The Missing Arduous Careers in Retirement Survey Data

Preliminary Version. Do not quote

Vincent Vandenberghe*

April 24, 2024

Abstract

In this paper, we highlight the propensity of people who had more(less) arduous careers to be under-(over)represented in cross-sectional health and retirement surveys as age rises. We show that this is the case with the SHARE wave 7 survey data. In line with the literature on the "missing poors", we argue that this could be evidence of more(less) attrition due to death among the workers who exerted more(less) arduous careers. We then quantify the attrition/mortality differential compatible with the under-(over)representation by arduousness quantile found in SHARE wave 7. Finally, we assess the plausibility of our results using a more traditional longitudinal perspective made possible by the use of SHARE wave 8.

Keywords: Career Arduousness, Longevity Differences, Missing Poors

JEL Codes: J81, J24, I10, J26

*Economics School of Louvain (ESL), IRES-LIDAM, Université catholique de Louvain (UCL), 3 place Montesquieu, B-1348 Belgium. Corresponding author: vincent.vandenberghe@uclouvain.be

1 Introduction

In advanced economies, jobs and careers still vary significantly in terms of their degree of arduousness (and other work-related dimensions). It has been shown that this may affect both physical and mental health in late-life years (Ravesteijn et al., 2018; Vandenberghe, 2023). Many works also posit and/or show that career arduousness affects `longevity`. The best way to assess the link between work arduousness and longevity is to use (preferably entire life) longitudinal data, as done by Mikkola et al. (2019) or Katikireddi et al. (2017). In this paper, we argue that there might be a second-best approach. It consists of *i)* using cross-sectional survey data augmented by retrospective modules informing about respondents' job history/career arduousness and *ii)* exploiting the fact that people who had more(less) arduous careers are `less(more) present` in these surveys as age rises. We posit here that such a gradient can be considered as a proxy of death-related attrition.

The exercise exposed hereafter can be related to the burgeoning literature on the "missing poors"(Kanbur and Mukherjee, 2007; Lefebvre et al., 2019; Lefebvre et al., 2023). These papers have demonstrated the risk with cross-sectional data of underestimating old-age poverty, essentially because people with the lowest level of income are missing from the data due to their propensity to die earlier. In the same vein, this paper shows that in data surveying older people – like the Survey of Health and Retirement in Europe (SHARE) – arduousness declines with the age of the respondents due to what we call "missing arduous career" problem. This paper should be seen as a contribution to the literature on arduous and hazardous jobs and careers (Bakker and Demerouti, 2007; Vermeer et al., 2016; Barnay, 2016; Bassanini and Caroli, 2015; Chen et al., 2017; Vandenberghe, 2023), particularly their consequences in terms of longevity differences.

The rest of this paper is organised as follows. The data on career arduousness assembled and used in this paper are presented in Section 2. Section 3 presents the cross-sectional evidence on the "missing arduous careers" as the age of respondents rises. That section also contains our estimates of the survival/mortality differences related to career arduousness that might generate the sort of cross-sectional pattern we find in SHARE wave 7. Section 5 uses the panel dimension of SHARE (i.e. transitions from wave 7 to wave 8 and the longitudinal information on mortality they contain) to assess the plausibility of Section 4 results. Section 6 concludes.

2 Data

2.1 SHARE wave 7- job history

The analysis of the career arduousness/longevity relationship at the core of this paper rests on a (quite important and time-consuming) preliminary work that consists of quantifying the arduousness of the `entire career` of SHARE respondents. That task is based on the use of the 7th wave of SHARE. This wave was assembled in 2018 across 28 European countries plus Israel (Table 1). It contains several "retrospective" modules that provide detailed data about the respondent's history. Extensive information is provided about job history at ISCO4 level.

In the 7th wave of SHARE, respondents are asked to retrace their complete job history by providing the starting/ending year of each of their successive jobs/occupations, and whether these were done on a full- or part-time basis. A participant's history is reported retrospectively and thus a long time after work happened (i.e. a retiree in 2018 must recall her work history since 1970 if she started working at age 20). This can lead to memory biases. To reduce this problem, the SHARE surveyors used a "Life History Calendar" approach to help the respondent report accurately. The Life History Calendar (LHC) method uses a calendar-like matrix to map out life events, providing visual cues to both the interviewer and interviewee regarding the onset, duration, sequencing, and co-occurrence of events. The calendar includes rows, which are categories of life events; these might include schools attended, jobs, living arrangements, dating relationships, and so on. Numerous innovations of the LHC provide benefits relative to data collection through traditional questionnaires. The LHC's columns encourage recall at the temporal level, while the rows encourage recall at the thematic level. The LHC has been tested extensively with respondents of varying ages and cultural backgrounds, including those with unstable lives and cognitive difficulties (DeHart, 2021). The LHC permits calculating the duration of their entire career, both in absolute years and in equivalent-full-time years. Also, the occupation title is reported for each of the successive jobs/occupations at ISCO-4 digits. We merge that information with arduousness indices that have been estimated separately for each ISCO-4 occupation (more on this below in Section 2.2). The combination of SHARE job history data and arduousness data puts us in a position to compute, *inter alia*, an average career arduousness index and examine how it correlates with a series of usual predictors (gender, educational attainment...) and also age.

Table 1: SHARE: Wave 7 respondents aged 50+. Count by country and gender

	Male	Female	Total
Austria (AUT)	800	1,099	1,899
Belgium (BEL)	1,184	1,281	2,465
Bulgaria (BGR)	273	402	675
Switzerland (CHE)	659	701	1,360
Cyprus (CYP)	106	109	215
Czech Rep. (CZE)	1,063	1,615	2,678
Germany (DEU)	1,188	1,290	2,478
Denmark (DNK)	685	822	1,507
Spain (ESP)	1,126	1,111	2,237
Estonia (EST)	1,329	2,240	3,569
Finland (FIN)	392	437	829
France (FRA)	676	858	1,534
Greece (GRC)	315	246	561
Croatia (HRV)	772	805	1,577
Hungary (HUN)	123	177	300
Hungary (HUN)	518	600	1,118
Italy (ITA)	1,040	858	1,898
Lithuania (LTU)	285	543	828
Luxembourg (LUX)	431	463	894
Latvia (LVA)	108	167	275
Malta (MLT)	188	160	348
Poland (POL)	559	647	1,206
Portugal (PRT)	355	373	728
Romania (ROU)	317	283	600
Slovakia (SVK)	340	388	728
Slovenia (SVN)	1,148	1,470	2,618
Sweden (SWE)	926	984	1,910
Total	16,906	20,129	37,035

Source: SHARE 2004-2020 (Wave 7).

2.2 O*NET and EWCS: how to quantify the arduousness of jobs and occupations

SHARE wave 7 provides a lot of information about people's careers. But it falls short of providing information about the `arduousness` of successive jobs/occupations. To overcome that limitation, we turn to O*NET from the US and also to the European Working Conditionx Survey (EWCS).

