# 

# APPENDICES FOR

# The growing American health penalty:

# International trends in the employment of

# older workers with poor health

**Table of contents for Appendices:**

Online Appendix A1. Further details on HRS/ELSA/SHARE data 2

Online Appendix A2. Missing data 8

Online Appendix A3. Descriptive statistics 9

Online Appendix A4. Additional material for Section 3 10

Online Appendix A5. Additional material for Section 4 17

Online Appendix A6. Additional material for Section 5 24

Online Appendix A7. Additional material on the PCA 27

Online Appendix A8. Further detail on sickness and disability policy reforms 29

References for Online Appendices 36

## Further details on HRS/ELSA/SHARE data

### Use of the Gateway to Global Aging Data RAND Harmonized dataset

Full detail of the Gateway to Global Aging Data (G2Aging)-constructed variables are given in the G2Aging codebooks.[[1]](#footnote-1) We summarize key points here, alongside full details of further changes that we made to enhance comparability.

### Employment variables

The employment variable in the G2Aging harmonized dataset is not comparable across surveys or over time. Most of the questions are roughly similar to one another: in HRS respondents were asked if they were doing work ‘at the present time’, in ELSA they were asked about working in the ‘last month’, and new interviewees in SHARE were asked about working in the past 4 weeks (for new interviewees). However, repeat interviewees in SHARE were instead asked if they had worked since the last interview, usually about two years previously.

It is however possible to make these questions comparable across time and surveys. To do this, we used information given in SHARE by repeat interviewees on their employment history since the last wave to create a measure of whether or not they had worked in the past month. This is then comparable to new SHARE interviewees and also to the ELSA/HRS measures. (Note that because only the month and not the full date of job starts or terminations are given, this strictly refers to employment in the present or past *calendar* month, rather than the past 30 days.)

We also use two further employment-related measures taken from the G2Aging harmonized file: (i) working hours, and (ii) current job tenure. In both cases we restrict our analysis to those who are presently employed using the measure just described. (While Zamarro and Lee, 2011 perform international comparisons with HRS-SHARE-ELSA data using self-defined labour market status, we avoid this as it depends upon culturally- and temporally-variable understandings of what ‘employment’, ‘unemployment’, ‘disability’, etc. mean).

### Disability benefit variables

While the G2Aging harmonized file does include a nominally comparable disability benefit variable, there are two errors in this variable that we correct for our analyses: (i) there are some coding errors (for income sources that are not truly disability benefits in ELSA & SHARE); (ii) the supplied derived variables do not make the SHARE variables consistent across different question versions in different waves. The supplied variables also refer to any disability benefit receipt in the past year, rather than present receipt.

We therefore use a revised version of the G2Aging variables, which are based on the same underlying questions, but which calculate the derived variables independently, as we now describe for each survey in turn.

#### HRS

For HRS, we use the same underlying questions in the RAND HRS Income and Wealth file. Unlike the main RAND HRS file, however, we create a dummy variable only for those who claimed disability benefits for *all* of the past year, to ensure that the measure captures people currently claiming disability benefits (a particularly important step given the analyses of simultaneous work-plus-benefits in Part 4 of the paper).[[2]](#footnote-2)

#### SHARE

The disability benefit variable in the G2Aging harmonized dataset is based upon derived variables supplied by the SHARE team in their generated variables dataset, which combine underlying questions (on receipt of multiple different sources of income in the past year) into income categories. There are however two problems here:

* The main questionnaire (in English) changes between waves. For example, in wave 1 respondents are asked about ‘Public disability insurance’ and ‘Public invalidity or incapacity pension’, whereas in waves 2-5 respondents are asked about ‘Main public disability insurance pension, or sickness benefits’ and ‘Secondary public disability insurance pension, or sickness benefits’. Moreover, at wave 6 respondents were asked separately about sickness benefits and disability benefits.
* The translated versions of these categories differ between countries. For example, sickness benefits were explicitly mentioned in waves 2-5 in Austria, Germany, Spain, Sweden, and Denmark, but not elsewhere.

The variables we use therefore aim for consistency within a country over time, rather than across countries (where sickness benefits are only inconsistently included), as follows:

|  |  |  |
| --- | --- | --- |
| **Country** | **Includes sick pay?** | **Inconsistencies** |
| Austria | Y | Sick pay not available w1, but included w2-6. |
| Germany | Y | Sick pay not available w1, but included w2-6. |
| Sweden | Y |  |
| Spain | Y | Sick pay not available w1, but included w2-6. |
| Italy | N |  |
| France | N |  |
| Denmark | Y |  |
| Switzerland | N |  |
| Belgium | N |  |

The full syntax used to generate the disability benefit variables is given in the replication file.

#### ELSA

The disability benefit variable in the G2Aging harmonized file combines all disability-related benefits together, including those that are not work-related and which in most other countries would not form part of the benefits system. This includes the extra-costs benefits Disability Living Allowance/Personal Independence Payment and Attendance Allowance, as well as Carer’s Allowance. In recent waves, the G2Aging data also mistakenly misses out claims of the latest version of the out-of-work disability benefit (Employment and Support Allowance). We therefore created new variables that focused on all out-of-work disability benefits: Incapacity Benefit/Employment and Support Allowance/Severe Disablement Allowance.

### Health variables

The health variables are as follows:

#### Self-reported global health

We included the G2Aging measure of global self-reported health, which was asked consistently across surveys (barring ELSA wave 3, which we exclude from analysis). This was dichotomized into fair/poor vs. excellent/very good/good health.

#### Motor skills

We used 10 binary measures of motor skills: walking 100m/one block, lifting 5kg, pulling/pushing large objects, climbing one flight of stairs, climbing several flights of stairs, stooping/kneeling/crouching, picking up a small coin, sitting for 2hrs, getting up from a chair, reaching above shoulder height. These are sometimes referred to as 'Nagi functions’ (Crimmins, et al., 2010), and which refer to *‘movements involving the upper and/or lower limbs, most of which require a degree of muscle strength but a few of which are more to do with dexterity and flexibility’* (Breeze and Lang, 2006).

While our analysis only requires the assumption that these variables are comparable over time within each country, for ease of interpretation we improve their cross-national comparability in several ways. Firstly, we assume that anyone who struggles walking up one flight of stairs would struggle with several flights of stairs (an assumption made by question filtering in HRS, but the questions were asked independently in ELSA/SHARE, and about 1% of respondents who report no problems with several flights of stairs report problems with a single flight). Secondly, a small proportion of HRS (but not ELSA/SHARE) responses are coded as ‘don’t do’, which was not listed on the response options but was allowed as a spontaneous response. For most motor skills, we assumed that ‘don’t do’ means that they cannot do the task for health reasons, because these tasks are effectively universal. However, for walking 100m or climbing stairs, we have treated ‘don’t do’ as missing, as it seems plausible that some HRS respondents genuinely do not do these tasks.

#### Functional disability

We used two measures of functional disability, one for any limitation in Activities of Daily Living (ADLs), one for any limitation in Instrumental Activities of Daily Living (IADLs). As we note in the main text, *‘The former are basic and universal physical tasks such as eating, the latter are mixtures of physical and cognitive competencies such as preparing a hot meal (Breeze and Lang, 2006)’*. The full list of measures underlying the scales are problems with:

|  |  |
| --- | --- |
| **ADLs** | **IADLs[[3]](#footnote-3)** |
| Dressing  Walking across a room  Bathing or showering  Eating  Getting in or out of bed  Using the toilet | Preparing a hot meal  Shopping for groceries  Making telephone calls  Taking medications  Managing money |

As for motor skills, a small proportion of HRS (but not ELSA/SHARE) responses are coded as ‘don’t do’. We generally assumed that ‘don’t do’ means that they cannot do the task for health reasons; but for the IADLs we used HRS’s follow-up question on the reasons for not doing the activity.

#### Doctor-diagnosed chronic disease

We use seven measures of self-reported lifetime doctor-diagnosed chronic diseases that are available in the G2Ageing harmonized file: high blood pressure, stroke, diabetes/high blood sugar, chronic lung disease (excluding asthma), heart problems, and arthritis.

#### Body-Mass Index

We use two measures of non-optimal Body-Max Index (BMI) – underweight and overweight – based on self-reported height and weight in HRS & SHARE, and measured height and weight in ELSA. Because the ELSA variable therefore has more missing responses (particularly as it was only collected in alternate waves), in our main analyses we use the temporally closest BMI observation for all respondents that have a valid BMI at any wave. This substantially increases the sample size and thereby the power of the analyses, but it does this at the cost of potential biases; we therefore conduct two sensitivity analyses, one that excludes BMI completely, and one that imputes BMI within a full multiple imputation analysis, as we detail below.

#### Mental health

Despite claims to the contrary (e.g. Riumallo-Herl, et al., 2014) and a common focus on recent negative feelings, the mental health measure in ELSA/HRS is not comparable to the measure in SHARE (Courtin, et al., 2015, Zamarro, et al., 2008). However, our analysis does not depend on cross-national comparability, but only the weaker assumption of comparability within countries over time – a much more plausible assumption in this case.

The mental health measure in SHARE is the 12-item Euro-D scale, while the mental health measure in HRS/ELSA is the 20-item CES-D scale. We use two versions of these scales in our analysis: a normalized continuous version, and a binary measure of poor mental health based on the standard cut-off for the scale measure (3+ in ELSA/HRS, 4+ in SHARE); see Crimmins et al (2010) and Courtin et al (2015) for a detailed description of these measures.

