

Emilie Courtin, Martin Knapp, Emily Grundy and Mauricio Avendano

Are different measures of depressive symptoms in old age comparable? An analysis of the CES-D and Euro-D scales in 13 countries

**Article (Accepted version)
(Refereed)**

Original citation: Courtin, Emilie, Knapp, Martin, Grundy, Emily and Avendano, Mauricio (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 (4). pp. 287-304. ISSN 1049-8931
DOI: [10.1002/mpr.1489](https://doi.org/10.1002/mpr.1489)

© 2015 John Wiley & Sons, Ltd.

This version available at: <http://eprints.lse.ac.uk/63332/>

Available in LSE Research Online: August 2016

LSE has developed LSE Research Online so that users may access research output of the School. Copyright © and Moral Rights for the papers on this site are retained by the individual authors and/or other copyright owners. Users may download and/or print one copy of any article(s) in LSE Research Online to facilitate their private study or for non-commercial research. You may not engage in further distribution of the material or use it for any profit-making activities or any commercial gain. You may freely distribute the URL (<http://eprints.lse.ac.uk>) of the LSE Research Online website.

This document is the author's final accepted version of the journal article. There may be differences between this version and the published version. You are advised to consult the publisher's version if you wish to cite from it.

Are Different Measures of Depressive Symptoms in Old age Comparable? An Analysis of the CES-D and Euro-D Scales in 13 Countries

Emilie Courtin¹, Martin Knapp¹, Emily Grundy¹, Mauricio Avendano-Pabon^{1,2}

1 London School of Economics and Political Science, Social Policy, London UK

2 Harvard School of Public Health, Harvard University, Boston, MA, USA

ABSTRACT

The Centre for Epidemiologic Studies of Depression (CES-D) and the Euro-D are commonly used depressive symptom scales but their comparability has not been assessed to date. This article aims to contribute to the literature comparing the drivers of depression in old age across countries by examining whether CES-D (in its eight-item short version) and Euro-D are comparable. Data from the Survey of Health, Ageing and Retirement in Europe (SHARE, $N=15,487$) covering 13 countries was used to examine the scales' distributional properties, systematic differences between population subgroups, sensitivity and specificity, and associations with established risk factors for depression in old age. CES-D and Euro-D were strongly correlated ($r=0.6819$ ($p<0.000$)). However, agreement between the two scales was moderate. There were systematic discrepancies in scores by demographic characteristics. CES-D captures a more extreme pool of depressed individuals than Euro-D. Although associations with risk factors are always in the same direction, they are often stronger for CES-D than Euro-D. Findings highlight the need to be cautious when comparing depression levels and associations with risk factors between surveys using different measures of depressive symptoms.

Competing interests: The authors have no competing interests.

Word count: 4,259 (excluding references, figures and tables)

Keywords: depression; old age; Euro-D; CES-D.

Introduction

Depression is the leading cause of years lived with disability and the fourth leading contributor to the global burden of disease worldwide (Alexopoulos, 2005; Djernes, 2006; Ferrari et al., 2013). Depression is also the most frequent cause of emotional suffering in later life (Beekman, Copeland, & Prince, 1999). A growing literature based on cross-national comparable data suggests that there are significant differences in the prevalence of later-life depressive symptoms across countries (Castro-Costa et al., 2007; Kok, Avendano-Pabon, Bago d'Uva, & Mackenbach, 2012; Missinne, Vandevive, Van de Velde, & Bracke, 2014; Ploubidis & Grundy, 2009). In addition, recent studies suggest that risk factors for depression may differ cross-nationally (Crimmins, Kim, & Solé-Auró, 2011; Di Gessa & Grundy, 2014; Lunau, Wahrendorf, Dragano, & Siegrist, 2013; Riumallo-Herl, Basu, Stuckler, Courtin, & Avendano-Pabon, 2014; Siegrist, Lunau, Wahrendorf, & Dragano, 2012). Most of these studies are based on data from harmonized longitudinal ageing studies, such as the Survey of Health, Ageing and Retirement in Europe (SHARE), the English Longitudinal Study of Ageing (ELSA) and the Health and Retirement Survey (HRS) in the United States (US). An advantage of these surveys is that they include broadly comparable measures of health, employment, social interactions and well-being (Banks, Nazroo, & Steptoe, 2012; Börsch-Supan, Hank, & Jürges, 2005; National Institute on Ageing, 2014). However, they use different measures to assess depressive symptoms. While SHARE uses the Euro-D scale of depressive symptoms, HRS and ELSA rely on a short version of the Centre for Epidemiological Studies Depression (CES-D) scale. Although measurement comparability is an essential prerequisite for robust comparisons across countries, it is as yet unclear how the CES-D scale compares to the Euro-D scale, and whether cross-national comparisons using these two different measures are valid.

In this paper, we exploit unique data from the second wave of SHARE, which administered both the CES-D and Euro-D scales to a sample of older Europeans in 13 countries. Our aim was to assess the comparability of the scales; their sensitivity and specificity to identify depression caseness; and to assess differences in the association of each scale with established risk factors for depression. To our knowledge, this is the first study examining the comparability of the CES-D and Euro-D measures of depressive symptoms.

Methods

Data Collection and Participants

SHARE is a longitudinal, nationally representative survey designed to provide comparable information on the health, employment and social conditions of Europeans aged 50+ in 13 countries (Austria, Belgium, Czechia, Denmark, France, Germany, Greece, Italy, Poland, Spain, Sweden, Switzerland and the Netherlands). Specific details about the survey are available elsewhere (Börsch-Supan & Jürges, 2005). Participants in each country were interviewed in 2004/5 and subsequently re-interviewed in 2006/7, 2008/9, 2010/11 and 2012/13 through face-to-face interviews using Computer-Assisted Personal Interviewing (CAPI) technology. Expert agencies translated items, with extensive pre-testing to ensure comparability. Response rates varied from country to country, but overall household response at enrolment was 62% (Börsch-Supan & Jürges, 2005). For our analysis, we used data from the second wave, which contained measures of both the Euro-D and CES-D for the same respondents. Only respondents with scores from the two scales were included in the analysis, resulting in a final sample of 15,487 respondents.

Measures

We compared two scales of depressive symptoms: (a) the eight-item version of the CES-D scale; and (b) the 12-item EURO-D scale. The original CES-D scale comprises 20 items (Radloff, 1977), but shorter versions are frequently used and have been shown to be reliable (Kohout, Berkman, Evans, & Cornoni-Huntley, 1993). An eight-item version of the CES-D scale was included in the second wave of SHARE, asking respondents whether they had experienced any of the following symptoms during the previous week: felt depressed, felt that everything was an effort, felt that their sleep was restless, were happy, felt lonely, enjoyed life, felt sad, or were unable to get going. Possible responses were *yes* or *no*. The score ranges from zero to eight, with higher scores indicating higher levels of depressive symptoms. A cut-off point of three is frequently used to define depression caseness (Han, 2002; Turvey, Wallace, & Herzog, 1999). The CES-D scale was originally designed to measure depressive symptom levels in the US but the validity of translated versions has been confirmed for European countries (Fuhrer & Rouillon, 1989; Goncalves & Fagulha, 2004; Missinne, et al., 2014; Van de Velde, Bracke, & Levecque, 2010; Van de Velde, Bracke, Levecque, & Meuleman, 2010).

The EURO-D scale was developed to collect harmonized data on late-life depressive symptoms in the 11 European countries which took part in the EURODEP study (Prince et al., 1999). Five existing depression measures (Geriatric Mental State-AGECAT, SHORT-CARE, CES-D, Zung Self-Rating Depression Scale, and the Comprehensive Psychopathological Rating Scale) were merged to form a 12-item scale (Prince, et al., 1999). The Euro-D has been evaluated as reliable and is highly correlated with other mental health measures (Prince, et al., 1999). Respondents were asked to report whether during the past month they experienced any of the following symptoms: depressed mood, pessimism,

suicidality, guilt, sleep, interest, irritability, appetite, fatigue, concentration, enjoyment and tearfulness. Possible responses were *yes* or *no*. The score ranges from zero to 12, with higher scores indicating higher levels of depressive symptoms (Prince, 2002). A threshold of four has been suggested for depression caseness (Castro-Costa, et al., 2007; Castro-Costa et al., 2008; Dewey & Prince, 2005).

Data Analysis

As the two scales include different numbers of items and consequently have different total scores, values were normalized to obtain a common metric for both ranging from 0 to 1. Normalized scales were obtained by dividing individual scores by the country-specific maximum value for each scale. For Euro-D, this value ranged from 0 to 11 or 12 depending on the country, while for CES-D this ranged from 0 to 8. Roughly, estimates from normalized scores can be translated into original scales by multiplying coefficients by the mid-range of the maximum value for each scale.

