Socioeconomic Status and Depressive Symptoms: An Individual-Participant Data Meta-Analysis on Range Restriction and Measurement

Supplemental Online Material

Contents	Page(s)

Supplemental Methods

Identifying Studies

The data for this project were pooled from a collection of 127 datasets on racial/ethnic disparities in depressive symptoms in the U.S. A detailed breakdown of study identification, systematic search, and eligibility is reported elsewhere. Briefly, a systematic search was conducted in ICPSR to identify open access datasets that were considered nationally-representative. Two search strings were used: String 1 = *depression "United States" -"great depression"*; String 2 = *"depressive symptoms" "United States" -"great depression"*.

Inclusion and Exclusion Criteria

Variables

IPD with a psychometric measure of depressive symptoms—that consisted of at least two or more items of depressive symptoms—were included. Clinician based diagnoses of depression, or by self-report, were excluded. In addition to these criteria, we excluded datasets with measures of depressive symptoms that employed skip logic because these measures only assess depressive symptoms after participants report feeling depressed or lost interest (e.g., CIDI). Additionally, IPD that did not include a codable measure for one of the three components of SES (i.e., income, education, or occupation) were excluded.

Geographic Location

IPD that recruited participants from the U.S. were included. This decision helped us manage some heterogeneity between correlations.

Research Design and Time Period

IPD were included if they randomly recruited participants from at least two or more U.S. states. Regarding the research design, IPD were excluded if they were collected from intervention designs. IPD were included regardless of the year the data was collected.

Participants

Participants within each dataset were included regardless of their age, sex, or race/ethnicity. Parents and children within the same dataset were included, in addition to siblings and twins. Participants missing data on depressive symptoms, components of SES, age, sex, or race/ethnicity were excluded during the data extraction process.

Data Items – Participant Level

Income

Annual income for each participant was extracted. The highest income was extracted when individual income, and household/family income was reported separately, and household/family or parental income was extracted for youth under 18 years old. Intervals were recoded as the midpoint of the interval (e.g., $20,000 - $29,999 = $24,999.50). Bottom- or top-coded intervals were recoded as their value (e.g., ≥ $80,000 = $80,000).

Years of Education

The number of completed years of education was extracted. For youth under the age 18, the highest level of education between both parents was included in the analysis. Intervals were recoded as the midpoint of the interval (e.g., 9 to 12 years = 10.5). Education measured by the highest level of degree completed was recoded to best represent the number of years to complete the degree (e.g., Bachelor's Degree = 16); categories that had no clear connection to the number of years completed were coded as missing (e.g., Vocational School). Bottom- or top-coded intervals were recoded as their value (e.g., College Graduate or Above = 16).

Occupational Status and Prestige

Occupational status and prestige scores were based on U.S. Census Occupational Categorization. The highest parental occupation was included for youth under the age 18. The Nam-Powers-Boyd Occupational Status Scale, which is based on the 2000 U.S. Census, was used to compute occupational status scores (Nam and Boyd, 2004). For occupational prestige, we used the prestige scale of Nakao and Treas (Nakao and Treas, 1994), which is based on the occupational categories from the 1980 U.S. Census. Both scales range from 0 (lowest) to 100 (highest). When a classification could not be matched to either of the two scales, it was coded as missing.

Depressive Symptoms

For each participant, we computed a mean score of depressive symptoms. Only mean scores with 50% or less of missing items were included to help make the scores more comparable across participants. Higher mean scores indicated a greater number of depressive symptoms.

Data Items – Study Level

Range Restriction

Range restriction was coded for each component of SES by examining the minimum and maximum value across participants within each IPD. For income, IPD were flagged for range restriction towards the lower end if they did not include participants with less than $20,000 and towards the upper end if they did not include, or distinguish between, participants with greater than $120,000 because The American Trends Panel reported that higher income families had annual incomes of $120,400 or more in 2018 (Horowitz et al., 2020). For years of education, IPD were flagged for range restriction towards the lower end if they did not include participants with

less than 12 years of education (e.g., high school degree), and towards the upper end if they did not include, or distinguish between, participants with greater than 16 years of education (e.g., bachelor's compared to master's degree). For occupational status and prestige, IPD were flagged for range restriction towards the lower end if they did not included participants with scores in the bottom quartile (scores < 25), and towards the upper end if they did not included participants with scores in the top quartile (scores > 75).

Measurement Instrument of Depressive Symptoms

Each type of measurement instrument used to assess depressive symptoms was dummy coded. The specific measurement instruments of depressive symptoms included in this study were based on the instruments used in nationally-representative, public-access datasets that were available on ICPSR. These instruments included the Behavior Problem Index (BPI; Peterson and Zill, 1986), CES-D (Radloff, 1977), K6+ Self-Reporting Measure (K6; Kessler et al., 2003), Mental Health Inventroy-5 (MHI-5; Stewart et al., 1988), Psychiatric Epidemiology Research Interview Demoralization Scale (PERID; Dohrenwend et al., 1980), Public Health Questionnaire-5 (PHQ-9; Kroenke et al., 2001), and Short Form-36 (SF3-6; Ware and Donald Sherbourne, 1992). A brief description of each measure and example symptoms are displayed in Supplemental Table 2.