## Missing data

The following table presents the prevalence of missing data by country-wave (using unweighted data), dividing between: (i) no missing data; (ii) just BMI missing (which is particularly the case for England); (iii) one other variable missing (primarily mental health/ADLs/mobility in the USA); and (iv) 2+ variables missing:

Table A1: *Missing data by wave.*

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | ***No missing data*** | | | | ***Just BMI missing*** | | | |
| Country | Wave 1 | Wave 2 | Wave 5 | Wave 6 | Wave 1 | Wave 2 | Wave 5 | Wave 6 |
| USA | 85.9 | 90.5 | 92.5 | 91.8 | 0.3 | 0.2 | 0.4 | 0.5 |
| England | 90.7 |  | 84.9 | 69.4 | 8.1 |  | 11.6 | 26.6 |
| Austria | 99.6 | 98.9 | 97.2 | 99.8 | 0.0 | 0.7 | 0.1 | 0.1 |
| Belgium | 99.0 | 98.9 | 99.2 | 98.5 | 0.5 | 0.7 | 0.3 | 0.5 |
| Denmark | 99.5 | 98.4 | 99.0 | 99.3 | 0.1 | 0.4 | 0.4 | 0.4 |
| France | 96.0 | 97.5 | 99.5 | 98.5 | 0.8 | 0.3 | 0.1 | 0.4 |
| Germany | 99.0 | 98.3 | 98.6 | 99.0 | 0.3 | 1.2 | 0.3 | 0.1 |
| Greece | 98.9 | 99.0 |  | 99.4 | 0.2 | 0.2 |  | 0.4 |
| Italy | 98.6 | 99.5 | 99.5 | 98.9 | 0.1 | 0.4 | 0.2 | 0.2 |
| Netherlands | 98.6 | 98.3 | 98.6 |  | 0.6 | 1.0 | 0.4 |  |
| Spain | 97.7 | 97.6 | 99.2 | 98.5 | 1.7 | 1.7 | 0.6 | 0.8 |
| Sweden | 99.7 | 97.7 | 98.9 | 98.2 | 0.2 | 0.7 | 0.6 | 0.3 |
| Switzerland | 99.6 | 99.3 | 100.0 | 99.8 | 0.4 | 0.2 | 0.0 | 0.0 |

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | ***One other variable missing*** | | | | ***2+ variables missing*** | | | |
| Country | Wave 1 | Wave 2 | Wave 5 | Wave 6 | Wave 1 | Wave 2 | Wave 5 | Wave 6 |
| USA | 12.2 | 8.2 | 5.9 | 6.7 | 1.7 | 1.2 | 1.3 | 1.1 |
| England | 0.3 |  | 0.2 | 0.4 | 1.0 |  | 3.3 | 3.6 |
| Austria | 0.0 | 0.2 | 0.1 | 0.0 | 0.4 | 0.2 | 2.7 | 0.1 |
| Belgium | 0.2 | 0.3 | 0.3 | 0.5 | 0.4 | 0.2 | 0.2 | 0.4 |
| Denmark | 0.1 | 1.0 | 0.4 | 0.4 | 0.3 | 0.2 | 0.2 | 0.0 |
| France | 0.7 | 0.8 | 0.1 | 0.4 | 2.5 | 1.4 | 0.3 | 0.8 |
| Germany | 0.4 | 0.4 | 0.8 | 0.8 | 0.3 | 0.1 | 0.4 | 0.1 |
| Greece | 0.6 | 0.5 |  | 0.1 | 0.3 | 0.4 |  | 0.1 |
| Italy | 0.1 | 0.0 | 0.1 | 0.7 | 1.1 | 0.1 | 0.2 | 0.2 |
| Netherlands | 0.1 | 0.4 | 1.0 |  | 0.8 | 0.3 | 0.1 |  |
| Spain | 0.2 | 0.3 | 0.1 | 0.8 | 0.5 | 0.4 | 0.1 | 0.0 |
| Sweden | 0.1 | 1.4 | 0.3 | 1.1 | 0.0 | 0.2 | 0.2 | 0.4 |
| Switzerland | 0.0 | 0.1 | 0.0 | 0.2 | 0.0 | 0.4 | 0.0 | 0.1 |

## Descriptive statistics

### Age and gender by country-wave

Descriptive statistics on sample size, retirement age, respondents’ age, and gender are given below. Note that the sample is defined as being between age 50 and the country’s retirement age (see main text), and hence the gender balance of the effective sample will depend on the respective retirement ages of men and women in the sample.

Table A2: *Distribution of age, gender and retirement age by country-wave.*

|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | ***Sample size*** | | ***Regular retirement age1*** | | ***Mean age of respondents*** | | | | ***Gender=male (%)*** | |
| Country | 2004-7 | 2012-15 | 2004-5 | 2014-152 | 2004-5 | 2006-7 | 2012-13 | 2014-15 | 2004-7 | 2012-15 |
| USA | 13,640 | 16,181 | 65.2 - 66.0 | 66.0 - 66.8 | 56.9 | 58.1 | 58.8 | 59.9 | 46.4 | 46.9 |
| England | 7,018 | 5,153 | 60.0 - 65.0 | 62.0 - 65.0 | 57.4 | 57.2 | 56.9 | 58.2 | 58.0 | 56.1 |
| Austria | 1,136 | 2,491 | 60.0 - 65.0 | 60.0 - 65.0 | 56.6 | 57.2 | 56.5 | 57.4 | 58.4 | 56.7 |
| Belgium | 3,648 | 5,934 | 63.0 - 65.0 | 65.0 | 56.8 | 56.9 | 56.9 | 57.1 | 51.3 | 49.6 |
| Denmark | 2,437 | 4,177 | 65.0 | 65.0 | 56.9 | 57.5 | 57.8 | 57.6 | 49.4 | 49.9 |
| France | 2,401 | 2,768 | 60.0 | 61.3 | 54.8 | 55.0 | 55.9 | 55.0 | 49.0 | 49.1 |
| Germany | 3,101 | 5,409 | 65.0 | 65.3 | 57.4 | 57.4 | 57.3 | 57.8 | 49.2 | 49.2 |
| Greece | 3,051 | 2,253 | 60.0 - 65.0 | 62.0 - 65.0 | 56.6 | 56.6 |  | 57.3 | 52.3 | 48.7 |
| Italy | 2,413 | 4,257 | 60.0 - 65.0 | 63.8 - 66.3 | 56.6 | 56.6 | 56.8 | 56.6 | 57.1 | 53.2 |
| Netherlands | 3,369 | 2,146 | 65.0 | 65.0 | 56.8 | 57.1 | 56.2 |  | 49.8 |  |
| Spain | 2,163 | 5,116 | 65.0 | 65.3 | 56.9 | 57.2 | 56.7 | 57.5 | 49.9 | 50.0 |
| Sweden | 3,139 | 3,052 | 65.0 | 65.3 | 57.1 | 57.4 | 57.8 | 58.1 | 50.4 | 50.8 |
| Switzerland | 1,315 | 2,577 | 63.0 - 65.0 | 64.0 - 65.0 | 56.5 | 57.0 | 57.3 | 57.9 | 53.6 | 50.9 |

***Note****: 1 Details of regular retirement age variable is given in the main paper; 2 Retirement age for Netherlands in the later period refers to 2012-13.*

## Additional material for Section 3

### Health-related employment gaps by wave, top tertile vs. bottom tertile

Table A3: Health-related employment gap by country

|  |  |  |  |
| --- | --- | --- | --- |
| **Country** | **Period** | **Gap (%)** | **95% CI** |
| USA | 2004-7 | 33.8 | (31.3, 36.4) |
| USA | 2012-5 | 41.9 | (39.3, 44.5) |
| England | 2004-7 | 32.0 | (28.2, 35.8) |
| England | 2012-5 | 28.8 | (24.6, 33.1) |
| Austria | 2004-7 | 21.8 | (14.4, 29.2) |
| Austria | 2012-5 | 32.1 | (25.9, 38.3) |
| Belgium | 2004-7 | 23.0 | (18.9, 27.2) |
| Belgium | 2012-5 | 25.1 | (20.5, 29.7) |
| Denmark | 2004-7 | 28.1 | (23.4, 32.7) |
| Denmark | 2012-5 | 28.2 | (24.4, 31.9) |
| France | 2004-7 | 21.3 | (15.9, 26.8) |
| France | 2012-5 | 25.3 | (20.1, 30.5) |
| Germany | 2004-7 | 23.9 | (19.5, 28.3) |
| Germany | 2012-5 | 26.5 | (22.6, 30.4) |
| Greece | 2004-7 | 11.3 | (6.6, 16.0) |
| Greece | 2012-5 | 12.6 | (7.3, 18.0) |
| Italy | 2004-7 | 16.8 | (11.4, 22.2) |
| Italy | 2012-5 | 19.7 | (14.4, 24.9) |
| Netherlands | 2004-7 | 28.4 | (24.2, 32.6) |
| Netherlands | 2012-5 | 25.1 | (16.6, 33.7) |
| Spain | 2004-7 | 32.5 | (26.6, 38.4) |
| Spain | 2012-5 | 31.2 | (22.8, 39.6) |
| Sweden | 2004-7 | 31.4 | (27.1, 35.8) |
| Sweden | 2012-5 | 24.0 | (19.2, 28.8) |
| Switzerland | 2004-7 | 15.8 | (9.7, 21.9) |
| Switzerland | 2012-5 | 10.9 | (6.2, 15.7) |

Figures are average marginal effects controlling for age and gender.

### Trends in employment gaps and rates

Table A4: Changes in absolute employment level among those with poor health (‘level’) and the health-related employment gap (‘gap’) by country.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Country** | **Level (%)** | **95% CI** | **Gap (%)** | **95% CI** |
| USA | -4.5 | (-7.3, -1.8) | 8.1 | (4.8, 11.4) |
| England | 5.6 | (1.2, 10.0) | -3.2 | (-8.6, 2.3) |
| Austria | 8.2 | (1.0, 15.3) | 10.3 | (0.8, 19.8) |
| Belgium | 10.4 | (6.0, 14.9) | 2.1 | (-3.8, 7.9) |
| Denmark | 8.6 | (3.8, 13.3) | 0.1 | (-5.6, 5.8) |
| France | 4.1 | (-1.9, 10.1) | 4.0 | (-3.3, 11.2) |
| Germany | 8.8 | (4.4, 13.1) | 2.6 | (-3.0, 8.3) |
| Greece | -1.1 | (-5.9, 3.8) | 1.3 | (-5.3, 8.0) |
| Italy | 10.3 | (4.7, 15.8) | 2.9 | (-4.3, 10.1) |
| Netherlands | 13.2 | (5.5, 21.0) | -3.2 | (-12.2, 5.7) |
| Spain | 5.6 | (-1.3, 12.5) | -1.3 | (-11.0, 8.4) |
| Sweden | 13.0 | (7.5, 18.6) | -7.4 | (-13.6, -1.3) |
| Switzerland | 11.4 | (5.4, 17.3) | -4.9 | (-12.2, 2.5) |

Figures are average marginal effects controlling for age and gender; trends refer to 2004-7 vs 2012-15.