First, to assess whether there were systematic differences in the response to each scale by the same respondent, a difference score (Euro-D *minus* CES-D) was used to summarize congruence between the two scales (Edwards, 2001). Ordinary least squares (OLS) models were then used to assess the predictors of incongruence (Buber & Engelhardt, 2011; Cairney & Krause, 2005). All multiple linear regression models included gender, age in three categories (50 to 60, 61 to 70, over 70), marital status (married or in a partnership; divorced, widowed or single), education in three categories (primary education or less; secondary education; post-secondary education), a measure of economic strain (household is able to make ends meet with difficulty or with some difficulty; household is able to make ends meet easily or fairly easily), the number of chronic illnesses (less than two chronic illnesses; two or more chronic illnesses), limitations in activities of daily living (less than one limitation with

activities of daily living; more than one limitation with activities of daily living) reported by the respondent; and country of residence.

The Euro-D and CES-D scales use different cut-off points to determine depression caseness, *i.e.* whether respondents are likely to be clinically depressed. We estimated Cohen's kappa scores (κ) to assess the level of agreement between the two scores in identifying depression caseness. Values range between 0 (agreement equivalent to chance) and 1 (perfect agreement) (Altman, 1991). Next, the sensitivity (proportion of depressed individuals identified as depressed) and the specificity (proportion of non-depressed individuals categorized as non-depressed) of the CES-D were estimated, using the Euro-D scale as reference. A perfect match would be described as 100% sensitive (all respondents classified as depressed by the Euro-D scale are classified as depressed by the CESD scale) and 100% specific (all respondents classified as non-depressed in the Euro-D scale are classified as non-depressed by the CESD scale). We used the recommended threshold for caseness for depression for both scales. In sensitivity analyses, we examined alternative cut-off points (e.g. a threshold of three for the Euro-D scale, as sometimes recommended in the literature). We found that results were not sensitive to the specific thresholds and therefore decided to report only the results for the optimal combination between two scores (four for Euro-D and three for CES-D). The Receiver Operating Characteristic (ROC) curve was also used to examine whether the CES-D scale identified the same depressed respondents as the Euro-D scale. The ROC curve (Hanley & McNeil, 1982) measures the overall ability of the CES-D scale to discriminate against the criterion of the Euro-D score. The area under the ROC curve measures accuracy: an area of 1 represents a "perfect match", while an area of 0.5

represents a “worthless match”. A value between 0.90 and 1 was considered as excellent, between 0.80 and 0.90 as good, and between 0.70 and 0.80 as fair (Hanley & McNeil, 1982).

Finally, the associations between each normalized depression score and selected well-established risk factors for depression were explored. Multiple linear regression models were first estimated to assess whether the associations of depressive symptoms with risk factors differed for the two scales. Coefficients report the association between each explanatory variable and the continuous outcome score. Logistic regression was then used to estimate the association between each risk factor and depressive symptomatology. Cross-equation tests were performed to determine whether the associations between depressive symptom scores and explanatory variables were statistically different across the two scales.

All analyses were conducted using Stata 13 (StataCorp, 2013).

Results

Descriptive statistics

The correlation between the two scores was 0.6819 ($p < 0.000$). The scales had a similar level of internal consistency as measured by the Cronbach’s alpha (8-item CES-D, $\alpha = 0.82$; 12-item Euro-D, $\alpha = 0.72$), which are indicative of high reliability in measuring depressing symptoms.

Distributions of the CES-D and Euro-D scores are presented in Figure 1. The Kernel density plot shows that the CES-D scale is more skewed to the left than the Euro-D scale (*i.e.* CES-D scores fall under lower depressive symptoms scores than Euro-D scores). Table 1 reports the summary statistics for the two scores for the entire sample and separately by

gender and age group. The mean is 0.20 for both scales, but the larger standard deviation of the CES-D (0.26) compared to the Euro-D (0.20) illustrates the wider spread and left tale of the CES-D scale. Paired t-tests indicated that differences in means were significant only for the youngest age group (higher depression scores using the Euro-D scale) and for respondents aged 70+ (higher depression scores using the CES-D scale).

<Figure 1 about here>

<Table 1 about here>

Intra-individual differences

The determinants of intra-individual differences in scores are formally explored in Table 2, which presents estimates from the linear regression analyses with the difference between the Euro-D and the CES-D score as dependent variable ($Mean=0.009$; $SD=0.192$). A value of 0 for males, for example, would indicate that males do not display different levels of incongruence as compared to females. In contrast, a positive value would indicate that males display a larger positive discrepancy between the Euro-D and CES-D scales than do females. Results show that most variables examined significantly predict the difference scores. Males, those over 70 years of age, those with lower educational attainment, those with 2+ chronic diseases, those with one or more limitations in Activities of Daily Living (ADLs), those divorced, widowed or single, and those experiencing economic strain are significantly more likely to have a negative discrepancy between the Euro-D and CES-D scales than their respective counterparts. Significant score differences were also found between countries. Respondents in Sweden, the Netherlands, France, Denmark, Greece, Switzerland, Belgium,

and Poland were more likely to score higher on the Euro-D than the CESD score compared to respondents in Austria (the reference country), while the opposite was true for Italy.

Predicted score differences can be estimated for different individual profiles to illustrate the magnitude of these differences. For example, an Austrian male respondent, aged over 70, single or widowed, with secondary education, reporting both more than one limitation in ADL and more than two chronic conditions will score 0.0844 points higher on the CES-D scale than on the Euro-D scale. In contrast, an Austrian female in the youngest age group without health limitations, higher levels of education and married or in a partnership will have a smaller score difference of 0.054 point.

<Table 2 about here>

Discriminability of the two scales

The Euro-D and the CES-D score use different cut-off points to screen for older people with depression. The level of agreement between the two scores, as measured by the kappa score, was moderate ($\kappa=0.529$, 95% CI 0.514-0.545). Table 3 reports the results of the sensitivity and specificity levels of the CES-D scale, taking the Euro-D scale as reference. Sensitivity was 62.6%, indicating that from all respondents identified as depressed by the Euro-D scale, 62.6% are also classified as depressed by the CES-D scale. This implies that 37.4% of respondents who were identified as depressed by the Euro-D went ‘undetected’ by the CES-D scale. In turn, the specificity level was 89.5%, indicating that the CES-D scale identifies as non-depressed 89.5% of respondents categorized as non-depressed by the Euro-D scale. Table 3 also displays a positive predictive value of 65.7%, which corresponds to the proportion of respondents identified by the CES-D scale as reporting high depressive symptoms who were also identified by the Euro-D scale as reporting high depressive

symptomatology.

<Table 3 about here>

The ROC curve in Figure 2 plots the false (discordant) positives (non-depressed individuals according to the Euro-D scale who were classified as depressed by the CES-D scale) against the true (concordant) positives for the cut-off points determined above. The area under the curve determines the accuracy of the CES-D cut-off point compared to that of the Euro-D scale, *i.e.* how well the scale separates the sample into those with and without high levels of depressive symptoms with the results of the Euro-D scores as reference. The area under the ROC curve for our sample is 0.7603 (95% CI 0.7522-0.7684), which is considered as fair.

<Figure 2 about here>

Associations with established risk factors

Table 4 summarizes results from models that assess whether associations between depressive symptoms scores and established risk factors for depression in old age vary by scale. Findings from the linear regressions indicate that all predictors were significantly associated with both scales, with the exception of the oldest age group (over 70) for the normalised Euro-D score. Associations with gender were larger for the Euro-D than the CES-D scale, but for most other variables, associations were stronger for the CES-D. The largest difference between coefficients was found for marital status. Compared to their married counterparts, divorced, widowed or single respondents scored on average 0.0689 points higher (95% CI 0.0583-0.0789) on the CES-D scale but only 0.0299 points higher (95% CI 0.0220-0.0371) on the Euro-D. The two health variables included in the models (number of

chronic diseases and of limitations in ADLs) were more strongly associated with the normalised CES-D score. For instance, having more than one limitation in ADLs was associated with scoring 0.176 points higher (95% CI 0.1624-0.2027) on the normalised CES-D scale compared to 0.126 points higher (95% CI 0.1154-0.1448) on the normalised Euro-D score. Similarly, reporting two or more chronic illnesses was associated with scoring 0.0868 points higher (95% CI 0.0809-0.1007) on the normalised CES-D score, versus 0.0692 (95% CI 0.0641-0.0787) on the Euro-D score. Cross-equation tests indicate that associations of depression with age, chronic diseases, marital status, ADLs, educational level and economic strain were significantly stronger for the CES-D than for the Euro-D scale. Only associations with gender were stronger for the Euro-D. In order to understand how these differences translate back into original scores, Appendix 1 displays the OLS estimates using the original scale rather than the normalized scores. In the first two columns, we report results of the main regression for the normalised Euro-D and CES-D scores. The next two columns present estimates using original scales.

<Table 4 about here>

Table 4 also displays the results of the logistic regression models, which summarize the association between explanatory variables and depressive symptomatology as ascertained by each scale. With the exception of three country dummies and age above 70 for the CES-D scale, all variables were significantly associated with the odds of being classified as depressed by the two scales. In line with the results from the linear regressions, the largest difference between the two scales was found for marital status. Being divorced, single or widowed is associated with higher odds of being classified as depressed by the CES-D scale (1.861, 95% CI 1.6750-2.0681) than by the Euro-D scale (1.319, 95% CI 1.1873-1.4648). Cross-equation

tests suggest no significant difference between the two scales in their associations with health measures and educational level. However, the association with gender, being over 70, marital status and economic strain was stronger for the CES-D score.

