



# The effects of U.S. county and state income inequality on self-reported happiness and health are equivalent to zero

Nicolas Sommet<sup>1</sup> · Andrew J. Elliot<sup>2</sup>

Accepted: 26 March 2022 / Published online: 28 April 2022  
© The Author(s) 2022

## Abstract

**Purpose** A popular idea in the social sciences is that contexts with high income inequality undermine people's well-being and health. However, existing studies documenting this phenomenon typically compare a small number of higher-level units (countries/regions). Here, we use local income inequality indicators and temporal designs to provide the most highly powered test to date of the associations between income inequality and self-reported happiness and health in the USA

**Method** We combined county-level income inequality data (county-level Gini coefficients) with the responses from the General Social Survey (GSS) Cross-sectional dataset (13,000+ participants from  $\approx 1000$  county-waves) and Panels ( $3 \times 3000$ + participants from  $3 \times \approx 500$  county-waves); we used the GSS happiness ("not too happy," "pretty happy," or "very happy") and health ("poor," "fair," "good," or "excellent") variables.

**Results** Multilevel-ordered logistic models and equivalence tests revealed that the within-county effects of income inequality on self-reported happiness and health were systematically equivalent to zero. Additional analyses revealed that the within-state effects were identical, that using alternative measures of state income inequality led to the same conclusions, and that lagged effects (between + 1 and + 12 years) were never significant and always equivalent to zero.

**Conclusion** The present work suggests that—at least in the USA—income inequality is likely neither associated with self-reported happiness nor with self-reported health.

**Keywords** Income inequality · Well-being · Health · Multilevel modeling · Equivalence testing

During the past few decades, the GDP per capita reached record levels in most developed economies [1]. However, economic growth has not benefited everyone equally. For instance, since the 1990s, the top 1% of American earners have captured half of the total real income growth and they have nearly doubled their income; over the same period, the real income of the bottom 99% grew by only 15% [2].

In *The Spirit Level* and, more recently, *The Inner Level*, Wilkinson and Pickett [3, 4] popularized the idea that contexts with high income inequality create status anxiety, which is not only purported to exact a cost in terms of well-being, but also to cause wear and tear on the body and erode

health. This idea has become widely accepted, and articles from high-impact journals regularly cite one or both of these books to describe the deleterious consequences of income inequality as a basic and established empirical fact (e.g., refs. [5–7]).

However, both *The Spirit Level* and *The Inner Level* elicited various methodological criticisms [8–11]. Arguably, the most important of these criticisms pertains to the use of small higher-level sample size: For instance, in *The Spirit Level*, the evidence of a link between income inequality and well-being-related or health-related outcomes is based on a series of bivariate correlations that never involve more than  $K = 50$  higher-level units of analysis (e.g., comparing  $K = 12$  countries for mental illness [p. 67] or comparing  $K = 50$  U.S. states for life expectancy [p. 83]). With such small higher-level sample sizes, the statistical power to detect a small-to-medium bivariate correlation of  $r = .20$  (with  $\alpha = .05$ ) is no more than 29% [12], meaning that the outcome of Wilkinson and Pickett's analyses is particularly sensitive to model specification, and that congruent false-positive findings are

✉ Nicolas Sommet  
nicolas.sommet@unil.ch

<sup>1</sup> Centre LIVES, University of Lausanne, Bâtiment Géopolis, Bureau #5785, Quartier UNIL-Mouline, Lausanne, Switzerland

<sup>2</sup> Department of Psychology, University of Rochester, Rochester, NY, USA

more likely to be observed and selected for publication (for research on power and type I error rates, see ref. [13]).

Importantly, this small- $K$  problem is widespread in the literature. The most influential existing studies compared only a small number of higher-level units when documenting negative effects<sup>1</sup> of national income inequality on well-being or health (e.g., Alesina et al. [14], Study 1,  $K = 13$  units; Lynch et al. [15],  $K = 22$  units; Oishi et al. [16],  $K = 27$  units [cumulated citations:  $\approx 4,000$ ; refs. [14–16]]). At the same time, several (less influential) studies comparing much larger numbers of higher-level units failed to replicate these findings, even when controlling for GDP or distinguishing developed from developing countries (e.g., Berg and Veenhoven [17],  $K = 119$  units; Jen et al. [18],  $K = 116$  units; Rözer and Karaka [19],  $K = 85$  units [cumulated citations:  $\approx 400$ ; refs. [17–19]; for a critical discussion on the limitation of such cross-national comparisons, see ref. [20]).

Equally important, existing reviews and meta-analyses do not solve the small- $K$  problem in the literature, as their conclusions are at least partially based on extant small- $K$  studies. For instance, an influential meta-analysis documented a small-sized negative association between income inequality and health [21], but neither reported the number of higher-level units in the meta-analytic sample, nor differentiated findings from small- $K$  studies from findings from large- $K$  studies (for a meta-analysis documenting a null association between income inequality and well-being, see ref. [22]). As another example, an influential literature review used the (disputed) “vote-counting procedure” to conclude that income inequality undermines health [23], but also neither mentioned the issue of higher-level sample size, nor differentiated small- $K$  from large- $K$  studies (for literature reviews focused on mental health and well-being but having similar shortcomings, see refs. [24, 25], respectively).

## Local-level and temporal design studies as a remedy to the small- $K$ problem

Beyond cross-national comparisons and meta-analyses, we believe that the gold standard for estimating the effects of income inequality is to combine the use of (1) local-level rather than broader-level income inequality indicators (e.g., U.S. county-based rather than national Gini coefficients) with (2) temporal rather than non-repeated cross-sectional designs (i.e., spanning multiple years rather than a single point in time).

Multilevel studies using this approach have two main advantages. First, they have better ecological and predictive validity. The reason is that people are rather accurate in estimating variation in local-level income inequality [26], whereas they tend to misperceive the mean level of broader-level income inequality [27]. This phenomenon is likely explained by “bounded rationality” (herein, the incapacity of the human brain to collect and process countless pieces of economic information at broad levels of geographic aggregation, such as the state or national level [28]).

Second, and perhaps more importantly, multilevel studies using local-level indicators and temporal designs solve the small- $K$  problem. The reason is that they make it possible to compare a large number of higher-level units over time (e.g., hundreds of U.S. county-years such as Queens County-2006, Queens County-2008, ..., Fairfax County-2008, etc.), thereby increasing the sample size at the level the effect is measured. Generally speaking, the number of higher-level units ( $K$ ), such as county-years, is much more critical than the number of lower-level units ( $N$ ), such as participants, to achieve adequate statistical power to detect the effect of a higher-level variable, such as county income inequality over time [29], and to produce reliable and replicable effects [30].