### Sensitivity analyses

#### Results of sensitivity analyses (exc. gender-stratified analysis)

Table A5: Sensitivity analyses for employment trend among those in poor health.

|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | ***Main (%)*** | | ***Aged 54+(%)*** | | ***Imputation (%)*** | | ***Exc. BMI (%)*** | | ***Logit (%)*** | |
| USA | -4.5 | *(-7.3 to -1.8)* | -4.0 | *(-6.9 to -1.0)* | -4.5 | *(-7.1 to -1.9)* | -4.3 | *(-7.0 to -1.5)* | -4.3 | *(-7.2 to -1.5)* |
| England | 5.6 | *(1.2 to 10.0)* | 4.8 | *(0.8 to 8.7)* | 7.7 | *(4.0 to 11.4)* | 7.6 | *(3.5 to 11.7)* | 5.4 | *(1.2 to 9.7)* |
| Austria | 8.2 | *(1.0 to 15.3)* | 5.3 | *(-1.0 to 11.6)* | 7.8 | *(0.7 to 14.9)* | 8.9 | *(1.8 to 16.1)* | 7.5 | *(0.8 to 14.1)* |
| Belgium | 10.4 | *(6.0 to 14.9)* | 9.7 | *(5.0 to 14.5)* | 10.0 | *(5.6 to 14.5)* | 9.8 | *(5.4 to 14.3)* | 9.8 | *(5.2 to 14.5)* |
| Denmark | 8.6 | *(3.8 to 13.3)* | 9.9 | *(4.5 to 15.3)* | 8.0 | *(3.3 to 12.7)* | 8.4 | *(3.7 to 13.1)* | 8.6 | *(3.8 to 13.5)* |
| France | 4.1 | *(-1.9 to 10.1)* | 9.7 | *(5.0 to 14.4)* | 5.3 | *(-0.6 to 11.3)* | 6.2 | *(0.2 to 12.2)* | 3.7 | *(-1.4 to 8.8)* |
| Germany | 8.8 | *(4.4 to 13.1)* | 11.4 | *(6.4 to 16.3)* | 8.5 | *(4.1 to 12.9)* | 8.3 | *(4.0 to 12.7)* | 9.2 | *(4.6 to 13.9)* |
| Greece | -1.1 | *(-5.9 to 3.8)* | -1.6 | *(-6.7 to 3.5)* | -0.5 | *(-5.4 to 4.3)* | -0.1 | *(-5.0 to 4.7)* | -1.7 | *(-6.8 to 3.4)* |
| Italy | 10.3 | *(4.7 to 15.8)* | 13.5 | *(8.5 to 18.6)* | 10.7 | *(5.2 to 16.2)* | 10.2 | *(4.6 to 15.8)* | 8.8 | *(3.4 to 14.2)* |
| Netherlands | 13.2 | *(5.5 to 21.0)* | 18.7 | *(13.5 to 24.0)* | 13.8 | *(6.2 to 21.5)* | 14.9 | *(7.2 to 22.6)* | 12.7 | *(4.6 to 20.8)* |
| Spain | 5.6 | *(-1.3 to 12.5)* | 5.7 | *(-1.4 to 12.8)* | 5.8 | *(-1.0 to 12.5)* | 4.8 | *(-2.0 to 11.5)* | 6.1 | *(-1.3 to 13.6)* |
| Sweden | 13.0 | *(7.5 to 18.6)* | 14.6 | *(8.8 to 20.5)* | 13.3 | *(7.7 to 18.8)* | 13.7 | *(8.2 to 19.2)* | 13.6 | *(8.2 to 19.1)* |
| Switzerland | 11.4 | *(5.4 to 17.3)* | 14.1 | *(7.5 to 20.8)* | 10.5 | *(4.5 to 16.5)* | 11.2 | *(5.2 to 17.2)* | 11.3 | *(5.6 to 17.0)* |

Table A6: Sensitivity analyses for trend in the health-related employment gap.

|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | ***Main (%)*** | | ***Aged 54+ (%)*** | | ***Imputation (%)*** | | ***Exc. BMI (%)*** | | ***Logit (%)*** | |
| USA | 8.1 | *(4.8 to 11.4)* | 7.1 | *(3.5 to 10.8)* | 8.4 | *(5.3 to 11.6)* | 7.9 | *(4.6 to 11.2)* | 7.2 | *(3.8 to 10.6)* |
| England | -3.2 | *(-8.6 to 2.3)* | -1.8 | *(-6.8 to 3.2)* | -4.4 | *(-8.9 to 0.04)* | -5.2 | *(-10.3 to -0.1)* | -2.9 | *(-8.4 to 2.5)* |
| Austria | 10.3 | *(0.8 to 19.8)* | 14.5 | *(5.6 to 23.4)* | 12.3 | *(2.9 to 21.8)* | 8.5 | *(-1.0 to 18.0)* | 9.9 | *(1.0 to 18.7)* |
| Belgium | 2.1 | *(-3.8 to 7.9)* | 4.6 | *(-2.3 to 11.4)* | 2.2 | *(-3.7 to 8.1)* | 3.2 | *(-2.6 to 8.9)* | 3.3 | *(-2.8 to 9.5)* |
| Denmark | 0.1 | *(-5.6 to 5.8)* | 2.2 | *(-4.6 to 9.0)* | 1.2 | *(-4.5 to 6.9)* | -0.2 | *(-6.0 to 5.6)* | 0.5 | *(-5.3 to 6.3)* |
| France | 4.0 | *(-3.3 to 11.2)* | 2.0 | *(-4.3 to 8.3)* | 3.4 | *(-3.8 to 10.6)* | 1.7 | *(-5.6 to 9.1)* | 5.1 | *(-1.7 to 11.8)* |
| Germany | 2.6 | *(-3.0 to 8.3)* | 1.9 | *(-4.6 to 8.3)* | 2.9 | *(-2.8 to 8.5)* | 4.2 | *(-1.5 to 10.0)* | 2.4 | *(-3.5 to 8.3)* |
| Greece | 1.3 | *(-5.3 to 8.0)* | 3.5 | *(-3.8 to 10.9)* | 0.8 | *(-5.9 to 7.6)* | 0.2 | *(-6.6 to 6.9)* | 2.1 | *(-4.8 to 9.0)* |
| Italy | 2.9 | *(-4.3 to 10.1)* | 4.4 | *(-2.6 to 11.5)* | 3.2 | *(-4.0 to 10.3)* | 3.8 | *(-3.5 to 11.1)* | 4.8 | *(-2.3 to 12.0)* |
| Netherlands | -3.2 | *(-12.2 to 5.7)* | -5.6 | *(-12.7 to 1.5)* | -3.2 | *(-12.1 to 5.7)* | -5.9 | *(-15.1 to 3.2)* | -0.7 | *(-9.9 to 8.6)* |
| Spain | -1.3 | *(-11.0 to 8.4)* | 3.1 | *(-7.4 to 13.6)* | -0.4 | *(-10.0 to 9.1)* | -0.2 | *(-9.7 to 9.3)* | -1.8 | *(-12.1 to 8.6)* |
| Sweden | -7.4 | *(-13.6 to -1.3)* | -6.8 | *(-13.4 to -0.2)* | -7.5 | *(-13.7 to -1.3)* | -8.2 | *(-14.3 to -2.0)* | -8.0 | *(-14.1 to -2.0)* |
| Switzerland | -4.9 | *(-12.2 to 2.5)* | -5.8 | *(-14.2 to 2.7)* | -3.5 | *(-11.0 to 3.9)* | -6.6 | *(-14.0 to 0.7)* | -3.9 | *(-11.4 to 3.5)* |

#### Results of gender-stratified sensitivity analysis

The main text notes that there are four countries in which we can have some confidence that trends differ by gender. In Austria, England, Spain, and Switzerland, the employment gap between those in poor and better health becomes worse for women than men over this period, but the main text notes that the nature of this varies by country. This is shown in the table below: in Austria there are rising gaps for both women and men, whereas in England, Spain, and Switzerland there are falling gaps among men but static or increasing gaps among women.

Table A7: Gender-stratified trend in the health-related employment gap.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | ***Women (%)*** | | ***Men (%)*** | |
| USA | 8.1 | (3.6 to 12.6) | 8.1 | (2.9 to 13.3) |
| England | -0.4 | (-8.8 to 8) | -5.7 | (-13 to 1.6) |
| Austria | 19.0 | (5 to 33.1) | 5.3 | (-7.2 to 17.8) |
| Belgium | 3.9 | (-4.7 to 12.5) | 2.0 | (-6.6 to 10.5) |
| Denmark | 3.2 | (-4.7 to 11.2) | -3.3 | (-11.8 to 5.2) |
| France | 3.4 | (-6.5 to 13.4) | 6.4 | (-5.5 to 18.2) |
| Germany | 3.5 | (-4.2 to 11.1) | 3.8 | (-4.6 to 12.2) |
| Greece | 3.8 | (-5.6 to 13.2) | 8.5 | (-2.1 to 19) |
| Italy | 8.1 | (-2.9 to 19.1) | 2.8 | (-6.9 to 12.6) |
| Netherlands | 0.6 | (-10.9 to 12.1) | -4.3 | (-19.8 to 11.2) |
| Spain | 16.2 | (2.8 to 29.6) | -14.6 | (-29.6 to 0.4) |
| Sweden | -4.8 | (-13.2 to 3.5) | -10.2 | (-19.9 to -0.4) |
| Switzerland | 6.8 | (-3.9 to 17.5) | -15.1 | (-26.3 to -4) |

#### Further details of multiple imputation

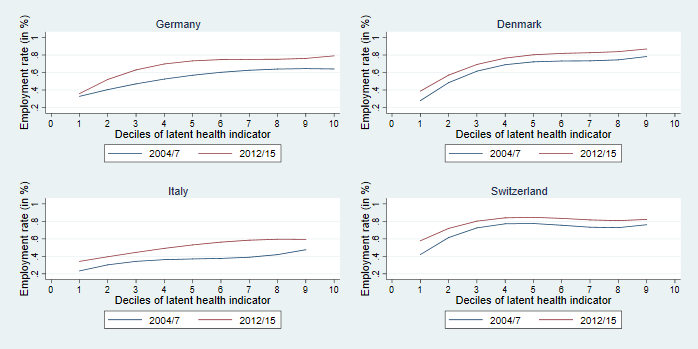
To deal with the missing data described in Online Appendix A2, we conducted multiple imputation by chained equations using the Stata multiple imputation suite. Following best practice, we constructed the imputation models to be as close to our main models as possible:

* Because we primarily treat the data as repeated cross-sections (rather than conducting longitudinal analyses at the individual level), we did not conduct a longitudinal imputation, but instead imputed responses based upon each respondent’s other responses at that wave only.
* To account for country-wave clustering, imputation was done separately for each country-wave (while not necessarily being efficient for small clusters, this is a reasonable choice if the sample sizes per cluster are large enough to make this stable; Grund, et al., 2018). The one exception was the ELSA data, for which pairs of waves were combined to enable BMI to be imputed (as BMI is only available in alternate waves).
* Imputation included all of the health variables listed in the main text, employment status, age, and gender. The only variables excluded when imputing any other variable were those that were collinear: climbing one flight of stairs vs. multiple flights of stairs; and the two ways of operationalizing mental health status, both of which were excluded from each other’s imputation model (while leaving both versions of each as predictors for other variables).
* Based on trace plots to check for convergence, we used 10 burn-in observations and 10 imputations.