Country-specific models

In supplementary analyses, separate models as presented in Table 4 were fitted for each country (see Appendix 2). These models revealed no clear systematic differences between countries. Although it is difficult to identify a common pattern, in most countries, associations between health measures and education tended to be systematically stronger for the CES-D than the Euro-D. Gender differences were larger for the Euro-D than for the CES-D scale in many countries, confirming results from Table 4. In most countries, the largest difference between the Euro-D and CES-D scores was for associations with gender and marital status.

Discussion

Our aim was to assess the comparability of the Euro-D and CES-D scales, two measures commonly used in ageing surveys. We found a high correlation between the two scores, but there are important differences in their properties. The CES-D scale is more skewed to the left resulting in a higher standard deviation compared to the Euro-D scale. Being male, as well as characteristics associated with social disadvantage (older age, divorced/widowed/single, low education, economic strain) and higher levels of physical limitations (two or more chronic diseases and one or more ADL limitations) are associated with significantly more negative discrepancies in assessments between the Euro-D and CES-

D scales. Sensitivity estimates suggest that the CES-D scale captures a more extreme pool of depressed individuals than the Euro-D scale. As a result, the association between risk factors and depressive symptom scores is often stronger for the CES-D than for the Euro-D scale. Our findings highlight the need for some caution in interpreting comparisons of levels and associations with risk factors between surveys using different measures of depressive symptoms.

Limitations and strengths

Our study is unique by assessing depressive symptoms using the Euro-D and CES-D scales for the same respondents in representative samples of Europeans in 13 countries. However, several limitations should be considered. The scales use different reference periods: Euro-D asks about depressive symptoms in the past month while CES-D asks about symptoms in the past week (Zamarro, Meijer, & Fernandes, 2008). On the other hand, both scales have been shown to have relatively high test-retest reliability over short to medium periods of assessment, suggesting that although differences in reference periods might contribute to differences in scores, they are unlikely to fully account for the systematic differences in distributions observed in our study (Larraga, Saz, Dewey, Marcos, & Lobo, 2006; Radloff, 1977). Second, the cross-sectional nature of CES-D and Euro-D measures in SHARE did not enable us to examine comparatively in score changes. In addition, our paper focused only on participants who responded to the questions used to build the two scores. Although focusing on these respondents was necessary in order to compare the scales, caution should be exercised when interpreting the results as respondents included in the study may not be representative of the full sample.

Despite these limitations, our study expands upon previous research by assessing the comparability of these two depression scales. An important question is why the same individual (presumably having a single underlying true depression state) reports different scores depending on the scale being used, resulting in more cases identified as depressed by the CES-D than by the Euro-D scale. A possible explanation is that the CES-D scale includes items not included in the Euro-D scale. In particular, the CES-D includes two positive affect items (happiness and enjoyment of life), while the Euro-D scale includes only negative affect items. The positive affect items in CES-D may lead to sharper identification of depressed individuals, as those reporting no positive affect are more likely to have higher number of symptoms in the CES-D, while those with no positive affect are not identified by the Euro-D scale, which only asks about negative affect. This may partly explain why individuals with the same level of depression report more depressive symptoms in the CESD than the Euro-D scale.

Implications for future research

Despite high levels of correlation between the two scales, caution in the interpretation of associations with risk factors is required. Our results highlight some differences in associations between CES-D and Euro-D with established risk factors for depression. This would suggest that differences in these associations reported in previous comparative studies are to some extent due to the use of different scales. On the other hand, an important question is whether differences between estimates for CES-D and Euro-D are clinically meaningful. In order to provide an estimate of the magnitude of differences in the associations between risk factors and depressive scores, we computed partial Eta-squared

(Richardson, 2011). This measure provides an estimate of the clinical significance of the results by comparing the relative sizes of the effects from different risk factors on depressive symptoms levels as measured by the two scales. Appendix 3 displays the estimates for the two scores side by side, together with their 95% CI. For all variables, we find a small to medium effect size, according to Cohen's criterion (Cohen, 1988). More importantly for the validity of previous comparative studies using these scales together, we find that the 95% CI overlap for most risk factors, but gender and marital status. Based on the results presented in Table 2 (differences in CES-D and Euro-D depressive symptoms scores), we estimated the effect sizes of the differences between the two scales for gender and marital status. The effect size of the difference between two scores for these two risk factors is of 0.006 and 0.0087 respectively, values which correspond to very small effect sizes according to Cohen's criterion. Together, these results suggest that while the relationship between risk factors and depressive symptoms sometimes differ between the CES-D and Euro-D scales, conclusions on the clinical significance of the effects are often very similar between the two scales. This adds to the argument in favour of the comparability of the two scales.

There may also be several alternatives to address the differences between the two measures of depressive symptoms. A first approach would be to identify the items that are similar across both scales (e.g., depression, sleep, energy to do things), as done in earlier studies for robustness check (Riumallo-Herl, et al., 2014). This approach is still limited by the fact that the internal consistency of the two measures is compromised by using selected items individually. A second approach might be to use a more comprehensive set of measures of wellbeing across surveys and to compare their findings with those of the depressive symptoms scale. For example, SHARE, ELSA and HRS include a 12-item or 19-item

version of the CASP scale of wellbeing (Control, Autonomy, Self-Realisation and Pleasure), specifically designed to measure wellbeing in old age (Wiggins, Higgs, Hyde, & Blane, 2004). Using this scale in combination with depression scores and self-reports of anti-depressant use, has been shown to provide a more comprehensive assessment of mental health in old age (Kruk & Reinhold, 2014; Ploubidis & Grundy, 2009). However, these extensive measures are costly and unlikely to be available always across different studies and across waves. In addition, the prescription of anti-depressives may differ substantially across countries, again introducing a source of potential bias. A third, and potentially more feasible approach derived from our study would be to achieve comparability by rescaling the Euro-D items to reflect the levels of depression as measured by the CES-D scale. Following Jürges and colleagues (Jürges, Avendano, & Mackenbach, 2008), it would be possible to make the two measures more comparable by imputing conditional probabilities. Assuming an individual with value x in the Euro-D scale has systematically lower depressive symptom levels than an individual with the same value in the CES-D scale, it is in principle possible to rescale down the Euro-D values to match the same levels of depressive symptoms captured with the CES-D scale. This would enhance cross-national comparisons of depressive symptoms across countries, and it would diminish the systematic tendency of the CES-D to show stronger associations with risk factors than the Euro-D scale.

In conclusion, we find that despite a high correlation between the two scales, there are differences in the way individuals report depressive symptoms when using the Euro-D and CES-D scales. Our results suggest that while direct comparisons of depressive symptoms levels between countries and using different measures should be avoided, studies that compare associations between risk factors and depressive symptoms across countries using

these measures can still be valid. Although the strength of associations differs, our study shows that associations between each scale and risk factors are often in the same direction and display similar levels of clinical significance. Rescaling one of the scales or using more comprehensive assessments of wellbeing may be helpful in minimizing bias. Our findings imply that both scales measure the same underlying concept and, with some adjustments, can be used in comparative studies of the determinants of depression in old age.

REFERENCES

- Alexopoulos, G. S. (2005). Depression in the elderly. *The Lancet*, 364, 1961-1971.
- Altman, D. G. (1991). *Practical statistics for medical research*. London: Chapman and Hall.
- Banks, J., Nazroo, J., & Steptoe, A. (2012). *The dynamics of ageing: evidence from the English Longitudinal Study of Ageing 2002-2010*. London: Institute for Fiscal Studies.
- Beekman, A., Copeland, J., & Prince, M. (1999). Review of community prevalence of depression in later life. *British Journal of Psychiatry*, 174, 307-311.
- Börsch-Supan, A., Hank, K., & Jürges, H. (2005). A New Comprehensive and International View on Ageing: Introducing the Survey of Health, Ageing and Retirement in Europe. *European Journal of Ageing* 2(4), 245-253.
- Börsch-Supan, A., & Jürges, H. (Eds.). (2005). *The Survey of Health, Ageing and Retirement in Europe - Methodology*. University of Mannheim: Mannheim Research Institute for the Economics of Aging.
- Buber, I., & Engelhardt, H. (2011). The association between age and depressive symptoms among older men and women in Europe. Findings from SHARE. *Comparative Population Studies*, 36(1), 103-126.
- Cairney, J., & Krause, N. (2005). The Social Distribution of Psychological Distress and Depression in Older Adults. *Journal of Aging and Health*, 17(6), 807-835.
- Castro-Costa, E., Dewey, M., Stewart, R., Banerjee, S., Hupper, F., Mendonca-Lima, C., et al. (2007). Prevalence of depressive symptoms and syndromes in later life in ten European countries: The SHARE study. *British Journal of Psychiatry*, 191(5), 393-401.
- Castro-Costa, E., Dewey, M., Stewart, R., Banerjee, S., Huppert, F., Mendonca-Lima, C., et al. (2008). Ascertaining late-life depressive symptoms in Europe: an evaluation of the survey version of the EURO-D scale in 10 nations. The SHARE project. *International Journal of Methods in Psychiatric Research*, 17(1), 12-29.
- Cohen, J. (1988). *Statistical power analysis for the behavioral sciences*. New York: Routledge Academic.
- Crimmins, E. M., Kim, J. K., & Solé-Auró, A. (2011). Gender differences in health: results from SHARE, ELSA and HRS. *European Journal of Public Health* 21(1), 81-91.
- Dewey, M., & Prince, M. (2005). Mental health. In A. Borsch-Supan (Ed.), *Health, Ageing and Retirement in Europe. First results from the Survey of Health, Ageing and Retirement in Europe* (pp. 108-117). Mannheim: Mannheim Research Institute for the Economics of Ageing.
- Di Gessa, G., & Grundy, E. (2014). The relationship between active ageing and health using longitudinal data from Denmark, France, Italy and England. *Journal of Epidemiology and Community Health*, 68(3), 261-267.
- Djernes, J. K. (2006). Prevalence and predictors of depression in populations of elderly: a review. *Acta Psychiatrica Scandinavica*, 113(5), 372-387.
- Edwards, J. R. (2001). Ten difference score myths. *Organizational Research Methods*, 4(3), 264-286.
- Ferrari, A., Charlson, F., Norman, R., Patten, S., Freedman, G., Murray, C., et al. (2013). Burden of depressive disorders by country, sex, age and year. Findings from the Global Burden of Disease study 2010. *PLOS Medicine*, 10(11).
- Fuhrer, R., & Rouillon, F. (1989). Version française de l'échelle CES-D (Center for Epidemiologic Studies-Depression scale). Description et traduction de l'échelle d'autoévaluation. *Psychiatrie Psychobiologie*, 4(3), 163-166.