The few existing multilevel studies using local-level indicators and temporal designs paint a very different picture than that offered by the prevailing narrative in the literature. On one hand, some of these studies document small deleterious within-cluster effects of local income inequality over time on well-being or health (e.g., see ref. [31], Study 1,  $K = 424$  German Länder-years; ref. [32],  $K = 216$  Brazilian state-years). On the other hand, many of the existing studies yield null effects of within-cluster local income inequality and are therefore inconclusive (e.g., ref. [33],  $K = 156$  South African district-years; ref. [34],  $K = 576$  Chinese county-years; ref. [35],  $K = 70$  Canadian province-years; ref. [36],  $K = 117$  Brazilian village-years).

Interestingly, there are virtually no published studies testing the within-cluster effects of local income inequality over time on well-being or health in the USA (for one exception using data from the 1970s, see ref. [37]). This is unfortunate for two reasons. First, the evidence of a deleterious influence of income inequality inferred from U.S. data is often considered to be more compelling than the evidence inferred from data from other countries (e.g., see ref. [38]). For instance, in response to the criticism that cross-national survey data fail to support a negative inequality-happiness association, Wilkinson and Pickett [39] dismissed international comparison as unreliable and argued that “in sub-national analyses, more equal societies, for example more equal U.S. states, are happier.” (p. 7). Second, small- $K$  studies using U.S. data are often used to formulate policy recommendations (e.g., see ref. [40]). For instance, Oishi et al. [41] reported analyses involving only  $K = 28$  years of U.S. data and concluded

<sup>1</sup> In this article, the word “effect” should be understood as a technical term that refers to the cluster-level fixed slope of income inequality in a multilevel model. It is not used to indicate a causal impact of income inequality.

that “progressive taxation could be an important policy tool [...] in combating the possible negative effects of inequality” (p. 9).

## Overview of the study and research questions

The present study combines U.S. local economic data with responses from the longest-running U.S. sociological survey, namely, the General Social Survey (GSS). Specifically, we combined county-level economic data (the most local level of geographic aggregation available) with the GSS Cross-sectional dataset (13,000+ participants from  $\approx 1000$  county-waves) and the three GSS Panels ( $3 \times 3000$ + participants from  $3 \times \approx 500$  county-waves) to answer the following two research questions:

*RQ1.* What is the within-county effect of income inequality over time on well-being (measured via self-reported happiness)?

*RQ2.* What is the within-county effect of income inequality over time on health (measured via self-reported health)?

To test these research questions, we used advanced analytical techniques, integrating multilevel modeling with equivalence testing [42, 43]. Such a novel analytical approach enabled us to reverse the null and alternative hypotheses, so that the burden of proof rested in proving equivalence. This not only enabled us to test whether the within-county effects of income inequality on self-reported happiness and health are different from zero, but also whether they are equivalent to zero.

## Method

### Institutional review board statement

This study received approval from the Ethics Committee from the region of the country of the first author (Req-2018-00113).

### Open science statement

All data exclusions and variables analyzed are reported. Raw data (economic raw data and instructions to retrieve the GSS data) and syntax files (Stata.do files) are available via the OSF (<https://osf.io/b9frp/>).

## Sample and procedure

Table 1 presents the sample demographic characteristics.

We pooled the responses from the GSS, a U.S. sociological survey that uses a stratified multi-stage probability sampling method to achieve national representativeness. We used both (i) the GSS Cross-sectional dataset, which is a cumulative cross-sectional repeated sample covering the years between 1972 and 2016 (response rate in 2016: 61.30%), and (ii) the GSS Panels, which consist of three longitudinal samples spanning 2006–08–10 (Panel 1), 2008–10–12 (Panel 2), and 2010–12–14 (Panel 3; attrition rate in 2010–12–14: 24.12%). We obtained the “GSS sensitive data files” from the National Opinion Research Center to identify participants’ county of residence (i.e., the most local level of geographic aggregation for which the annual economic estimates provided by the U.S. Census Bureau were available).

Regarding the GSS Cross-sectional dataset, given that annual county income inequality estimates were not available prior to 2006, we focused on the last six waves of assessment (2006–08–10–12–14–16). The final GSS Cross-sectional dataset comprised 13,266 participants from 279 counties (i.e., 1,088 county-wave income inequality estimations). The sample size was sufficient to detect an effect of income inequality with the smallest effect size of interest ( $\beta = .05$ ) with a power higher than .99 (for the sensitivity analysis, see *Supplementary Materials*, p. 2).

Regarding the GSS Panels, given the focus on longitudinal effects, we retained (i) participants with nonmissing values for at least two waves and (ii) cases pertaining to participants who did *not* move out of their county over the course of the study (93.56% of the samples). Panel 1 comprised 3242 cases nested in 1168 participants from 145 counties (431 county-waves), Panel 2 comprised 3444 cases nested in 1247 participants and 165 counties (490 county-waves), and Panel 3 comprised 3301 cases nested in 1,185 participants and 168 counties (498 county-waves). The sample sizes were again sufficient to detect the smallest effect size of interest with a power of at least .81 (see *Supplementary Materials*, p. 2).

## Variables

Table 1 (right side) presents the descriptive statistics.

### County income inequality

The 1-year annual estimates of the county-level Gini coefficient from the U.S. Census Bureau were used. These estimates represent the distributions of household income for a given county in a given year. They may range from 0 (perfect

**Table 1** Description of the GSS Cross-sectional data and Panel 1–3 demographic characteristics and descriptive statistics