O*NET is a rolling survey about working conditions by occupation that contains over 180 variables. Those variables are included in different modules. Here, we concentrate on the `Work Context` module. Items composing the version of the module used here have been collected in 2021. They explicitly describe working conditions (e.g. exposition to contaminants, spending time bending or twisting the body, working in very hot or cold temperatures…), structural job characteristics (e.g. consequence of error, time pressure, freedom to decide), and interpersonal/managerial relationship at work (e.g. contact with others, responsibility for other's health and safety, face-to-face discussions). We use a principal component (PC) analysis to get a summary indicator of occupation arduousness. More information (1st and 2nd principal components, eigenvalues and loading factors) is reported in the Appendix 1.1. Only the 1st PC is used in the paper to quantify each occupation's arduousness. We show in that table that it correlates with working conditions items associated with arduousness (e.g. "Exposed to Contaminants", "Pace (of work) determined by the speed of Equipment", "Sounds noise levels are distracting or uncomfortable"…). We also show that the 2nd Principal component correlates more with managerial vs. non-managerial work content: a dimension that is a priori less relevant in an exercise centred on the health impact of arduousness. In the Appendix 1.1, Figure 4 presents our O*NET 1st principal component (PC) at ISCO 2 level. We see that typical manual/outdoor occupations (e.g. building and related trades works) translate into high arduousness PC values, while more intellectual/indoor occupations (e.g. Business and Administration) display much lower values.

It is important to stress what we do with these occupation-specific arduousness data. Once injected into SHARE, they are used to compute, for each respondent, career arduousness indices. For instance, we compute the weighted average of all O*NET-estimated PC for his/her consecutive ISCO 4-digit occupations self-reported in SHARE wave 7. The weights for that average reflect the duration (in years) of the successive occupation spells. Note that the years have been multiplied by .5 if the occupation was declared always part-time, 1 if always full-time and .75 when variable. Hereafter, we mostly use the entire (average) career arduousness index. But we also consider the arduousness of the main job declares.

One objection to using O*NET is that the resulting career arduousness indices rest on data assembled in the US, reflecting working conditions in jobs as they exist in the US; whereas SHARE is about health and career history in Europe. Working conditions by occupation are likely to be similar, but they may also diverge to an extent. As a robustness check, we compare the results we get when using the US O*NET-based measure of career arduousness to the ones delivered by a European measure, namely the one we find in EWCS. Refer to Appendix 1.2 for more information on EWCS and the way we use it.

2.3 Descriptive statistics: Career arduousness and the other variables used

The first line of Table 2 describes the respondents' *average* career arduousness that has been standardized. It is computed as the weighted[1] average O*NET arduousness index characterizing the successive ISCO4 occupations (reported in SHARE wave 7). The next lines describe the other variables mobilised in the analysis below. There is of course the age of the respondents, but also their gender and educational attainment.

Table 2: SHARE, O*NET: career arduousness* and controls. Descriptive statistics

	mean	sd	min	max
Ardu. (car. av.)^a	0.06	0.99	-1.67	2.73
Ardu. (car. cumul.)^a	0.09	0.99	-5.99	8.55
Ardu. car. av. (EU-EWCS ISCO2)^a	-16.48	1.04	-18.73	-14.20
Female	0.54	0.50	0.00	1.00
Age (in years)	67.98	9.45	50.00	102.00
Education (ISCED level)^d	3.07	1.41	0.00	6.00

Source: SHARE 2018 (Wave 7), O*NET 2021, EWCS 2015
*: Based on O*NET (ISCO4) unless specified otherwise.
a: Average career arduousness (weighted)- Standardized
b: Cumulative career arduousness (weighted)- Standardized
c: Average career arduousness (weighted) (EU/EWCS-based index.)
d: ISCED1997 classification of educational attainment [0:no degree 6: tertiary long].

3 Results based on cross-sectional SHARE wave 7

3.1 Regression results: evidence of missing arduous careers

In Table 3, we expose the regression results that hint at the "missing arduous career" phenomena. The reference age band is 50-54. The estimated dummy coefficient bands are estimates of the (conditional) career arduousness differences characterising the respondents belonging to the older age bands. Quite invariably, across the three columns (each of them corresponding to a different way of quantifying career arduousness), we observe the same pattern: older respondents display a significantly lower career arduousness. We interpret this as evidence that individuals with more arduous careers (who still contribute to the 50-54 average arduousness reference value) die faster than their peers and thus contribute

[1] Where the weights correspond to the duration of the different job spells (themselves weighted to account for the part-time vs. full-time nature of the spell. Weight is .5 if people worked part-time, .75 if they shifted between part-time to full-time, and 1 otherwise.

proportionally less to the estimates of the average arduousness of older age bands on display Table 3. Another result visible in Table 3 is that the drop in arduousness (and the underlying death attrition differential) starts materializing only from the age of 60.

In what follows, we focus on comparing the career arduousness of respondents aged 55-59 and those aged 80-84. We also work exclusively with the O*Net-based quantification of career corresponding to the first line of Table 2 or the first column of Table 3.

The first thing we do is to retrieve the full career arduousness distributions underpinning the age bands 55-59, 80-84 coefficient Table 3. For that, we resort to the residuals of the regression of arduousness on the control variables, but without the age band dummies. The resulting (conditional) distributions by age band (55-59 vs 80-84) are displayed in Figure 1. They logically confirm what was visible in Table 3, namely that the age 80-84 distribution is shifted to the left (thus to lower career arduousness level).

Table 3: Econometric evidence that average career arduousness is lower among older SHARE respondents (ref. 50-54)

	Ardu. (car. av.)[a]	Ardu. (car. mainjob)[b]	Ardu. (car. av.)[c] EU-EWCS)[c]
$[55-59[$	0.0030	-0.0161	-0.0041
	(0.0215)	(0.0224)	(0.0230)
$[60-64[$	-0.0054	-0.0152	-0.0558**
	(0.0211)	(0.0219)	(0.0225)
$[64-70[$	-0.0551***	-0.0344	-0.1145***
	(0.0211)	(0.0219)	(0.0226)
$[70-74[$	-0.0864***	-0.0432*	-0.1528***
	(0.0215)	(0.0223)	(0.0230)
$[75-80[$	-0.0712***	-0.0397*	-0.1774***
	(0.0225)	(0.0233)	(0.0240)
$[80-84[$	-0.0995***	-0.0587**	-0.2011***
	(0.0242)	(0.0252)	(0.0259)
Gender	Yes	Yes	Yes
Educ.[d]	Yes	Yes	Yes
N	35,086	35,086	35,086

Source: SHARE 2018 (Wave 7), O*NET 2021, EWCS 2015
[a]: Average career arduousness (weighted)- Standardized
[b]: Cumulative career arduousness (weighted)- Standardized
[c]: Average career arduousness (weighted) (EU/EWCS-based index.)
[d]: ISCED1997 classification of educational attainment [0:no degree 6: tertiary long].