Having imputed the data for each country-wave separately, the imputation files were combined for analysis. We then constructed our latent health measure using PCA for each imputation separately (so that ‘poor health’ refers to the bottom tertile within each country-wave for each imputation). We then ran analyses across the 10 imputed datasets that take into account the imputation-related uncertainty using Rubin’s rules, using the Stata multiple imputation suite. Full syntax for the multiple imputation is given in our replication files.

### Distribution of employment along the whole health distribution

Figure A1: Employment across the whole distribution of health, 2004-7 to 2012-15 (remaining countries).



Figures show kernel-weighted local polynomial smoothing of a regression of employment on health, controlling for age and gender.

## Additional material for Section 4

### Trends in disability benefit receipt

The table corresponding to Figure 5 in the text is as follows:

Table A8: *Trends in disability benefit receipt across the whole older working-age population*

|  |  |  |
| --- | --- | --- |
| **Country** | ***Trend*** | ***95% CI*** |
| USA | 1.5% | (0.6%, 2.4%) |
| England | -3.4% | (-4.7%, -2.1%) |
| Austria | 2.4% | (0.1%, 4.7%) |
| Belgium | 1.5% | (0.0%, 3.1%) |
| Denmark | -2.9% | (-4.9%, -1.0%) |
| France | 1.5% | (0.3%, 2.7%) |
| Germany | 2.1% | (0.7%, 3.5%) |
| Italy | -2.7% | (-4.2%, -1.2%) |
| Netherlands | -1.2% | (-4.0%, 1.7%) |
| Spain | -2.1% | (-4.4%, 0.1%) |
| Sweden | -8.5% | (-11.2%, -5.8%) |
| Switzerland | -2.3% | (-4.4%, -0.3%) |

### Trends in disability benefit receipt in Sweden

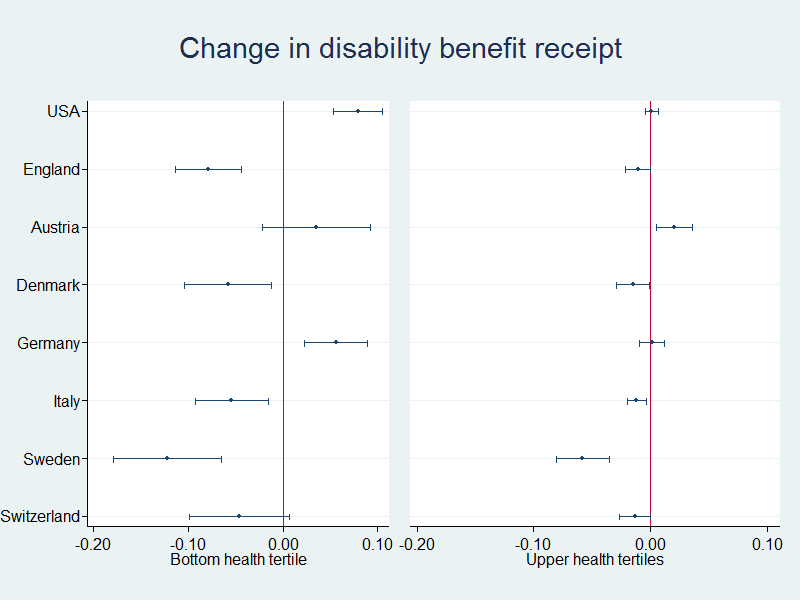
In the main text, we see a particularly sharp decline in disability benefit claims in Sweden. We carefully examined the data to check that this is not an artefactual trend due to changing questions in SHARE, and this does not appear to be the case:

* Reported claim rates in Sweden in 2004 and 2006-7 are very similar, so the trend is not due to any changes in SHARE question wording between wave 1 and wave 2.
* OECD claim data (taken from the OECD SOC R database 2007- and from OECD (2010) for 2004-7) show an almost identical proportion decrease in disability benefit claims among the whole Swedish population using administrative data (for 2013 vs. 2004, a 28% relative decline in SHARE for people between 50 and the retirement age, and a 31% decline in the administrative data for all ages). It is however worth noting that the timing of this decline is slightly different in the SHARE data than the administrative data (in SHARE the change occurs between 2011 and 2013, while in the administrative data it occurs between 2007 and 2012).

It is also worth noting here that because of estimation problems for these logit models, we control for age, age2, and gender using constant coefficients across all countries (rather than country-specific coefficients as in the models in the rest of the text). We have however re-run these models using OLS rather than logit models (for which there are no estimation problems in using country-specific coefficients), and these produce virtually identical results.

### Trends in disability benefit receipt split by health tertile

Figure A2: Disability benefit claim rates split by health, 2004-5 to 2014-15.

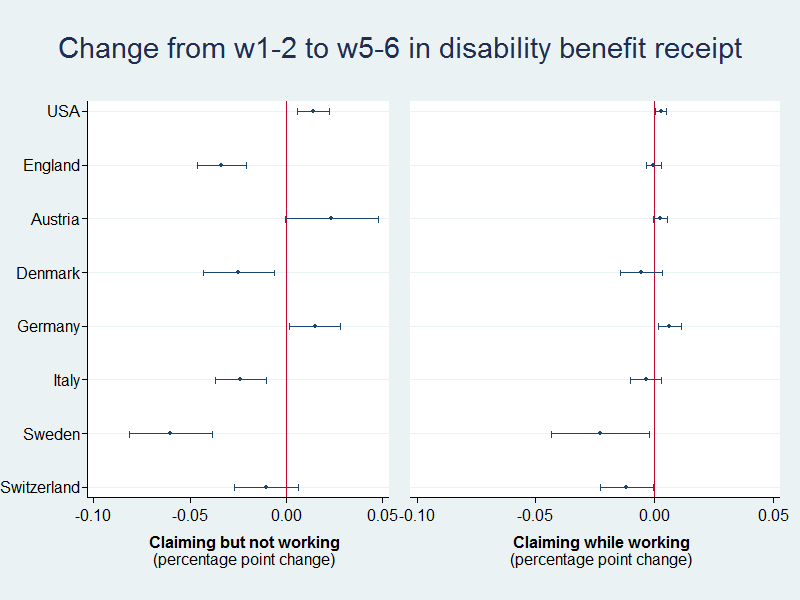


Figures show trends among the bottom health tertile (left panel) and the other two health tertiles (right panel) of our health index.

### Trends in disability benefit receipt split by employment status

Figure A3 below sheds light on how DI benefit enrolment changed with time, distinguishing between working and nonworking claimants. In all countries that reduced disability benefit claims, the reduction was stronger among inactive recipients than among those who worked. This is particularly true of Sweden, where benefit enrolment rates decreased by approximately 7 percentage points for nonworking persons. Switzerland is the only exception in this respect, as the decrease in benefit rates was approximately equal for both groups. In contrast, the increase in DI benefit rates observed in Austria, Germany, and the United States was basically concentrated on higher shares of non-working benefit claimants.

Figure A3: Changes in benefit claim rates split by employment status, 2004-5 to 2014-15.

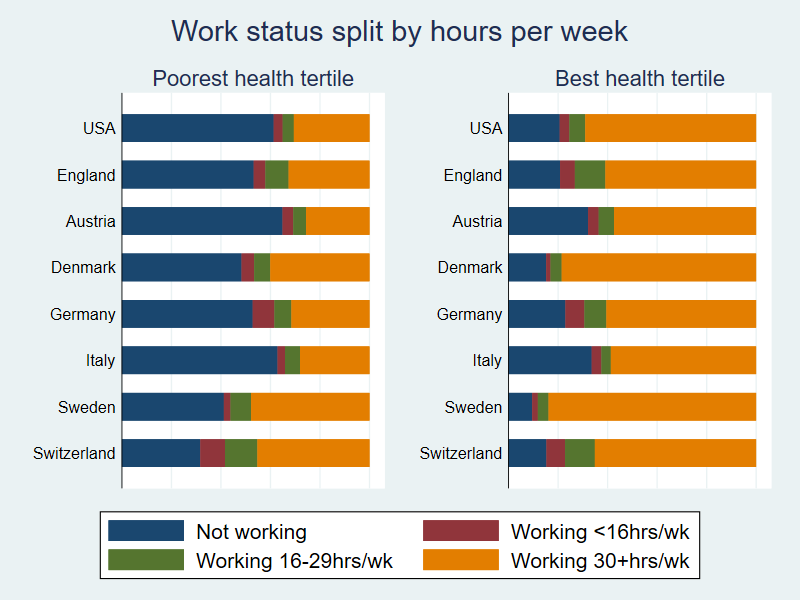
****

### Hours of work

#### Comparison of hours of work by health status, pooled across waves

In all countries, the share of workers who work fewer than 30 hours per week is greater among those with poor health status than among the healthier workers. This is however true only if we base our comparison on the number of workers who are employed. As Figure A4 below shows, among the total sample of working-aged persons there is (surprisingly) little variation in part-time employment between health tertiles within countries. It is interesting to note that Sweden – which has the largest shares of people combining benefit claim and paid work activity – does not have high numbers of workers with reduced working hours.

Figure A4: Work status split by hours per week worked, all waves.



#### Trends in hours of work

The table below corresponds to Figure 6 in the main text:

Table A9: Trend in hours of work by country, 2004-7 to 2012-15

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | ***Not working*** | | ***Working <15hrs/wk*** | | ***Working 15-29hrs/wk*** | | ***Working 30+ hrs/wk*** | |
| Country | *Trend (%)* | *95% CI* | *Trend (%)* | *95% CI* | *Trend (%)* | *95% CI* | *Trend (%)* | *95% CI* |
| USA | 5.5 | (2.6, 8.3) | 0.6 | (-0.4, 1.7) | -0.6 | (-1.8, 0.5) | -5.6 | (-8.4, -2.9) |
| England | -5.6 | (-9.9, -1.4) | -0.8 | (-2.9, 1.4) | 0.1 | (-2.9, 3.1) | 5.5 | (1.8, 9.3) |
| Austria | -6.8 | (-13.0, -0.5) | 0.4 | (-2.6, 3.3) | 3.9 | (1.1, 6.7) | 2.9 | (-2.8, 8.5) |
| Denmark | -8.3 | (-13.1, -3.6) | 0.7 | (-1.5, 2.9) | 0.7 | (-1.7, 3.2) | 6.7 | (2.1, 11.4) |
| Germany | -9.5 | (-13.9, -5.2) | 4.4 | (1.9, 6.8) | 2.8 | (0.6, 5.1) | 2.5 | (-1.6, 6.7) |
| Italy | -8.7 | (-13.8, -3.6) | -2.5 | (-4.1, -0.8) | 2.2 | (-0.5, 4.8) | 9.0 | (4.1, 13.9) |
| Sweden | -13.4 | (-19.0, -7.7) | -0.8 | (-2.3, 0.7) | -5.0 | (-8.2, -1.8) | 19.1 | (13.5, 24.7) |
| Switzerland | -12.5 | (-18.5, -6.5) | -2.8 | (-6.8, 1.2) | 4.5 | (0.4, 8.7) | 9.9 | (3.8, 15.9) |

Figures are average marginal effects controlling for age and gender.