	Cross-sectional	Panel 1	Panel 2	Panel 3
<b>Demographic characteristics</b>				
<b>Individuals</b>				
Percentage of women	55.56%	59.59%	55.09%	57.72%
Age	47.67 (17.40)	49.70 (16.70)	49.38 (16.84)	49.99 (16.66)
Percentage of White respondents	72.43%	72.26%	75.06%	75.86%
Education (number of year of school completed)	13.67 (3.14)	13.82 (3.05)	13.72 (3.04)	13.93 (3.02)
Annual household Income (constant 1986 USD, thousands)	34.89 (34.99)	36.85 (34.45)	35.75 (35.31)	36.64 (37.20)
Percent of workers (part- or full-time)	59.76%	59.38%	61.48%	58.41%
<b>Counties</b>				
Number of inhabitants (in hundreds of thousands)	7.83 (0.42)	8.41 (0.23)	8.06 (0.23)	8.16 (0.25)
Poverty headcount ratio	13.24% (1.54)	12.91% (1.46)	13.78% (1.69)	14.52% (1.05)
Unemployment rate	7.57% (2.14)	7.82% (2.37%)	8.87% (2.16)	9.20% (1.77)
Median annual income (USD, thousands)	59.18 (0.40)	56.57 (2.64)	57.57 (2.20)	58.19 (2.63)
Percentage of poorly educated (below 9th grade)	5.31 (0.62)	5.76 (0.67)	5.51 (0.62)	5.27 (0.63)
<b>Descriptive statistics</b>				
County income inequality (Gini coefficient)	.45 (.01)	.45 (.01)	.45 (.01)	.46 (.01)
Self-reported happiness: # of responses (% of missing values)	12,008 (0.10%)	3,231 (0.36%)	3,430 (0.38%)	3,296 (0.15%)
Percent of nonmissing values falling in each category				
Very happy	29.28%	30.05%	27.35%	27.00%
Pretty happy	56.18%	57.01%	57.06%	58.83%
Not too happy	14.54%	12.94%	15.60%	14.17%
Self-reported health: # of responses (% of missing value)	9,192 (0.26%)	2,190 (0.14%)	2,363 (0.12%)	2,065 (0.08%)
Percent of nonmissing values falling in each category				
Excellent	26.29%	27.21%	25.05%	24.36%
Good	47.38%	47.03%	47.95%	48.67%
Fair	20.69%	20.73%	21.58%	22.03%
Poor	5.64%	5.02%	5.42%	4.94%

*SDs* are given in parentheses; *SDs* for the county-level variables are within-county *SDs*; “% of missing values” correspond to the percentage of “no answer” or “I don’t know”

equality: Each household in the county has an equal share of income) to 1 (perfect inequality: Only one household in the county has *all* of the income).

### Self-reported happiness

The GSS variable “General happiness” was used. Participants reported whether they were “not too happy” (coded “1”), “pretty happy” (coded “2”), or “very happy” (coded “3”). For a study demonstrating the validity (acceptable temporal satiability and concurrent validity) of this single-item measure, see ref. [44].

### Self-reported health

The GSS variable “Condition of health” was used. Participants reported whether their health was “poor” (coded “1”), “fair” (coded “2”), “good” (coded “3”), or “excellent” (coded “4”). For a study demonstrating the validity

(acceptable temporal satiability and predictive validity) of this single-item measure, see refs [45, 46].

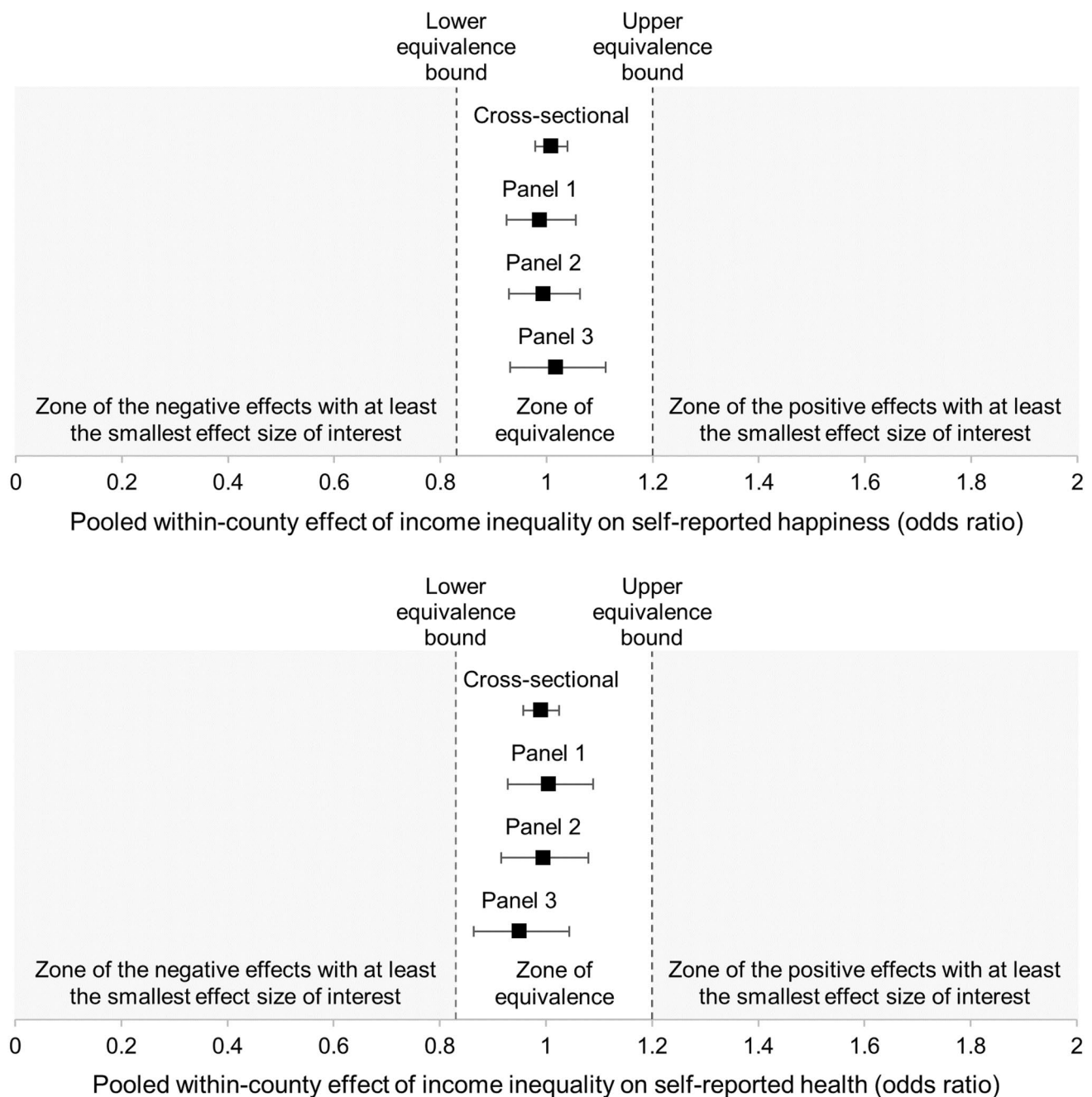
## Results

Figure 1 presents the main findings, and Tables S1–S2 presents the full set of results and the multilevel regression equations.