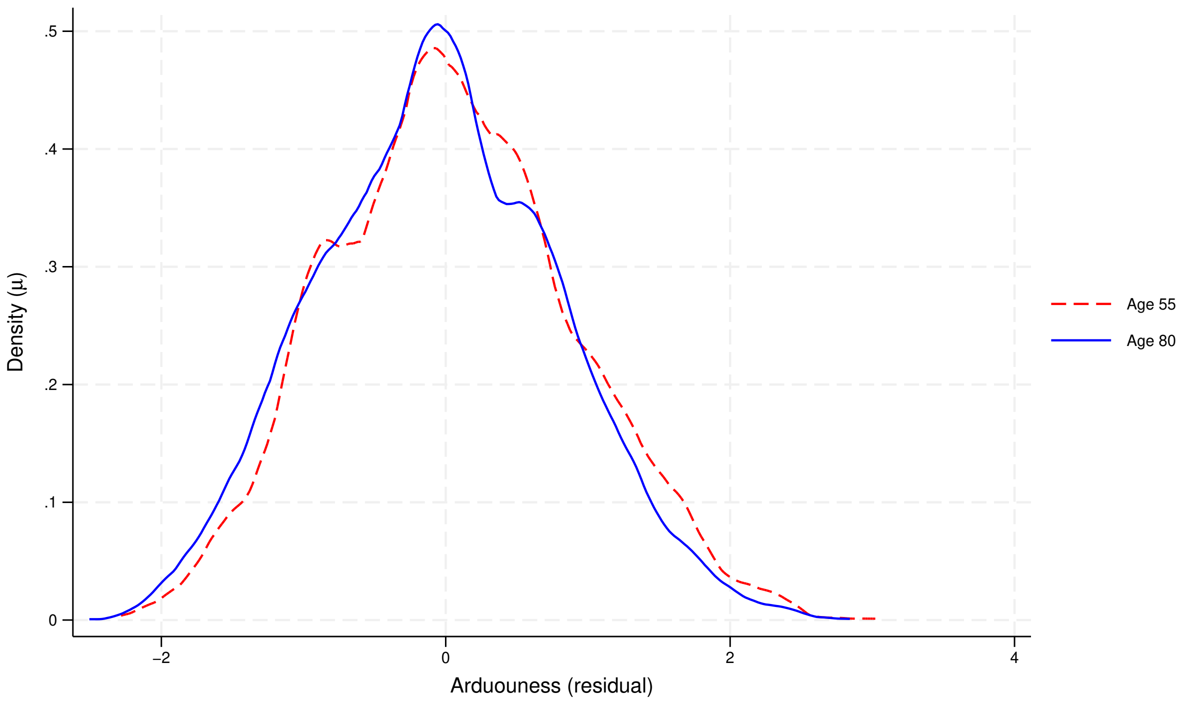

Figure 1: Missing Arduousness: conditional density distributions
Source: SHARE 2018 (Wave 7), O*NET 2021. Our calculus.

3.2 Inferring longevity differences

How could we infer arduousness-related longevity differences from the above density distributions?

Consider that the one for age band $[55-59[$ corresponds to a vector of weights (i.e. arduousness sample proportions) $\mu^{55} \equiv (\mu_1^{55}, \ldots, \mu_K^{55})$, with career arduousness ranging from (quantiles) $k = 1, \ldots K$, and $\sum_{k=1}^{K} \mu_k = 1$. The equivalent observed distribution for age band $[80-84[$ is noted μ^{80}.

Each age band/arduousness-quantile-specific weight μ_k^a corresponds to the ratio of the number of observed individuals forming the age/arduousness cell to the total number of individuals belonging to the age band: $\mu_K^a \equiv N_k^a / N^a$, $a = [55-59[, \ldots, [80-84[$.

$$\mu^{55} \equiv \begin{pmatrix} \mu_1^{55} \\ \mu_2^{55} \\ \vdots \\ \mu_K^{55} \end{pmatrix} = \begin{pmatrix} N_1^{55}/N^{55} \\ N_2^{55}/N^{55} \\ \vdots \\ N_K^{55}/N^{55} \end{pmatrix} \tag{1}$$

Assuming that differences between the two distributions/vectors of weights are primarily driven by survival differences (s_k), one can then write the age 80-84 weights as

$$\mu^{80} \equiv \begin{pmatrix} \mu_1^{80} \\ \mu_2^{80} \\ \vdots \\ \mu_K^{80} \end{pmatrix} = \begin{pmatrix} s_1 N_1^{55}/N^{80} \\ s_2 N_2^{55}/N^{80} \\ \vdots \\ s_K N_K^{55}/N^{80} \end{pmatrix} \tag{2}$$

If we consider the ratio between two specific (above-observed) densities/weights

$$\frac{\mu_k^{80}}{\mu_k^{55}} = s_k \frac{N^{55}}{N^{80}} = \frac{s_k}{\overline{s}}$$

or equivalently (3)

$$s_k = \frac{\mu_k^{80}}{\mu_k^{55}} \overline{s}, \quad k = 1 \ldots K$$

where $\overline{s} \equiv \frac{N^{80}}{N^{55}}$ the overall survival rate between the 55-59 and 80-84. In other words, the survival rate specific to individuals exposed to arduousness k is equal to the product of *i)* the relative observed/arduousness-specific weights stemming from the above-estimated density distributions by *ii)* the overall survival rate $\overline{s}$ informs. The latter can be estimated using official survival tables. Results in terms of survival between 55-59 and 80-84 are reported in Figure 2. They show a significantly higher(lower) survival rate for individuals with low(high) career arduousness. More detailed estimates based on SHARE-wave7 cross-sectional evidence and the algebra exposed above are reported in Table 4.