### Job tenure

Table A10: Job tenure by health (in years)

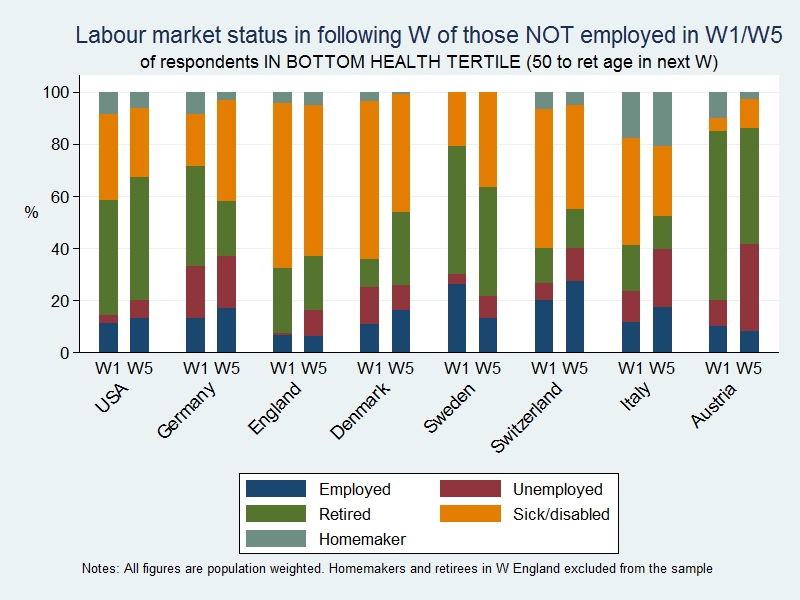
|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Country** | ***Poorest health*** | *95% CI* | ***Best  health*** | *95% CI* |
| USA | 11.5 | (11.0, 12.1) | 13.0 | (12.6, 13.4) |
| England | 10.0 | (9.4, 10.6) | 11.4 | (10.8, 12.0) |
| Austria | 22.3 | (20.7, 23.8) | 21.9 | (20.7, 23.1) |
| Belgium | 24.4 | (23.2, 25.7) | 24.6 | (23.7, 25.4) |
| Denmark | 14.8 | (14.0, 15.7) | 16.1 | (15.4, 16.9) |
| France | 22.1 | (21.0, 23.2) | 23.0 | (22.1, 23.9) |
| Germany | 17.0 | (16.1, 18.0) | 19.8 | (19.0, 20.6) |
| Greece | 24.0 | (22.9, 25.1) | 24.7 | (23.9, 25.5) |
| Italy | 24.3 | (23.0, 25.7) | 27.2 | (26.5, 28.0) |
| Netherlands | 17.6 | (16.1, 19.0) | 18.4 | (17.4, 19.4) |
| Spain | 21.5 | (19.2, 23.7) | 23.5 | (22.1, 24.9) |
| Sweden | 17.2 | (16.1, 18.3) | 16.6 | (15.6, 17.6) |
| Switzerland | 16.8 | (15.9, 17.8) | 18.0 | (17.1, 18.9) |

Table A11: Job tenure across countries, 2012-15

|  |  |  |
| --- | --- | --- |
| **Country** | ***Level*** *(years)* | ***95% CI*** |
| USA | 11.8 | 10.9 to 12.7 |
| England | 8.2 | 6.9 to 9.4 |
| Austria | 21.7 | 19.7 to 23.8 |
| Denmark | 14.1 | 13.1 to 15.1 |
| Germany | 17.1 | 16.1 to 18.2 |
| Italy | 23.8 | 22.2 to 25.3 |
| Sweden | 17.0 | 15.6 to 18.5 |
| Switzerland | 17.1 | 15.8 to 18.4 |

### Detailed labor market transitions between waves

Figure A5: Labor market status at follow-up wave of non-workers at baseline wave (for wave 1→2 and wave 5→6)



Notes: All figures are population weighted. Homemakers and retirees in baseline (W1/W5) excluded from the sample.

## Additional material for Section 5

### Macroeconomic indicators for full sample of countries

The main text primarily focuses on a geographically and institutionally dispersed subsample of countries that show different trends (Austria, Sweden, Denmark, England, Germany, Italy, and the US). However, we briefly refer to employment trends in Greece, and the macroeconomic data for the full sample of countries is as follows:

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
|  | **GDP**1 *(base year=100*) | | **Employment rate**2 **(15-64),** % | | **Employment rate3 (50-64)**, % | |
|  | 2007 -2015 | 2004 -2015 | 2008 -2015 | 2004 -2015 | 2008 -2015 | 2004 -2015 |
| USA | 113.5 | 119.6 | -2.2 | -2.5 | -0.6 | 1.6 |
| UK | 109.3 | 116.2 | 1.2 | 1.1 | 4.1 | 5.5 |
| Austria | 107.0 | 115.3 | 0.3 | 5.8 | 6.9 | 18.5 |
| Belgium | 109.3 | 115.1 | -0.6 | 1.3 | 7.5 | 12.0 |
| Denmark | 103.9 | 109.5 | -4.4 | -2.5 | 3.4 | 2.4 |
| France | 106.1 | 110.8 | -1.1 | 0.4 | 6.3 | 6.3 |
| Germany | 112.2 | 116.2 | 3.9 | 9.7 | 9.6 | 20.1 |
| Greece | 69.8 | 81.0 | -10.6 | -8.5 | -8.3 | -5.2 |
| Italy | 93.9 | 96.5 | -2.3 | -1.5 | 9.2 | 13.9 |
| Netherlands | 104.8 | 113.3 | -3.1 | 1.0 | 4.8 | 11.8 |
| Spain | 96.6 | 108.3 | -6.7 | -3.3 | 0.4 | 4.9 |
| Sweden | 117.7 | 123.6 | 1.2 | 3.1 | 3.4 | 4.4 |
| Switzerland | 115.0 | 124.6 | -0.3 | 1.8 | 1.6 | 4.6 |

Table 12: Changes in macroeconomic indicators 2004-2015

Notes: 1 GDP change refers to cumulative change since base year (indexing 2004 or 2007 values to 100); 2 Employment rate change refers to percentage point changes since base year (2004 or 2008; we use 2007 as the base year for GDP and 2008 for employment as there is a lag between GDP-based recession measures and employment consequences); 3 Employment rate for the United States refers to age 55-64 instead of 50-64. Source: European Commission annual macro-economic database (AMECO) Spring 2018, supplemented by OECD data for the US.

### The nature of work

The main text mentions in brief several caveats about the use of self-reported data on the quality of work; in this section, we discuss each of these in further detail.

***Trends among older workers***

The main text shows trends in self-reported quality of work among the whole population. If we look specifically at older workers, however, we see a slightly more negative picture (particularly for physical demands) as follows:

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| ***Trend in the whole population*** | | | | | |
| **Country** | **Job  strain** | **Job  autonomy** | **Long  hours** | **Inflexible hours** | **Physical demands** |
| USA | -2.3 | 5.1 | 2.4 | -2.3 | 25.8 |
| UK | -7.7 | 1.2 | -0.1 | -4.2 | -1.5 |
| Austria | -2.5 | 6.2 | 0.3 | 2.3 | -1.4 |
| Denmark | -5.0 | 3.1 | -1.3 | -3.2 | -1.0 |
| Germany | -16.3 | 4.3 | -0.7 | -15.9 | -1.6 |
| Italy | -6.0 | -1.5 | -0.4 | -1.1 | -5.1 |
| Sweden | 2.3 | -4.7 | 0.4 | 3.4 | -0.2 |
| ***Trend among 50-64 year olds*** | | | | | |
| **Country** | **Job  strain** | **Job  autonomy** | **Long  hours** | **Inflexible hours** | **Physical demands** |
| USA | -2.0 | 7.8 | 0.2 | -3.5 | 28.4 |
| UK | -6.2 | -6.5 | -2.6 | -5.9 | 5.4 |
| Austria | -10.6 | 5.8 | 1.1 | -3.1 | -4.6 |
| Denmark | 0.3 | -1.0 | -2.0 | -2.8 | 0.4 |
| Germany | -16.7 | 10.2 | -0.2 | -15.1 | -3.1 |
| Italy | 1.9 | -6.0 | -2.4 | 7.5 | 9.8 |
| Sweden | 1.4 | -5.6 | 0.8 | 1.5 | -0.3 |

Table 13: Trends in the quality of work 2005-2015, for older workers vs. the whole population.

Source: OECD Job Quality Database 2005-2015 based on the European Working Conditions Survey (US uses rescaled International Social Survey Programme data); no data are available for Switzerland. 1 Job strain is an OECD measure of high job demands (proxied by physical demands, long hours and inflexible hours) with low job resources (proxied by autonomy & learning opportunities, training & learning, and opportunity for career advancement); see OECD (2014) for details.

However, it is unclear that we should be interested in trends in job quality specifically among older workers – this will reflect not only the changing quality of jobs in the labor market, but also the changing patterns of which workers continue to work in this age group. For example, if workers in lower socioeconomic status jobs continue working for longer, then the apparent quality of work among older workers will decline – but this is a compositional effect, rather than reflecting a substantive change in the nature of work. It is for this reason that the main text focuses on trends among the whole population.

***Trends in the US***

As the table above shows, there is a very large increase in physical demands in the US 2005-2015. This is not an error in the OECD database but is mirrored in the raw weighted ISSP data, in which respondents saying they ‘always’ or ‘often’ do ‘hard physical work’ rises from 23.9% in 2005 to 47.1% in 2015. While there are slight changes in the ISSP questionnaire, these do not appear to explain this considerable shift as there are few signs of changes on this scale in other countries. Nor is there a clear change in survey mode or routing that could explain this. While there is some evidence that reported physical demands in countries such as the UK and Norway have slightly risen in recent decades despite the declining role of manufacturing jobs (Felstead, et al., 2007:28,86, Green, 2009:26, Handel, 2012:47, Olsen, et al., 2010:233-234), our view is that the size of this change appears to be implausible. Nevertheless, readers who find the trend more believable may view this as a possible explanation of the anomalous trends in the US.

***Comparison with European Social Survey trends***

While our intention is not to conduct a systematic review of all of the different data sources on work quality, it is worth noting that the OECD database trends do not fully align with those shown in the European Social Survey. If we look at the four countries for which trends in work autonomy are available in both datasets (comparing the results above to Gallie, 2013), we see:

* Concordance in Denmark, where work control increased in both datasets;
* Concordance in Germany, where work control increased in both datasets;
* Discordance in Sweden, where work control decreased in the OECD/EWCS data but was unchanged in ESS;
* Discordance in the UK, where work control slightly declined in the ESS data but slightly improved in the OECD/EWCS data.