### Main analyses: multilevel modeling

#### Overview of the analytical strategy

Given the hierarchical structure of the data, we built a series of multilevel models in which participants were nested in counties (for the GSS Cross-sectional dataset) or in which wave-specific observations were nested in participants, themselves nested in counties (for the GSS Panels).



**Fig. 1** Equivalence Tests. Pooled within-county effects of income inequality on self-reported happiness (upper panel) and health (lower panel) in Panels 1–3. *Notes:* Error bars are 90% CIs; the fact that the

90% CIs fall within the lower and upper equivalence bounds means that the effects of income inequality are equivalent to zero

For each of the samples, we used standardized self-reported happiness and health as the focal outcomes (in two separate multilevel models), and we used standardized county income inequality as our focal predictor. Moreover, we controlled for year fixed effects (wave dummies) rather than using county-year as a level (as adding another level of analysis was too computationally demanding).

Our multilevel models had two distinctive features. First, given that both of our outcomes were ordinal, we used *ordered* logistic (rather than linear) modeling. Note that approximate likelihood-ratio tests of proportionality of odds from naïve one-level models showed that the proportional odds assumption of the ordered logistic regression was neither violated in the GSS Cross-sectional dataset ( $p \geq .358$ ) nor the GSS Panels



( $ps \geq .310$ ). Second, given that we aimed to estimate the effects of income inequality over time, we used county-specific mean centering [47]. Simply put, we subtracted the county-specific mean of income inequality from each wave-specific income inequality estimate so as to obtain an unambiguous estimation (uncontaminated by between-county variation) of the *pooled within-county effect of income inequality over time*.

### The effects of local income inequality on self-reported happiness and health are not different from zero

Multilevel analyses led to two sets of findings. First, addressing RQ1, when county income inequality increased by +1 SD, self-reported happiness was found to change by +0.87%,  $p = .641$ , in the cross-sectional dataset, and by −1.27%,  $p = .753$  (Panel 1), −0.58,  $p = .886$  (Panel 2), and +1.75%,  $p = .702$  (Panel 3) in the longitudinal datasets. Second, addressing RQ2, when county income inequality increased by +1 SD, self-reported health was found to change by −0.98%,  $p = .638$ , in the cross-sectional dataset, and by +0.53%,  $p = .913$  (Panel 1), +0.55%,  $p = .913$  (Panel 2), and +5.32%,  $p = .336$  (Panel 3) in the longitudinal datasets. Taken together, these findings suggest that—across all samples—the pooled within-county effects of income inequality over time were *not different from zero*.

### Follow-up analyses: equivalence testing

#### Overview of the analytical strategy

Because absence of evidence is not evidence of absence, we performed equivalence tests to determine if the within-county effects of income inequality on self-reported happiness and health were indeed absent, that is, if these effects were statistically smaller than the smallest negative effect of interest, as well as than the smallest positive effect of interest [42].

Before conducting these equivalence tests, we needed to define the smallest effect size of interest. A correlation of  $r = |0.05|$  is usually considered a trivially small effect size, which corresponds to a negative effect of OR = 0.83 and a positive effect of OR = 1.20 [48]. For each of the samples and each of our two outcomes, we compared the odds ratio associated with the standardized effect of county income inequality to OR = 0.83 (our lower equivalence bound) and OR = 1.20 (our higher equivalence bound) using a one-sided postestimation Wald test.

### The effects of local income inequality on self-reported happiness and health are equivalent to zero

Equivalence tests led to two sets of findings. First, addressing RQ1, the within-county effect of income inequality

on self-reported happiness was found to fall within the equivalence bounds in the cross-sectional dataset,  $\chi^2s$  (1,  $N = 12,008$ )  $\geq 86.49$ ,  $ps \leq .001$ , and the longitudinal datasets,  $\chi^2s$  (1,  $N = 3231$ )  $\geq 17.57$ ,  $ps \leq .001$  (Panel 1),  $\chi^2s$  (1,  $N = 3429$ )  $\geq 18.28$ ,  $ps \leq .001$  (Panel 2), and  $\chi^2s$  (1,  $N = 3296$ )  $\geq 13.14$ ,  $ps \leq .001$  (Panel 3). Second, addressing RQ2, the within-county effect of income inequality on self-reported health was also found to fall within the equivalence bounds in the cross-sectional dataset,  $\chi^2s$  (1,  $N = 9192$ )  $\geq 68.90$ ,  $ps \leq .001$ , and the longitudinal datasets,  $\chi^2s$  (1,  $N = 2190$ )  $\geq 13.19$ ,  $ps \leq .001$  (Panel 1),  $\chi^2s$  (1,  $N = 2362$ )  $\geq 12.32$ ,  $ps \leq .001$  (Panel 2), and  $\chi^2s$  (1,  $N = 2065$ )  $\geq 5.11$ ,  $ps \leq .012$  (Panel 3). Taken together, these findings suggest that—across all samples—the pooled within-county effects of income inequality over time were *equivalent to zero*.

### Repeating the multilevel analyses and equivalence tests while including control variables

We repeated the analyses while controlling for the same a priori-selected participant-based sociodemographic variables (sex, age, race, education, income, work status) and county-based contextual variables (population, poverty, unemployment, median income, level of education) used in extant research [49, 50]. Out of the 24 individual focal effects, 23 remained similar (the effect of county income inequality in Panel 3 on self-reported health was no longer significantly different from the lower equivalence bound; for the full results, see *Supplementary Materials*, p. 6 and Tables S3–S4).

### Additional analyses: state income inequality and self-reported happiness and health

An argument that could be brought up to explain the absence of effects of income inequality within counties is that income inequality measured at this level of geographic aggregation poorly reflect stratification in society [23, 51]. This argument is regularly used to account for the fact that studies measuring income inequality in small areas do not or weakly predict well-being or health, whereas studies measuring income inequality in larger areas such as states seemingly have more predictive strength [52]. However, we believe that a more parsimonious explanation of these differences lays in the fact that studies measuring income inequality in small areas have larger  $Ks$ , are more powered, and give more reliable estimates (close to zero), whereas studies measuring income inequality at the state level have smaller  $Ks$ , are less powered, and give more biased estimates (in one direction or the other). Because the outcome of small- $K$  studies are more affected by flexible statistical practices, they tend to be over-represented in the pool of false-positives and “contaminate”

the literature [13]. In additional analyses, we took advantage of the fact that the GSS datasets cover many U.S. states over several years (resulting in reasonably large- $K$  samples) to test the effects of state income inequality on self-reported happiness and health.