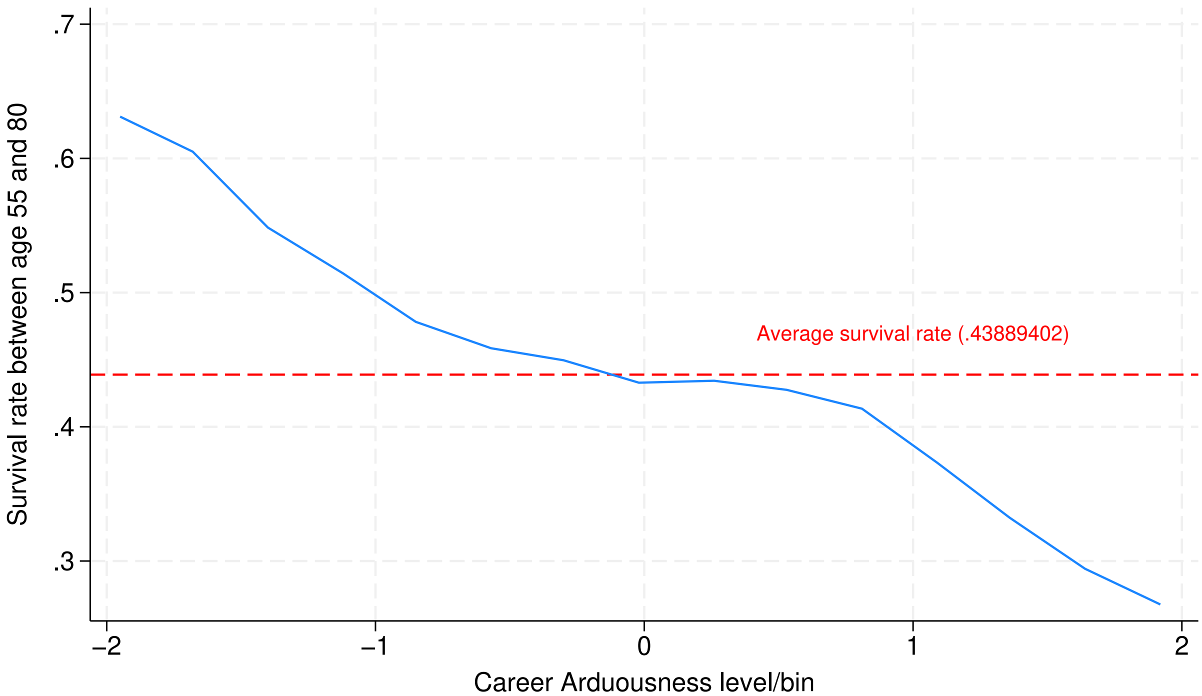

Figure 2: Missing Arduousness: inferred survival-mortality differences between the age of 55-59 and 80-84

Source: SHARE 2018 (Wave 7), O*NET 2021. Our calculus.

Table 4: Missing Arduousness: survival rate differential between high and low career (ref.) arduousness individuals[a]. Estimated between 55-59 and various older age bands

	60-64	65-69	70-74	75-79	**80-84**
Survival handicap	-0.162	-0.260	-0.263	-0.319	-0.363

Source: SHARE 2018 (Waves 7), O*NET 2021. Our calculus.
[a]: Corresponding (resp.) to the upper and lower quartile of the career arduousness variable.

4 Assessment based on SHARE longitudinal follow-up of Wave 7 respondents

In this section, we exploit the SHARE waves that have been implemented after wave 7 and inform about the dead/alive status of wave 7 respondents. The aim is to *i)* see if we also find a relationship between arduousness and mortality/survival, *ii)* evaluate whether the arduousness survival gap estimated in the previous section using cross-sectional evidence is backed by survival/mortality estimates derived from a longitudinal perspective. Results are reported in Table 5. The first(second) line corresponds to the mortality rate recorded during a follow-up of 2.5 years, as the individuals have experienced less(most) arduous careers. The third line informs of the mortality rate difference. To get an idea of the impact on the survival/mortality handicap between 55-59 and 80-84, these values need to be annualised and then compounded. At the bottom of the table, we report a 0.202 handicap for the individuals with the most arduous careers. The equivalent figure derived from the analysis of cross-section data is .363.

Table 5: Mortality rate, by age band [55-54/80-84] & career arduousness quartile. Estimates based on the follow-up of SHARE wave 7 respondents during 2.5 years

	Age band					
Arduousness	55-59	60-64	65-69	70-74	75-79	80-84
First quartile (Q1)	0.010	0.014	0.023	0.041	0.061	0.127
Fourth quartile (Q4)	0.011	0.032	0.042	0.053	0.089	0.147
Interquartile diff. (Q4-Q1)	0.000	0.018	0.019	0.011	0.028	0.020
Cumulative 55-59/80-84 diff.[a]					0.202	

Source: SHARE 2018-2020 (Waves 7,8,9), O*NET 2021. Our calculus.
[a]: Estimation of the cumulative mortality handicap between the ages of 55-59 and 80-84 of individuals forming the highest quartile of career arduousness compared to those forming the lowest quartile.

5 Conclusion

In this paper, we highlight the propensity of people who had more(less) arduous careers to be under-(over)represented in cross-sectional retirement survey data as age rises. We show that this is the case with the SHARE wave 7 survey data. Echoing the literature on the "missing poors", we posit that this is evidence of more(less) death attrition among the workers who exerted more(less) arduous careers. We also quantify the mortality differential compatible with the under-(over)representation we find in SHARE wave 7. Our best estimate points at a cumulative mortality handicap of .363 (36 percentage points) between the ages of 50-54 and 80-84 for the individuals who had the most arduous career compared to those with the least arduous ones (i.e. top v.s. lowest quartile). Such a result aligns with the one obtained using the (limited) longitudinal evidence available in SHARE concerning the propensity of wave 7 respondents to die during the subsequent 2.5 years. However, that longitudinal analysis points to a lower cumulative mortality handicap, namely .202 instead of .363. The tentative conclusion is that the existence of a career arduousness-related mortality handicap can probably be reliably inferred from cross-sectional data. However, whether these data are enough to estimate the exact magnitude of the handicap remains an open question.

Acknowledgement

This project has received financial support from the Belgian French-speaking Community (convention ARC n°18/23-088 on "Sustainable, adequate and safe pensions: financial architecture, social justice and governance").

This paper uses data from SHARE Waves 1, 2, 4, 5, 6, 7 (DOIs: 10.6103/SHARE.w1.710, 10.6103/SHARE.w2.710, 10.6103/SHARE.w4.710, 10.6103/SHARE.w5.710, 10.6103/SHARE.w6.710, 10.6103/SHARE.w7.711, 10.6103/SHARE.w8cabeta.001), see Börsch-Supan et al. (2013) for methodological details. The SHARE data collection has been funded by the European Commission through FP5 (QLK6-CT-2001-00360), FP6 (SHARE-I3: RII-CT-2006-062193, COMPARE: CIT5-CT-2005-028857, SHARELIFE: CIT4-CT-2006-028812), FP7 (SHARE-PREP: GA N°211909, SHARE-LEAP: GA N°227822, SHARE M4: GA N°261982, DASISH: GA N°283646) and Horizon 2020 (SHARE-DEV3: GA N°676536, SHARE-COHESION: GA N°870628, SERISS: GA N°654221, SSHOC: GA N°823782) and by DG Employment, Social Affairs & Inclusion. Additional funding from the German Ministry of Education and Research, the Max Planck Society for the Advancement of Science, the U.S. National Institute on Aging (U01 AG09740-13S2, P01 AG005842, P01 AG08291, P30 AG12815, R21 AG025169, Y1-AG-4553-01, IAG BSR06-11, OGHA 04-064, HHSN271201300071C) and from various national funding sources is gratefully acknowledged (www.share-project.org).

