NBER WORKING PAPER SERIES

ESTIMATING TRENDS IN MENTAL DISTRESS USING SURVEY DATA

Christopher J. Ruhm

Working Paper 35480
http://www.nber.org/papers/w35480

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
July 2026

I thank Brandyn Churchill, Richard Frank, Thomas Guterbock, Ronald Kessler, Tom McGuire, Maya Rossin-Slater, Sebastian Tello-Trillo, and Jane Waldfogel for helpful comments and suggestions on earlier versions of this paper. The views expressed herein are those of the author and do not necessarily reflect the views of the National Bureau of Economic Research.

NBER working papers are circulated for discussion and comment purposes. They have not been peer-reviewed or been subject to the review by the NBER Board of Directors that accompanies official NBER publications.

© 2026 by Christopher J. Ruhm. All rights reserved. Short sections of text, not to exceed two paragraphs, may be quoted without explicit permission provided that full credit, including © notice, is given to the source.

Estimating Trends In Mental Distress Using Survey Data
Christopher J. Ruhm
NBER Working Paper No. 35480
July 2026
JEL No. I1, I10, I19

ABSTRACT

One way of examining how mental health has changed over time is to use information available in large national surveys. While these survey-based measures are valuable in principle, and have been widely employed in prior research, a number of challenges may limit their practical usefulness or require considerable care in their application. This study uses data on U.S. 18-64 year olds from five major household surveys over the period 1999-2019 to investigate three specific questions. First, do different measures or the same indicator from different data sets consistently show how mental distress has changed over time? Second, how does the performance of single-item measures compare to those of psychometrically validated screeners based on multiple questions? Third, what are the benefits or potential costs of using different functional forms of these measures? The answers to these questions are as follows: 1) the estimated trends vary, often dramatically, across measures and, for the same indicator, across data sets and, as a result, some proxies provide a misleading understanding of how mental health has changed over time; 2) the single-item measures do as well or better than those based on multiple questions; 3) no single functional form demonstrates clear superiority over the others. After demonstrating these points, lessons are elucidated for researchers attempting to use survey data to examine trends in mental health.

Christopher J. Ruhm
University of Virginia
Frank Batten School of Leadership
and Public Policy
and NBER
ruhm@virginia.edu

It is often important to understand how the mental health of populations or sub-populations has changed over time. For example, alarm has been raised that the increased use of smartphones and social media, since around 2010, has adversely affected the mental wellbeing of youths (Keyes et al. 2019; Haidt 2024), and that growing "despair" increased the death rates of prime-age adults at the beginning of the 21st century (Case and Deaton 2017, 2022).

One way of attempting to identify these trends is to directly examine potential consequences of worsening mental health, such as changes in rates of successful or attempted suicides, the use of mental health services, or prescriptions for psychiatric medications. While useful, this approach has at least two limitations. First, these outcomes may be affected by difficult-to-account-for confounding factors, such as changes in the means of causing self-harm or in the economic environment (e.g. the availability of health insurance) that affect service use. This issue may be partially mitigated if these potential consequences exhibit common trends even while having different spurious correlates. Second, most such measures will be obtained in ways that permit controls for only a relatively limited set of personal or geographic characteristics. For example, suicide data comes from death certificates that include a somewhat sparse collection of demographic variables and in many cases, even when such information is available, often yield samples for subgroups that are too small for reliable analysis.[1]

A potentially useful alternative is to employ information on mental distress available from large national surveys providing data for repeated cross-sections that are designed, with the use of survey weights, to provide representative samples. A possible advantage is that many of these indicators have been validated against direct measures of mental

[1] Large underlying populations are also frequently needed to avoid confidentiality issues; since suicides are a relatively rare event, this may preclude analysis of smaller geographic area or population groups.

disorders (e.g. based on DSM diagnoses).[2] A second is that relationships can be ascertained between individual mental health and a wide variety of potential determinants or consequences, including those not of obvious interest *a priori* but that turn out to be valuable to examine.[3]

While the survey-based measures are valuable in principle, and have been widely utilized in prior research, several complications may limit their practical usefulness or require considerable care in their application. First, the thresholds used to delineate psychological distress in the validation studies may not be the most appropriate for ascertaining population trends. In response to this, prior researchers have sometimes used a variety of cutpoints, often determined somewhat arbitrarily. Second, there is likely to be non-response, at both the survey and question level, that may vary over time and introduce poorly understood biases. Third, the surveys sometimes undergo redesigns that may affect responses to the questions used for determining mental health. Fourth, the sample sizes could be too small to produce precise estimates, particularly for measures of serious mental distress that occur relatively infrequently. Fifth, it is at least possible that the same indicator will show heterogenous patterns across surveys.

Three primary questions are examined here. First, how similar or divergent are the trends in observed mental distress across specific measures and, for given proxies, across surveys? Disparate findings obviously suggest difficulties in using such metrics to determine how mental health has changed over time. Second, how does the predictive performance of indicators based on psychometrically validated multi-item measures

[2] Terms like mental distress, psychological distress, and poor mental health will be used somewhat interchangeably here for convenience. "Psychological distress" is the designation often used for the multi-item measures developed for use in survey samples or as psychiatric screeners.

[3] For example, using such measures, the effects of depression on future employment and earnings have been examined (Germinario et al. 2022), as has the frequency of mental health problems among adults diagnosed with osteoarthritis (Alwhaibi and Alhawassi 2025).

compare to that of variables based responses to a single question examining aspects of overall mental health? The single item measures have the advantage of being simpler to administer, potentially promoting wider and potentially more consistent use in large household surveys but these benefits could be more than offset if they are less accurate. Third, are some functional forms more useful than others? Most commonly, previous researchers have focused on dichotomous measures where scores above a specified threshold are interpreted as indicating serious distress. In addition to use such proxies, the current analysis considers lower thresholds, that also incorporate moderate mental distress, and quasi-continuous indicators that exploit information on the full range of scores from the multiple or single item measures.

These issues are examine through the analysis of six distinct measures of mental distress, including two coming from multiple sources, using data from five major U.S. surveys that, with weighting, are designed to be nationally representative. The focus is on non-elderly adults (18-64 year olds) over the period 1999-2019, or sub-periods with available data. This age range is chosen to be comparable across surveys, some of which only provide information for adults, and with seniors excluded since: the determinants of their mental health may differ from younger adults; mortality selection will be more problematic for them; and they are less often able to be surveyed because of physical health problems or because they are more likely to be in nursing homes which are excluded from most household surveys. In future analyses, it would be useful to provide corresponding examinations of youths and seniors. The first year studied is 1999, so as to focus on relatively recent trends, and the last is 2019 to avoid the COVID-19 pandemic, during which both actual mental health and its measurement are likely to have been uniquely affected. The resulting trends will be compared to those obtained for several

potential consequences of poor mental health to shed light on the plausibility of disparate findings across measures or surveys.

Three primary findings are highlighted. First, estimated mental health trends vary, sometimes dramatically, between indicators and, for the same measure, across data sets. As a result, analyses naively relying on these variables may provide a misleading understanding of how psychological distress has changed over time. In particular, while mental health has almost certainly worsened during the analysis period, some proxies suggest that it has improved. Dramatic differences in specific types of non-response provide a likely source of the erroneous results, even though these measures come from an important and generally high quality data source. Second, variables based on a single question inquiring about overall mental health status generally do as well or better than validated multi-item measures in correlating the trends in mental distress with other indicators of changes over time in poor mental health. Third, there is not a clear "winner" when it comes to choice of functional form, with threshold-based dichotomous measures of serious mental distress sometimes being marginally preferred, but at the cost of being less precisely estimated than quasi-continuous measures or binary classifications that include more moderate distress. The discussion concludes by elucidating lessons for researchers examining mental health trends for other age groups or time periods, or using alternative data sets.