Measurement Error

Measurement error was examined using two coefficients: alpha (α) (Cronbach, 1951) and omega (ω) (McDonald, 1999). Both reliability statistics were included because alpha can sometimes bias the true reliability whereas omega can be more robust(Dunn et al., 2014).

Year of Data Collection

The year of data collection for each dataset was coded as a potential confounding methodological characteristic.

Analysis Plan

We conducted a multi-level mixed-effects meta-analysis using the metaSEM package (Cheung, 2015). Correlations between components of SES and depressive symptoms were transformed from *r* to Fisher's *z* metric prior to analysis (Borenstein et al., 2021). We used a multi-level model because several correlations were extracted from the same dataset, making them conditionally dependent. Within each dataset, we also extracted correlations for specific subsamples of participants by age, sex, and race/ethnicity. Clustering the effect sizes more accurately estimates the variance at the different levels of analysis (Cheung, 2014). In our multi-level model, we defined level-one as the sampling variability of the correlation, level-two as the correlations (*r*), level-three as the IPD from which the correlations were extracted from. We chose to cluster by IPD given that some IPD, although from the same parent-study, may have different measures of depressive symptoms and reliability estimates.

Moderation analyses were conducted separately for each component of SES. First, a single moderator model was specified for each predictor of the correlation: range restriction, measure of depressive symptoms, measurement error, and year of data collection. Then, significant moderators were combined into a multiple moderator model to examine the contribution of each moderator over and above the other. For depressive symptoms, an intercept-free model was specified in order to estimate the correlation for each measure and its 95% confidence interval (95% CI). Variation across measures was examined by comparing the intercept free moderation model to a model with just the intercept. A chi-square result with a *p*-value < .05 was evidence of moderation by measure of depressive symptoms. We also conducted

a moderation test for depressive symptoms by including the CES-D as the reference measure because it was the most frequently used measure across the included IPD and it has the least amount of overlap in symptom items (Fried, 2017). In this model, the intercept was the estimated correlation for the CES-D and the slopes were the difference in the magnitude of the correlation for each measurement instrument. A slope estimate with a p-value < .05 was evidence that the correlation for a specific measure was significantly different from the CES-D. The latter interpretation also applies to interpretating range restriction moderation whereas the slope for measurement reliability and year of data collection suggests the difference in the magnitude of correlation for each unit increase in the moderator. As a post-hoc exploratory test, the number of items used for the CES-D instrument was also tested as a continuous moderator among datasets that included the CES-D. For each test, we reported the intercept and slopes and their 95% CI.

Range Restriction – Characteristics

Income

From 122 datasets, we extracted 2,573 correlations. Average income across participants and datasets was $43,640.51 (range of mean income $11,217.00–$88,533.85). All datasets included participants with incomes below $20,000, thus showing no evidence of a lower-end restricted range. Some datasets had an upper-end restricted range: 44 did not include or delineate incomes above $120,000. These datasets were flagged for range restriction and compared with the remaining 78 datasets.

Years of Education

From 122 datasets, we extracted 2,486 correlations. Average years of education across participants and datasets was 13.03 years (range of mean years of education 10.33–15.02). All datasets included participants with less than 12 years of education, thus showing no evidence of a

lower-end restricted range. Some datasets had an upper-end restricted range: 17 did not include or delineate participants with more than 16 years of education. These datasets were flagged for range restriction and compared with the remaining 105 datasets.

Occupational Status and Prestige

From 61 datasets, we extracted 1,181 correlations. The average occupational status and prestige score across participants and datasets was 51.72 for status (range of mean status 25.38–64.68) and 43.86 for prestige (range of mean prestige 32.26–51.32). All datasets included participants with bottom-quartile occupational status/prestige scores, thus showing no evidence of a lower-end restricted range. No evidence was seen of an upper-end restricted range for occupational status; however, some datasets had an upper-end restricted range for occupational prestige: 11 excluded participants with prestige scores in the top-quartile. These were flagged for range restriction and compared with the remaining 50 datasets.

Measurement of Depressive Symptoms – Characteristics

The most frequently used instrument across datasets was the CES-D (60 datasets; 1,031 correlations), followed by the K6 (29 datasets, 958 correlations); BPI (12 datasets, 213 correlations); SF-36 (9 datasets, 174 correlations); MHI-5 (8 datasets, 133 correlations); PHQ-9 (2 datasets, 36 correlations); and PERID (1 dataset, 21 correlations). For 2 datasets, we could not identify the instrument used. Across datasets and correlations, the average α was $\bar{\alpha} = 0.80$ (range of $\bar{\alpha}$ 0.64–0.92) and the average ω was $\bar{\omega} = 0.81$ (range of $\bar{\omega}$ 0.64–0.92). The average reliability for each measure was: CES-D ($\bar{\alpha} = 0.80$, $\bar{\omega} = 0.81$); K6 ($\bar{\alpha} = 0.86$, $\bar{\omega} = 0.86$); BPI ($\bar{\alpha} = 0.69$, $\bar{\omega} = 0.71$); SF-36 ($\bar{\alpha} = 0.79$, $\bar{\omega} = 0.80$); MHI-5 ($\bar{\alpha} = 0.77$, $\bar{\omega} = 0.79$); PHQ-9 ($\bar{\alpha} = 0.79$, $\bar{\omega} = 0.80$); and PERID ($\bar{\alpha} = 0.79$, $\bar{\omega} = 0.83$).