This paper includes information from O*NET Resource Center by the U.S. Department of Labor, Employment and Training Administration (USDOL/ETA). Used under the CC BY 4.0 license. O*NET is a trademark of USDOL/ETA. We have modified all or some of this information. USDOL/ETA has not approved, endorsed, or tested these modifications.

Finally, this paper used data from the European Foundation for the Improvement of Living and Working Conditions. (2020). European Working Conditions Survey Integrated Data File, 1991-2015. [data collection]. 8th Edition. UK Data Service. SN: 7363, http://doi.org/10.5255/UKDA-SN-7363-8.

References

Bakker, A. B. and E. Demerouti (2007). "The job demands-resources model: State of the art". In: *Journal of managerial psychology* 22.3, pp. 309–328.

Barnay, T. (2016). "Health, work and working conditions: a review of the European economic literature". In: *The European Journal of Health Economics* 17.6, pp. 693–709.

Bassanini, A. and E. Caroli (2015). "Is Work Bad for Health? The Role of Constraint versus Choice". In: *Annals of Economics and Statistics* 119, pp. 13–37.

Börsch-Supan, A., M. Brandt, C. Hunkler, T. Kneip, J. Korbmacher, F. Malter, B. Schaan, S. Stuck, and o. b. o. t. S. C. C. T. Zuber Sabrina (2013). "Data resource profile: The survey of health, ageing and retirement in europe (SHARE)". In: *International Journal of Epidemiology* 42.4, pp. 992–1001.

Chen, Y., S. Li, Q. Xia, and C. He (2017). "The Relationship between Job Demands and Employees' Counterproductive Work Behaviors: The Mediating Effect of Psychological Detachment and Job Anxiety". In: *Frontiers in Psychology* 8, pp. 1–15.

DeHart, D. (2021). "Life history calendar method". In: *The encyclopedia of research methods in criminology and criminal justice*. John Wiley & Sons, Ltd, pp. 68–71.

Kanbur, R. and D. Mukherjee (2007). "Premature mortality and poverty measurement". In: *Bulletin of Economic Research* 59.4, pp. 339–359.

Katikireddi, S. V., A. H. Leyland, M. McKee, K. Ralston, and D. Stuckler (2017). "Patterns of mortality by occupation in the UK, 1991–2011: a comparative analysis of linked census and mortality records". In: *The Lancet Public Health* 2.11, e501–e512.

Lefebvre, M., P. Pestieau, and G. Ponthiere (2019). "Missing poor and income mobility". In: *Journal of Comparative Economics* 47.2, pp. 330–366.

Lefebvre, M., P. Pestieau, and G. Ponthiere (2023). "Counting the missing poor in pre-industrial societies". In: *Cliometrica, Journal of Historical Economics and Econometric History* 17.1, pp. 155–183.

Mikkola, T. M., M. B. Von Bonsdorff, M. K. Salonen, H. Kautiainen, L. Ala-Mursula, S. Solovieva, E. Viikari-Juntura, and J. G. Eriksson (2019). "Physical heaviness of work and sitting at work as predictors of mortality: a 26-year follow-up of the Helsinki Birth Cohort Study". In: *BMJ open* 9.5, e026280.

Ravesteijn, B., H. v. Kippersluis, and E. v. Doorslaer (2018). "The wear and tear on health: What is the role of occupation?" In: *Health Economics* 27.2, e69–e86.

Vandenberghe, V. (2023). "Career arduousness and instability: Both matter for health beyond 50". In: *LABOUR* 37.3, pp. 343–384.

Vermeer, N., M. Mastrogiacomo, and A. Van Soest (2016). "Demanding occupations and the retirement age". In: *Labour Economics* 43, pp. 159–170.

Appendix

1.1 O*NET Principal Components, Load Factors

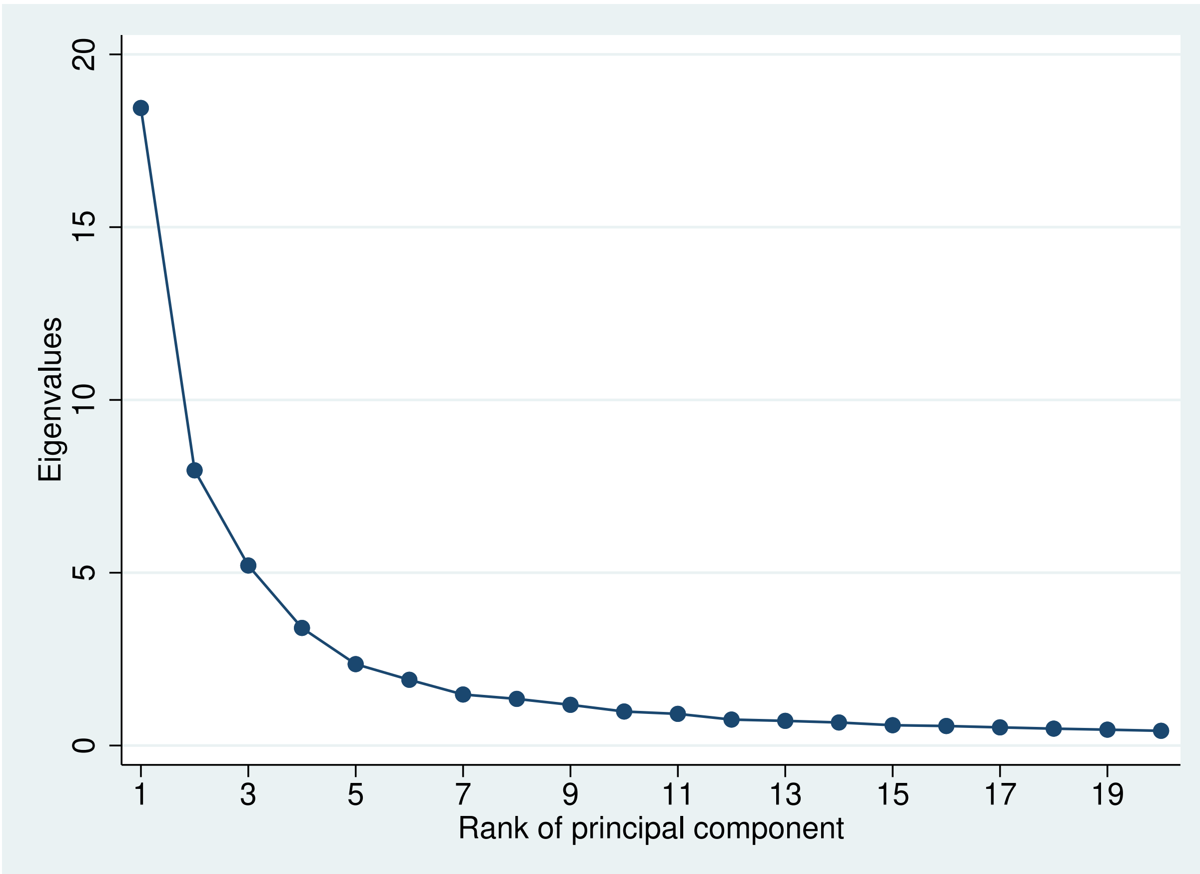

Figure 3: O*NET arduousness items (ISCO4): proportion of variance explained by first (and following) principal components (i.e. eigenvalues)

Source: O*NET 2020, Work Context Items.