1. Measures and Data Sources

1.1 Mental Distress Measures

Six indicators of mental distress were examined: Kessler 6 (K6) and Patient Health Questionnaire 2 and 9 (PHQ2 and PHQ9) scales, Mental Component Summary (MCS) scores, the number of poor mental health days in the last 30 (MH DAYS), and low levels

of self-rated overall mental health (POOR MH). The K6, PHQ2, PHQ2, and MCS combine information from multiple questions, while MH DAYS and POOR MH use responses from a single inquiry. Data came from five large national surveys, chosen because they provided information for at least one measure over most or all of the analysis period: the National Health Interview Survey (*NHIS*), Medical Expenditure Panel Survey (*MEPS*), National Health and the Nutrition Examination Survey (*NHANES*), Behavior Risk Factor Surveillance System (*BRFSS*), and National Survey on Drug Use and Health (NSDUH).[4]

The Kessler K6, one of the most widely used measures of mental distress, was based on responses to 6 questions indicating aspects of negative mental health during the last 30 days related to feeling sad, nervous, restless, hopeless, worthless, or that everything was an effort (R. C. Kessler et al. 2002; Ronald C. Kessler et al. 2003). Answers were coded between 0 for "none of the time" and 4 for "all of the time". Aggregating the responses yielded a range between 0 and 24, with higher scores indicating greater distress. K6 data were obtained from the 2004-2018 *NHIS*, the (mostly biennial) 2008-2019 *NSDUH*, and from self-administered questionnaire responses completed by adults in the 2004-2019 *MEPS*.[5] Non-response is a significant issue for the *MEPS* questionnaire responses in some years.[6]

[4] Information on these data sources is available from Blewett et al. (2024, 2025), National Center for Health Statistics (2022), Centers for Disease Control and Prevention (2021), and the Substance Abuse and Mental Health Services Administration (2020).

[5] The *NSDUH* also contained K6 scores before 2008 but these were not fully comparable to those in later years since they were based on the respondent's worst month during the prior year, rather than the last month.

[6] Other information in *MEPS* comes from face-to-face interviews in the core survey with almost complete response. By contrast, what are referred to here as *MEPS* questionnaire responses come from the self-administered questionnaire that individuals had to complete and sometimes mail back, resulting in greater and more variable non-response.

The PHQ2 scale was constructed from two questions indicating depressed mood and anhedonia (reduced ability to experience pleasure) during the last two weeks, each scored between 0 for "not at all" and 3 for "nearly every day", and summed to provide a range from 0 to 6 (Kroenke et al. 2003). PHQ9 scores were based on the same two questions plus seven additional inquiries related to aspects of poor mental health (Kroenke et al. 2001). The scoring of each response was the same as for the PHQ2, yielding an aggregate score between 0 and 27. PHQ2 data were available from the *MEPS* self-administered questionnaire from 2004-2019, with the same limitations as the K6 scores from that source, and the PHQ2 and PHQ9 in the *NHANES* from 2005-2006 through 2015-2016 and 2017-2020 (pre-pandemic).[7]

The MCS used responses to a subset of 12 items from the original 36-item Short-Form Health (SF-36) Survey (Ware et al. 1996), related to aspects of physical and mental health during the last four weeks. The latter were weighted more heavily in computing the MCS,[8] which was normalized to have a mean of 50 and standard deviation of 10 for the 1998 U.S. population, with higher scores indicating better mental health (Ware et al. 2001). The MCS was available from the *MEPS* self-administered questionnaire from 2000-2016 and, for this analysis, was reverse coded so that higher values indicated worse mental health.

MH DAYS were obtained from a *BRFSS* question on the number of days in the last 30 that "your mental health, which includes stress, depression, and problems of emotion" was not good. These data were available for all years except 2002, when they were only included in an optional modules not used by all states.

---

[7] *NHANES* data were available in two-year panels, with relevant information starting in 2005-2006. The last panel covered January 2017 through March 2020, just prior to the beginning of the COVID-19 pandemic; this is referred to as the 2017-2020 panel.

[8] There is an analogous physical component summary score that placed greater weight on the physical health questions.

Poor MH came from a *MEPS* core survey question, included starting in 2000, asking whether the respondent's mental health was "excellent", "very good", "good", "fair", or "poor", with responses coded from 1 to 5. Self-rated mental health has been shown in earlier work to be correlated to other symptoms or measures of mental distress (Fleishman and Zuvekas 2007; Ahmad et al. 2014).

The value of the individual question or sum of the values from multiple questions was calculated for each measure and is referred to as the "Distress Score" (*DS*).[9] Annual average distress scores were calculated, using appropriate sample weights, to obtain nationally representative estimates. Two additional indicators were computed, "mental distress" (*MD*) and "serious mental distress" (*SMD*), based on two principles. First, standard cutpoints were used for the multiple item psychometrically validated measures where these exist (e.g. K6 $\geq$ 13 for serious mental distress). Second, respecting the first criterion and to the extent permitted by the data, *MD* referred to scores at approximately the 80th percentile of the distribution and *SMD* to those at roughly the 90th percentile.[10] One exception was that the threshold for *MD* when proxied by POOR MH was at around the 70th percentile because there was no response between this and the 93rd percentile.

For clarity of exposition, terms like "measure", "indicator", and "proxy" are used below to define the category of variables used to evaluate mental distress (i.e. K6, PHQ2, PHQ9, MCS, MH DAYS, and POOR MH) and "form" or "functional form" to denote the specific metric (i.e. *DS, MD*, and *SMD*) within that proxy. Also, although these are

[9] The exception is the MCS where the normed scores were directly provided in the *MEPS*.

[10] The percentiles were calculated using weighted values from all years with available data. Since the distress scores were integer values, rather than continuous, the percentiles delineating *MD* and *SMD* were often approximate rather than exact. For the K6 scale, the cutoffs were based on the *NSDUH* data since the audio computer-assisted self-interviewing methods used in that survey are believed to provide more accurate prevalence estimates than the computer-assisted personal interviewing employed in the *NHIS*, or the self-administered *MEPS* questionnaires (Hedden et al. 2012).

referred to as mental of "distress" some measures capture broader manifestations of poor mental health.

Table 1 provides information on the data sources, years of availability, cutpoints, and corresponding percentiles for the *MD* and *SMD* functional forms. Appendix A provides additional details on the mental distress variables and examples of previous research utilizing them. Heterogeneity in sample sizes of the five data sets is worth noting. The *BRFSS* contained 125-330 thousand usable observations annually; the *NHIS*, *NSDUH*, and *MEPS* 20-25, 35-45, and 45-65 thousand respectively (around 25 thousand for the *MEPS* questionnaire variables); while the *NHANES* biennial sample sizes were just 3 to 5 thousand. These differences affect precision, with small sample sizes resulting in greater fluctuations across years.

1.2 Potential Consequences of Mental Distress

Four potential consequences of poor mental health (*CMH*) were used to examine the plausibility and reliability of results for the mental distress measures described above: suicides (SUICIDE); emergency department visits with a diagnosis of suicide ideation or attempts (SUICIDE ED), having visited or talked to a mental health professional (MH VISITS), and the use of prescribed psychotropic medications (MH DRUGS). All of these were defined as rates for 18-64 year olds, measured over one-year periods.

Trends in these variables were not necessarily determined exclusively by changes in mental distress. For instance, suicides partly depend on the means for self-harm, such as the availability of guns (Yip et al. 2012; Martínez-Alés et al. 2022; Currie 2025); the coding of emergency department visits was modified in 2016 with the switch from the ICD-9-CM to ICD-10-CM systems (Hedegaard et al. 2018; Sebastião et al. 2021; Corredor-Waldron and Currie 2024), and there were also relevant coding, screening, and

reimbursement changes in other years (Choi et al. 2025); and visits to mental health professionals and the use of prescribed medications may have been influenced by financial and regulatory considerations, such as improved health insurance coverage under the Affordable Care Act (Hockenberry et al. 2019; Breslau et al. 2020; Ortega 2023). For these reasons, these are described as potential, rather than actual, consequences of mental distress. Nevertheless, if they exhibit common trends, despite being affected by different confounding factors, they likely provide useful information on changes in mental health.[11]

Annual suicide rates, per 100,000 were obtained for all analysis years from the *CDC Wonder Multiple Cause of Death* data (National Center for Health Statistics 2021). Suicide emergency department visits per 100,000 were available starting in 2006 from the Health Care Utilization Project Nationwide Emergency Department Sample (Agency for Health Care Research and Quality 2025). MH Visits came from 1999-2018 *NHIS* sample adult responses to a question asking whether the respondent had "received counseling or therapy in the past 12 months from a mental health professional such as a psychiatrist, psychologist, psychiatric nurse, or clinical social worker". MH DRUGS was obtained from the core *MEPS* survey and referred to the use of antidepressants, central nervous system stimulants, and benzodiazepines and related drugs.

Additional details on the potential consequences of poor mental health are provided in Table 2 and in Appendix A.