These comparisons may simply reflect sampling fluctuations, but even if so, this draws attention to the potential variation in trends across surveys, particularly when point estimates rather than confidence intervals are used {see also \Baumberg, 2011 #1570}{Geiger, 2017 #3118}.

## Additional material on the PCA

Section 2.3 in the main paper explains our analytical approach, including that we estimate the PCA separately for each country (but pool waves within each country, to ensure that the health index has a consistent weighting within countries over time). However, it is worth noting that the PCA weightings across countries are actually very consistent, as shown in the table below.

In only two cases do weights differ moderately – in both cases these are for motor skills in England vs. all other countries (where the weight given to picking up a small coin is 0.27, compared to 0.13-0.20 elsewhere; and where the weight given to getting up from a chair is 0.16, compared to 0.25-0.27 elsewhere). There are no *a priori* reasons why we should expect these motor skills measures to behave differently in England than elsewhere (there are some differences in measures between countries as we explain in Online Appendix A1, but none that concern these two measures specifically in England), but it is worth noting that other studies have also found greater differences in the clustering of related health measures in ELSA compared to HRS or SHARE (Chan, et al., 2012).

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | USA | UK | Austria | Belgium | Denmark | France | Germany | Greece | Italy | NL | Spain | Sweden | Switzerland | Isreal | Czech R | Poland | **Range** |
| Difficulties: |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| walk 100m | 0.25 | 0.28 | 0.25 | 0.27 | 0.26 | 0.26 | 0.26 | 0.28 | 0.27 | 0.27 | 0.26 | 0.25 | 0.23 | 0.25 | 0.27 | 0.25 | 0.05 |
| lift 5kg | 0.28 | 0.28 | 0.27 | 0.27 | 0.27 | 0.27 | 0.27 | 0.25 | 0.27 | 0.28 | 0.29 | 0.29 | 0.28 | 0.28 | 0.27 | 0.28 | 0.03 |
| push large objects | 0.27 | 0.28 | 0.26 | 0.27 | 0.26 | 0.27 | 0.26 | 0.25 | 0.27 | 0.28 | 0.28 | 0.27 | 0.25 | 0.27 | 0.25 | 0.28 | 0.04 |
| climb 1 flight of stairs | 0.26 | 0.28 | 0.27 | 0.26 | 0.27 | 0.26 | 0.24 | 0.26 | 0.27 | 0.28 | 0.27 | 0.25 | 0.25 | 0.27 | 0.25 | 0.24 | 0.04 |
| climb several flights of stairs | 0.26 | 0.25 | 0.28 | 0.28 | 0.30 | 0.29 | 0.29 | 0.26 | 0.28 | 0.29 | 0.29 | 0.30 | 0.29 | 0.28 | 0.28 | 0.28 | 0.05 |
| stoop, kneel or crouch | 0.25 | 0.27 | 0.27 | 0.26 | 0.27 | 0.29 | 0.27 | 0.28 | 0.27 | 0.28 | 0.29 | 0.25 | 0.26 | 0.27 | 0.28 | 0.29 | 0.04 |
| pick up small coin | 0.15 | 0.27 | 0.20 | 0.14 | 0.16 | 0.14 | 0.17 | 0.19 | 0.17 | 0.15 | 0.15 | 0.19 | 0.13 | 0.14 | 0.15 | 0.12 | 0.15 |
| sit for 2 hours | 0.23 | 0.25 | 0.20 | 0.20 | 0.22 | 0.22 | 0.23 | 0.21 | 0.21 | 0.22 | 0.23 | 0.22 | 0.22 | 0.24 | 0.22 | 0.23 | 0.05 |
| get up from chair | 0.26 | 0.16 | 0.25 | 0.25 | 0.26 | 0.26 | 0.26 | 0.26 | 0.26 | 0.26 | 0.28 | 0.26 | 0.25 | 0.27 | 0.27 | 0.27 | 0.12 |
| reach above shoulder | 0.22 | 0.24 | 0.21 | 0.20 | 0.20 | 0.20 | 0.23 | 0.22 | 0.22 | 0.21 | 0.24 | 0.22 | 0.21 | 0.22 | 0.23 | 0.24 | 0.04 |
| Any of 6 ADL | 0.27 | 0.26 | 0.25 | 0.26 | 0.27 | 0.26 | 0.25 | 0.26 | 0.24 | 0.25 | 0.24 | 0.26 | 0.27 | 0.26 | 0.27 | 0.27 | 0.04 |
| Any of 5 IADL | 0.23 | 0.21 | 0.23 | 0.24 | 0.24 | 0.22 | 0.21 | 0.22 | 0.22 | 0.22 | 0.17 | 0.24 | 0.22 | 0.25 | 0.22 | 0.22 | 0.08 |
| high blood pressure | 0.12 | 0.09 | 0.11 | 0.09 | 0.08 | 0.10 | 0.11 | 0.13 | 0.12 | 0.07 | 0.09 | 0.07 | 0.10 | 0.10 | 0.11 | 0.12 | 0.06 |
| stroke | 0.10 | 0.07 | 0.12 | 0.09 | 0.09 | 0.08 | 0.12 | 0.11 | 0.09 | 0.10 | 0.08 | 0.10 | 0.07 | 0.12 | 0.11 | 0.13 | 0.06 |
| diabetes | 0.11 | 0.08 | 0.13 | 0.09 | 0.08 | 0.10 | 0.10 | 0.10 | 0.10 | 0.10 | 0.10 | 0.07 | 0.11 | 0.13 | 0.12 | 0.10 | 0.06 |
| chronic lung disease | 0.13 | 0.10 | 0.11 | 0.14 | 0.12 | 0.10 | 0.11 | 0.11 | 0.11 | 0.09 | 0.09 | 0.11 | 0.12 | 0.08 | 0.12 | 0.09 | 0.06 |
| ever had heart problems | 0.12 | 0.08 | 0.10 | 0.10 | 0.10 | 0.12 | 0.11 | 0.10 | 0.12 | 0.09 | 0.11 | 0.10 | 0.09 | 0.12 | 0.12 | 0.12 | 0.04 |
| cancer | 0.04 | 0.03 | 0.06 | 0.05 | 0.04 | 0.06 | 0.06 | 0.07 | 0.07 | 0.04 | 0.04 | 0.03 | 0.05 | 0.04 | 0.06 | 0.08 | 0.06 |
| arthritis | 0.17 | 0.18 | 0.13 | 0.18 | 0.15 | 0.17 | 0.15 | 0.20 | 0.17 | 0.15 | 0.17 | 0.15 | 0.17 | 0.14 | 0.15 | 0.17 | 0.06 |
| depression caseness (CESD) | 0.21 | 0.20 | 0.20 | 0.21 | 0.20 | 0.20 | 0.21 | 0.20 | 0.22 | 0.19 | 0.21 | 0.20 | 0.23 | 0.20 | 0.20 | 0.21 | 0.04 |
| depression symptoms (CESD) | 0.24 | 0.21 | 0.23 | 0.24 | 0.23 | 0.23 | 0.24 | 0.22 | 0.24 | 0.21 | 0.23 | 0.23 | 0.25 | 0.23 | 0.23 | 0.23 | 0.04 |
| self-reported health | 0.24 | 0.26 | 0.27 | 0.27 | 0.27 | 0.27 | 0.27 | 0.27 | 0.24 | 0.27 | 0.25 | 0.27 | 0.29 | 0.22 | 0.24 | 0.23 | 0.06 |
| underweight | 0.01 | 0.01 | 0.02 | 0.04 | 0.02 | 0.00 | 0.03 | -0.01 | 0.00 | 0.04 | 0.00 | 0.03 | 0.02 | 0.02 | 0.04 | 0.04 | 0.05 |
| obese | 0.10 | 0.10 | 0.12 | 0.09 | 0.09 | 0.12 | 0.11 | 0.12 | 0.09 | 0.10 | 0.10 | 0.10 | 0.12 | 0.10 | 0.10 | 0.06 | 0.06 |

## Further detail on sickness and disability policy reforms

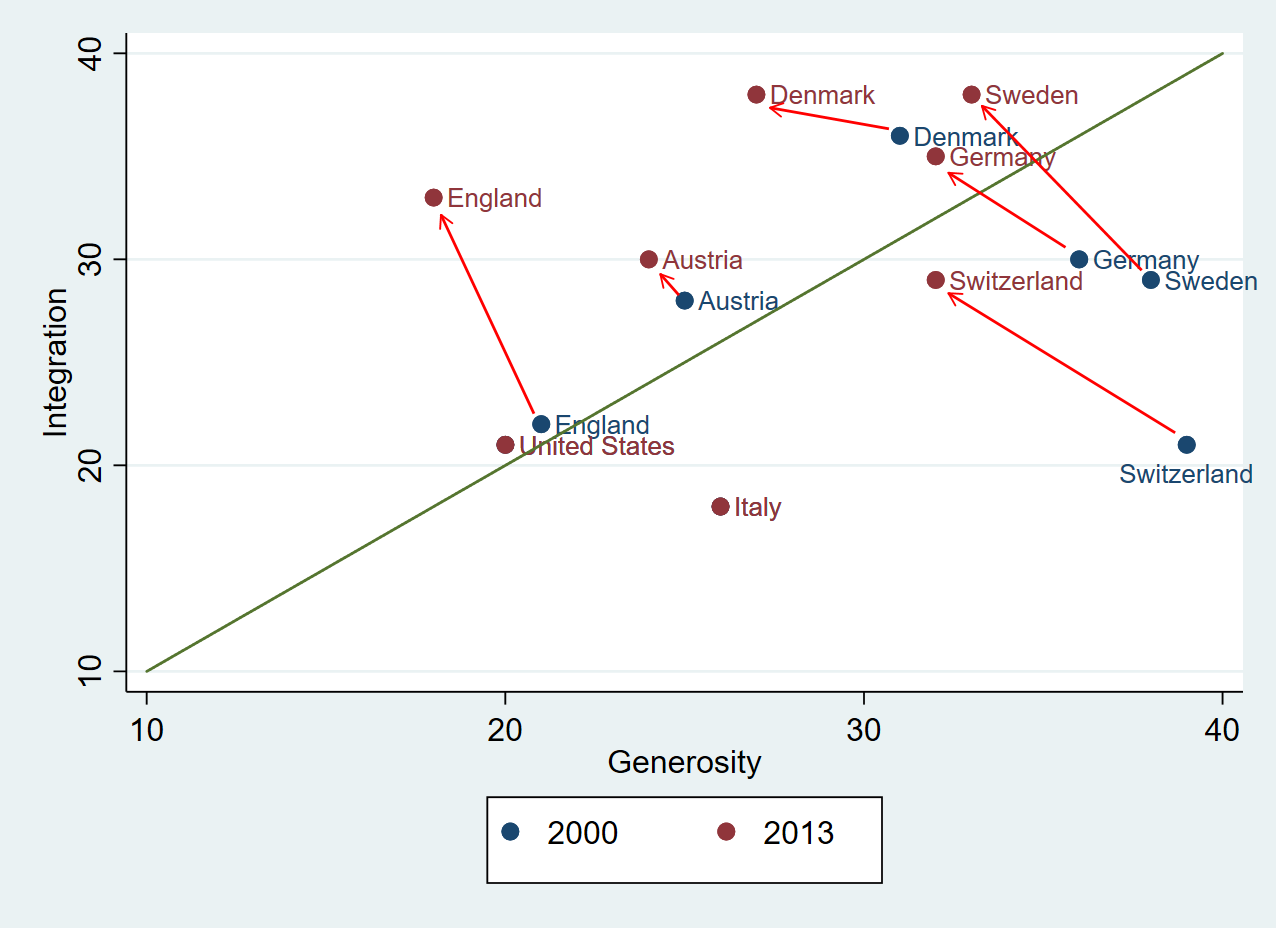
We here summarize changes in the field of sickness and disability policies in the countries that we focus on in the main text, using two approaches: (i) quantitative indicators provided by the OECD and (ii) qualitative information from the wider literature. Inevitably, when we look at many countries and an extended period of time, we are confronted with numerous policy adaptations and legislative changes. Rather than to give a complete and exhaustive account of these changes, the aim of the following paragraphs is to provide a picture of the heterogeneous reform activities and the main reform trajectories in the countries under scrutiny. This overview draws partially on the extension of the OECD scores and the related information collected in (Böheim and Leoni, 2018).