References

Borenstein, M., Hedges, L.V., Higgins, J.P.T., Rothestein, H.R., 2021. Introduction to meta-analysis, 2nd ed. Wiley, Hoboken, NJ.

Cheung, M.W.-L., 2015. metaSEM: an R package for meta-analysis using structural equation modeling. Front. Psychol. 5, 1–7. https://doi.org/10.3389/fpsyg.2014.01521

Cheung, M.W.-L., 2014. Modeling dependent effect sizes with three-level meta-analyses: A structural equation modeling approach. Psychol. Methods 19, 211–229. https://doi.org/10.1037/a0032968

Cronbach, L.J., 1951. Coefficient alpha and the internal structure of tests. Psychometrika 16, 297–334. https://doi.org/10.1007/BF02310555

Dohrenwend, B.P., Shrout, P.E., Egri, G., Mendelsohn, F.S., 1980. Nonspecific psychological distress and other dimensions of psychopathology. Arch. Gen. Psychiatry 37, 1229–1236. https://doi.org/10.1001/archpsyc.1980.01780240027003

Dunn, T.J., Baguley, T., Brunsden, V., 2014. From alpha to omega: A practical solution to the pervasive problem of internal consistency estimation. Br. J. Psychol. 105, 399–412. https://doi.org/10.1111/bjop.12046

Fried, E.I., 2017. The 52 symptoms of major depression: Lack of content overlap among seven common depression scales. J. Affect. Disord. 208, 191–197. https://doi.org/10.1016/j.jad.2016.10.019

Horowitz, J.M., Igielnik, R., Kochhar, R., 2020. The American Trends Panel survey methodoy [WWW Document]. Pew Soc. Trends. URL https://www.pewsocialtrends.org/2020/01/09/methodology-27/ (accessed 1.4.21).

Kessler, R.C., Barker, P.R., Colpe, L.J., Epstein, J.F., Gfroerer, J.C., Hiripi, E., Howes, M.J., Normand, S.-L.T., Manderscheid, R.W., Walters, E.E., Zaslavsky, A.M., 2003. Screening for serious mental illness in the general population. Arch. Gen. Psychiatry 60, 184. https://doi.org/10.1001/archpsyc.60.2.184

Kroenke, K., Spitzer, R.L., Williams, J.B.W., 2001. The PHQ-9: Validity of a brief depression severity measure. J. Gen. Intern. Med. 16, 606–613. https://doi.org/10.1046/j.1525-1497.2001.016009606.x

McDonald, R.P., 1999. Test theory: a unified treatment. L. Erlbaum Associates, Mahwah, N.J.

Nakao, K., Treas, J., 1994. Updating occupational prestige and socioeconomic scores: How the new measures measure up. Sociol. Methodol. 24, 1–72. https://doi.org/10.2307/270978

Nam, C.B., Boyd, M., 2004. Occupational Status in 2000; Over a Century of Census-Based Measurement. Popul. Res. Policy Rev. 23, 327–358. https://doi.org/10.1023/B:POPU.0000040045.51228.34

Peterson, J.L., Zill, N., 1986. Marital disruption, parent-child relationships, and behavior problems in children. J. Marriage Fam. 48, 295–307. https://doi.org/10.2307/352397

Radloff, L.S., 1977. The CES-D scale: a self-report depression scale for research in the general population. Appl. Psychol. Meas. 1, 385–401.

Stewart, A.L., Hays, R.D., Ware, J.E., 1988. The MOS Short-Form General Health Survey: Reliability and validity in a patient population. Med. Care 26, 724–735. https://doi.org/10.1097/00005650-198807000-00007

Ware, J.E., Donald Sherbourne, C., 1992. The MOS 36-Item Short-Form Health Survey (SF-36): I. Conceptual framework and item selection. Med. Care 30, 473–483. https://doi.org/10.1097/00005650-199206000-00002