2. Methods

In most specifications, individual data were weighted and then aggregated to obtain nationally representative average annual *DS, MD*, and *SMD* values. Standard errors and

[11] Bidirectional causation is generally not a concern here. For example, physical health problems are likely to cause mental health to deteriorate but this investigation does not focus on why mental distress is changing, only on how it is trending and what are the best ways of measuring it.

95 percent confidence intervals (CIs) were sometimes calculated, accounting for complex survey design. Annual values of the potential consequences of mental distress were similarly computed. In some cases, metal distress patterns were estimated by regressing individual (not aggregate) values on linear time trends, incorporating population weights, and with the associated confidence intervals again accounting for complex survey design. For selected indicators, trends were additionally calculated using varying starting and ending years, to evaluate the sensitivity of the findings to the timing of analysis. In addition to descriptive patterns, correlations between the three functional forms for a specific indictor were computed.

Most analysis focused on "normalized" variables, calculated as the percentage difference versus the 2008 mean, with the base year chosen since it was the first period with available data for all measures and potential consequences. To simplify exposition, dates denote the later year for the biennial *NHANES* measures (e.g. those from 2005-2006 are referred to as 2006) and those for 2017 through March 2020 as 2018. Sensitivity and robustness checks were performed as described below.

Normalized *MD* and *SMD* for a given measure can be directly compared but *DS* cannot, since the transformation from raw numbers to the share of persons with moderate or severe distress is non-linear.[12] An alternative, standardized measures were calculated as actual values minus the 2008 year means divided by the standard deviation of the variable over the period of data availability. Standardized variables have the advantage that *DS* can be directly compared to the *MD* and *SMD* but the disadvantage that they obscure

[12] For example, annual average POOR MH scores varied narrowly from a low of 1.909 in 1999 to a high of 2.086 in 2019 – a disparity of 9.3%, whereas the percentages of persons with *SMD* in these two years was 4.96% and 7.93%, a difference of 59.9%.

differences over time in the magnitude of the overall changes in mental distress across measures or functional forms.

Associations between the mental distress measures and potential consequences of poor mental health were obtained using the following bootstrapping procedure to account for the sometimes substantial confidence intervals on the distress indicators. First, year-specific means and survey-corrected standard errors were computed for the aggregated annual distress measures. Second, 1000 draws were taken, with replacement, with each mental distress proxy treated as a random variable with the year-specific means and standard errors calculated as above. Third, after each draw, the realized value of the distress variables were correlated with the four potential consequences of poor mental health, and the average of the four correlations was also calculated.[13] Finally, the empirical mean and 95% confidence intervals from the 1000 draws were computed.

3. Results

3.1 Trends in Mental Distress

Figure 1 documents trends in normalized *SMD*, *DS,* and *MD*. Two key results are immediately apparent. First, the changes over time varied substantially across measures. Most indicated worsening mental health, although often with considerable year-over-year variability, but the *MEPS* self-administered questionnaire indicators (K6, PHQ2, and MCS) suggested *reductions* in mental distress after 2011 or 2012. Because of this, even when using the same mental health proxy, the patterns differed sharply across surveys. For instance, serious mental distress measured using the K6 *rose* from 77% to 126% of the

[13] When examining the *NHANES* PHQ2 and PHQ9 measures, which were available at 2-year intervals (e.g. 2005-2006), values of the potential consequences of poor mental health were averaged over the corresponding two years.

2008 mean from 1999-2018 in the *NHIS* and to 151% of the 2008 average in the *NSDUH*, but it *fell* from 106% to 74% of the 2008 value from 2004-2019 in the *MEPS* (Figure 1A).[14]

Second, some findings were sensitive to the choice of analysis periods. For example, *NHIS* K6 data from 1999-2013 suggested strong secular *SMD* growth – rising from 77% to 124% of the 2008 mean – whereas there appeared to be no trend using data from 2002-2012. Serious mental distress, based on POOR MH, was virtually identical in 2004 and 2016– 103% and 104% of the 2008 average – but with a substantial increase from 2006-2019 – from 101% to 126%. These issues are examined further below, to the extent the data allow.

The patterns were generally similar for *DS* and *MD*, as for *SMD*, with increasing mental distress using most measures, which contrasts with declines for the *MEPS* questionnaire K6, PHQ2, and MCS indicators (Figures 1B and 1C).

Differences in mental health trends across measures or time periods might partially reflect imprecision of the estimates. This was examined in Appendix Figure A3, which shows annual means and associated 95% confidence intervals (CIs) for the three functional forms of each mental distress measure.[15] The confidence intervals were sometimes wide, particularly for data sets with small samples and for severe forms of mental distress, which had relatively low prevalence. For example, the 95% CIs for *SMD* from the *NHANES* PHQ2 were 5.5%-7.8% in 2006 and 7.7%-9.8% in 2018, making it difficult to rule out either increases or decreases in mental distress over time. The same applied, to a lesser extent, for the PHQ9. Conversely, most other indicators were fairly precisely measured

[14] The tables in Appendix B provide the specific numerical values for this and later Figures.

[15] The figure shows actual rather than normalized values. Normalized data yields the same figures, except for transformation of the Y-axis values.

and, importantly, the downwards trend in mental distress for the three *MEPS* questionnaire measures did not result from wide confidence intervals.

Figure 2 provides additional evidence of differential trends across proxies, showing the coefficients and confidence intervals for regressions of individual (rather than aggregate) mental distress data on a linear time trend. Once again, the *MEPS* questionnaire measures suggest that mental health has improved over time, with narrow confidence intervals, in contrast to all the other indicators which showed worsening psychological distress. These estimates were again fairly precise in most cases, except for the extremely wide confidence intervals for the two *NHANES* PHQ measures, reflecting small sample sizes. CIs were frequently wider for *SMD* and, to a lesser extent, *MD* than for *DS*.

3.2 Trends in Potential Consequences of Mental Health

The results just described showed sharply disparate estimated trends in different survey-based measures of mental health, with most indicating worsening distress but with the three *MEPS* questionnaire variables (K6, PHQ2, and MCS) suggesting improvements in later years.

To shed light on which of these patterns was more likely to be correct, *CMH* trends were next examined. Figure 4 summarizes the main results, with the variables again normalized as a percentage of 2008 values. As discussed, none of the observed changes need to have exclusively reflected the evolution of mental health, so the results for any single potential consequence should be interpreted with caution. However, since the spurious factors influencing each of them were likely to have differed, consistent patterns reduced the chances that the observed changes were dominated by confounding influences and suggest that, in combination, they provided a useful guide for assessing the reasonableness of the survey-based mental distress indicators.

The data provide clear evidence that the potential negative consequences of poor mental health increased over time. SUICIDE and MH DRUGS were 86% and 65% of 2008 values in 1999, compared with 119% and 114% in 2019. MH Visits grew from 74% of the 2008 baseline in 1999 to 135% in 2018. Data on SUICIDE ED became available in 2006 and from that year until 2019 rose from 81% to 165% of the 2008 average. Moreover, the correlations between these measures were high: ranging from 0.78 between MH VISITS and MH DRUGS to 0.99, for SUICIDE and SUICDE ED (Appendix Table A1).

This uniformity of results provides strong evidence that psychological distress grew during the first two decades of the 21st century, at least for prime-age adults. The findings were also consistent with previous studies indicating that suicide rates, suicide attempts, and emergency room visits increased during at least portions of the analysis period (Owens et al. 2017, 2020; Martínez-Alés et al. 2022; Bommersbach et al. 2024; Garnett and Curtin 2024; Marcotte and Hansen 2024; Ting et al. 2012), as did mental health visits and treatments (Mojtabai and Jorm 2015; Wang et al. 2023). However, it is noteworthy that MH DRUGS was least strongly associated with the other three *CMH* indicators, raising the possibility that some results below may be sensitive to whether or not it is included in the comparison set.

3.3 Explaining Anomalous Mental Distress Trends in the MEPS

The *MEPS* questionnaire measure (the K6, PHQ2, and MCS) results indicating that mental health improved after 2011 or 2012 were almost certainly anomalous since they contrasted sharply with the growth in mental distress observed using all other proxies, including the K6 and PHQ2 from other surveys, and also with increases in the potential consequences of poor mental health.

This was likely caused by changes in non-response that were unique to the *MEPS* questionnaire variables. *MEPS* participants were interviewed in-person multiple times over 2.5 years. However, the K6, PHQ2, and MCS did not come from these "core" interviews but rather from separate self-administered questionnaires (SAQ) completed once per year (twice per respondent) that were submitted separately. For the other household surveys, as well as the core *MEPS*, data could be missing due to either survey or item non-response. However, for these proxies there was the extra possibility of missing data due failure to return the SAQ (Kindratt 2022).