Table 6: O*NET arduousness items (ISCO4): Loading factors for 1^{st} and 2^{nd} Principal Component[a]

	Load factors	
	Principal Comp.#1	Principal Comp.#2
Consequence_of_Error	0.10	0.18
Contact_With_Others	-0.06	0.23
Coordinate_or_Lead_Others	-0.03	0.25
Cramped_Work_Space_Awkward_Positions	0.20	0.07
Deal_With_External_Customers	-0.08	0.21
Deal_With_Physically_Aggressivity	0.01	0.16
Deal_With_Unpleasant_or_Angry_	-0.01	0.18
Degree_of_Automation	-0.00	0.00
Duration_of_Typical_Work_Week	0.00	0.12
Electronic_Mail	-0.16	0.15
Exposed_to_Contaminants	0.20	0.02
Exposed_to_Disease_or_Infections	-0.01	0.15
Exposed_to_Hazardous_Condition	0.18	0.05
Exposed_to_Hazardous_Equipment	0.20	0.02
Exposed_to_High_Places	0.17	0.07
Exposed_to_Minor_Burns_Cuts_Bi	0.21	-0.00
Exposed_to_Radiation	0.04	0.11
Exposed_to_Whole_Body_Vibrations	0.17	0.04
Extremely_Bright_or_Inadequate	0.19	0.07
Face-to-Face_Discussions	-0.03	0.22
Freedom_to_Make_Decisions	-0.07	0.16
Frequency_of_Conflict_Situations	-0.03	0.25
Frequency_of_Decision_Making	0.02	0.25
Impact_of_Decisions_on_Coworkers	0.00	0.27
Importance_of_Being_Exact_or_Accurate	-0.01	0.08
Importance_of_Repeating_Same_Task	0.01	0.07
In_an_Enclosed_Vehicle_or_Equipment	0.10	0.14
In_an_Open_Vehicle_or_Equipment	0.18	0.03
Indoors_Environmentally_Controlled	-0.17	0.06
Indoors_Not_Environmentally_Controlled	0.18	0.04
Letters_and_Memos	-0.12	0.21
Level_of_Competition	-0.04	0.10
Outdoors_Exposed_to_Weather	0.15	0.10
Outdoors_Under_Cover	0.14	0.11
Pace_Determined_by_Speed_of_Equipment	0.16	-0.07
Physical_Proximity	0.05	0.12
Public_Speaking	-0.10	0.11
Responsibility_for_Outcomes_an	0.06	0.21
Responsible_for_Others_Health_	0.15	0.16
Sounds_Noise_Levels_Are_Distraction	0.18	0.04
Spend_Time_Bending_or_Twisting	0.20	-0.03
Spend_Time_Climbing_Ladders_Scaffolds	0.17	0.04
Spend_Time_Keeping_or_Regaining_Balance	0.19	0.03
Spend_Time_Kneeling_Crouching_	0.18	0.00
Spend_Time_Making_Repetitive_M	0.10	-0.11
Spend_Time_Sitting	-0.17	0.04
Spend_Time_Standing	0.17	-0.05
Spend_Time_Using_Your_Hands_to_Handle_objects	0.16	-0.08
Spend_Time_Walking_and_Running	0.17	-0.01
Structured_versus_Unstructured	-0.10	0.15
Telephone	-0.11	0.23
Time_Pressure	0.02	0.10
Very_Hot_or_Cold_Temperatures	0.20	0.04
Wear_Common_Protective_or_Safety_Equipment	0.19	0.03
Wear_Specialized_Protective_or_Safety_Equipment	0.16	0.08
Work_Schedules	0.09	-0.01
Work_With_Work_Group_or_Team	-0.01	0.21

Source: O*NET 2021, Work Context Items. [a] Only the 1^st Principal component is used in this paper to compute career arduousness $CAR_{i,j}^{ard}$ in equation (1). It clearly correlates with items associated with arduousness (e.g. Exposed to Contaminants, Pace (of work) determined by speed of Equipment, Sounds noise levels are distracting or uncomfortable.... The second Principal component correlates more with managerial vs non-managerial work content, a dimension that is a priori less relevant in an exercise centred on the health impact of arduousness.