- Goncalves, B., & Fagulha, T. (2004). The portuguese version of the Center for Epidemiologic Studies Depression scale (CES-D). *European Journal of Psychological Assessment, 20*(4), 339–348.
- Han, B. (2002). Depressive Symptoms and Self-Rated Health in Community-Dwelling Older Adults: A Longitudinal Study. *Journal of the American Geriatrics Society, 50*(9), 1549-1556.
- Hanley, J. A., & McNeil, B. J. (1982). The meaning and use of the area under a receiver operating characteristic (ROC) curve. *Radiology, 143*(1), 29–36.
- Jürges, H., Avendano, M., & Mackenbach, J. (2008). Are different measures of self-rated health comparable? An assessment in five European countries. *European Journal of Epidemiology, 23*(12), 773-781.
- Kohout, F. J., Berkman, L. F., Evans, D. A., & Cornoni-Huntley, J. (1993). Two Shorter Forms of the CES-D Depression Symptoms Index. *Journal of Aging and Health, 5*(2), 179-193.
- Kok, R., Avendano-Pabon, M., Bago d'Uva, T., & Mackenbach, J. P. (2012). Can reporting heterogeneity explain differences in depressive symptoms across Europe? *Social Indicators Research, 105*(2), 191-210.
- Kruk, K. E., & Reinhold, S. (2014). The effect of children on depression in old age. *Social Science & Medicine, 100*(0), 1-11.
- Larraga, L., Saz, P., Dewey, M., Marcos, G., & Lobo, A. (2006). Validation of the Spanish version of the EURO-D Scale: an instrument for detecting depression in older people. *International Journal of Geriatric Psychiatry, 21*(12), 1199–1205.
- Lunau, T., Wahrendorf, M., Dragano, N., & Siegrist, J. (2013). Work stress and depressive symptoms in older employees: impact of national labour and social policies. *BMC Public Health, 13*(1), 1086.
- Missinne, S., Vandevive, C., Van de Velde, S., & Bracke, P. (2014). Measurement equivalence of the CES-D 8 depression-scale among the ageing population in eleven European countries. *Social Science Research, 46*, 38–47.
- National Institute on Ageing. (2014). *Growing older in America. The Health and Retirement Study*. Washington: US Department of Health and Human Services.
- Ploubidis, G. B., & Grundy, E. (2009). Later-life mental health in Europe: a country-level comparison. *Journal of Gerontology: Social Sciences, 64B*(5), 666-676.
- Prince, M. (2002). The development of the EURO-D scale. In J. Copeland, M. Abou-Saleh & D. Blazer (Eds.), *Principles and practice of Geriatric Psychiatry* (pp. 159-160). London: John Wiley & Sons.
- Prince, M., Reischies, F., Beekman, A. T., Fuhrer, R., Jonker, C., Kivela, S. L., et al. (1999). Development of the EURO-D scale--a European, Union initiative to compare symptoms of depression in 14 European centres. *British Journal of Psychiatry, 174*(4), 330-338.
- Radloff, L. S. (1977). The CES-D Scale. A Self-Report Depression Scale for Research in the General Population. *Applied Psychological Measurement, 1*(3), 385-401.
- Richardson, J. T. E. (2011). Eta squared and partial eta squared as measures of effect size in educational research. *Educational Research Review, 6*(2), 135-147.
- Riumallo-Herl, C., Basu, S., Stuckler, D., Courtin, E., & Avendano-Pabon, M. (2014). Job loss and depression during the 2008 Great Recession in the US and Europe. *International Journal of Epidemiology, 43*(5), 1508-1517.

- Siegrist, J., Lunau, T., Wahrendorf, M., & Dragano, N. (2012). Depressive symptoms and psychosocial stress at work among older employees in three continents. *Globalization and Health, 8*(27).
- StataCorp. (2013). *Stata Statistical Software: Release 13*. College Station, TX: StataCorp LP.
- Turvey, C. L., Wallace, R. B., & Herzog, R. (1999). A Revised CES-D Measure of Depressive Symptoms and a DSM-Based Measure of Major Depressive Episodes in the Elderly. *International Psychogeriatrics, 11*(02), 139-148.
- Van de Velde, S., Bracke, P., & Levecque, K. (2010). Gender differences in depression in 23 European countries. Cross-national variation in the gender gap in depression. *Social Science & Medicine, 71*(2), 305-313.
- Van de Velde, S., Bracke, P., Levecque, K., & Meuleman, B. (2010). Gender differences in depression in 25 European countries after eliminating measurement bias in the CES-D 8. *Social Science Research, 39*(3), 396-404.
- Wiggins, R. D., Higgs, P. F. D., Hyde, M., & Blane, D. B. (2004). Quality of life in the third age: Key predictors of the CASP-19 measure. *Ageing and Society, 24*(5), 693-708.
- Zamarro, G., Meijer, E., & Fernandes, M. (2008). Mental health and cognitive ability. Comparison between SHARE, ELSA and HRS. In A. Börsh-Supan, A. Brugiavini, H. Jürges, A. Kapteyn, J. P. Mackenbach, J. Siegrist & G. Weber (Eds.), *First results from the Survey of Health, Ageing and Retirement in Europe (2004-2007). Starting the longitudinal dimension* (pp. 40-48). Mannheim: Mannheim Research Institute for the Economics of Ageing.

TABLES

Table 1. Summary statistics of the normalised CES-D and Euro-D scores (N=15,487)

	<i>Euro-D</i>			<i>CES-D</i>			<i>Comparison</i>	
	N	Mean	Standard deviation	N	Mean	Standard deviation	Paired t-test	P value
Entire sample	15,487	0.20	0.20	15,487	0.20	0.26	0.5192	0.6036
Females	8,445	0.23	0.20	8,445	0.23	0.28	-1.3634	0.1728
Males	7,042	0.15	0.17	7,042	0.15	0.22	0.9244	0.3553
50 to 60	6,372	0.18	0.18	6,372	0.17	0.24	3.8248	0.0001
61 to 70	4,812	0.18	0.18	4,812	0.18	0.25	1.4147	0.1572
Over 70	4,303	0.24	0.21	4,303	0.25	0.25	-6.6647	0.0000

Table 2. Estimated differences in CES-D and Euro-D depressive symptoms scores (fully adjusted model)

	<i>Score difference</i>
Male	-0.0116** (0.00382)
61 to 70 (ref.: 50 to 60)	0.00184 (0.00448)
Over 70	-0.0102* (0.00503)
Divorced, widowed or single	-0.0374*** (0.00418)
Secondary education (ref.: primary education or less)	0.00889 (0.00504)
Post-secondary education	0.0122* (0.00569)
Economic strain	-0.0384*** (0.00453)
2+ chronic diseases	-0.0176*** (0.00399)
1+ limitations in ADLs	-0.0494*** (0.00778)
Germany (ref.: Austria)	-0.00997 (0.0107)
Sweden	0.0506*** (0.0103)
Netherlands	0.0648*** (0.0101)
Spain	0.0220 (0.0115)
Italy	-0.0226* (0.0112)
France	0.0598*** (0.0108)
Denmark	0.0737*** (0.00998)
Greece	0.0288** (0.0111)
Switzerland	0.0411*** (0.0114)
Belgium	0.0418*** (0.0103)
Czechia	-0.0113 (0.0107)
Poland	0.0395*** (0.0118)
Constant	0.00805 (0.0104)
Observations	10,536
R-squared	0.068

Robust standard errors in parentheses, *** $p < 0.001$, ** $p < 0.01$, * $p < 0.05$.