Supplemental Table 1. PRISMA-IPD Checklist

PRISMA-IPD Section/topic	Item No	Checklist item	Reported on page
Title			
Title	1	Identify the report as a systematic review and meta-analysis of individual participant data.	1
Abstract			
Structured summary	2	Provide a structured summary including as applicable:	2
		Background: state research question and main objectives, with information on participants, interventions, comparators and outcomes.	
		Methods: report eligibility criteria; data sources including dates of last bibliographic search or elicitation, noting that IPD were sought; methods of assessing risk of bias.	
		Results: provide number and type of studies and participants identified and number (%) obtained; summary effect estimates for main outcomes (benefits and harms) with confidence intervals and measures of statistical heterogeneity. Describe the direction and size of summary effects in terms meaningful to those who would put findings into practice.	
		Discussion: state main strengths and limitations of the evidence, general interpretation of the results and any important implications.	
		Other: report primary funding source, registration number and registry name for the systematic review and IPD meta-analysis.	
Introduction			
Rationale	3	Describe the rationale for the review in the context of what is already known.	4-6
Objectives	4	Provide an explicit statement of the questions being addressed with reference, as applicable, to participants, interventions, comparisons, outcomes and study design (PICOS). Include any hypotheses that relate to particular types of participant-level subgroups.	6
Methods			
Protocol and registration	5	Indicate if a protocol exists and where it can be accessed. If available, provide registration information including registration number and registry name. Provide publication details, if applicable.	7
Eligibility criteria	6	Specify inclusion and exclusion criteria including those relating to participants, interventions, comparisons, outcomes, study design and characteristics (e.g. years when conducted, required minimum follow-up). Note whether these were applied at the study or individual level i.e. whether eligible participants were included (and ineligible participants excluded) from a study that included a wider population than specified by the review inclusion criteria. The rationale for criteria should be stated.	7, supplemental material
Identifying studies -	7	Describe all methods of identifying published and unpublished studies including, as applicable: which bibliographic databases were searched with dates of coverage; details of any hand searching including of conference proceedings;	7, supplem

information sources		use of study registers and agency or company databases; contact with the original research team and experts in the field; open adverts and surveys. Give the date of last search or elicitation.	ental material
Identifying studies - search	8	Present the full electronic search strategy for at least one database, including any limits used, such that it could be repeated.	supplem ental material
Study selection processes	9	State the process for determining which studies were eligible for inclusion.	supplem ental material
Data collection processes	10	Describe how IPD were requested, collected and managed, including any processes for querying and confirming data with investigators. If IPD were not sought from any eligible study, the reason for this should be stated (for each such study).	7
		If applicable, describe how any studies for which IPD were not available were dealt with. This should include whether, how and what aggregate data were sought or extracted from study reports and publications (such as extracting data independently in duplicate) and any processes for obtaining and confirming these data with investigators.	
Data items	11	Describe how the information and variables to be collected were chosen. List and define all study level and participant level data that were sought, including baseline and follow-up information. If applicable, describe methods of standardising or translating variables within the IPD datasets to ensure common scales or measurements across studies.	8, supplem ental
IPD integrity	A1	Describe what aspects of IPD were subject to data checking (such as sequence generation, data consistency and completeness, baseline imbalance) and how this was done.	NA
Risk of bias assessment in individual studies.	12	Describe methods used to assess risk of bias in the individual studies and whether this was applied separately for each outcome. If applicable, describe how findings of IPD checking were used to inform the assessment. Report if and how risk of bias assessment was used in any data synthesis.	NA
Specification of outcomes and effect measures	13	State all treatment comparisons of interests. State all outcomes addressed and define them in detail. State whether they were pre-specified for the review and, if applicable, whether they were primary/main or secondary/additional outcomes. Give the principal measures of effect (such as risk ratio, hazard ratio, difference in means) used for each outcome.	8
Synthesis methods	14	Describe the meta-analysis methods used to synthesise IPD. Specify any statistical methods and models used. Issues should include (but are not restricted to): • Use of a one-stage or two-stage approach. • How effect estimates were generated separately within each study and combined across studies (where applicable).	9, supplem ental

		• Specification of one-stage models (where applicable) including how clustering of patients within studies was accounted for. • Use of fixed or random effects models and any other model assumptions, such as proportional hazards. • How (summary) survival curves were generated (where applicable). • Methods for quantifying statistical heterogeneity (such as I^2 and τ^2). • How studies providing IPD and not providing IPD were analysed together (where applicable). • How missing data within the IPD were dealt with (where applicable).	
Exploration of variation in effects	A2	If applicable, describe any methods used to explore variation in effects by study or participant level characteristics (such as estimation of interactions between effect and covariates). State all participant-level characteristics that were analysed as potential effect modifiers, and whether these were pre-specified.	9
Risk of bias across studies	15	Specify any assessment of risk of bias relating to the accumulated body of evidence, including any pertaining to not obtaining IPD for particular studies, outcomes or other variables.	NA
Additional analyses	16	Describe methods of any additional analyses, including sensitivity analyses. State which of these were pre-specified.	9
Results			
Study selection and IPD obtained	17	Give numbers of studies screened, assessed for eligibility, and included in the systematic review with reasons for exclusions at each stage. Indicate the number of studies and participants for which IPD were sought and for which IPD were obtained. For those studies where IPD were not available, give the numbers of studies and participants for which aggregate data were available. Report reasons for non-availability of IPD. Include a flow diagram.	10, Figure 1
Study characteristics	18	For each study, present information on key study and participant characteristics (such as description of interventions, numbers of participants, demographic data, unavailability of outcomes, funding source, and if applicable duration of follow-up). Provide (main) citations for each study. Where applicable, also report similar study characteristics for any studies not providing IPD.	10, supplemental
IPD integrity	A3	Report any important issues identified in checking IPD or state that there were none.	NA
Risk of bias within studies	19	Present data on risk of bias assessments. If applicable, describe whether data checking led to the up-weighting or down-weighting of these assessments. Consider how any potential bias impacts on the robustness of meta-analysis conclusions.	NA
Results of individual studies	20	For each comparison and for each main outcome (benefit or harm), for each individual study report the number of eligible participants for which data were obtained and show simple summary data for each intervention group (including, where applicable, the number of events), effect estimates and confidence intervals. These may be tabulated or included on a forest plot.	Table 1
Results of syntheses	21	Present summary effects for each meta-analysis undertaken, including confidence intervals and measures of statistical heterogeneity. State whether the analysis was pre-specified, and report the numbers of studies and participants and, where applicable, the number of events on which it is based.	