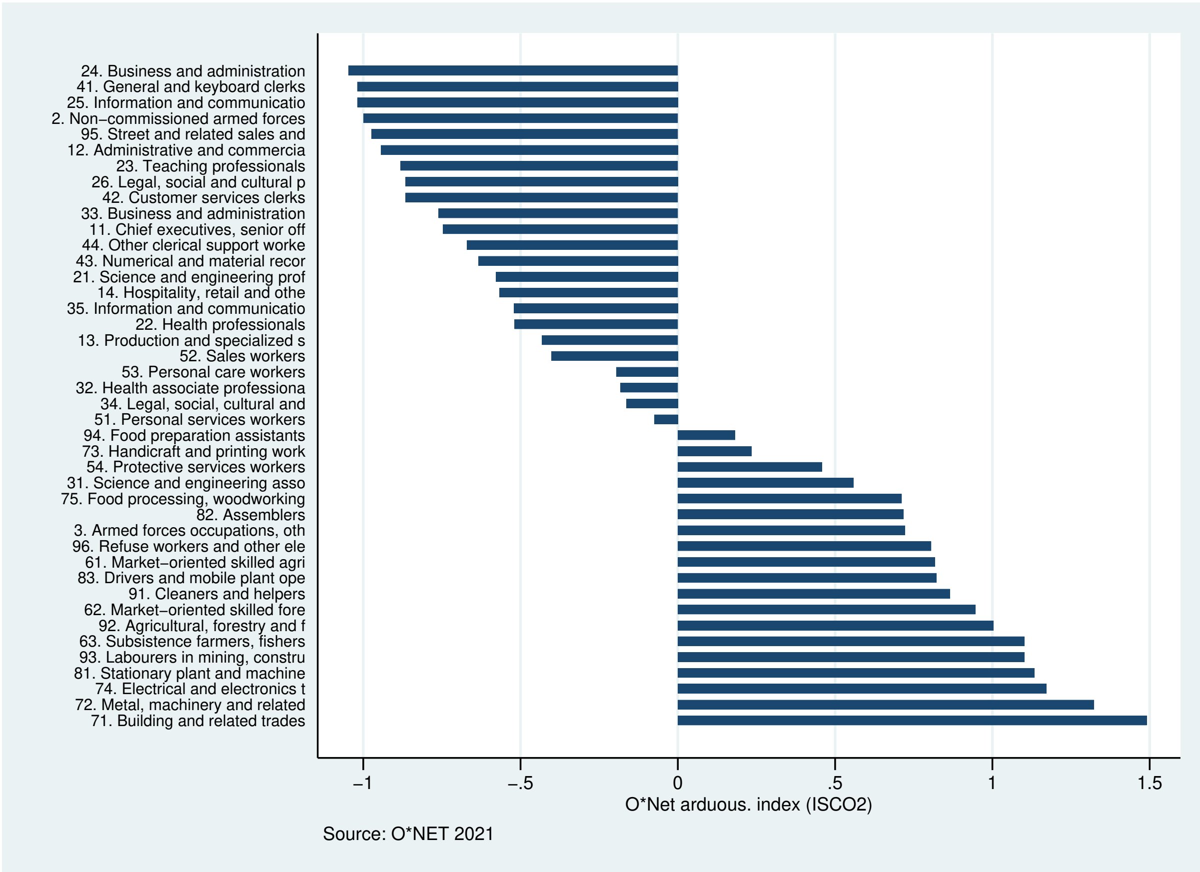

Figure 4: O*NET career arduousness indices (ISCO 2)

Indices reported on the X axis are First Principal Components of items forming the O*NET *Work Context* module. More information (1st and 2nd principal components, eigenvalues and loading factors) is available in the Appendix (Figure 5, Table 15)

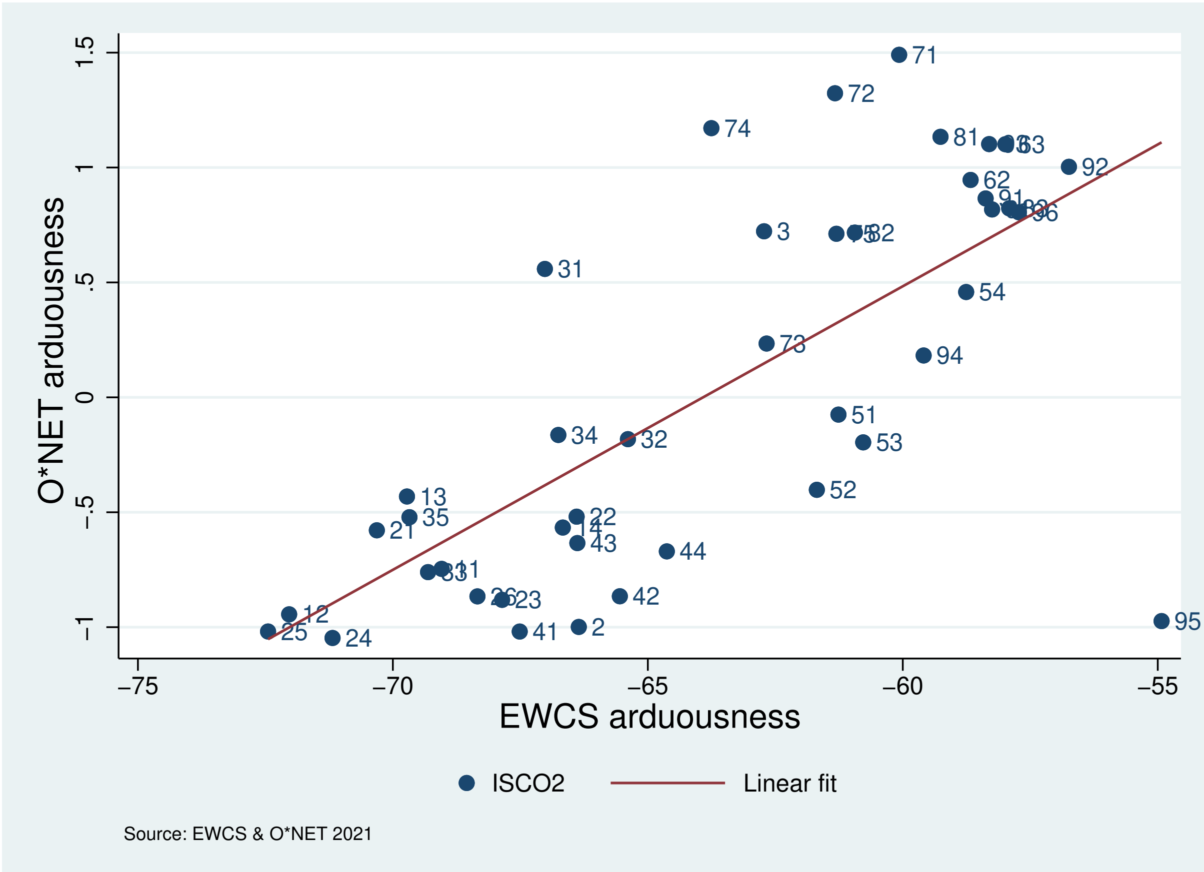

Figure 5: O*NET vs. EWCS arduousness indices at ISCO 2 level

0 Armed forces occupations 1 Commissioned armed forces officers 2 Non-commissioned armed forces officers 3 Armed forces occupations, other ranks 10 Managers 11 Chief executives, senior officials and legislators 12 Administrative and commercial managers 13 Production and specialised services managers 14 Hospitality, retail and other services managers 20 Professionals 21 Science and engineering professionals 22 Health professionals 23 Teaching professionals 24 Business and administration professionals 25 Information and communications technology professionals 26 Legal, social and cultural professionals 30 Technicians and associate professionals 31 Science and engineering associate professionals 32 Health associate professionals 33 Business and administration associate professionals 34 Legal, social, cultural and related associate professionals 35 Information and communications technicians 40 Clerical support workers 41 General and keyboard clerks 42 Customer services clerks 43 Numerical and material recording clerks 44 Other clerical support workers 50 Services and sales workers 51 Personal services workers 52 Sales workers 53 Personal care workers 54 Protective services workers 60 Skilled agricultural, forestry and fishery workers 61 Market-oriented skilled agricultural workers 62 Market-oriented skilled forestry, fishery and hunting workers 63 Subsistence farmers, fishers, hunters and gatherers 70 Craft and related trades workers 71 Building and related trades workers (excluding electricians) 72 Metal, machinery and related trades workers 73 Handicraft and printing workers 74 Electrical and electronics trades workers 75 Food processing, woodworking, garment and other craft and related trades workers 80 Plant and machine operators and assemblers 81 Stationary plant and machine operators 82 Assemblers 83 Drivers and mobile plant operators 90 Elementary occupations 91 Cleaners and helpers 92 Agricultural, forestry and fishery labourers 93 Labourers in mining, construction, manufacturing and transport 94 Food preparation assistants 95 Street and related sales and services workers 96 Refuse workers and other elementary workers.