Overall survey response rates fell over time for all five household surveys (Meyer et al. 2015; Czajka and Beyler 2016; Williams and Brick 2018), but there was no evidence of a discontinuity around 2012 in the *MEPS* (see Appendix C), making it unlikely that this was the major source of the anomalous declines in mental distress in the later analysis years using the questionnaire measures. By contrast, reductions in SAQ completion rates probably played the key role. To illustrate, Figure 5A shows that the share of missing *MEPS* K6, PHQ2, and MCS responses increased markedly after 2012, around the time that measured mental distress began its sharp decline, and Figure 5B shows that this coincided with substantial reductions in SAQ return rates.[16] There was no corresponding rise in missing values conditional on return of the SAQ (Appendix Figure A4).[17] As further evidence that changes in SAQ completion rates were crucial, no corresponding reduction in mental distress was observed for POOR MH, which used information from the *MEPS* core survey.

[16] This analysis begins in 2002 because that is the first year that the *IPUMS MEPS* reports who was eligible to return the SAQ.

[17] SAQ return rates were virtually identical in the first and second years individuals participated in the *MEPS*, belying the likelihood that survey fatigue caused these results (Figure 5B).

A crude adjustment for declining SAQ completion rates was obtained by regressing the *MEPS* K6, PHQ2, and MCS measures ssed on the percentage of non-returned questionnaires and then computing predicted values were computed assuming a 100% return rate. These "adjusted" measures, shown in Figure 5C for *SMD* and Appendix Figure A5 for *DS* and *MD*, indicate that the post-2012 decline would have been largely eliminated, although increases in mental distress were still not observed. The last result was unsurprising since the correction method does not account for the likelihood that non-respondents became increasingly comprised of individuals in poor mental health in later years.

Appendix C details examinations of additional potential sources of decrease in mental distress in the *MEPS* questionnaire measures including declining overall response rates to household surveys and changes in the design of the *MEPS* survey over time. Neither of these seems likely to play a substantial role, although modifications to the survey could have contributed to some degree.

3.4 Correlations Between Mental Distress Measures and Potential Consequences

Correlations between the measures and potential consequences of mental distress were next computed to demonstrate how well the survey-based indicators align with actual trends in mental health. Point estimates and confidence intervals were determined using the bootstrapping procedure described earlier. As expected, given the results documented above, the *MEPS* questionnaire variables were always strongly *inversely* associated with *CMH*: average correlation coefficients were below -0.5 in most cases (Appendix Figure A6), providing further evidence that these variables do not reliably indicate how mental health changed during the beginning of the 21st century.

Conversely, the other mental distress indicators were almost always strongly positively associated with the potential consequences of poor mental health (Figure 6). Of particular note, the correlation coefficients averaged 0.87, 0.86, and 0.90 for *DS*, *MD*, and *SMD* based on MH DAYS, with miniscule confidence intervals. K6 from the *NHIS* or *NSDUH*, and POOR MH, also had fairly high average associations, in the 0.6 to 0.8 range, that were frequently not statistically distinguishable from each other, albeit measured with less precision. These results were sometimes dependent on the choice of functional forms or particular potential consequences examined. For example, the average correlations for POOR MH *SMD* were 0.78 compared to 0.68 for *MD*. Mean associations between *NSDUH* K6 *DS*, *MD*, and *SMD* and the four potential consequences were 0.71, 0.75, and 0.79; however, these were considerably affected by the weak correlations with MH DRUGS. For the other three potential consequences (SUICIDE, SUICIDE ED, and MH VISITS) the correlation coefficients averaged 0.80, 0.86, and 0.89. The PHQ2 and PHQ9 proxies were relatively weakly related to *CMH*. For example, average associations were 0.36 for *MD* and 0.42 for *SMD* from the PHQ9, and with the small *NHANES* sample sizes resulting in extremely large CIs (0.12 to 0.59 and 0.12 to 0.67).

Table 3 aggregates the mean correlations for a given mental distress measure across functional forms (top panel) or, for a given functional form, across measures (lower panel), with the *MEPS* questionnaire variables excluded from these averages given the erroneous results for them. The upper panel highlights the high correlations between MH DAYS and *CMH*, followed by a fairly tight grouping of the K6 proxies from the *NHIS* and *NSDUH* as well as POOR MH, and then with much lower and less precisely associations for the two *NHANES* PHQ measures.

Across the three functional forms, correlations between MH DAYS and the four potential consequences of poor mental health averaged 0.88, and this varied only slightly –

from 0.86 to 0.89 – when one of the *CMH* variables was excluded. The corresponding overall average correlations ranged from 0.71 to 0.75 for the K6 indicators and POOR MH. However, as discussed, the K6 *NSDUH* associations were sensitive to whether or not *MH DRUGS* was included in the *CMH* comparison set. PHQ2 and PHQ9 were much less strongly associated with *CMH*: the correlation coefficients averaged 0.60 and 0.41, and with wide confidence intervals.

The bottom panel of Table 3 shows that the choice of functional forms was much less important. The average correlation coefficients between the six mental distress measures and four *CMH* indicators were higher for *SMD* than *MD* or *DS* but the differences were not large – 0.70 versus 0.68 and 0.67. All of the mean correlations increased when excluding the low performing *PHQ* indicators but the ranking remained the same and the gaps between the *SMD* average and those for *MD* and *DS* remained modest (0.80 versus 0.77 and 0.76).

3.5 Sensitivity to the Choice of Analysis Periods

The final stage of the analysis evaluated the extent to which the estimated mental distress trends varied with the timing of analysis by using individual data to regress the normalized measures on linear time trends for samples with various starting and ending dates covering a minimum of 10 years. This analysis focused on K6 from the *NHIS*, MH DAYS, and POOR MH, since these had data available for sufficiently lengthy periods needed for this purpose. The earliest year included was extended back to 1997 to expand the available timeframe. Starting years varied from 1997-2010 (1997-2009 for the *NHIS* K6) with 2019 (2018) as the ending year, and final years from 2006-2019 (2006-2018) where the initial year was 1997. The trend coefficients show estimated annual percentage point changes, relative to the 2008 baseline.

The estimated trends could vary with the choice of analysis windows because the underlying mental health trends were themselves evolving over time. However, small shifts in timing should produce only modest and gradual changes in the estimates, with substantial instability suggesting either that the estimates were imprecise or that the mental health measures lacked reliability. Figure 7 details how the estimated trend coefficients varied by analysis period. Appendix Figures A7-A9 provide the same information, but with confidence intervals also shown.

A primary result is that the K6 *NHIS* measures were more sensitive to timing than those for either MH DAYS or POOR MH. This was particularly true for serious mental distress where, for example, the estimated annual increase was 0.82 (95% CI: -0.05-1.68) percentage points from 1997-2006 versus 0.26 (-0.55-1.07) points from 1997-2007, and 2.36 (95% CI: 1.40-3.33) points from 2007-2018 compared to 1.64 (95% CI: 0.38-2.91) points from 2009-2018. Some of this likely reflected the somewhat modest *NHIS* sample sizes; however, the relatively greater robustness of the estimates for POOR MH was noteworthy, given their similarly wide confidence intervals. By contrast, volatility to the timing of the analysis was fairly minimal for MH DAYS, with narrow confidence intervals.[18]

4. **Discussion**

This study focuses on three questions relevant for estimating mental health trends using data from major household surveys. First, is the evidence on changes over time consistent across measures and data sets, and also in line with other indicators of mental health trends? Second, how does the performance of single item measures compare to

[18] Figure A10 explored whether trends in the potential consequences of poor mental health were themselves evolving by displaying corresponding trend coefficients for alternative analysis windows. These were constructed using aggregated, rather than individual, data and with SUICIDE ED excluded given its limited period of data availability. These estimates reveal little evidence that the trends changed over time, except for the possibility of particularly large increases for MH DRUGS in earlier sample years.

those of psychometrically validated screeners based on multiple aspects of psychological health? Third, what are the benefits or potential costs of focusing on binary thresholds indicating serious mental distress versus less restrictive cutoffs, that include milder distress, or quasi-continuous scores that do not impose a cutpoint.