Policies are likely to require time until they are implemented and further time to actually change behavior. For this reason, we focus on institutional settings and reforms for the period 2000 to 2013, i.e. leaving a time-lag with respect to the moment at which data were collected by SHARE, ELSA, and HRS.

### Policy Detail Part I: Evidence from quantitative indicators

The figure below shows the change in the two policy dimensions categorized by the OECD (generosity and integration). All countries shifted their policy focus in the same direction, i.e. they reduced the generosity (i.e. compensation dimension) of their systems and strengthened efforts to activate and integrate workers with health problems in the labor market. Only the United States and Italy did not implement any reforms that affected these policy scores. The strongest changes can be observed in the Sweden, Switzerland, and England. Germany and Denmark carried out some significant changes, too, whereas Austria recorded only moderate changes. Broadly speaking, we may distinguish a group which consists of the Scandinavian countries together with Germany, which at the end of the period under scrutiny is characterized by high levels of social protection in combination with strong activation and integration policies. The United States and England have much lower generosity levels and, at least in the case of the US, also lower levels of integration policies. Switzerland and Austria take an intermediate position, whereas Italy can be singled out for its combination of moderately high compensation and very low integration policies.

Figure A6: Changes in sickness and disability policy in selected countries, 2000 to 2013.

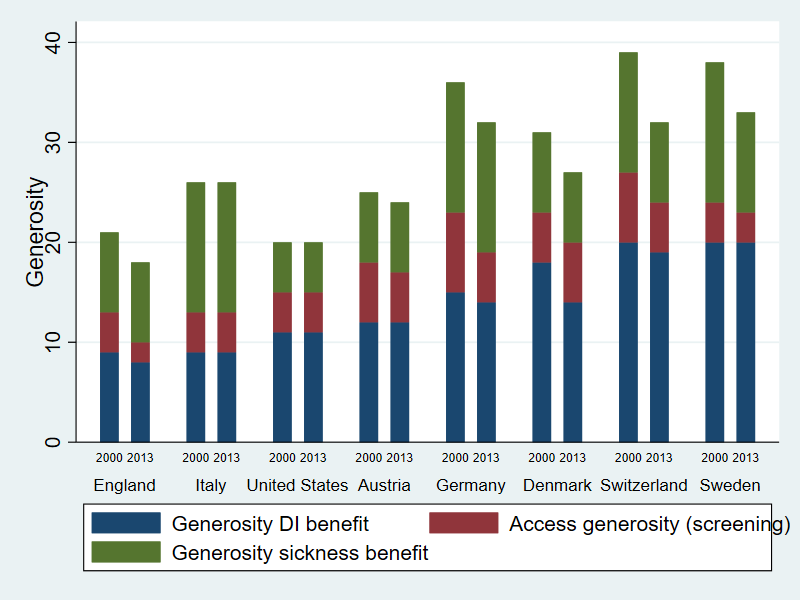


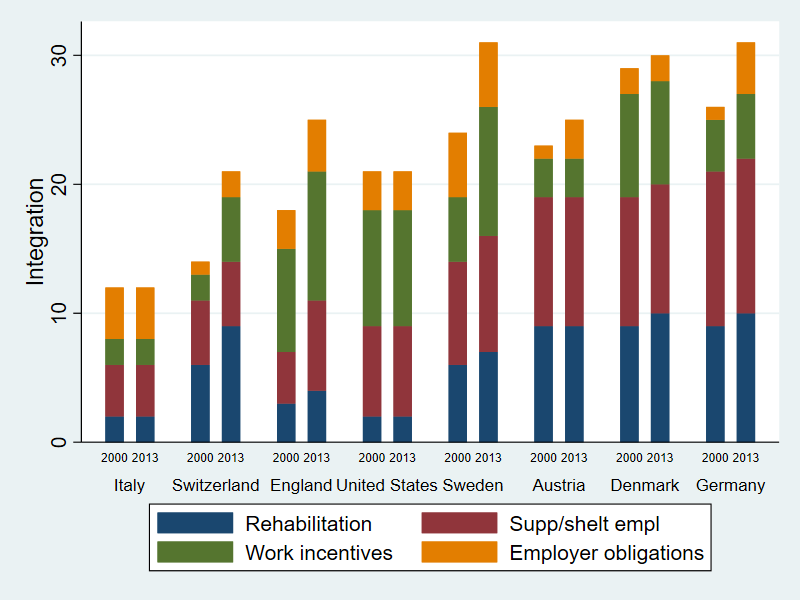
Source: authors’ calculations based on OECD (2003, 2010) and Böheim & Leoni (2018). Each scale is based on 10 sub-components ranging from 0 to 5, with a maximum of 50 points.

The next graphs show that the reductions in the compensation dimension were partially driven by reductions in benefit generosity. This concerned mainly DI benefits (in England, Germany, Denmark, and Sweden) and to a lesser extent sickness benefits (Switzerland and Sweden). Retrenchment did however often take the form of tighter screening for the application of benefits (in England, Austria, Germany, Switzerland, and Sweden).

With respect to the integration dimension, we can observe substantial differences with respect to how countries organize rehabilitation. Rehabilitation policies play a subordinate role in the Anglo-Saxon countries and in Italy, while they are more prominent and have been further expanded over time in the remaining countries. The Anglo-Saxon countries, on the other hand, are characterized by more conditional levels of social protection as well as by stronger emphasis on ‘workfarist’ policy elements which are aimed at activating working-aged persons primarily through incentives and sanctions. Work incentives as well as employer obligations were however also strengthened in several Continental European countries.

Figure A7: Policy scores and changes in sub-components of the two dimensions.





Source: authors’ calculations based on OECD (2003, 2010); Böheim & Leoni (2018).

### Policy Detail Part II: Qualitative summaries of policy change

The aim of this section is to give a brief description of salient reforms in the field of sickness and disability policy which were carried out since 2000, thus complementing the quantitative evidence discussed in the previous section.

In **Switzerland,** Disability Insurance (*Invalidenversicherung*) underwent several significant revision rounds. In 2003/4, as part of the fourth revision of the Disability Insurance Act stricter medical assessment criteria and early disability risk identification mechanisms were introduced. The fifth revision, which became effective in 2008, reinforced the emphasis on vocational rehabilitation, added a new focus on job retention, and implemented a paradigm shift in the disability insurance (OECD, 2014). Not unlike the fourth revision (2004), the fifth revision focused primarily on reducing the number of new claims, and it resulted in significant changes in the way in which the benefit system is being accessed. Together with a set of early intervention measures, substantial wage subsidies for employers hiring a claimant were introduced and vocational rehabilitation was strengthened. The sixth revision of DI, which intended to address the benefit caseload, was only partially implemented because the second half of the reform was not approved by Parliament. Vocational rehabilitation was however further strengthened in 2013.

In reaction to high sickness absence rates and rising disability benefit rolls, **Sweden** began to strengthen its activation policies already in the 1990s, mainly by expanding supported employment programs and by introducing options to suspend benefit receipt and return to work. More radical reforms, including strong retrenchment elements, were implemented as part of the welfare reforms of 2008. Sweden emphasized the incentives to take up work while also improving the employment support for individuals with disabilities. The principal reform component was the establishment of a new timeline for the provision of rehabilitation services, with additional work capacity assessments during sickness (OECD, 2013). The law which was enforced in 2008 requires workers who are sick more than six months, to accept an adequate job offer by the social security service. Sweden also limited the duration of sickness benefit receipt and introduced comprehensive guidelines for sickness absence certification in 2008 (Försäkringskassan, 2014, Skånér, et al., 2011). These guidelines contain general recommendations and provide doctors with reference values for the length and extent of sickness absence for different diagnoses. At the same time the use of partial sick-leave, which had already been introduced in the 1970s, was boosted in an effort to increase the focus on residual work ability and to avoid long absences from the workplace.

Not unlike Sweden, also **Denmark** focused on stricter sickness absence monitoring and a procedure of early risk identification and intervention to lower the inflow in disability benefits. After the introduction of a sickness absence monitoring process in the late 1990s, accompanied by a reduction in disability benefit generosity, a major reform was decided in 2003 and implemented over a transitional period. Among other things, Denmark changed its approach to assessing entitlement for disability benefits, with a stronger focus on residual work capacity. Whereas in other countries, most notably the Netherlands, the employer has become a major stake-holder in the monitoring and activation process, in Denmark it is mainly the municipal job centre, which is responsible for all people seeking to stay in or return to employment, irrespective of the type of benefit they receive. Denmark further intensified its sickness absence monitoring in 2010. In 2014 it also restricted the access to disability pension benefits for those aged under 40 and increased the requirements for older persons to take part in a rehabilitation program before becoming eligible for the benefit.