		When exploring variation in effects due to patient or study characteristics, present summary interaction estimates for each characteristic examined, including confidence intervals and measures of statistical heterogeneity. State whether the analysis was pre-specified. State whether any interaction is consistent across trials.	11-12, Tables 2-5
		Provide a description of the direction and size of effect in terms meaningful to those who would put findings into practice.	
Risk of bias across studies	22	Present results of any assessment of risk of bias relating to the accumulated body of evidence, including any pertaining to the availability and representativeness of available studies, outcomes or other variables.	NA
Additional analyses	23	Give results of any additional analyses (e.g. sensitivity analyses). If applicable, this should also include any analyses that incorporate aggregate data for studies that do not have IPD. If applicable, summarise the main meta-analysis results following the inclusion or exclusion of studies for which IPD were not available.	NA
Discussion			
Summary of evidence	24	Summarise the main findings, including the strength of evidence for each main outcome.	12-16
Strengths and limitations	25	Discuss any important strengths and limitations of the evidence including the benefits of access to IPD and any limitations arising from IPD that were not available.	16
Conclusions	26	Provide a general interpretation of the findings in the context of other evidence.	17
Implications	A4	Consider relevance to key groups (such as policy makers, service providers and service users). Consider implications for future research.	12-16
Funding			
Funding	27	Describe sources of funding and other support (such as supply of IPD), and the role in the systematic review of those providing such support.	Title Page

Supplemental Table 2. A Brief Description and Example Symptoms for Each Measurement Instrument of Depressive Symptoms Included in the Meta-Analysis

Measure	Description	Example Symptoms
Behavior Problem Index (BPI)	The BPI (Peterson and Zill, 1986) measures children's behavior problems. Only items that assessed depressive symptoms were utilized and coded in this study.	Too fearful or anxious Felt worthless Unhappy, sad, or depressed Worried too much
Center for Epidemiologic Studies Depression (CES-D)	The CES-D (Radloff, 1977) contains up to 20 items and measures four domains of depressive symptoms: depressed affect, positive affect, somatic symptoms, and interpersonal problems. All items (if used) were utilized and coded in this study.	I felt depressed I was happy My sleep was restless I could not get "going"
K6+ Self-Reporting Measure (K6)	The K6 (Kessler et al., 2003) contains six items and measures non-specific psychological distress. All items (if used) were utilized and coded in this study.	Nervous Hopeless Restless or fidgety Worthless
Mental Health Inventory-5 (MHI-5)	The MHI-5 (Stewart et al., 1988) contains five items from the mental health subscale of the SF-36 (Ware and Sherbourne, 1992), and measures general mental health (e.g., depression and anxiety). All five items were utilized and coded in this study.	Been a very nervous person Felt calm and peaceful Felt downhearted and blue Been a happy person
Psychiatric Epidemiology Research Interview Demoralization Scale (PERID)	The PERID (Dohrenwend, 1980) is a psychiatric symptom checklist for use in the general population. Only items that assessed depressive symptoms were utilized and coded in this study.	Been bothered by feelings of sadness or depression Felt very bad or worthless Felt completely hopeless about everything Wondered if anything is worthwhile
Public Health Questionnaire-9 (PHQ-9)	The PHQ-9 (Kroenke et al., 2001) contains nine items and measures each diagnostic criteria for depressive disorder from the Diagnostic and Statistical Manual of Mental Disorders, Fourth Edition. All items were utilized and coded in this study.	Little interest or pleasure in doing things Feeling down, depressed, or hopeless Feeling tired or having little energy Trouble concentrating on things
Short Form-36 (SF-36)	The SF-36 (Ware and Sherbourne, 1992) measures health across multiple domains. Only nine-items that assess depressive symptoms from the mental health and energy/fatigue subscales were utilized and coded in this study.	Been a very nervous person? Felt downhearted and blue? Been a happy person Felt tired

Supplemental Table 3. Characteristics Across the 59 Inter-university Consortium for Political and Social Research (ICPSR) Datasets