With respect to the first question, this analysis identifies important divergences in the results across data sets and measures. Specifically, while most proxies provide evidence of increased mental distress over time, and this corresponds to the patterns observed for other indicators of mental health, the three indicators based self-administered questionnaire data from the *MEPS* provide almost certainly erroneous indications of improvements in mental health after around 2012. A key lesson for researchers is that even clinically validated measures from high quality surveys cannot be relied upon uncritically, but rather that additional plausibility checks will often be needed.

Second, at least for the 1999-2019 timespan and 18-64 year old population examined here, the measures based on a single general question do as well or better than the validated multiple-item variables. Indeed, for this investigation, the most predictive indicator is the number of poor mental health days in the last 30 from the *BRFSS*. It has the highest and most precisely estimated positive correlations with the potential consequences of poor mental health, and yields trend estimates that are robust to the choice of functional forms (*DS, MD*, and *SMD*) and relatively so to changes in the timing of analysis.

Self-rated overall mental health from the core *MEPS* survey also fares reasonably well, with estimates that are relatively unaffected by small modifications in the analysis period and yielding correlations with *CMF* that were only slightly below those using K6 measures from the *NHIS* or *NSDUH* – though all of these remain lower than for MH DAYS. By contrast, the *NHIS* K6 yields trend estimates that were sometimes quite sensitive to the

timespan studied, and the *NHANES* PHQ2 and PHQ9 measures were relatively weakly correlated with the potential consequences of mental health and, presumably due to relatively small sample sizes, produced wide confidence intervals. These conclusions hold even when the K6, PHQ2, and MCS data from the *MEPS* were excluded. Since they were strongly *negatively* correlated with *CMH*, including them would have dramatically worsen results for the multi-item measures relative to MH DAYS or POOR MH.

Third, no functional form demonstrated conclusive superiority. *SMD* was slightly more highly correlated with *CMH* than *DS* or *MD*, but the differences were marginal, frequently reversed, and almost never statistically significant. Further, any modest gains from using *SMD* often came at the cost of reduced precision – producing wider confidence intervals and more sensitivity to small changes in timing of the analysis.

It is interesting to speculate why the single-item variables appear to be at least as of useful as the multi-item measures specifically designed to indicate mental distress. One potential explanation is that former capture aspects of mental health not accounted for by the K6 or PHQ scales. In the same way, there is a substantial literature demonstrating the superiority, in at least some contexts, of single overall self-rated health variables compared to those based on specific health conditions or diagnoses (Idler and Benyamini 1997; Benyamini 2011).

A second possibility is that the greater complexity of the multi-question scales makes it more difficult to include them in the largest household surveys and, when they are incorporated, this occurs in ways that lead to less reliable results.[19] Results for the mental health variables in the *MEPS* are particularly instructive in this regard. The single item

[19] Compare sample sizes in the hundreds of thousands annually for the *BRFSS*, from which MH DAYS was obtained, to the three to five thousand for the biennial *NHANES* samples used to construct the PHQ2 and PHQ9 scales. Smaller samples resulted in larger confidence intervals.

POOR MH scale, which came from the core questionnaire completed by around 50 thousand persons each year, provided useful results. Conversely, data for the K6, PHQ2, and MCS measures came from the supplementary self-administered questionnaire which yielded only around half as many responses and greater year-by-year variability in them. As discussed, the resulting trend estimated were inconsistent with changes in the potential consequences of poor mental health observed here and in other research (Owens et al. 2020; Martínez-Alés et al. 2022; Bommersbach et al. 2024; Wang et al. 2023; Marcotte and Hansen 2024), as well as patterns of the same mental distress measures obtained from other data sets.

Previous studies using survey data to examine trends in mental health during the 21st century have obtained mixed results. Some indicate rising mental distress (Blanchflower and Oswald 2020; Daly 2022; Udupa et al. 2023; Ruhm 2025), others decreases in it (Brody et al. 2018; Olfson et al. 2019), and still others obtain varying findings, no trend, or differing effects over time (Mojtabai and Jorm 2015; Tomitaka et al. 2019; Yu et al. 2020; Iranpour et al. 2022).

Many of these ambiguities may be explained by the choice of measures and, to a lesser degree, time periods analyzed. To provide an extreme example, Olfson et al. (2019), using *MEPS* K6 data, concluded that serious mental distress *declined* from 4.9% of adults in 2009-2010 to 3.7% in 2014-2015. However, as discussed above, this decrease was illusory and reflected problems with measuring mental health using variables from the *MEPS* self-administered questionnaire. Investigations using the *NHANES* PHQ9 have usually not obtained strong evidence of trends in serious depression since 2006 (Brody et al. 2018; Yu et al. 2020; Iranpour et al. 2022; Udupa et al. 2023) but, as highlighted above, this largely reflected large confidence intervals due to the relatively small *NHANES* sample sizes.

A key lesson for researchers studying trends in mental health using survey data is that single item indicators are often equally or more useful than scales based on multiple questions, even when the latter have been psychometrically validated. Interestingly, the gains from focusing on threshold-based measures of severe distress are marginal at best and these may be measured less precisely than proxies incorporating milder distress or using the full information from the underlying scores. More generally, the results highlight three important issues that researchers should examine. First, are the sample sizes large enough to provide sufficiently precise estimates? Second, are the findings robust to small changes in the analysis period? Third, are the results consistent with measures of mental distress obtained from other surveys and with external (often non-survey based) indicators of mental health? This analysis provides partial answers to these questions for U.S. 18-64 year olds from 1999-2019, and supplies guidance for future researchers examining other groups, time periods, or locations.

References

Agency for Health Care Research and Quality. 2025. "HCUPnet Data Tools – Healthcare Cost and Utilization Project (HCUPnet)." https://datatools.ahrq.gov/hcupnet/.

Ahmad, Farah, Anuroop K. Jhajj, Donna E. Stewart, Madeline Burghardt, and Arlene S. Bierman. 2014. "Single Item Measures of Self-Rated Mental Health: A Scoping Review." *BMC Health Services Research* 14 (September): 398. https://doi.org/10.1186/1472-6963-14-398.

Alwhaibi, Monira, and Tariq M. Alhawassi. 2025. "Mental Health and Health-Related Quality of Life among Adults with Osteoarthritis: A National Population-Based Study." *PLOS One* 20 (6): e0325780. https://doi.org/10.1371/journal.pone.0325780.

Benyamini, Yael. 2011. "Why Does Self-Rated Health Predict Mortality? An Update on Current Knowledge and a Research Agenda for Psychologists." *Psychology & Health* 26 (11): 1407–13. https://doi.org/10.1080/08870446.2011.621703.

Blanchflower, David G., and Andrew J. Oswald. 2020. "Trends in Extreme Distress in the United States, 1993–2019." *American Journal of Public Health* 110 (10): 1538–44. https://doi.org/10.2105/AJPH.2020.305811.

Blewett, Lynn A., Julia A. Rivera Drew, Daniel Backman, et al. 2024. "IPUMS Health Surveys: Medical Expenditure Panel Survey: Version 2.4." Version 2.4. Minneapolis, MN: IPUMS. https://doi.org/10.18128/D071.V2.4.

Blewett, Lynn A., Julia A. Rivera Drew, Andrew Fenelon, et al. 2025. "IPUMS Health Surveys: National Health Interview Survey, Version 8.1." IPUMS. https://doi.org/10.18128/D071.V8.1.

Bommersbach, Tanner J., Mark Olfson, and Taeho Greg Rhee. 2024. "National Trends in Emergency Department Visits for Suicide Attempts and Intentional Self-Harm." *The American Journal of Psychiatry* 181 (8): 741–52. https://doi.org/10.1176/appi.ajp.20230397.

Breslau, Joshua, Bing Han, Julie Lai, and Hao Yu. 2020. "Impact of the Affordable Care Act Medicaid Expansion on Utilization of Mental Health Care." *Medical Care* 58 (9): 757–62. https://doi.org/10.1097/MLR.0000000000001373.

Brody, Debra J., Laura A. Pratt, and Jeffery P. Hughes. 2018. "Prevalence of Depression Among Adults Aged 20 and Over: United States, 2013-2016." *NCHS Data Brief*, no. 303 (February): 1–8.

Case, Anne, and Angus Deaton. 2017. "Mortality and Morbidity in the 21st Century." *Brookings Papers on Economic Activity* 2017 (1): 397–476. https://doi.org/10.1353/eca.2017.0005.

Case, Anne, and Angus Deaton. 2022. "The Great Divide: Education, Despair, and Death." *Annual Review of Economics* 14 (1): 1–21. https://doi.org/10.1146/annurev-economics-051520-015607.