**Germany** carried out a major pension reform in 2001/2002, which included a harmonization of two different types of disability pension benefit as well as measures to slow down benefit inflow. In 2001, it expanded the Social Code (with the ‘Ninth Book’, *Neuntes Buch Sozialgesetzgebung*), introducing the right to rehabilitation and strengthening the rights of persons with disability. This new legislation to promote the self-determination and inclusion of people with disability included also a focus on prevention. As part of this preventive focus, in 2004 a sickness absence monitoring that entails a stronger involvement of employers in the rehabilitation process was introduced. The law institutionalised disability management (*Betriebliches Eingliederungsmanagment*), mandating firms to contact their employees who are on sick leave for a longer period of time (six weeks) and to offer them support for their return to work. Germany also boosted supported employment in 2008, when it introduced a legal definition for this form of activation measure. The new law led to the creation of specific supported employment programs by the Public Employment Service (*Agentur für Arbeit*) as part of vocational training measures (*Maßnahmen zur betrieblichen Qualifizierung*).

Despite having an institutional set-up with strong parallels to the German one, **Austria** implemented less extensive reforms in the period under scrutiny. With a lag of several years, in 2011, it passed a law that resembles the German disability management regulation. The new law instituted an informational and consulting service of early intervention, which is activated when workers show an indication or risk for long term sickness or permanent reduction of the work ability. However, whereas in Germany employers are mandated to take action, in Austria it is the social insurance agency that offers consulting services, but only on a voluntary basis. The Austrian reform has thus not altered substantially the incentives for firms and workers to get involved in early intervention and prevention. A more far-reaching reform of the Austrian DI system, including the abolition of disability pensions for younger workers and implementation of a more stringent rehabilitation process, came into effect in 2014. This reform has the potential to have a substantial impact on the labor market integration of workers with health problems in the long run. The measures however apply only to individuals born in 1964 or later, which means that they did hardly affect the population aged 50 years and above surveyed in the sixth SHARE wave.

Substantial changes in the integration dimension for the **United Kingdom** are related to the implementation of the New Deal for Disabled People (NDDP) and the Pathways to Work program. The NDDP was rolled out at a national level in 2001, Pathways to Work was piloted in 2003 and rolled out nationally in 2005 (Burkhauser, et al., 2014, DWP, 2007). These programs were aimed at supporting people on incapacity-related benefits in securing employment, slowing the inflow of disability beneficiaries, and boosting outflows for recent beneficiaries. Mandatory work-focused interviews were a central component of these programs. The UK carried out a further, far-reaching reform in 2008, with the introduction of the Employment and Support Allowance (ESA). ESA replaced three different benefits and introduced a new and stricter assessment procedure ‒ the Work Capability Assessment (Burkhauser, et al., 2014, Morris, 2016). The United Kingdom also attempted a large-scale reassessment of the work capacity of its invalidity benefit claimants. The reassessment, however, led to a wave of appeals, a large backlog of claims awaiting assessment, political controversy and, ultimately, it resulted in few reactivations (Gaffney, 2015).

In the two remaining countries that are included in our data set, reform activity has been less intense:

* This is particularly true of the **United States**, where, in spite of rising benefit enrolment, hardly any action was taken. Morris (2016) points out that this lack of reform activity can at least partly be explained with the structural arrangements of the Social Security Disability Insurance (SSDI).
* The OECD policy scores did not change in our period of observation for **Italy**, too. This lack of changes, at least with respect to the activation dimension, is partially attributable to the fact that in Italy measures and services to support persons with disabilities are implemented on a regional or local level and are therefore more diverse and difficult to survey. At the national level, several rounds of pension system reform were implemented since the early 1990s, mainly with the aim to reduce benefit generosity and to enhance fiscal sustainability (Brugiavini and Peracchi, 2012, Marano, 2006, Natali and Stamati, 2013). The most recent, substantial reforms took place in 2004/5 and – under pressure from the crisis – in 2011. The sickness and disability insurance programs were however not directly addressed by these reforms. A legislative change that affected the activation/integration dimension was however introduced in 2013, when Italy mandated employers to provide reasonable accommodation at workplace for workers with disabilities (Legislative Decree 76/2013).

## References for Online Appendices

1. Böheim, R., Leoni, T., 2018. Sickness and disability policies: Reform paths in OECD countries between 1990 and 2014. International Journal of Social Welfare 27**,** 168-185.

2. Breeze, E., Lang, I., 2006. Physical functioning in a community context. In: Banks, J., and et al, (Eds.), Living in the 21st Century: older people in England (The 2006 English Longitudinal Study of Ageing, Wave 3). Institute for Fiscal Studies, London, pp. 57-117.

3. Brugiavini, A., Peracchi, F., 2012. Health Status, Welfare Programs Participation, and Labor Force Activity in Italy. In: Wise, D. A., (Ed., Social Security Programs and Retirement around the World. University of Chicago Press, Chicago, pp. 175–215.

4. Burkhauser, R.V., Daly, M.C., McVicar, D., Wilkins, R., 2014. Disability benefit growth and disability reform in the US: lessons from other OECD nations. IZA Journal of Labor Policy 3**,** 1-30.

5. Chan, K.S., Kasper, J.D., Brandt, J., Pezzin, L.E., 2012. Measurement Equivalence in ADL and IADL Difficulty Across International Surveys of Aging: Findings From the HRS, SHARE, and ELSA. The Journals of Gerontology Series B: Psychological Sciences and Social Sciences 67B**,** 121-132.

6. Courtin, E., Knapp, M., Grundy, E., Avendano-Pabon, M., 2015. Are different measures of depressive symptoms in old age comparable? An analysis of the CES-D and Euro-D scales in 13 countries. International Journal of Methods in Psychiatric Research 24**,** 287-304.

7. Crimmins, E.M., Kim, J.K., Solé-Auró, A., 2010. Gender differences in health: results from SHARE, ELSA and HRS. European Journal of Public Health 21**,** 81-91.

8. DWP, 2007. New Deal for Disabled People: Third synthesis report – key findings from the evaluation, Research Report No 430. Department of Work and Pensions (DWP), London.

9. Felstead, A., Gallie, D., Green, F., Zhou, Y., 2007. Skills at Work, 1986 to 2006. ESRC Centre on Skills, Knowledge and Organisational Performance (SKOPE).

10. Fernandes, M., Zamarro, G., Meijer, E., 2008. 2.2 Health Comparisons. In: Börsch-Supan, A., Brugiavini, A., Jürges, H., Kapteyn, A., Mackenbach, J., Siegrist, J., and Weber, G., (Eds.), First Results from the Survey of Health, Ageing and Retirement in Europe (2004-2007): Starting the Longitudinal Dimension. SHARE, Mannheim.

11. Försäkringskassan, 2014. Analyzing the variation in the level of sickness absence, Social Insurance Report 17. Försäkringskassan (Social Insurance Office), Stockholm.

12. Gaffney, D., 2015. Retrenchment, Reform, Continuity: Welfare Under The Coalition. National Institute Economic Review 23**,** 45-53.

13. Gallie, D., 2013. Economic Crisis, Quality of Work and Social Integration: Topline Results from Rounds 2 and 5 of the European Social Survey, ESS Topline Results Series issue 3. European Social Survey.

14. Green, F., 2009. Employee involvement, technology and job tasks, NIESR Discussion Paper No. 326. National Institute of Economic and Social Research.

15. Grund, S., Lüdtke, O., Robitzsch, A., 2018. Multiple Imputation of Missing Data for Multilevel Models:Simulations and Recommendations. Organizational Research Methods 21**,** 111-149.

16. Handel, M.J., 2012. Trends in Job Skill Demands in OECD Countries, OECD Social, Employment and Migration Working Papers, No. 143. OECD Publishing, Paris.

17. Marano, A., 2006. Pension Reforms in Italy: Principles and Consequences. Revue française des affaires sociales 5**,** 223-252.

18. Morris, Z., 2016. Constructing the need for retrenchment: disability benefits in the United States and Great Britain. Policy & Politics 44**,** 609-626.

19. Natali, D., Stamati, F., 2013. Reforming pensions in Europe: a comparative country analysis, ETUI Working Paper 2013.08. ETUI, Florence.

20. OECD, 2003. Transforming Disability into Ability: Policies to Promote Work and Income Security for Disabled People. OECD, Paris.

21. OECD, 2010. Sickness, Disability and Work: breaking the barriers. A synthesis of findings across OECD countries. OECD, Paris.

22. OECD, 2013. Mental Health and Work: Sweden. OECD, Paris.

23. OECD, 2014. Mental Health and Work: Switzerland. OECD, Paris.

24. OECD, 2014. OECD Employment Outlook 2014. OECD Publishing, Paris.

25. Olsen, K., Kalleberg, A.L., Nesheim, T., 2010. Perceived job quality in the United States, Great Britain, Norway and West Germany, 1989-2005. European Journal of Industrial Relations 16**,** 221-240.

26. Riumallo-Herl, C., Basu, S., Stuckler, D., Courtin, E., Avendano, M., 2014. Job loss, wealth and depression during the Great Recession in the USA and Europe. International Journal of Epidemiology 43**,** 1508-1517.

27. Skånér, Y., Nilsson, G.H., Arrelöv, B., Lindholm, C., Hinas, E., Wilteus, A.L., Alexanderson, K., 2011. Use and usefulness of guidelines for sickness certification: results from a national survey of all general practitioners in Sweden. BMJ Open 1.

28. Zamarro, G., Lee, J., 2011. Harmonization of Cross-National Studies of Aging to the Health and Retirement Study, RAND Working Paper. RAND.

29. Zamarro, G., Meijer, E., Fernandes, M., 2008. 2.3 Mental Health and Cognitive Ability. In: Börsch-Supan, A., Brugiavini, A., Jürges, H., Kapteyn, A., Mackenbach, J., Siegrist, J., and Weber, G., (Eds.), First Results from the Survey of Health, Ageing and Retirement in Europe (2004-2007): Starting the Longitudinal Dimension. SHARE, Mannheim.

1. These are:

   RAND HRS Data Documentation Version P, by Delia Bugliari et al, August 2016. Labor & Population Program of the RAND Center for the Study of Aging.

   Harmonized ELSA dataset and Codebook, Version E as of April 2017 developed by the Gateway to Global Aging Data;

   Harmonized SHARE dataset and Codebook, Version D.2 as of September 2017 developed by the Gateway to Global Aging Data.

   (Gateway to Global Aging Data also request acknowledgement of their funding by the National Institute on Ageing (R01 AG030153, RC2 AG036619, R03 AG043052, R03AG043052). [↑](#footnote-ref-1)
2. In HRS it would in fact be possible to include a measure of current disability benefit receipt, but this is not available in SHARE, and for consistency we have instead focussed on full-year claims. [↑](#footnote-ref-2)
3. We do not use the final IADL measure, “having problems in using a map”. The responses have many more missing entries than the other IADL measures, the measure does not group consistently with other IADL measures (Breeze and Lang, 2006), and seems not to be comparable between HRS and SHARE (Fernandes, et al., 2008). [↑](#footnote-ref-3)