Dataset	# DS	N	M_{Inc} (range)	M_{Edu} (range)	M_{Status} (range)	$M_{Prestige}$ (range)	DS Measure	$\bar{\alpha}$ (range)	$\bar{\omega}$ (range)
Aging, Status, and Sense of Control (ASOC), 1995, 1998, 2001 [United States]	3	4,858	49.44 (0-800)	13.51 (1-20)	58.84 (1-100)	46.26 (17-86)	CES-D	0.79 (0.45-0.91)	0.83 (0.75-0.91)
Americans' Changing Lives: Waves I, II, III, IV, and V, 1986, 1989, 1994, 2002, and 2011	5	11,729	41.10 (0-4,000)	12.21 (0-17)	52.77 (1-100)	44.62 (17-94)	CES-D	0.82 (0.66-0.96)	0.83 (0.68-0.96)
Collaborative Psychiatric Epidemiology Surveys (CPES), 2001-2003 [United States]	1	5,587	35.67 (0-200)	--	--	--	CES-D	0.74 (0.58-0.84)	0.74 (0.59-0.85)
Early Childhood Longitudinal Study [United States]: Kindergarten Class of 1998-1999, Kindergarten-Eighth Grade Full Sample	3	32,639	54.90 (0-200)	12.72 (1-20)	--	--	CES-D	0.85 (0.68-0.96)	0.86 (0.56-0.96)
General Social Survey, 1972-2016 [Cumulative File]	1	925	21.60 (1-25)	13.78 (0-20)	53.80 (1-100)	44.03 (17-86)	CES-D	0.74 (0.51-0.87)	0.77 (0.64-0.87)
Health and Retirement Study (HRS)	15	271,359	23.28 (-1-6,530)	12.41 (0-17)	--	--	CES-D	0.80 (0.53-0.98)	0.81 (0.64-0.98)
Midlife in the United States (MIDUS 2), 2004-2006	1	3,903	71.52 (0-300)	14.28 (4-20)	63.57 (1-100)	49.22 (17-86)	K6	0.80 (0.47-0.92)	0.86 (0.81-0.91)
Midlife in the United States (MIDUS 3), 2013-2014	1	2,689	88.53 (0-300)	14.54 (4-20)	64.68 (1-100)	50.91 (17-86)	K6	0.85 (0.67-0.96)	0.87 (0.79-0.96)
Midlife in the United States (MIDUS Refresher), 2011-2014	1	2,430	86.31 (0-300)	14.93 (4-20)	64.37 (1-100)	51.32 (17-86)	K6	0.84 (0.72-0.92)	0.84 (0.65-0.92)
National Comorbidity Survey: Reinterview (NCS-2), 2001-2002	1	4,939	57.65 (0-91)	13.16 (6-16)	--	--	--	0.88 (0.74-0.95)	0.88 (0.71-0.95)
National Health and Nutrition Examination Survey (NHANES), 2005-2006	1	4,619	42.49 (3-75)	12.69 (8-16)	--	--	PHQ-9	0.79 (0.64-0.87)	0.80 (0.68-0.89)
National Health and Nutrition Examination Survey (NHANES), 2007-2008	1	5,191	41.30 (3-101)	12.50 (8-16)	--	--	PHQ-9	0.80 (0.70-0.89)	0.80 (0.62-0.90)
National Health Interview Survey, 1997	1	37,093	34.78 (3-75)	12.58 (0-22)	--	--	K6	0.85 (0.77-0.92)	0.85 (0.77-0.93)
National Health Interview Survey, 1998	1	32,794	36.86 (3-75)	12.66 (0-22)	--	--	K6	0.86 (0.77-0.91)	0.86 (0.75-0.92)
National Health Interview Survey, 1999	1	31,304	37.87 (3-75)	12.70 (0-22)	--	--	K6	0.86 (0.67-0.94)	0.86 (0.77-0.94)
National Health Interview Survey, 2000	1	32,906	38.29 (3-75)	12.70 (0-22)	--	--	K6	0.85 (0.55-0.94)	0.84 (0.48-0.94)