We used the 30 waves (1973–2016) of the GSS Cross-sectional dataset for which participants' state of residence was provided. The sample comprised 59,220 participants from 50 states + Washington D.C. (1131 state-waves income inequality estimations), yielding power higher than .99 to detect the smallest effect size of interest). In parallel, we used the three GSS Panels (focusing on participants with nonmissing values for at least two waves and who did not move out of their state over the course of the study). The smallest sample (Panel 1) comprised 4155 cases nested in 1477 participants and 40 states (120 state-waves, yielding power of .77 to detect the smallest effect size of interest).

We replicated our main analysis while substituting county-level Gini coefficients from the U.S. Census Bureau with state-level Gini coefficient from the Frank-Sommeiller-Price (FSP) Series [53] as the focal predictor (for the description of the variable and the rationale for using it, see *Supplementary Materials*, p. 9). Figure 2 presents the main findings, and Tables S5–S6 present the full set of results and the multilevel regression equations.

First, addressing RQ1, when state income inequality increased by +1 SD, self-reported happiness was found to change by +0.66%,  $p = .804$ , in the cross-sectional dataset, by -5.64%,  $p = .128$ , in Panel 1, by +2.80%,  $p = .464$ , in Panel 2, and by -8.98%,  $p = .205$  in Panel 3. These estimates were found to fall within the equivalence bounds in the cross-sectional dataset ( $ps \leq .001$ ), Panel 1 ( $ps \leq .001$ ), and Panel 2 ( $ps \leq .001$ ), but not in Panel 3 (where the estimate was not different from the lower equivalence bound,  $p = .080$ ).

Second, addressing RQ2, when state income inequality increased by +1 SD self-reported health was found to change by +0.07%,  $p = .982$ , in the cross-sectional dataset, by -3.55%,  $p = .430$ , in Panel 1, by +5.22,  $p = .256$  in Panel 2, and by -6.36%,  $p = .458$ , in Panel 3. These estimates were found to fall within the equivalence bounds in the cross-sectional dataset ( $ps \leq .001$ ), Panel 1 ( $ps \leq .001$ ), and Panel 2 ( $ps \leq .001$ ), but not in Panel 3 (where the estimate was not different from the lower equivalence bound,  $p = .086$ ).

In addition, we took advantage of the fact that the FSP Series includes three other types of income inequality estimates besides the Gini coefficient (Theil's entropy index, Atkinson's index, and relative mean deviation) and covers a large number of years (1917-present) to: (i) replicate the findings using alternative measures of state income inequality (more than 90% of the findings were similar) and (ii) test for lagged effects of state income inequality on self-reported happiness and health (ranging from 1 to 12 years;

see ref. [54]) in the GSS Cross-sectional data (0% of the lagged effects were significant and 100% were equivalent to zero; for the description of the findings, see *Supplementary Materials*, pp. 11–14, including Figures S1–S2).

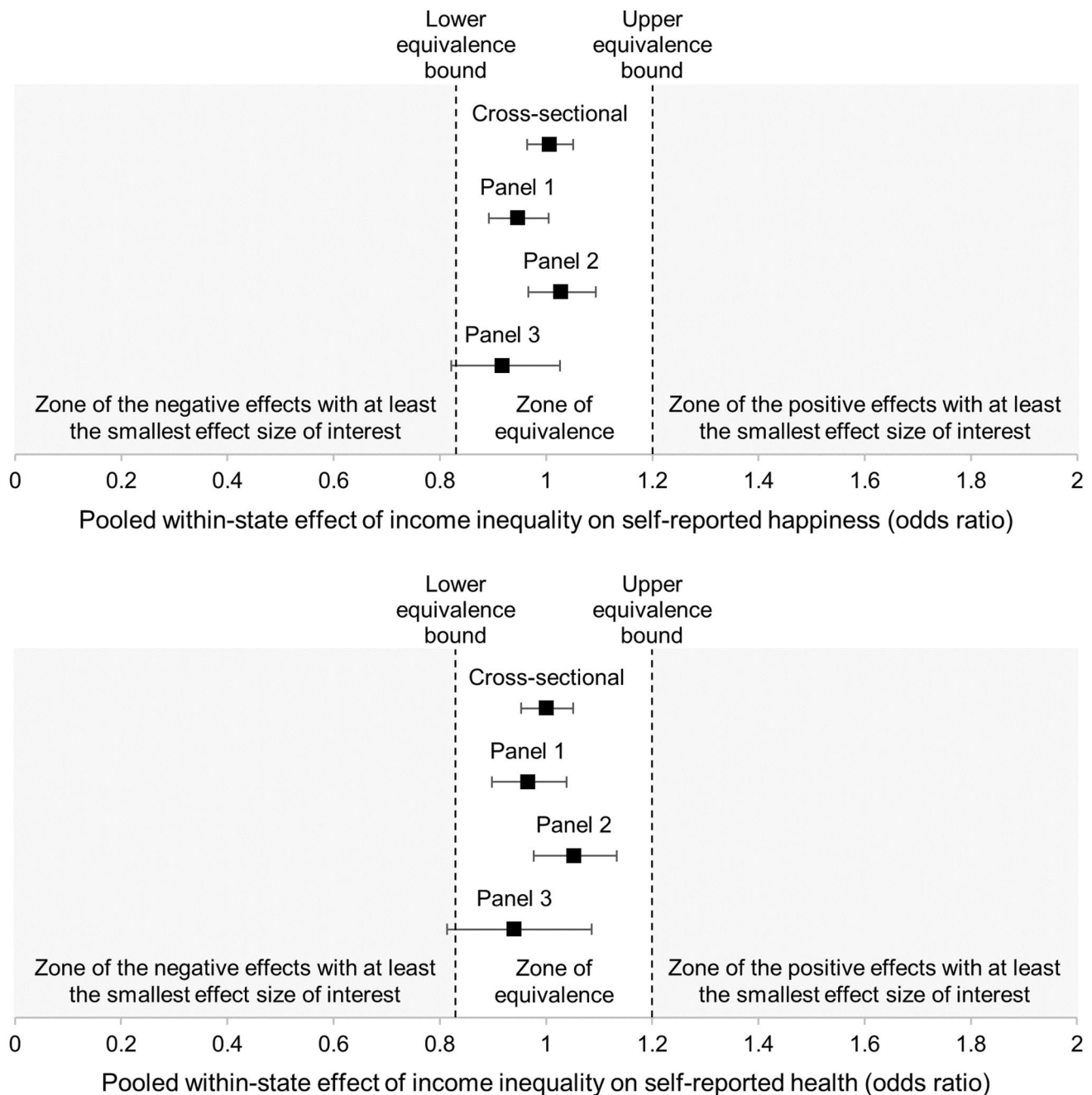
## Discussion

The idea that income inequality is a powerful societal determinant of well-being and health is commonly presented as a well-established fact (for narrative reviews, see refs. [23, 24, 55]), even though the extant foundational research used small- $K$  samples [3, 16, 56]. In the present research, we used local income inequality indicators *and* repeated cross-sectional or longitudinal designs (rather than single-point designs) to produce the most highly powered test to date of the association between income inequality and well-being or health in the USA.