Centers for Disease Control and Prevention (CDC). 2021. "Behavioral Risk Factor Surveillance System (BRFSS), 1999–2019." U.S. Department of Health and Human Services, Centers for Disease Control and Prevention. https://www.cdc.gov/brfss/.

Choi, Han, Adriana Corredor-Waldron, Janet Currie, and Chris Felton. 2025. *What Can Trends in Emergency Department Visits Tell Us About Child Mental Health?* No. W33550. National Bureau of Economic Research. https://doi.org/10.3386/w33550.

Corredor-Waldron, Adriana, and Janet Currie. 2024. "To What Extent Are Trends in Teen Mental Health Driven by Changes in Reporting?: The Example of Suicide-Related Hospital Visits." *Journal of Human Resources* 59 (S): S14–40. https://doi.org/10.3368/jhr.0423-12854R1.

Currie, Janet. 2025. "Presidential Address: Investing in Children to Address the Child Mental Health Crisis." *American Economic Review* 115 (5): 1369–99. https://doi.org/10.1257/aer.115.5.1369.

Czajka, John L., and Amy Beyler. 2016. *Background Paper Declining Response Rates in Federal Surveys: Trends and Implications*. Mathematica Policy Research. https://aspe.hhs.gov/sites/default/files/private/pdf/255531/Decliningresponserates.pdf.

Daly, Michael. 2022. "Prevalence of Psychological Distress Among Working-Age Adults in the United States, 1999-2018." *American Journal of Public Health* 112 (7): 1045–49. https://doi.org/10.2105/AJPH.2022.306828.

Fleishman, John A., and Samuel H. Zuvekas. 2007. "Global Self-Rated Mental Health: Associations with Other Mental Health Measures and with Role Functioning." *Medical Care* 45 (7): 602–9. https://doi.org/10.1097/MLR.0b013e31803bb4b0.

Garnett, Matthew F., and Sally C. Curtin. 2024. "Suicide Mortality in the United States, 2002-2022." *NCHS Data Brief*, no. 509 (September). https://doi.org/10.15620/cdc/160504.

Germinario, Giuseppe, Vikesh Amin, Carlos A. Flores, and Alfonso Flores-Lagunes. 2022. "What Can We Learn about the Effect of Mental Health on Labor Market Outcomes under Weak Assumptions? Evidence from the NLSY79." *Labour Economics* 79 (December): 102258. https://doi.org/10.1016/j.labeco.2022.102258.

Haidt, Jonathan. 2024. *The Anxious Generation: How the Great Rewiring of Childhood Is Causing an Epidemic of Mental Illness*. Penguin Press.

Hedden, Sarra, Joe Gfroerer, Peggy Barker, et al. 2012. "Comparison of NSDUH Mental Health Data and Methods with Other Data Sources." *CBHSQ Data Review*.

Hedegaard, Holly, Michael Schoenbaum, Cynthia Claassen, Alex Crosby, Kristin Holland, and Scott Proescholdbell. 2018. "Issues in Developing a Surveillance Case Definition for Nonfatal Suicide Attempt and Intentional Self-Harm Using International Classification of Diseases, Tenth Revision, Clinical Modification (ICD-10-CM) Coded Data." *National Health Statistics Reports*, no. 108 (February): 1–19.

Hockenberry, Jason M., Peter Joski, Courtney Yarbrough, and Benjamin G. Druss. 2019. "Trends in Treatment and Spending for Patients Receiving Outpatient Treatment of Depression in the United States, 1998-2015." *JAMA Psychiatry* 76 (8): 810–17. https://doi.org/10.1001/jamapsychiatry.2019.0633.

Idler, E. L., and Y. Benyamini. 1997. "Self-Rated Health and Mortality: A Review of Twenty-Seven Community Studies." *Journal of Health and Social Behavior* 38 (1): 21–37.

Iranpour, Sohrab, Siamak Sabour, Fatemeh Koohi, and Hossein Mozafar Saadati. 2022. "The Trend and Pattern of Depression Prevalence in the U.S.: Data from National Health and Nutrition Examination Survey (NHANES) 2005 to 2016." *Journal of Affective Disorders* 298 (Pt A): 508–15. https://doi.org/10.1016/j.jad.2021.11.027.

Kessler, R. C., G. Andrews, L. J. Colpe, et al. 2002. "Short Screening Scales to Monitor Population Prevalences and Trends in Non-Specific Psychological Distress." *Psychological Medicine* 32 (6): 959–76. https://doi.org/10.1017/s0033291702006074.

Kessler, Ronald C., Peggy R. Barker, Lisa J. Colpe, et al. 2003. "Screening for Serious Mental Illness in the General Population." *Archives of General Psychiatry* 60 (2): 184–89. https://doi.org/10.1001/archpsyc.60.2.184.

Keyes, Katherine M., Dahsan Gary, Patrick M. O'Malley, Ava Hamilton, and John Schulenberg. 2019. "Recent Increases in Depressive Symptoms among US Adolescents: Trends from 1991 to 2018." *Social Psychiatry and Psychiatric Epidemiology* 54 (8): 987–96. https://doi.org/10.1007/s00127-019-01697-8.

Kindratt, Tiffany B. 2022. *Big Data for Epidemiology: Applied Data Analysis Using National Health Surveys*. Mavs Open Press. https://uta.pressbooks.pub/bigdataforepidemiology/.

Kroenke, Kurt, Robert L. Spitzer, and Janet B. W. Williams. 2001. "The PHQ-9: Validity of a Brief Depression Severity Measure." *Journal of General Internal Medicine* 16 (9): 606–13. https://doi.org/10.1046/j.1525-1497.2001.016009606.x.

Kroenke, Kurt, Robert L. Spitzer, and Janet B. W. Williams. 2003. "The Patient Health Questionnaire-2: Validity of a Two-Item Depression Screener." *Medical Care* 41 (11): 1284–92. https://doi.org/10.1097/01.MLR.0000093487.78664.3C.

Marcotte, Dave E., and Benjamin Hansen. 2024. "The Re-emerging Suicide Crisis in the U.S.: Patterns, Causes and Solutions." *Journal of Policy Analysis and Management* 43 (2): 582–612. https://doi.org/10.1002/pam.22526.

Martínez-Alés, Gonzalo, Tammy Jiang, Katherine M. Keyes, and Jaimie L. Gradus. 2022. "The Recent Rise of Suicide Mortality in the United States." *Annual Review of Public Health* 43 (April): 99–116. https://doi.org/10.1146/annurev-publhealth-051920-123206.

Meyer, Bruce D., Wallace KC Mok, and James X. Sullivan. 2015. "Household Surveys in Crisis." *Journal of Economic Perspectives* 29 (4): 199–226.

Mojtabai, Ramin, and Anthony F. Jorm. 2015. "Trends in Psychological Distress, Depressive Episodes and Mental Health Treatment-Seeking in the United States:

2001-2012." *Journal of Affective Disorders* 174 (March): 556–61. https://doi.org/10.1016/j.jad.2014.12.039.

National Center for Health Statistics. 2021. "Multiple Cause of Death Files, 1999–2019." U.S. Department of Health and Human Services, Centers for Disease Control and Prevention. https://wonder.cdc.gov/mcd-icd10.html.

National Center for Health Statistics (NCHS). 2022. "National Health and Nutrition Examination Survey, 1999–2020." U.S. Department of Health and Human Services, Centers for Disease Control and Prevention. https://www.cdc.gov/nchs/nhanes/.

Olfson, Mark, Shuai Wang, Melanie Wall, Steven C. Marcus, and Carlos Blanco. 2019. "Trends in Serious Psychological Distress and Outpatient Mental Health Care of US Adults." *JAMA Psychiatry* 76 (2): 152–61. https://doi.org/10.1001/jamapsychiatry.2018.3550.

Ortega, Alberto. 2023. "Medicaid Expansion and Mental Health Treatment: Evidence from the Affordable Care Act." *Health Economics* 32 (4): 755–806. https://doi.org/10.1002/hec.4633.

Owens, Pamela L., Kathryn R. Fingar, Kevin C. Heslin, Ryan Mutter, and Chelsea L. Booth. 2017. "Emergency Department Visits Related to Suicidal Ideation, 2006–2013." *Healthcare Cost and Utilization Project (HCUP) Statistical Brief #220*, January. http://www.ncbi.nlm.nih.gov/books/NBK442036/.