Notes:

¹Chronic diseases include high blood pressure, high blood cholesterol, stroke, diabetes, chronic lung disease, asthma, arthritis, osteoporosis, cancer, stomach or duodenal ulcer, Parkinson disease, cataract and hip fracture.

²Activities of daily living include putting on shoes and socks, walking across a room, bathing or showering, eating, getting in and out of bed, using the toilet, including getting up and down.

³Low education levels correspond to having a high school degree or lower qualifications.

⁴Economic strain is defined as reporting difficulties to make ends meet in the past month.

Table 3. Cross-tabulation of the depression thresholds of the Euro-D and CES-D scores and associated sensitivity and specificity values

	<i>Depressed Euro-D</i>	<i>Non-depressed Euro-D</i>	<i>Total</i>
Depressed CES-D	2,389	1,413	3,802
Non-depressed CES-D	1,258	10,427	11,685
Total	3,647	11,840	15,487

	<i>%</i>	<i>95% CI</i>
Sensitivity	62.8	61.3-64.4
Specificity	89.2	88.7-89.8
Positive predicted value	65.5	63.9-67
Negative predicted value	88.1	87.5-88.6

Table 4. Linear regressions (fully adjusted) for the normalized CES-D and Euro-D depressive symptoms scores and odds ratios for caseness for depression

	<i>Linear regressions</i>		<i>Cross-equation tests</i>		<i>Logistic regressions</i>		<i>Cross-equation tests</i>	
	Normalised CES-D	Normalised Euro-D	Chi-squared	P-value	CES-D	Euro-D	Chi-squared	P-value
Male	-0.0510*** (0.00467)	-0.0626*** (0.00353)	2.57	0.1087	0.608*** (0.0325)	0.470*** (0.0251)	20.88	0.0000
61 to 70 (ref.: 50 to 60)	-0.0188*** (0.00544)	-0.0170*** (0.00402)	3.95	0.0470	0.816** (0.0523)	0.749*** (0.0473)	0.00	0.9946
Over 70	0.0116 (0.00622)	0.00138 (0.00458)	63.24	0.0000	1.012 (0.0665)	0.860* (0.0567)	18.72	0.0000
2+ chronic illnesses	0.0868*** (0.00499)	0.0692*** (0.00370)	120.68	0.0000	2.295*** (0.121)	2.249*** (0.118)	0.36	0.5501
1+ limitations with ADLs	0.176*** (0.0101)	0.126*** (0.00746)	152.44	0.0000	3.3365*** (0.260)	2.973*** (0.228)	2.51	0.1128
Divorced, single or widowed	0.0653*** (0.00523)	0.0279*** (0.00381)	148.24	0.0000	1.861*** (0.100)	1.319*** (0.0707)	41.67	0.0000
Secondary education (ref.: primary education or less)	-0.0313*** (0.00638)	-0.0224*** (0.00462)	2.73	0.0986	0.760*** (0.0469)	0.755*** (0.0458)	1.48	0.2244
Post-secondary education	-0.0348*** (0.00704)	-0.0226*** (0.00531)	56.66	0.0000	0.650*** (0.0523)	0.715*** (0.0555)	1.61	0.2039
Economic strain	0.0942*** (0.00571)	0.0558*** (0.00411)	262.06	0.0000	1.841*** (0.132)	1.659*** (0.114)	10.86	0.0010
Germany (ref.: Austria)	0.0347** (0.0133)	0.0247** (0.00934)	35.22	0.0000	1.408* (0.207)	1.245 (0.199)	49.92	0.0000
Sweden	-0.0158 (0.0130)	0.0348*** (0.00958)	46.11	0.0000	0.739* (0.113)	1.074 (0.168)	0.18	0.6673
Netherlands	-0.0310* (0.0127)	0.0338*** (0.00931)	84.18	0.0000	0.632** (0.1000)	1.236 (0.195)	9.98	0.0016
Spain	0.0274 (0.0158)	0.0494*** (0.0109)	4.68	0.0305	1.379* (0.210)	2.004*** (0.313)	0.25	0.6198
Italy	0.0790***	0.0563***	129.64	0.0000	1.977***	2.433***	1.32	0.2508

	(0.0148)	(0.0101)			(0.280)	(0.357)		
France	0.00220	0.0620***	44.90	0.0000	1.095	2.181***	40.03	0.0000
	(0.0137)	(0.00946)			(0.158)	(0.320)		
Denmark	-0.0658***	0.00785	98.76	0.0000	0.387***	0.915	12.94	0.0003
	(0.0127)	(0.00942)			(0.0674)	(0.151)		
Greece	-0.0409**	-0.0120	12.57	0.0004	0.829	1.004	15.18	0.0001
	(0.0137)	(0.00976)			(0.122)	(0.154)		
Switzerland	-0.0179	0.0232*	9.80	0.0017	0.734	1.239	0.45	0.5012
	(0.0145)	(0.0106)			(0.136)	(0.225)		
Belgium	-0.00145	0.0404***	17.93	0.0000	1.047	1.532**	2.18	0.1396
	(0.0133)	(0.00945)			(0.152)	(0.230)		
Czechia	0.00954	-0.00174	103.79	0.0000	1.244	1.261	46.65	0.0000
	(0.0139)	(0.00947)			(0.176)	(0.190)		
Poland	0.0581***	0.0976***	7.43	0.0064	1.694***	3.708***	43.19	0.0000
	(0.0153)	(0.0104)			(0.244)	(0.548)		
Constant	0.133***	0.141***			0.127***	0.141***		
	(0.0136)	(0.00955)			(0.0196)	(0.0221)		
Observations	10,536	10,536			10,536	10,536		
R-squared	0.236	0.220			0.159	0.144		

Robust standard errors in parentheses, *** $p < 0.001$, ** $p < 0.01$, * $p < 0.05$.

FIGURES

Figure 1. Kernel density plot of the normalised CES-D and Euro-D scales

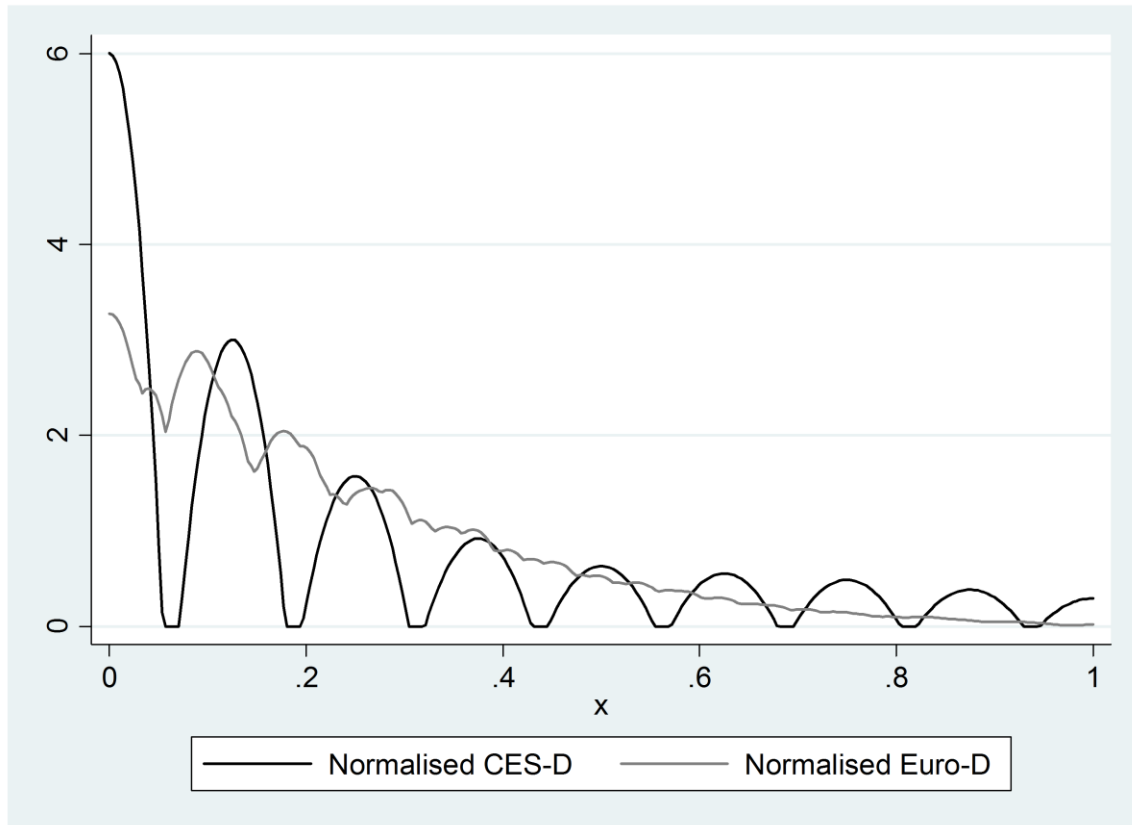
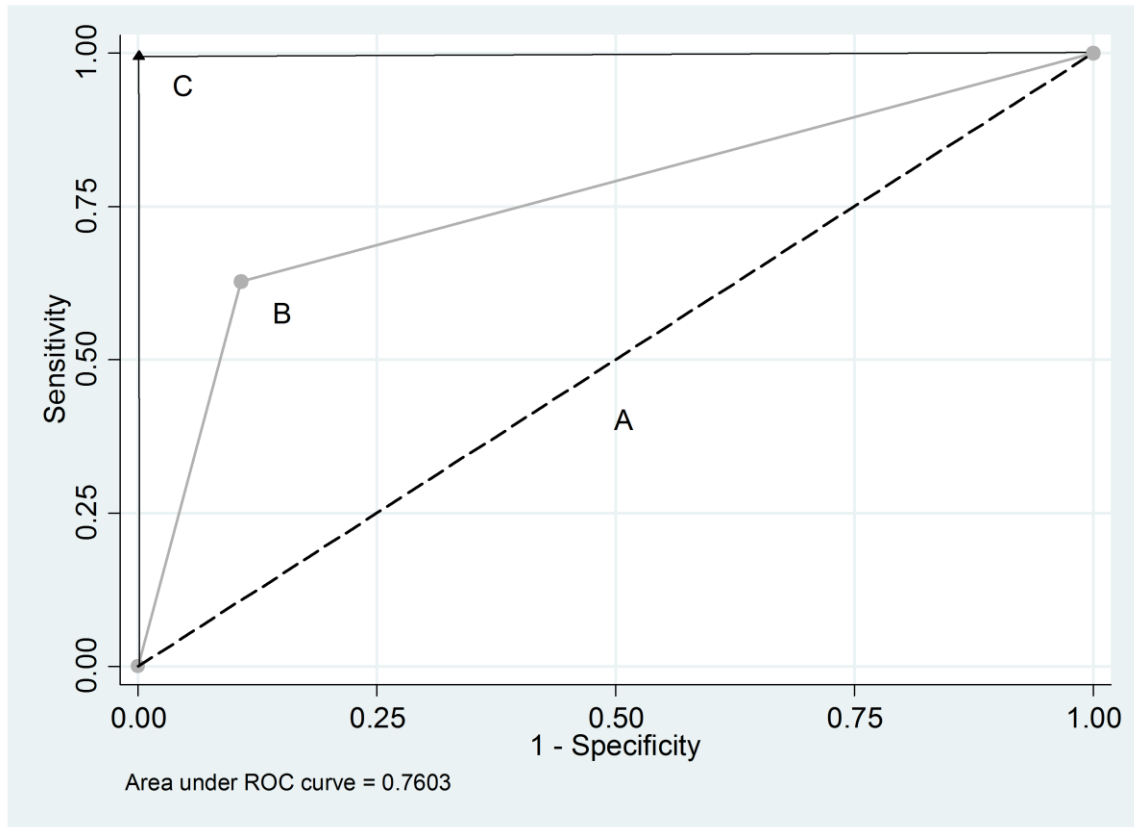


