University of Chicago Press.

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

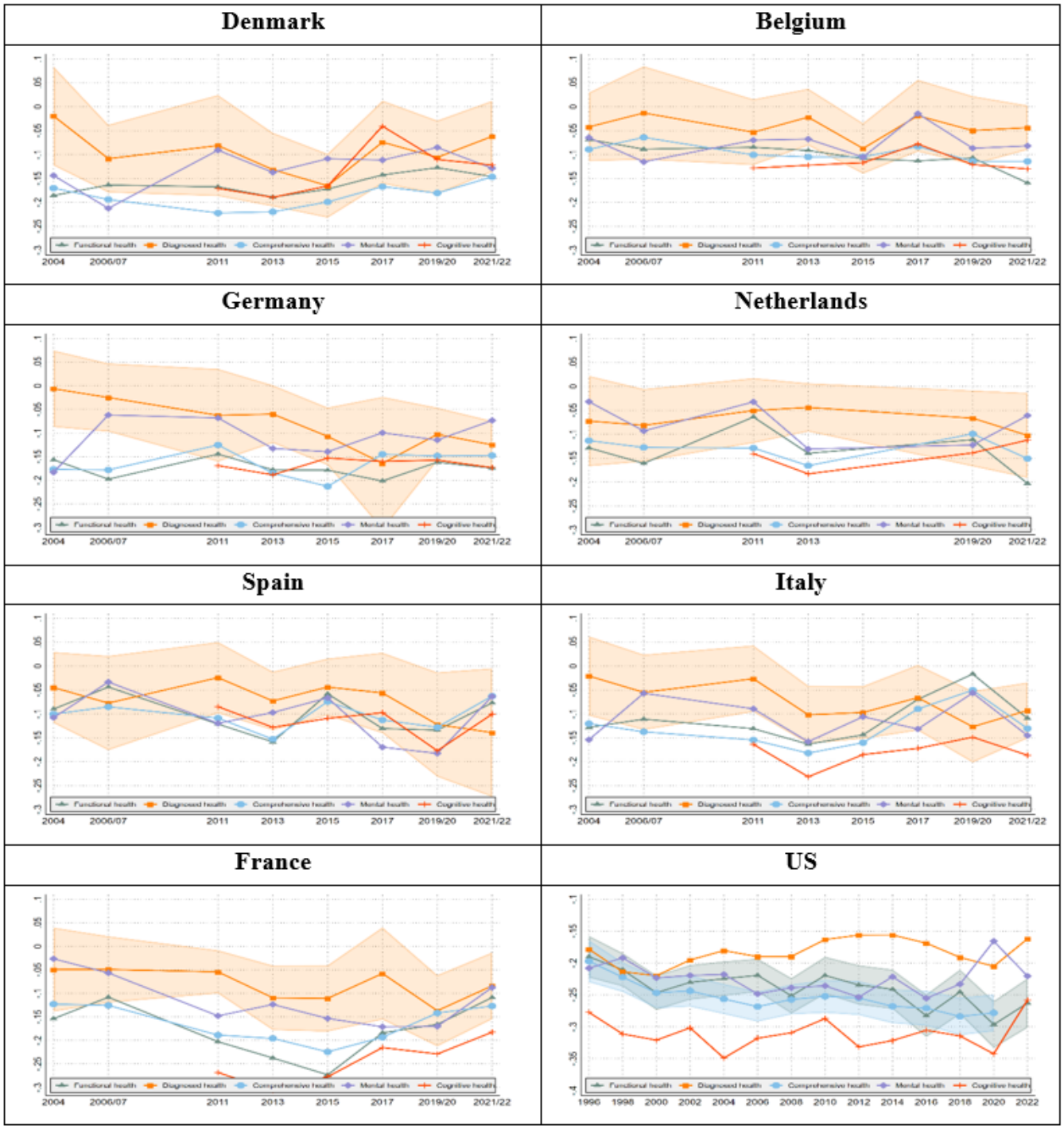


Source. Period life tables from the Social Security Administration (2024) and the German Federal Statistical Office (2015).

Figure 2: Development of the health-income gradient 2004-2022 for comprehensive health



Source: Authors' contributions**Figure 3: Development of concentration indexes 2004-2022 for five health measures**



Source: Authors' contributions