Owens, Pamela L., Kimberly W. McDermott, Rachel N. Lipari, and Megan M. Hambrick. 2020. "Emergency Department Visits Related to Suicidal Ideation or Suicide Attempt, 2008–2017." *Healthcare Cost and Utilization Project (HCUP) Statistical Brief #263*, September. http://www.ncbi.nlm.nih.gov/books/NBK563021/.

Ruhm, Christopher J. 2025. "Mental Health and Mortality Trends in the United States." *Journal of Health Economics* 102 (August): 103015. https://doi.org/10.1016/j.jhealeco.2025.103015.

Sebastião, Yuri V., Gregory A. Metzger, Deena J. Chisolm, Henry Xiang, and Jennifer N. Cooper. 2021. "Impact of ICD-9-CM to ICD-10-CM Coding Transition on Trauma Hospitalization Trends among Young Adults in 12 States." *Injury Epidemiology* 8 (1): 4. https://doi.org/10.1186/s40621-021-00298-x.

Substance Abuse and Mental Health Services Administration. 2020. "National Survey on Drug Use and Health (NSDUH): Combined Public-Use File (2002-2019)." https://www.samhsa.gov/data/data-we-collect/nsduh-national-survey-drug-use-and-health/datafiles/2002-2019.

Ting, Sarah A., Ashley F. Sullivan, Edwin D. Boudreaux, Ivan Miller, and Carlos A. Camargo. 2012. "Trends in US Emergency Department Visits for Attempted Suicide and Self-Inflicted Injury, 1993-2008." *General Hospital Psychiatry* 34 (5): 557–65. https://doi.org/10.1016/j.genhosppsych.2012.03.020.

Tomitaka, Shinichiro, Yohei Kawasaki, Kazuki Ide, Maiko Akutagawa, Yutaka Ono, and Toshi A. Furukawa. 2019. "Distribution of Psychological Distress Is Stable in Recent Decades and Follows an Exponential Pattern in the US Population." *Scientific Reports* 9 (1): 11982. https://doi.org/10.1038/s41598-019-47322-1.

Udupa, Nikhila S., Jean M. Twenge, Cooper McAllister, and Thomas E. Joiner. 2023. "Increases in Poor Mental Health, Mental Distress, and Depression Symptoms among U.S. Adults, 1993-2020." *Journal of Mood and Anxiety Disorders* 2 (August): 100013. https://doi.org/10.1016/j.xjmad.2023.100013.

Wang, Junzhe, Yang Qiu, and Xiaozhou Zhu. 2023. "Trends of Mental Health Care Utilization among US Adults from 1999 to 2018." *BMC Psychiatry* 23 (1): 665. https://doi.org/10.1186/s12888-023-05156-2.

Ware, J., M. Kosinski, and S. Keller. 2001. "SF-36 Physical and Mental Health Summary Scales: A User's Manual 1994." https://www.researchgate.net/profile/John-Ware-6/publication/292390260_SF-36_Physical_and_Mental_Health_Summary_Scales_a_User's_Manual/links/5af580264585157136caee31/SF-36-Physical-and-Mental-Health-Summary-Scales-a-Users-Manual.pdf.

Ware, J., M. Kosinski, and S. D. Keller. 1996. "A 12-Item Short-Form Health Survey: Construction of Scales and Preliminary Tests of Reliability and Validity." Medical Care, March.

Williams, Douglas, and J. Michael Brick. 2018. "Trends in US Face-to-Face Household Survey Nonresponse and Level of Effort." *Journal of Survey Statistics and Methodology* 6 (2): 186–211.

Yip, Paul SF, Eric Caine, Saman Yousuf, Shu-Sen Chang, Kevin Chien-Chang Wu, and Ying-Yeh Chen. 2012. "Means Restriction for Suicide Prevention." *The Lancet* 379 (9834): 2393–99. https://doi.org/10.1016/S0140-6736(12)60521-2.

Yu, Binbin, Xueyuan Zhang, Chunpeng Wang, Mengzi Sun, Lina Jin, and Xin Liu. 2020. "Trends in Depression among Adults in the United States, NHANES 2005-2016." *Journal of Affective Disorders* 263 (February): 609–20. https://doi.org/10.1016/j.jad.2019.11.036.

Table 1. Mental Distress Measures

Measure (Abbreviation)	Source/Years	Thresholds (Percentile)
Kessler 6 (K6)	NHIS, 1999-2018	*MD*: ≥8: (≥89.8) *SMD*: ≥13: (≥96.6)
	NSDUH, 2008-2019	*MD*: ≥8: (≥81.8) *SMD*: ≥13: (≥94.0)
	MEPS Questionnaire, 2004-2019	*MD*: ≥6: (≥87.2) *SMD*: ≥13: (≥95.4)
Mental Component Summary core (MCS)*	MEPS Questionnaire, 2000-2016	*MD*: ≥56.18: (≥80.0) *SMD*: ≥62.79: (≥90.0)
Patient Health Questionnaire-2 (PHQ-2)	MEPS Questionnaire, 2004-2019	*MD*: ≥2: (≥81.1) *SMD*: ≥3: (≥92.2)
	NHANES, 2005-2020	*MD*: ≥2: (≥81.9) *SMD*: ≥3: (≥91.9)
Patient Health Questionnaire-9 (PHQ-9)	NHANES, 2005-2020	*MD*: ≥6: (≥81.1) *SMD*: ≥10: (≥91.5)
Poor Mental Health Days in Last 30 (MH DAYS)	BRFSS, 1999-2019, except 2002	*MD*: ≥6: (≥82.0) *SMD*: ≥16: (≥91.5)
Self-rated Mental Health (POOR MH)	MEPS Core, 1999-2019	*MD*: ≥3: (≥70.3) *SMD*: ≥4: (≥93.4)

Note: Table shows time periods the mental distress measures are available, the data source, and thresholds for levels of mental distress, as well as percentiles covered by these thresholds for 1999-2019, or sub-periods with available data. Samples are weighted so as to be nationally representative and refer to 18-64 year olds. For all measures, higher values indicate worse mental distress. Abbreviations: NHIS – National Health Interview Survey, MEPS – Medical Expenditure Panel Survey, NHANES – National Health and Nutrition Examination Survey, BRFSS – Behavior Risk Factor Surveillance System, GSS – General Social Survey. For the MEPS, one measure is taken from the core household sample while the others are taken from the subsample completing a self-administered questionnaire. MCS is reverse coded by subtracting 100 from its original value. NHANES data are analyzed over 2-year intervals from 2005-2006 through 2015-2016, and then 2017-2020 (pre-pandemic). POOR MH is coded 1 through 5 for responses of "excellent", "very good", "good", "fair", and "poor". *MD* and *SMD* refer to mental distress and serious mental distress.

Table 2. Potential Consequences of Mental Distress

Measure (Abbreviation)	Source/Years	Details
Suicide Death Rates (SUICIDE)	MCOD, 1999-2019	ICD-10 Codes *U03,X60-X84,Y87.0
Emergency Department Visits with Suicidal Ideation/ Attempt (SUICIDE ED)	HCUP NEDS, 2006-2019	2006-2015: CCS Code 662 2016-2019: CCSR Code MBD012
Saw/Talked to Mental Health Professional in Last 12 Months (MH VISITS)	NHIS, 1999-2018	
Prescribed Psychotropic Medication Use in Last 12 Months (MH DRUGS)	MEPS, 1999-2019	MULTC1S1 Codes: 67, 71, 249

Note: Table shows potential consequences of mental distress. Abbreviations: MCOD – CDC Wonder Multiple Cause of Death files, HCUP NEDS, Healthcare Cost and Utilization Project National Emergency Department Sample, NHIS – National Health Interview Survey, MEPS – Medical Expenditure Panel Survey, ICD-10 – International Classification of Diseases 10th Revision, CCS – Clinical Classification Software, CCSR – Clinical Classification Software Refined, MULTC1S1 – Multum Lexicon Therapeutic subclass codes. SUICIDE and SUICIDE ED are per 100,000. MULTCI1S1 codes are: 67 – Anxiolytics, Sedatives, and Hypnotics; 71 – CNS Stimulants, and 249 – Antidepressants.