Figure 2. Receiver Operating Characteristic (ROC) curve for the CES-D scale of depressive symptoms



Notes:

A - Line of zero discrimination (theoretical)

B - ROC-curve for the CES-D score in our sample (0.7603)

C - Perfect discrimination (theoretical)

APPENDIX

Appendix 1. Linear regressions (fully adjusted) for the normalized and original CES-D and Euro-D scores

	<i>Normalised CES-D</i>	<i>Normalised Euro-D</i>	<i>Original CES-D</i>	<i>Original Euro-D</i>
Male	-0.0510*** (0.00467)	-0.0626*** (0.00353)	-0.428*** (0.0377)	-0.732*** (0.0405)
61 to 70	-0.0188*** (0.00544)	-0.0170*** (0.00402)	-0.174*** (0.0440)	-0.202*** (0.0462)
Over 70	0.0116 (0.00622)	0.00138 (0.00458)	0.0421 (0.0503)	-0.0179 (0.0527)
2+ chronic illnesses	0.0868*** (0.00499)	0.0692*** (0.00370)	0.727*** (0.0403)	0.819*** (0.0425)
1+ limitations with ADLS	0.176*** (0.0101)	0.126*** (0.00746)	1.461*** (0.0821)	1.492*** (0.0859)
Divorced, single or widowed	0.0653*** (0.00523)	0.0279*** (0.00381)	0.549*** (0.0421)	0.335*** (0.0438)
Secondary education	-0.0313*** (0.00638)	-0.0224*** (0.00462)	-0.318*** (0.0514)	-0.323*** (0.0533)
Post-secondary education	-0.0348*** (0.00704)	-0.0226*** (0.00531)	-0.367*** (0.0566)	-0.333*** (0.0606)
Economic strain	0.0942*** (0.00571)	0.0558*** (0.00411)	0.417*** (0.0397)	0.389*** (0.0437)
Germany (ref.: Austria)	0.0347** (0.0133)	0.0247** (0.00934)	0.314** (0.108)	0.148 (0.109)
Sweden	-0.0158 (0.0130)	0.0348*** (0.00958)	-0.128 (0.105)	0.0439 (0.108)
Netherlands	-0.0310* (0.0127)	0.0338*** (0.00931)	-0.232* (0.103)	0.239* (0.108)
Spain	0.0274 (0.0158)	0.0494*** (0.0109)	0.348** (0.128)	0.708*** (0.131)
Italy	0.0790*** (0.0148)	0.0563*** (0.0101)	0.797*** (0.119)	0.819*** (0.121)
France	0.00220 (0.0137)	0.0620*** (0.00946)	0.0712 (0.112)	0.790*** (0.114)
Denmark	-0.0658*** (0.0127)	0.00785 (0.00942)	-0.504*** (0.103)	-0.0428 (0.110)
Greece	-0.0409** (0.0137)	-0.0120 (0.00976)	-0.0849 (0.109)	-0.108 (0.113)
Switzerland	-0.0179 (0.0145)	0.0232* (0.0106)	-0.0925 (0.117)	0.157 (0.122)
Belgium	-0.00145 (0.0133)	0.0404*** (0.00945)	0.0591 (0.108)	0.346** (0.110)
Czechia	0.00954 (0.0139)	-0.00174 (0.00947)	0.230* (0.112)	0.116 (0.113)
Poland	0.0581*** (0.0153)	0.0976*** (0.0104)	0.710*** (0.123)	1.401*** (0.124)
Constant	0.133*** (0.0136)	0.141*** (0.00955)		
Observations	10,536	10,536	10,536	10,536

R-squared	0.236	0.220	0.219	0.226
-----------	-------	-------	-------	-------

Note: Robust standard errors in parentheses; *** $p < 0.001$, ** $p < 0.01$, * $p < 0.05$

Appendix 2. Linear regressions (fully adjusted) for the normalized CES-D and Euro-D depressive symptoms scores and odds ratios for caseness for depression per country

Austria	<i>Linear regressions</i>		<i>Logistic regressions</i>	
	<i>Normalised</i>		<i>CES-D</i>	<i>Euro-D</i>
	<i>Normalised CES-D</i>	<i>Euro-D</i>		
Male	-0.0112 (0.0240)	-0.0432** (0.0157)	0.999 (0.262)	0.390** (0.119)
61 to 70	0.00259 (0.0264)	0.0229 (0.0170)	1.533 (0.500)	1.404 (0.495)
Over 70	0.0386 (0.0270)	0.0425* (0.0189)	2.461** (0.780)	1.670 (0.599)
2+ chronic illnesses	0.0837*** (0.0240)	0.0662*** (0.0165)	2.068** (0.486)	2.161** (0.569)
1+ limitations with ADLs	0.154*** (0.0432)	0.122*** (0.0291)	3.049*** (0.909)	3.893*** (1.201)
Divorced, single or widowed	0.0474* (0.0239)	0.0149 (0.0152)	1.421 (0.385)	0.978 (0.290)
Secondary education	-0.0834* (0.0373)	0.000159 (0.0235)	0.762 (0.219)	0.831 (0.263)
Tertiary education	-0.101* (0.0392)	8.56e-05 (0.0260)	0.373** (0.142)	0.957 (0.372)
Economic strain	0.0752*** (0.0215)	0.0467** (0.0149)	2.348** (0.677)	2.047* (0.668)
Constant	0.127** (0.0434)	0.0740* (0.0292)	0.0679*** (0.0326)	0.0806*** (0.0451)
Observations	473	473	473	473
R-squared	0.168	0.171	0.142	0.134
Belgium				
	<i>Normalised</i>		<i>CES-D</i>	<i>Euro-D</i>
	<i>Normalised CES-D</i>	<i>Euro-D</i>		
Male	-0.0805*** (0.0157)	-0.0921*** (0.0120)	0.555** (0.102)	0.390*** (0.0707)
61 to 70	-0.0304 (0.0184)	-0.0253 (0.0143)	0.820 (0.170)	0.767 (0.156)
Over 70	-0.0527** (0.0185)	-0.0475*** (0.0143)	0.630* (0.137)	0.576* (0.125)
2+ chronic illnesses	0.0882*** (0.0161)	0.0591*** (0.0123)	2.180*** (0.379)	1.653** (0.289)
1+ limitations with ADLs	0.214*** (0.0314)	0.147*** (0.0247)	4.337*** (1.073)	4.248*** (1.041)
Divorced, single or widowed	0.0596*** (0.0170)	0.0246 (0.0126)	1.965*** (0.353)	1.249 (0.225)
Secondary education	-0.0328 (0.0199)	-0.00956 (0.0148)	0.848 (0.173)	1.029 (0.206)
Tertiary education	-0.0374 (0.0216)	-0.0109 (0.0163)	0.754 (0.182)	0.780 (0.184)
Economic strain	0.0325* (0.0151)	0.0157 (0.0119)	1.400 (0.264)	1.356 (0.243)
Constant	0.175***	0.213***	0.174***	0.290***