In contrast with the scientific Zeitgeist of the past two decades, we not only documented that the effects of income inequality over time on self-reported happiness and health were systematically *not significantly different from zero* (using regular null hypothesis testing), but also found that these effects were systematically *equivalent to zero* (using equivalence testing to shift the burden of proof). Moreover, the findings were highly similar when substituting county with state income inequality, using alternative measures of state income inequality, and testing lagged effects of state income inequality.

Reducing economic inequality is the tenth of seventeen Sustainable Development Goals set by the United Nations to be achieved by 2030 [57]. The belief that income inequality impairs well-being or health is often used to argue that achieving this Sustainable Development Goal is particularly urgent (e.g., see ref. [58]). As an example, Heinz et al. [59] recently claimed in *JAMA Psychiatry* that the negative association between of income inequality and mental health should push policymakers to “consider different measures [to reduce] income inequality, including progressive taxation policies and basic universal income schemes” (p. E2).

In light of the present findings and limited meta-analytic evidence [21, 22, 60, 61], the relationship between income inequality and well-being or health may not be substantial or robust enough to serve as a compelling argument supporting policies aimed at reducing income inequality. Obviously, it is possible that future research using a longitudinal large- $K$  sample will document an effect of income inequality on some other type of health-relevant outcome (for relevant research using a U.S. sample spanning 2001 to 2014 that documents a null relationship between local income inequality and life expectancy, see ref. [62]). In the meantime, however, we believe that in the U.S., legislators and political organizations desiring to address the structural antecedents



**Fig. 2** Equivalence Tests. Pooled within-state effects of income inequality on self-reported happiness (upper panel) and health (lower panel) in Panels 1–3. *Notes:* Error bars are 90% CIs; the fact that the

90% CIs fall within the lower and upper equivalence bounds means that the effects of income inequality are equivalent to zero

of population health would probably be better advised to shift from the issue of income inequality to a focus on more robust determining factors such as financial scarcity [63], unemployment [64], or social isolation [65].

Obviously, the fact that U.S. data does not appear to support the idea of a link between income inequality and well-being nor health does *not* mean that income inequality is acceptable. First, the effects of income inequality

may be limited to vulnerable groups of individuals (e.g., individuals facing financial scarcity [50]) or moderated by contextual variables (e.g., neighborhood social capital [66]). Second, income inequality is likely to predict other undesirable individual outcomes besides well-being and health, such as risk-taking behaviors [5]. Third, even if income inequality is not corrosive to well-being or health, it does not follow that there is no admissible *moral*



argument for reducing income inequality. As an illustration, in contemporary discourse on income inequality, one of the most influential moral arguments contends that inequality of outcome is only acceptable under conditions of strict equality of opportunity *and* when it benefits all members of society [67].

The present work has three main limitations. First, our study is based on data from a Western, educated, industrialized, rich, and democratic (WEIRD) country [68]. It has been documented that the psychological effects of income inequality differ between developed and developing economies [22]. Research using large-*K* data from non-WEIRD countries is therefore warranted. Second, as is often the case with secondary data, our study was based on single-item outcome variables. Despite our large number of higher-level units, this may create measurement error and exert an influence on coefficient estimates. Research with multi-item instruments is therefore warranted. Third, the GSS has recently experienced a decline in response rates (as have the other sociological surveys [69]). However, the GSS still matches the population benchmarks quite well with regard to sex, age, education, race/ethnicity, and education, showing that there is no apparent problem in terms of demographic representativeness and that nonresponse bias remains limited [70].

Despite these limitations, the present paper suggests that—at least in the USA—it is likely that the overall associations between income inequality and self-reported happiness/health are equivalent to zero.

**Supplementary Information** The online version contains supplementary material available at <https://doi.org/10.1007/s11136-022-03137-8>.

**Author contributions** NS: Conceptualization, Methodology, Software, Formal analysis, Investigation, Resources, Data Curation, Writing—Original Draft; Visualization; AJE: Conceptualization, Methodology, Writing—Reviewing and Editing, Supervision.

**Funding** Open access funding provided by University of Lausanne. This work was funded by a SNSF Ambizione fellowship granted to N. Sommet (#PZ00P1\_185979); this publication is based on research conducted at the Swiss National Center of Competence in Research LIVES—Overcoming vulnerability: Life course perspectives (NCCR LIVES), which is financed by the Swiss National Science Foundation.

**Data availability** Raw data (economic raw data and instructions to retrieve the GSS data) are available via the OSF (<https://osf.io/b9frp/>).

**Code availability** Syntax files (Stata.do files to build the dataset and reproduce the findings) are also available via the OSF (<https://osf.io/b9frp/>).

## Declarations

**Conflict of interest** The authors declare no competing interests.

**Ethical approval** This study received approval from the Ethics Committee from the canton of Vaud, Switzerland (CER-VD; Req-2018-00113).