Dataset	# DS	N	M_{Inc} (range)	M_{Edu} (range)	M_{Status} (range)	$M_{Prestige}$ (range)	DS Measure	$\bar{\alpha}$ (range)	$\bar{\omega}$ (range)
National Health Interview Survey, 2001	1	33,760	39.13 (3-75)	12.77 (0-22)	--	--	K6	0.85 (0.72-0.93)	0.85 (0.73-0.93)
National Health Interview Survey, 2002	1	31,568	39.84 (3-75)	12.81 (0-22)	--	--	K6	0.85 (0.64-0.93)	0.86 (0.75-0.93)
National Health Interview Survey, 2003	1	30,918	39.48 (3-75)	12.80 (0-22)	--	--	K6	0.85 (0.71-0.94)	0.85 (0.73-0.94)
National Health Interview Survey, 2007	1	23,734	45.98 (18-101)	13.60 (8-18)	52.12 (5-98)	43.10 (17-74)	K6	0.85 (0.75-0.92)	0.85 (0.74-0.92)
National Health Interview Survey, 2008	1	21,613	47.71 (18-101)	13.95 (11-18)	52.27 (5-98)	43.18 (17-74)	K6	0.81 (0.48-0.92)	0.82 (0.53-0.92)
National Health Interview Survey, 2009	1	27,842	47.57 (18-101)	13.97 (11-18)	51.84 (5-98)	42.96 (17-74)	K6	0.82 (0.58-0.92)	0.83 (0.62-0.91)
National Health Interview Survey, 2010	1	27,395	46.32 (18-101)	13.99 (11-18)	51.80 (5-98)	43.02 (17-74)	K6	0.84 (0.43-0.98)	0.85 (0.60-0.98)
National Health Interview Survey, 2011	1	34,078	46.66 (18-101)	13.81 (8-18)	51.97 (5-98)	43.07 (17-74)	K6	0.85 (0.64-0.94)	0.85 (0.66-0.93)
National Health Interview Survey, 2012	1	35,324	47.16 (18-101)	13.83 (8-18)	52.02 (5-98)	43.05 (17-74)	K6	0.85 (0.74-0.91)	0.85 (0.75-0.92)
National Health Interview Survey, 2013	1	34,574	47.75 (18-101)	13.88 (8-18)	52.46 (5-98)	43.32 (17-74)	K6	0.84 (0.66-0.90)	0.84 (0.60-0.90)
National Household Survey on Drug Abuse, 1985	1	7,976	11.22 (0-51)	10.97 (0-17)	--	--	CES-D	0.74 (0.61-0.83)	0.74 (0.61-0.83)
National Household Survey on Drug Abuse, 1993	1	26,386	32.74 (0-336)	11.39 (0-17)	--	--	CES-D	0.77 (0.29-0.90)	0.79 (0.54-0.91)
National Household Survey on Drug Abuse, 1994	1	4,331	32.48 (0-205)	11.28 (0-17)	--	--	CES-D	0.79 (0.60-0.88)	0.79 (0.55-0.89)
National Longitudinal Study of Adolescent to Adult Health (Add Health), 1994-2008 [Public Use]	4	18,188	45.96 (0-999)	13.43 (0-22)	50.85 (1-100)	43.15 (17-87)	CES-D	0.83 (0.52-0.91)	0.83 (0.53-0.91)
National Longitudinal Survey of Youth (NLSY)	34	143,941	53.10 (0-1,070)	13.44 (1-20)	48.48 (1-100)	42.75 (17-86)	BPI, CES-D, MHI-5	0.73 (0.10-0.93)	0.74 (0.36-0.94)
National Social Life, Health and Aging Project (NSHAP): Wave 3	1	4,199	54.83 (13-101)	11.74 (0-20)	--	--	CES-D	0.82 (0.77-0.85)	0.82 (0.78-0.86)
National Social Life, Health, and Aging Project (NSHAP): Wave 1, [United States], July 2005-March 2006	1	2,913	44.68 (13-101)	10.33 (0-20)	--	--	CES-D	0.81 (0.76-0.85)	0.81 (0.77-0.86)