	(0.0261)	(0.0196)	(0.0497)	(0.0757)
Observations	978	978	978	978
R-squared	0.193	0.178	0.129	0.105
Czechia				
	<i>Normalised CES-D</i>	<i>Normalised Euro-D</i>	<i>CES-D</i>	<i>Euro-D</i>
Male	-0.0233 (0.0168)	-0.0401*** (0.0112)	0.754 (0.126)	0.666* (0.119)
61 to 70	-0.0358 (0.0192)	-0.0273* (0.0128)	0.736 (0.144)	0.790 (0.166)
Over 70	0.0336 (0.0224)	0.0264 (0.0153)	1.194 (0.229)	1.335 (0.275)
2+ chronic illnesses	0.101*** (0.0169)	0.0719*** (0.0116)	2.168*** (0.349)	2.297*** (0.403)
1+ limitations with ADLs	0.164*** (0.0365)	0.131*** (0.0282)	2.961*** (0.796)	2.418** (0.651)
Divorced, single or widowed	0.101*** (0.0179)	0.0441*** (0.0117)	2.041*** (0.331)	1.698** (0.293)
Secondary education	-0.0245 (0.0249)	-0.0350* (0.0169)	0.815 (0.159)	0.807 (0.165)
Tertiary education	-0.0831** (0.0295)	-0.0542** (0.0200)	0.470* (0.141)	0.488* (0.154)
Economic strain	0.0607* (0.0256)	0.0100 (0.0159)	2.081* (0.666)	1.058 (0.308)
Constant	0.110** (0.0365)	0.154*** (0.0238)	0.127*** (0.0515)	0.184*** (0.0688)
Observations	940	940	940	940
R-squared	0.175	0.180	0.112	0.0975
Switzerland				
	<i>Normalised CES-D</i>	<i>Normalised Euro-D</i>	<i>CES-D</i>	<i>Euro-D</i>
Male	-0.0663*** (0.0182)	-0.0740*** (0.0151)	0.397** (0.121)	0.343*** (0.0992)
61 to 70	-0.0410 (0.0214)	-0.0504** (0.0173)	0.719 (0.251)	0.643 (0.207)
Over 70	-0.0219 (0.0251)	-0.0276 (0.0209)	0.847 (0.276)	0.726 (0.235)
2+ chronic illnesses	0.0304 (0.0213)	0.0408* (0.0179)	1.418 (0.391)	1.527 (0.420)
1+ limitations with ADLs	0.0356 (0.0702)	0.0634 (0.0508)	1.008 (0.574)	1.150 (0.572)
Divorced, single or widowed	0.0498* (0.0211)	0.0386* (0.0168)	1.749* (0.496)	1.825* (0.494)
Secondary education	-0.0200 (0.0350)	-0.0441 (0.0277)	0.815 (0.314)	0.426* (0.158)
Tertiary education	-0.0336 (0.0365)	-0.0328 (0.0302)	0.617 (0.276)	0.580 (0.235)
Economic strain	0.0698*** (0.0186)	0.0518*** (0.0153)	2.853*** (0.900)	1.742* (0.477)
Constant	0.144*** (0.0404)	0.195*** (0.0324)	0.120*** (0.0598)	0.315** (0.141)

Observations	497	497	497	497
R-squared	0.113	0.156	0.110	0.103
Germany				
	<i>Normalised CES-D</i>	<i>Normalised Euro-D</i>	<i>CES-D</i>	<i>Euro-D</i>
Male	-0.0206 (0.0162)	-0.0523*** (0.0120)	0.797 (0.162)	0.633* (0.139)
61 to 70	-0.0271 (0.0169)	-0.0181 (0.0124)	0.614* (0.143)	0.613* (0.147)
Over 70	0.0348 (0.0208)	-0.0162 (0.0147)	1.241 (0.302)	0.674 (0.173)
2+ chronic illnesses	0.105*** (0.0170)	0.0855*** (0.0118)	3.491*** (0.701)	3.085*** (0.666)
1+ limitations with ADLs	0.207*** (0.0361)	0.149*** (0.0270)	3.696*** (1.070)	3.391*** (0.977)
Divorced, single or widowed	0.0754*** (0.0185)	0.00761 (0.0124)	2.130*** (0.444)	0.987 (0.221)
Secondary education	-0.132 (0.124)	-0.0883 (0.0618)	0.315 (0.287)	0.271 (0.182)
Tertiary education	-0.172 (0.124)	-0.110 (0.0623)	0.160* (0.148)	0.204* (0.141)
Economic strain	0.0644*** (0.0142)	0.0183 (0.0112)	1.911** (0.417)	1.300 (0.302)
Constant	0.228 (0.123)	0.237*** (0.0627)	0.323 (0.300)	0.572 (0.409)
Observations	782	782	782	782
R-squared	0.255	0.204	0.196	0.108
Denmark				
	<i>Normalised CES-D</i>	<i>Normalised Euro-D</i>	<i>CES-D</i>	<i>Euro-D</i>
Male	-0.00116 (0.0130)	-0.0174 (0.0118)	1.064 (0.281)	0.812 (0.180)
61 to 70	-0.0389** (0.0132)	-0.0668*** (0.0122)	0.668 (0.212)	0.310*** (0.0826)
Over 70	0.00590 (0.0168)	-0.0427** (0.0151)	0.928 (0.298)	0.419** (0.121)
2+ chronic illnesses	0.0293* (0.0124)	0.0522*** (0.0111)	1.628 (0.436)	2.782*** (0.632)
1+ limitations with ADLs	0.151*** (0.0362)	0.113*** (0.0267)	3.074** (1.117)	3.222*** (1.026)
Divorced, single or widowed	0.0368** (0.0139)	0.00697 (0.0125)	1.989** (0.526)	1.019 (0.240)
Secondary education	-0.0174 (0.0192)	-0.00357 (0.0173)	0.629 (0.212)	0.937 (0.291)
Tertiary education	0.00802 (0.0203)	0.00259 (0.0176)	1.025 (0.352)	1.378 (0.447)
Economic strain	0.0378** (0.0122)	0.0518*** (0.0113)	1.693* (0.419)	1.975** (0.434)
Constant	0.0643** (0.0233)	0.136*** (0.0210)	0.0520*** (0.0236)	0.104*** (0.0402)
Observations	826	826	826	826

R-squared	0.118	0.118	0.0830	0.0911
Spain				
	<i>Normalised CES-D</i>	<i>Normalised Euro-D</i>	<i>CES-D</i>	<i>Euro-D</i>
Male	-0.110*** (0.0227)	-0.0885*** (0.0165)	0.430*** (0.0848)	0.410*** (0.0815)
61 to 70	-0.0107 (0.0268)	-0.0173 (0.0199)	0.958 (0.235)	0.661 (0.168)
Over 70	0.0715* (0.0302)	0.0353 (0.0207)	1.376 (0.348)	1.054 (0.253)
2+ chronic illnesses	0.140*** (0.0244)	0.0928*** (0.0174)	2.344*** (0.463)	2.534*** (0.494)
1+ limitations with ADLs	0.172*** (0.0405)	0.119*** (0.0299)	2.944*** (0.908)	2.367** (0.685)
Divorced, single or widowed	0.0389 (0.0246)	0.0179 (0.0175)	1.376 (0.275)	1.017 (0.204)
Secondary education	-0.00136 (0.0275)	-0.0150 (0.0190)	0.882 (0.207)	0.698 (0.163)
Tertiary education	-0.0506 (0.0327)	-0.0250 (0.0257)	0.535 (0.208)	0.877 (0.297)
Economic strain	0.0889** (0.0292)	0.0324 (0.0204)	2.152* (0.758)	2.000* (0.626)
Constant	0.112** (0.0377)	0.182*** (0.0273)	0.160*** (0.0652)	0.251*** (0.0943)
Observations	646	646	646	646
R-squared	0.219	0.198	0.136	0.120
France				
	<i>Normalised CES-D</i>	<i>Normalised Euro-D</i>	<i>CES-D</i>	<i>Euro-D</i>
Male	-0.0882*** (0.0165)	-0.0760*** (0.0119)	0.392*** (0.0732)	0.401*** (0.0674)
61 to 70	0.00629 (0.0200)	0.00972 (0.0138)	1.101 (0.227)	0.896 (0.168)
Over 70	-0.0306 (0.0210)	-0.0275 (0.0148)	0.752 (0.165)	0.526** (0.111)
2+ chronic illnesses	0.0961*** (0.0182)	0.0558*** (0.0123)	2.306*** (0.398)	1.813*** (0.291)
1+ limitations with ADLs	0.138*** (0.0348)	0.109*** (0.0219)	2.720*** (0.750)	2.809*** (0.717)
Divorced, single or widowed	0.114*** (0.0185)	0.0425*** (0.0126)	2.575*** (0.451)	1.621** (0.260)
Secondary education	-0.0590** (0.0196)	-0.0305* (0.0139)	0.644* (0.122)	0.661* (0.117)
Tertiary education	-0.0391 (0.0226)	-0.0435** (0.0162)	0.727 (0.173)	0.477** (0.108)
Economic strain	0.0441* (0.0173)	0.0576*** (0.0128)	1.390 (0.301)	1.939** (0.403)
Constant	0.149*** (0.0264)	0.196*** (0.0196)	0.184*** (0.0568)	0.365*** (0.104)
Observations	944	944	944	944
R-squared	0.194	0.183	0.149	0.120