1.2 European Working Condition Survey(EWCS)

Since 1991, Europe has been monitoring working conditions across Europe through its European Working Conditions Survey (EWCS). The survey's primary aim is like the one pursued by O*NET, namely to measure working conditions across European countries on a harmonised basis. We use the 1991-2015 combined version of the survey. More precisely, we exploit six of the indices that have been developed by the authors of the survey and added to the raw data. The six job quality indices we use are: Physical environment, Work intensity, Working time quality, Social environment, Skills and discretion, and Job prospect. Each of these 6 job quality indices is measured on a scale from 0 to 100. We inverse the sign of each of these indices because we are interested in arduousness (while these indices quantify the "quality" of jobs) and compute their average by occupation. A limitation with EWCS is that respondents' occupation is only available at the ISCO 2-digit level, while O*NET information exists at ISCO 4 digit. But again, the major advantage of EWCS in this paper is that the underlying observations of occupations come from Europe and might thus be more in line with what SHARE respondents have experienced throughout their professional lives. Our regression analysis systematically includes EWCS-based measures of career arduousness to see if these deliver results that deviate from O*NET-based ones. A first, purely descriptive, comparison of O*NET and EWCS is reported in Figure 6. We simply plot the arduousness values delivered by EWCS against those stemming from O*NET. As stated above, the comparison can only be done for ISCO 2 digit. It hints at a strong correlation but also differences for some ISCO 2 occupations.

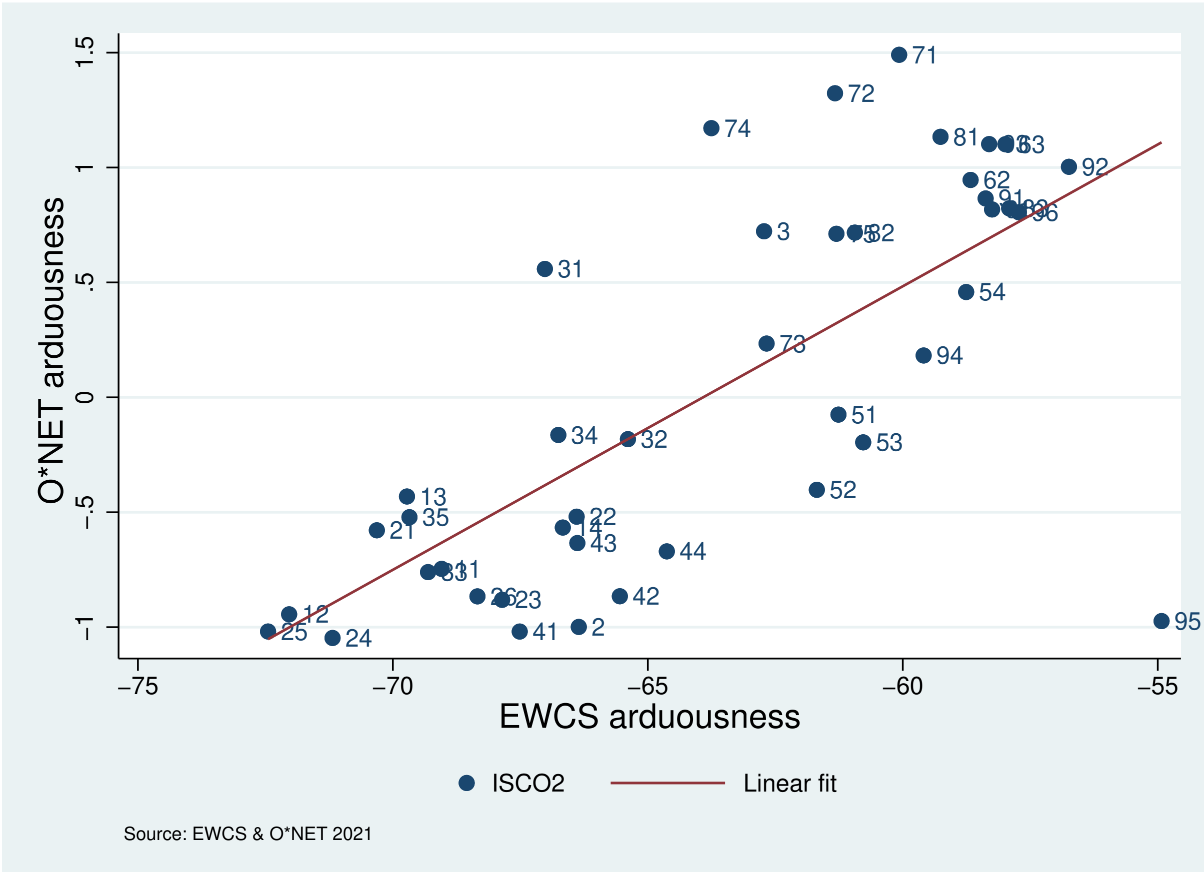

Figure 6: O*NET vs. EWCS arduousness indices at ISCO 2 level

0 Armed forces occupations 1 Commissioned armed forces officers 2 Non-commissioned armed forces officers 3 Armed forces occupations, other ranks 10 Managers 11 Chief executives, senior officials and legislators 12 Administrative and commercial managers 13 Production and specialised services managers 14 Hospitality, retail and other services managers 20 Professionals 21 Science and engineering professionals 22 Health professionals 23 Teaching professionals 24 Business and administration professionals 25 Information and communications technology professionals 26 Legal, social and cultural professionals 30 Technicians and associate professionals 31 Science and engineering associate professionals 32 Health associate professionals 33 Business and administration associate professionals 34 Legal, social, cultural and related associate professionals 35 Information and communications technicians 40 Clerical support workers 41 General and keyboard clerks 42 Customer services clerks 43 Numerical and material recording clerks 44 Other clerical support workers 50 Services and sales workers 51 Personal services workers 52 Sales workers 53 Personal care workers 54 Protective services workers 60 Skilled agricultural, forestry and fishery workers 61 Market-oriented skilled agricultural workers 62 Market-oriented skilled forestry, fishery and hunting workers 63 Subsistence farmers, fishers, hunters and gatherers 70 Craft and related trades workers 71 Building and related trades workers (excluding electricians) 72 Metal, machinery and related trades workers 73 Handicraft and printing workers 74 Electrical and electronics trades workers 75 Food processing, woodworking, garment and other craft and related trades workers 80 Plant and machine operators and assemblers 81 Stationary plant and machine operators 82 Assemblers 83 Drivers and mobile plant operators 90 Elementary occupations 91 Cleaners and helpers 92 Agricultural, forestry and fishery labourers 93 Labourers in mining, construction, manufacturing and transport 94 Food preparation assistants 95 Street and related sales and services workers 96 Refuse workers and other elementary workers.