**Consent to Participate** Nonapplicable (secondary data).

**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>

## References

1. World Bank. (2021). GDP per capita (current US\$). Retrieved April 26, 2022, from <https://data.worldbank.org/indicator/NY.GDP.PCAP.CD>.
2. Saez, E. (2020). Striking it richer: The Evolution of Top Incomes in the United States (Updated with 2018 estimates). Retrieved April 26, 2022, from <https://elsa.berkeley.edu/~saez/saez-UStop incomes-2018.pdf>
3. Wilkinson, R., & Pickett, K. (2010). *The spirit level: Why equality is better for everyone*. Penguin.
4. Wilkinson, R., & Pickett, K. (2018/2020). *The inner level: How more equal societies reduce stress, restore sanity and improve everyone's well-being*. Penguin Books.
5. Payne, B. K., Brown-Iannuzzi, J. L., & Hannay, J. W. (2017). Economic inequality increases risk taking. *Proceedings of the National Academy of Science USA*, 114(18), 4643–4648.
6. Sprong, S., Jetten, J., Wang, Z., Peters, K., Mols, F., Verkuyten, M., Bastian, B., Ariyanto, A., Autin, F., & Ayub, N. (2019). “Our country needs a strong leader right now”: Economic inequality enhances the wish for a strong leader. *Psychological Science*, 30(11), 1625–1637.
7. Fournier, M. A. (2020). Dimensions of human hierarchy as determinants of health and happiness. *Current Opinion in Psychology*, 33, 110–114.
8. Caan, W. (2019). The inner level. *Journal of Public Mental Health*, 18(3), 207–208.
9. Leigh, A., Jencks, C., & Smeeding, T. M. (2009). Health and economic inequality. In B. Nolan & W. Salverda (Eds.), *The Oxford handbook of economic inequality* (pp. 384–405). Oxford University Press.
10. Snowden, C. (2010). *The spirit level delusion: fact-checking the left's new theory of everything*. Little Dice.
11. Saunders, P., & Evans, N. (2011). *Beware false prophets*. Centre for Independent Studies.
12. Faul, F., Erdfelder, E., Lang, A.-G., & Buchner, A. (2007). G\* Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods*, 39(2), 175–191.
13. Button, K. S., Ioannidis, J. P., Mokrysz, C., Nosek, B. A., Flint, J., Robinson, E. S., & Munafò, M. R. (2013). Power failure: Why small sample size undermines the reliability of neuroscience. *Nature Reviews: Neurosciences*, 14(5), 365–376.

14. Alesina, A., Di Tella, R., & MacCulloch, R. (2004). Inequality and happiness: Are Europeans and Americans different? *Journal of Public Economics*, 88(9–10), 2009–2042.
15. Lynch, J., Smith, G. D., Hillemeier, M., Shaw, M., Raghunathan, T., & Kaplan, G. (2001). Income inequality, the psychosocial environment, and health: Comparisons of wealthy nations. *The Lancet*, 358(9277), 194–200.
16. Oishi, S., Kesebir, S., & Diener, E. (2011). Income inequality and happiness. *Psychological Science*, 22(9), 1095–1100.
17. Berg, M., & Veenhoven, R. (2010). Income inequality and happiness in 119 nations. In B. Greve (Ed.), *Social policy and happiness in Europe* (pp. 174–194). Edward Elgar.
18. Jen, M. H., Jones, K., & Johnston, R. (2009). Global variations in health: Evaluating Wilkinson's income inequality hypothesis using the World Values Survey. *Social Science & Medicine*, 68(4), 643–653.
19. Rözer, J., & Kraaykamp, G. (2013). Income inequality and subjective well-being: A cross-national study on the conditional effects of individual and national characteristics. *Social Indicators Research*, 113(3), 1009–1023.
20. Dorling, D., & Barford, A. (2009). The inequality hypothesis: Thesis, antithesis, and a synthesis? *Health & Place*, 15(4), 1166–1169. discussion 1163.
21. Kondo, N., Sembajwe, G., Kawachi, I., Van Dam, R. M., Subramanian, S., & Yamagata, Z. (2009). Income inequality, mortality, and self rated health: meta-analysis of multilevel studies. *BMJ*, 339, b4471.
22. Ngamaba, K. H., Panagioti, M., & Armitage, C. J. (2018). Income inequality and subjective well-being: A systematic review and meta-analysis. *Quality of Life Research*, 27(3), 577–596.
23. Pickett, K. E., & Wilkinson, R. G. (2015). Income inequality and health: A causal review. *Social Science & Medicine*, 128, 316–326.
24. Buttrick, N. R., Heintzelman, S. J., & Oishi, S. (2017). Inequality and well-being. *Current Opinion in Psychology*, 18, 15–20.
25. Tibber, M. S., Walji, F., Kirkbride, J. B., & Huddy, V. (2021). The association between income inequality and adult mental health at the subnational level—A systematic review. *Social Psychiatry and Psychiatric Epidemiology*, 57, 1–24.
26. Johnston, C. D., & Newman, B. J. (2016). Economic inequality and US public policy mood across space and time. *American Politics Research*, 44(1), 164–191.
27. Gimpelson, V., & Treisman, D. (2018). Misperceiving inequality. *Economics & Politics*, 30(1), 27–54.
28. Schneider, S. M. (2016). Income inequality and subjective well-being: Trends, challenges, and research directions. *Journal of Happiness Studies*, 17(4), 1719–1739.
29. Arend, M. G., & Schäfer, T. (2019). Statistical power in two-level models: A tutorial based on Monte Carlo simulation. *Psychological Methods*, 24(1), 1–19.
30. Altmejd, A., Dreber, A., Forsell, E., Huber, J., Imai, T., Johannesson, M., Kirchler, M., Nave, G., & Camerer, C. (2019). Predicting the replicability of social science lab experiments. *PLoS ONE*, 14(12), e0225826.
31. Cheung, F. (2018). Income redistribution predicts greater life satisfaction across individual, national, and cultural characteristics. *Journal of Personality and Social Psychology*, 115(5), 867–882.
32. Vincens, N., & Stafström, M. (2015). Income inequality, economic growth and stroke mortality in Brazil: Longitudinal and regional analysis 2002–2009. *PLoS ONE*, 10(9), e0137332.
33. Adjaye-Gbewonyo, K., Avendano, M., Subramanian, S., & Kawachi, I. (2016). Income inequality and depressive symptoms in South Africa: A longitudinal analysis of the National Income Dynamics Study. *Health & Place*, 42, 37–46.
34. Bakkeli, N. Z. (2016). Income inequality and health in China: A panel data analysis. *Social Science & Medicine*, 157, 39–47.
35. Latif, E. (2015). Income inequality and health: Panel data evidence from Canada. *The BE Journal of Economic Analysis & Policy*, 15(2), 927–959.
36. Undurraga, E. A., Nica, V., Zhang, R., Mensah, I. C., & Godoy, R. A. (2016). Individual health and the visibility of village economic inequality: Longitudinal evidence from native Amazonians in Bolivia. *Economics & Human Biology*, 23, 18–26.
37. Fiscella, K., & Franks, P. (1997). Poverty or income inequality as predictor of mortality: Longitudinal cohort study. *BMJ*, 314(7096), 1724–1727.
38. Lynch, J., Smith, G. D., Harper, S. A., Hillemeier, M., Ross, N., Kaplan, G. A., & Wolfson, M. (2004). Is income inequality a determinant of population health? Part 1. A systematic review. *Milbank Quarterly*, 82(1), 5–99.
39. Wilkinson, R., & Pickett, K. (2010). *Professors Richard Wilkinson and Kate Pickett, authors of The Spirit Level, reply to critics*. Equality Trust.
40. Institute for Policy Studies. (2019). FACTS. Inequality and Health. Retrieved April 26, 2022, from <https://inequality.org/facts/inequality-and-health/>
41. Oishi, S., Kushlev, K., & Schimmack, U. (2018). Progressive taxation, income inequality, and happiness. *American Psychologist*, 73(2), 157–168.
42. Quertemont, E. (2011). How to statistically show the absence of an effect. *Psychologica Belgica*, 51(2), 109–127.
43. Lakens, D., Scheel, A. M., & Isager, P. M. (2018). Equivalence testing for psychological research: A tutorial. *Advances in Methods and Practices in Psychological Science*, 1(2), 259–269.
44. Abdel-Khalek, A. M. (2006). Measuring happiness with a single-item scale. *Social Behavior and Personality*, 34(2), 139–150.
45. Burström, B., & Fredlund, P. (2001). Self rated health: Is it as good a predictor of subsequent mortality among adults in lower as well as in higher social classes? *Journal of Epidemiology & Community Health*, 55(11), 836–840.
46. Zajacova, A., & Dowd, J. B. (2011). Reliability of self-rated health in US adults. *American Journal of Epidemiology*, 174(8), 977–983.
47. Enders, C. K., & Tofighi, D. (2007). Centering predictor variables in cross-sectional multilevel models: A new look at an old issue. *Psychological Methods*, 12(2), 121–138.
48. Chen, H., Cohen, P., & Chen, S. (2010). How big is a big odds ratio? Interpreting the magnitudes of odds ratios in epidemiological studies. *Communications in Statistics Simulation and Computation*, 39(4), 860–864.
49. Sommet, N., Weissman, D., & Elliot, A. J. (2021). Income inequality predicts competitiveness and cooperativeness at school. *Journal of Educational Psychology*. <https://doi.org/10.1037/edu0000731>
50. Sommet, N., Morselli, D., & Spini, D. (2018). Income inequality affects the psychological health of only the people facing scarcity. *Psychological Science*, 29(12), 1911–1921.
51. Wilkinson, R. G., & Pickett, K. E. (2006). Income inequality and population health: A review and explanation of the evidence. *Social Science & Medicine*, 62(7), 1768–1784.
52. Kondo, N., van Dam, R. M., Sembajwe, G., Subramanian, S., Kawachi, I., & Yamagata, Z. (2012). Income inequality and health: The role of population size, inequality threshold, period effects and lag effects. *Journal of Epidemiology and Community Health*, 66(6), e11.
53. Frank, M., Sommeiller, E., Price, M., & Saez, E. (2015). Frank-Sommeiller-Price series for top income shares by US states since 1917.
54. Zheng, H. (2012). Do people die from income inequality of a decade ago? *Social Science & Medicine*, 75(1), 36–45.