Greece				
	<i>Normalised CES-D</i>	<i>Normalised Euro-D</i>	<i>CES-D</i>	<i>Euro-D</i>
Male	-0.0677*** (0.0155)	-0.0869*** (0.0124)	0.447*** (0.0897)	0.339*** (0.0728)
61 to 70	-0.0514** (0.0178)	-0.0186 (0.0135)	0.494** (0.119)	0.785 (0.196)
Over 70	-0.0144 (0.0217)	0.0252 (0.0169)	0.656 (0.158)	1.156 (0.285)
2+ chronic illnesses	0.0689*** (0.0163)	0.0454*** (0.0137)	2.024*** (0.374)	1.667* (0.339)
1+ limitations with ADLs	0.200*** (0.0386)	0.165*** (0.0307)	3.639*** (1.080)	3.413*** (0.991)
Divorced, single or widowed	0.0822*** (0.0168)	0.0306* (0.0130)	2.061*** (0.395)	1.430 (0.284)
Secondary education	-0.0534** (0.0178)	-0.0379** (0.0137)	0.683 (0.145)	0.646 (0.147)
Tertiary education	-0.0702** (0.0213)	-0.0304 (0.0175)	0.361** (0.112)	0.609 (0.182)
Economic strain	0.0798*** (0.0207)	0.0403* (0.0162)	2.283* (0.863)	1.907 (0.728)
Constant	0.122*** (0.0291)	0.149*** (0.0222)	0.141*** (0.0600)	0.143*** (0.0632)
Observations	955	955	955	955
R-squared	0.203	0.204	0.150	0.136
Italy				
	<i>Normalised CES-D</i>	<i>Normalised Euro-D</i>	<i>CES-D</i>	<i>Euro-D</i>
Male	-0.0897*** (0.0193)	-0.0864*** (0.0133)	0.536*** (0.0855)	0.339*** (0.0728)
61 to 70	0.00986 (0.0238)	0.0112 (0.0157)	1.070 (0.201)	0.785 (0.196)
Over 70	0.0349 (0.0276)	0.0347 (0.0186)	1.174 (0.254)	1.156 (0.285)
2+ chronic illnesses	0.139*** (0.0200)	0.109*** (0.0136)	2.835*** (0.455)	1.667* (0.339)
1+ limitations with ADLs	0.229*** (0.0347)	0.170*** (0.0256)	4.496*** (1.367)	3.413*** (0.991)
Divorced, single or widowed	0.0805*** (0.0224)	0.0325* (0.0153)	1.616** (0.278)	1.430 (0.284)
Secondary education	-0.0679** (0.0225)	-0.0374* (0.0147)	0.674* (0.118)	0.646 (0.147)
Tertiary education	-0.0836** (0.0305)	-0.0576* (0.0234)	0.583 (0.172)	0.609 (0.182)
Economic strain	0.0987** (0.0327)	0.0429 (0.0256)	2.287* (0.789)	1.907 (0.728)
Constant	0.161*** (0.0412)	0.166*** (0.0311)	0.179*** (0.0711)	0.143*** (0.0632)
Observations	895	895	895	895
R-squared	0.234	0.263	0.146	0.163
Netherlands				

	<i>Normalised CES-D</i>	<i>Normalised Euro-D</i>	<i>CES-D</i>	<i>Euro-D</i>
Male	-0.0346** (0.0128)	-0.0568*** (0.0111)	0.638 (0.147)	0.339*** (0.0728)
61 to 70	-0.0301* (0.0145)	-0.00637 (0.0132)	0.420** (0.140)	0.785 (0.196)
Over 70	-0.00847 (0.0184)	-0.0220 (0.0142)	0.802 (0.219)	1.156 (0.285)
2+ chronic illnesses	0.0617*** (0.0158)	0.0645*** (0.0133)	2.577*** (0.598)	1.667* (0.339)
1+ limitations with ADLs	0.192*** (0.0425)	0.107*** (0.0309)	4.827*** (1.540)	3.413*** (0.991)
Divorced, single or widowed	0.0710*** (0.0156)	0.0722*** (0.0130)	2.262*** (0.517)	1.430 (0.284)
Secondary education	-0.0352 (0.0232)	-0.0264 (0.0188)	0.609 (0.167)	0.646 (0.147)
Tertiary education	-0.0380 (0.0261)	-0.0322 (0.0210)	0.633 (0.216)	0.609 (0.182)
Economic strain	0.0413*** (0.0123)	0.0179 (0.0110)	2.028** (0.501)	1.907 (0.728)
Constant	0.104*** (0.0255)	0.167*** (0.0215)	0.0849*** (0.0322)	0.143*** (0.0632)
Observations	859	859	859	859
R-squared	0.175	0.172	0.158	0.110

Poland

	<i>Normalised CES-D</i>	<i>Normalised Euro-D</i>	<i>CES-D</i>	<i>Euro-D</i>
Male	-0.0654** (0.0205)	-0.0822*** (0.0139)	0.689* (0.114)	0.339*** (0.0728)
61 to 70	0.0215 (0.0266)	-0.00303 (0.0171)	1.162 (0.242)	0.785 (0.196)
Over 70	0.0229 (0.0283)	0.0237 (0.0184)	1.314 (0.282)	1.156 (0.285)
2+ chronic illnesses	0.155*** (0.0216)	0.113*** (0.0141)	2.824*** (0.483)	1.667* (0.339)
1+ limitations with ADLs	0.171*** (0.0273)	0.0953*** (0.0176)	3.073*** (0.599)	3.413*** (0.991)
Divorced, single or widowed	0.102*** (0.0235)	0.0268 (0.0155)	1.872*** (0.327)	1.430 (0.284)
Secondary education	-0.0102 (0.0233)	-0.0306* (0.0155)	0.819 (0.153)	0.646 (0.147)
Tertiary education	-0.0769** (0.0291)	-0.0508* (0.0213)	0.571* (0.156)	0.609 (0.182)
Economic strain	0.0778* (0.0382)	0.0494 (0.0293)	2.319 (1.063)	1.907 (0.728)
Constant	0.127** (0.0433)	0.221*** (0.0325)	0.122*** (0.0602)	0.143*** (0.0632)
Observations	825	825	825	825
R-squared	0.227	0.230	0.152	0.142

Sweden

	<i>Normalised CES-D</i>	<i>Normalised Euro-D</i>	<i>CES-D</i>	<i>Euro-D</i>
--	-------------------------	--------------------------	--------------	---------------

	<i>Euro-D</i>			
Male	-0.0171 (0.0133)	-0.0374** (0.0118)	0.767 (0.148)	0.339*** (0.0728)
61 to 70	-0.0366* (0.0159)	-0.0347** (0.0134)	0.539* (0.135)	0.785 (0.196)
Over 70	0.0135 (0.0201)	0.00739 (0.0162)	0.988 (0.247)	1.156 (0.285)
2+ chronic illnesses	0.0529*** (0.0148)	0.0567*** (0.0121)	1.796** (0.362)	1.667* (0.339)
1+ limitations with ADLs	0.173*** (0.0352)	0.149*** (0.0288)	4.202*** (1.143)	3.413*** (0.991)
Divorced, single or widowed	0.0262 (0.0141)	0.0157 (0.0125)	1.359 (0.257)	1.430 (0.284)
Secondary education	-0.0153 (0.0169)	9.00e-05 (0.0148)	0.856 (0.209)	0.646 (0.147)
Tertiary education	0.00583 (0.0178)	0.00229 (0.0150)	1.174 (0.295)	0.609 (0.182)
Economic strain	0.0564*** (0.0133)	0.0561*** (0.0114)	1.970*** (0.396)	1.907 (0.728)
Constant	0.102*** (0.0216)	0.139*** (0.0184)	0.0990*** (0.0326)	0.143*** (0.0632)
Observations	916	916	916	916
R-squared	0.127	0.147	0.0980	0.103

Robust standard errors in parentheses

*** p<0.001, ** p<0.01, * p<0.05

Appendix 3. Effect sizes for the Euro-D and CES-D scores per risk factor