Dataset	# DS	N	M_{Inc} (range)	M_{Edu} (range)	M_{Status} (range)	$M_{Prestige}$ (range)	DS Measure	$\bar{\alpha}$ (range)	$\bar{\omega}$ (range)
National Social Life, Health, and Aging Project (NSHAP): Wave 2 and Partner Data Collection, [United States], 2010-2011	1	3,107	48.99 (13-101)	10.98 (0-20)	--	--	CES-D	0.79 (0.55-0.89)	0.82 (0.79-0.89)
National Survey of America's Families (NSAF), 1997	2	45,423	38.36 (-5-201)	13.20 (8-18)	--	--	SF-36	0.76 (0.34-0.84)	0.78 (0.65-0.85)
National Survey of America's Families (NSAF), 1999	2	43,826	50.43 (-9-321)	13.42 (8-18)	--	--	SF-36	0.79 (0.64-0.91)	0.80 (0.65-0.92)
National Survey of America's Families (NSAF), 2002	2	41,330	55.95 (-10-338)	13.34 (8-18)	--	--	SF-36	0.75 (0.44-0.86)	0.77 (0.53-0.87)
National Survey of Families and Households, Wave 1: 1987-1988, [United States]	1	12,722	23.31 (0-975)	12.40 (0-20)	51.52 (1-100)	42.70 (17-86)	CES-D	0.92 (0.86-0.95)	0.92 (0.86-0.95)
National Survey of Families and Households, Wave 2: 1992-1994, [United States]	2	15,129	49.85 (0-1,000)	12.85 (0-20)	57.59 (1-100)	45.27 (17-86)	CES-D	0.91 (0.78-0.95)	0.92 (0.80-0.95)
National Survey of Families and Households, Wave 3: 2001-2003, [United States]	2	6,734	18.34 (0-1,000)	13.31 (0-20)	62.86 (1-100)	47.56 (17-86)	CES-D	0.89 (0.82-0.94)	0.89 (0.79-0.94)
National Survey of Functional Health Status, 1990	1	2,342	--	12.78 (1-17)	--	--	SF-36	0.88 (0.81-0.93)	0.88 (0.81-0.94)
National Survey on Drug Use and Health, 2002	1	36,129	40.47 (1-75)	12.95 (5-16)	--	--	K6	0.90 (0.80-0.96)	0.90 (0.80-0.95)
National Survey on Drug Use and Health, 2004	1	37,056	39.50 (1-75)	12.90 (5-16)	--	--	K6	0.89 (0.52-0.96)	0.91 (0.84-0.96)
National Survey on Drug Use and Health, 2005	1	37,003	40.64 (1-75)	12.90 (5-16)	--	--	K6	0.92 (0.87-0.98)	0.92 (0.88-0.98)
National Survey on Drug Use and Health, 2006	1	36,712	41.36 (1-75)	12.91 (5-16)	--	--	K6	0.91 (0.83-0.97)	0.91 (0.79-0.97)
National Survey on Drug Use and Health, 2007	1	37,449	42.28 (1-75)	12.95 (5-16)	--	--	K6	0.91 (0.82-0.95)	0.91 (0.85-0.95)
National Survey on Drug Use and Health, 2008	1	37,371	42.63 (1-75)	12.92 (5-16)	--	--	K6	0.85 (0.68-0.93)	0.86 (0.70-0.94)
National Survey on Drug Use and Health, 2009	1	37,599	42.31 (1-75)	13.01 (5-16)	--	--	K6	0.85 (0.69-0.90)	0.86 (0.75-0.91)
National Survey on Drug Use and Health, 2010	1	38,796	41.82 (1-75)	13.06 (5-16)	--	--	K6	0.86 (0.72-0.95)	0.86 (0.80-0.95)
National Survey on Drug Use and Health, 2011	1	39,003	41.4 (1-75)	13.07 (5-16)	--	--	K6	0.84 (0.59-0.91)	0.85 (0.66-0.91)

Dataset	# DS	N	M_{Inc} (range)	M_{Edu} (range)	M_{Status} (range)	$M_{Prestige}$ (range)	DS Measure	$\bar{\alpha}$ (range)	$\bar{\omega}$ (range)
National Survey on Drug Use and Health, 2012	1	37,763	41.78 (1-75)	13.09 (5-16)	--	--	K6	0.86 (0.64-0.93)	0.86 (0.66-0.93)
National Survey on Drug Use and Health, 2013	1	37,263	42.34 (1-75)	13.13 (5-16)	--	--	K6	0.87 (0.75-0.93)	0.87 (0.75-0.93)
National Survey on Drug Use and Health, 2014	1	41,453	45.48 (1-75)	13.32 (5-16)	--	--	K6	0.85 (0.47-0.91)	0.86 (0.75-0.91)
New Family Structures Study	1	2,787	56.11 (5-200)	13.87 (0-20)	--	--	CES-D	0.86 (0.92-0.91)	0.87 (0.82-0.90)
Physical Violence in American Families, 1985	1	5,163	26.93 (0-51)	13.14 (5-16)	61.08 (3-100)	48.10 (19-86)	PERID	0.79 (0.65-0.90)	0.83 (0.76-0.95)
Religion, Aging, and Health Survey, 2001, 2004 [United States]	2	2,428	24.81 (5-80)	11.40 (1-25)	--	--	CES-D	0.87 (0.82-0.90)	0.87 (0.83-0.91)
Teenage Attitudes and Practices Survey, 1989: [United States]	1	8,479	32.08 (1-50)	13.20 (0-17)	--	--	--	0.71 (0.66-0.78)	0.72 (0.67-0.78)
United States National Health Measurement Study, 2005-2006	1	3,731	44.57 (10-75)	13.79 (0-20)	--	--	SF-36	0.85 (0.62-0.94)	0.86 (0.66-0.94)
Violence and Threats of Violence Against Women and Men in the United States, 1994-1996	1	15,488	41.09 (0-100)	13.58 (0-18)	--	--	SF-36	0.75 (0.55-0.89)	0.74 (0.44-0.87)

Note. # DS = number of participant-level (IPD) files extracted from each dataset; N = sample size; M_{Inc} = mean income reported in thousands; M_{Edu} = mean years of education; M_{Status} = mean occupational status; $M_{Prestige}$ = mean occupational prestige; DS = depressive symptoms; $\bar{\alpha}$ = average alpha coefficient; $\bar{\omega}$ = average omega coefficient; range = range of values across participants or datasets; BPI = Behavior Problem Index; CES-D = Center for Epidemiologic Studies Depression Scale; SF-36 = Short Form-36; K6 = K6+ Self-Reporting Measure; PERID = Psychiatric Epidemiology Research Interview Demoralization Scale; PHQ-9 = Public Health Questionnaire-5; MHI-5 = Mental Health Inventory-5; Double dash (--) indicates missing data.