55. Peterson, E. (2017). Is economic inequality really a problem? A review of the arguments. *Social Sciences*, 6(4), 147.
56. Kennedy, B. P., Kawachi, I., Glass, R., & Prothrow-Stith, D. (1998). Income distribution, socioeconomic status, and self rated health in the United States: Multilevel analysis. *BMJ*, 317(7163), 917–921.
57. United Nations. (2015). *Transforming our world: The 2030 agenda for sustainable development*. United Nations.
58. Kuhn, H. (2020). Reducing inequality within and among countries: Realizing SDG 10—A developmental perspective. In M. Kaltenborn & M. Krajewski (Eds.), *Sustainable development goals and human rights* (Vol. 5, pp. 137–153). Springer.
59. Heinz, A., Zhao, X., & Liu, S. (2020). Implications of the association of social exclusion with mental health. *JAMA Psychiatry*, 77(2), 113–114.
60. Ribeiro, W. S., Bauer, A., Andrade, M. C. R., York-Smith, M., Pan, P. M., Pingani, L., Knapp, M., Coutinho, E. S. F., & Evans-Lacko, S. (2017). Income inequality and mental illness-related morbidity and resilience: A systematic review and meta-analysis. *The Lancet Psychiatry*, 4(7), 554–562.
61. Patel, V., Burns, J. K., Dhingra, M., Tarver, L., Kohrt, B. A., & Lund, C. (2018). Income inequality and depression: A systematic review and meta-analysis of the association and a scoping review of mechanisms. *World Psychiatry*, 17(1), 76–89.
62. Chetty, R., Stepner, M., Abraham, S., Lin, S., Scuderi, B., Turner, N., Bergeron, A., & Cutler, D. (2016). The association between income and life expectancy in the United States, 2001–2014. *JAMA*, 315(16), 1750–1766.
63. Sommet, N., & Spini, D. (2022). Financial scarcity undermines health across the globe and the life course. *Social Science & Medicine*, 292, 114607.
64. McKee-Ryan, F., Song, Z., Wanberg, C. R., & Kinicki, A. J. (2005). Psychological and physical well-being during unemployment: A meta-analytic study. *Journal of Applied Psychology*, 90(1), 53–76.
65. Teo, A. R., Choi, H., & Valenstein, M. (2013). Social relationships and depression: Ten-year follow-up from a nationally representative study. *PLoS ONE*, 8(4), e62396.
66. Huang, J., & Fang, Y. (2021). Income inequality, neighbourhood social capital and subjective well-being in China: Exploration of a moderating effect. *International Journal of Environmental Research and Public Health*, 18(13), 6799.
67. Rawls, J. (1971). *A theory of justice*. Harvard University Press.
68. Henrich, J. (2020). The weirdest people in the world: How the West became psychologically peculiar and particularly prosperous: Farrar, Straus and Giroux. *Journal of Intelligent & Fuzzy Systems*, 40, 1713.
69. National Research Council. (2013). *Nonresponse in social science surveys: A research agenda*. National Academies Press.
70. Morgan, S. L. (2020). Response rates and representativeness: A benchmark comparison of the general social surveys to the American Community Surveys, 2012–2018.

**Publisher's Note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.