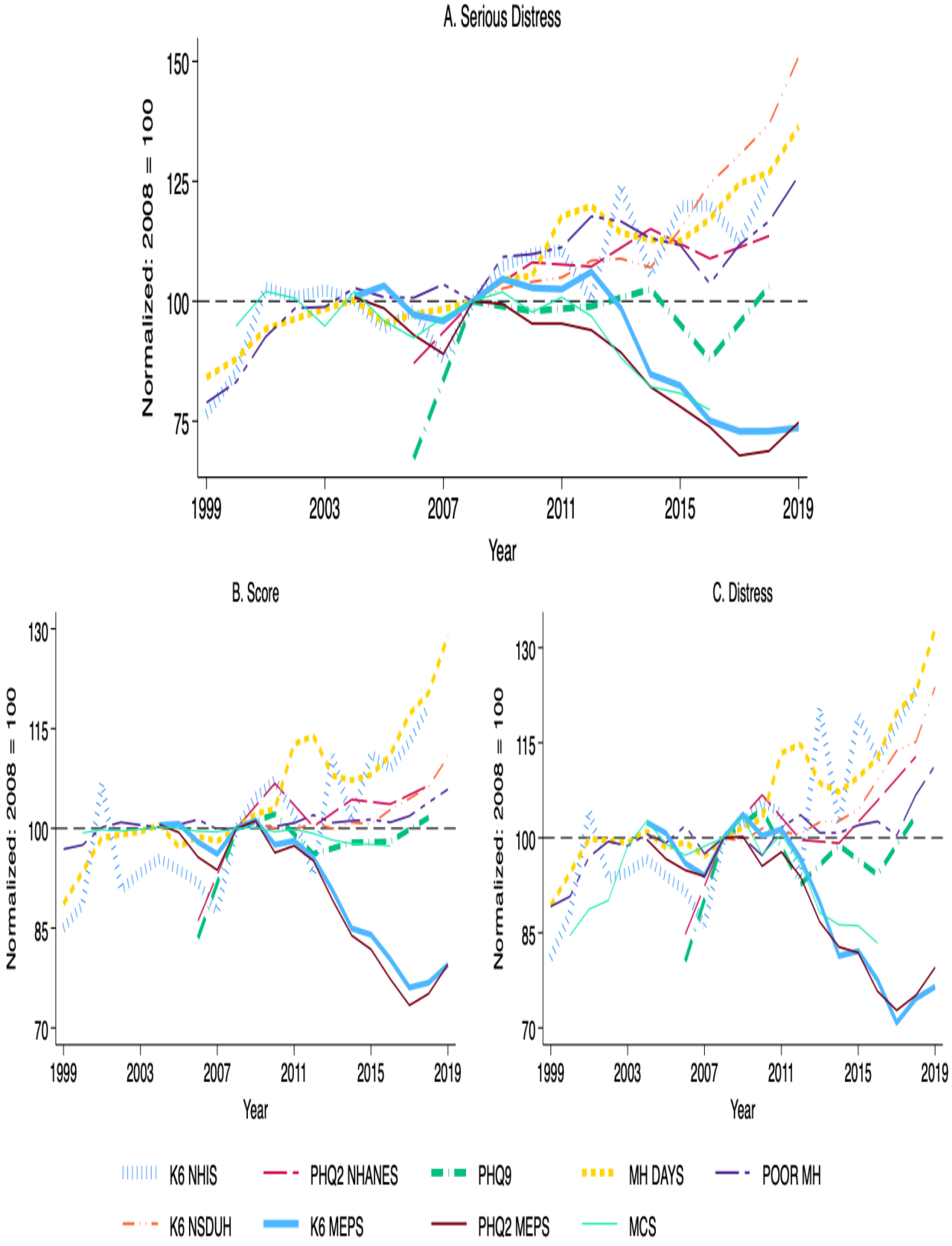
Table 3: Average Correlation Coefficients Between Selected Mental Distress Measures and Potential Consequences of Poor Mental Health

Mental Distress Measure	Potential Consequences of Poor Mental Health: All	Leave-One-Out Minimum	Leave-One-Out Maximum
MH DAYS	0.875	0.862	0.898
K6 NHIS	0.753	0.742	0.774
K6 NSDUH	0.748	0.707	0.851
POOR MH	0.710	0.686	0.724
PHQ2	0.595	0.573	0.651
PHQ9	0.406	0.371	0.495

Functional Form	Mental Distress Measures Included: All	No PHQ2 or PHQ9
SMD	0.700	0.795
MD	0.671	0.759
DS	0.673	0.761

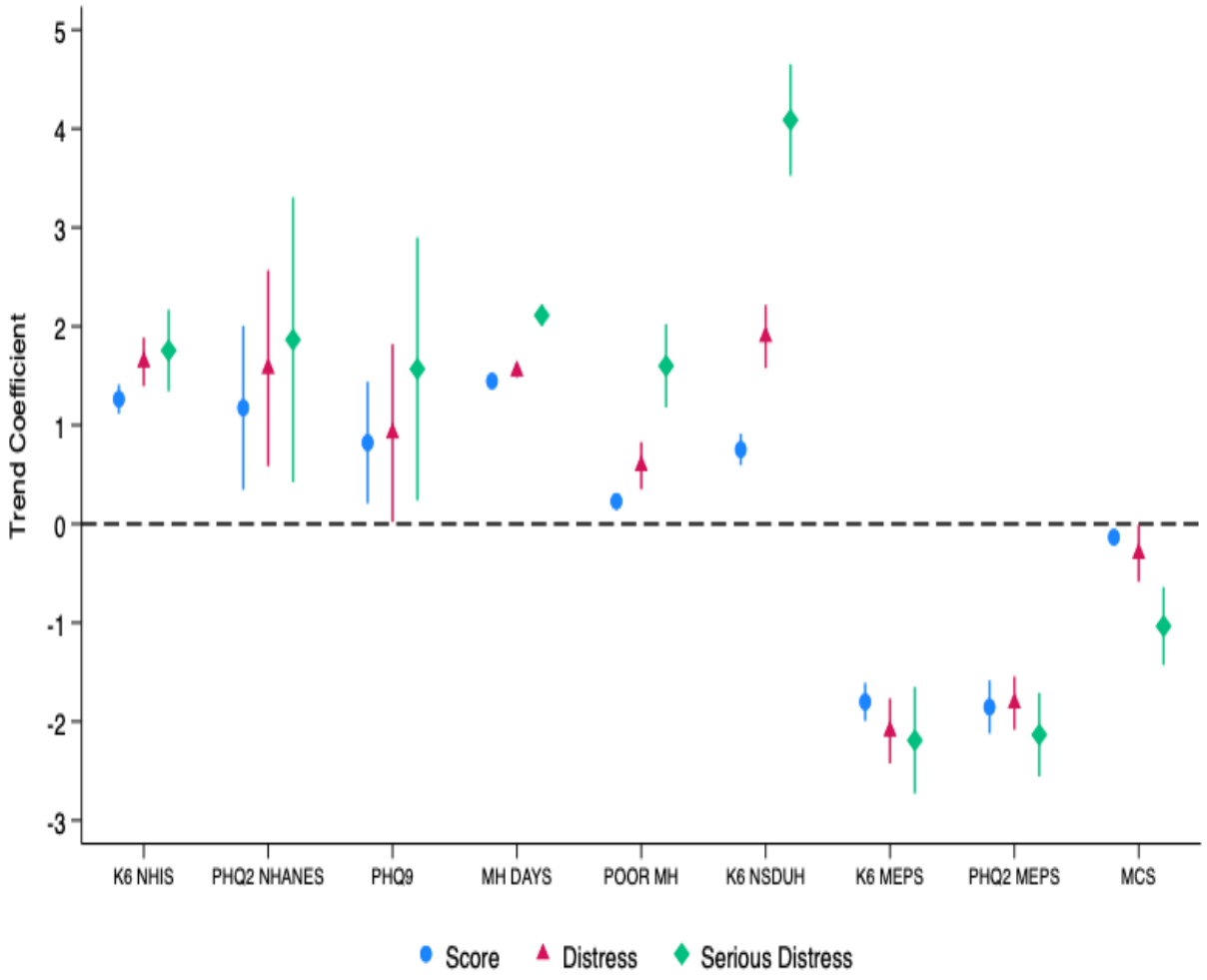
Note: Table shows average correlation coefficients between potential consequences of poor mental health and all functional forms of specified mental distress measure (top panel) or specific functional forms of all mental distress measures, excluding the *MEPS* Questionnaire variables, or all of these measures other than PHQ2 and PHQ9 (lower panel). The Leave-One-Out results in the upper panel indicate minimum and maximum average confidence intervals of the correlation coefficients where one of the potential consequences of poor mental health is excluded.

Figure 1. Normalized Mental Distress Trends



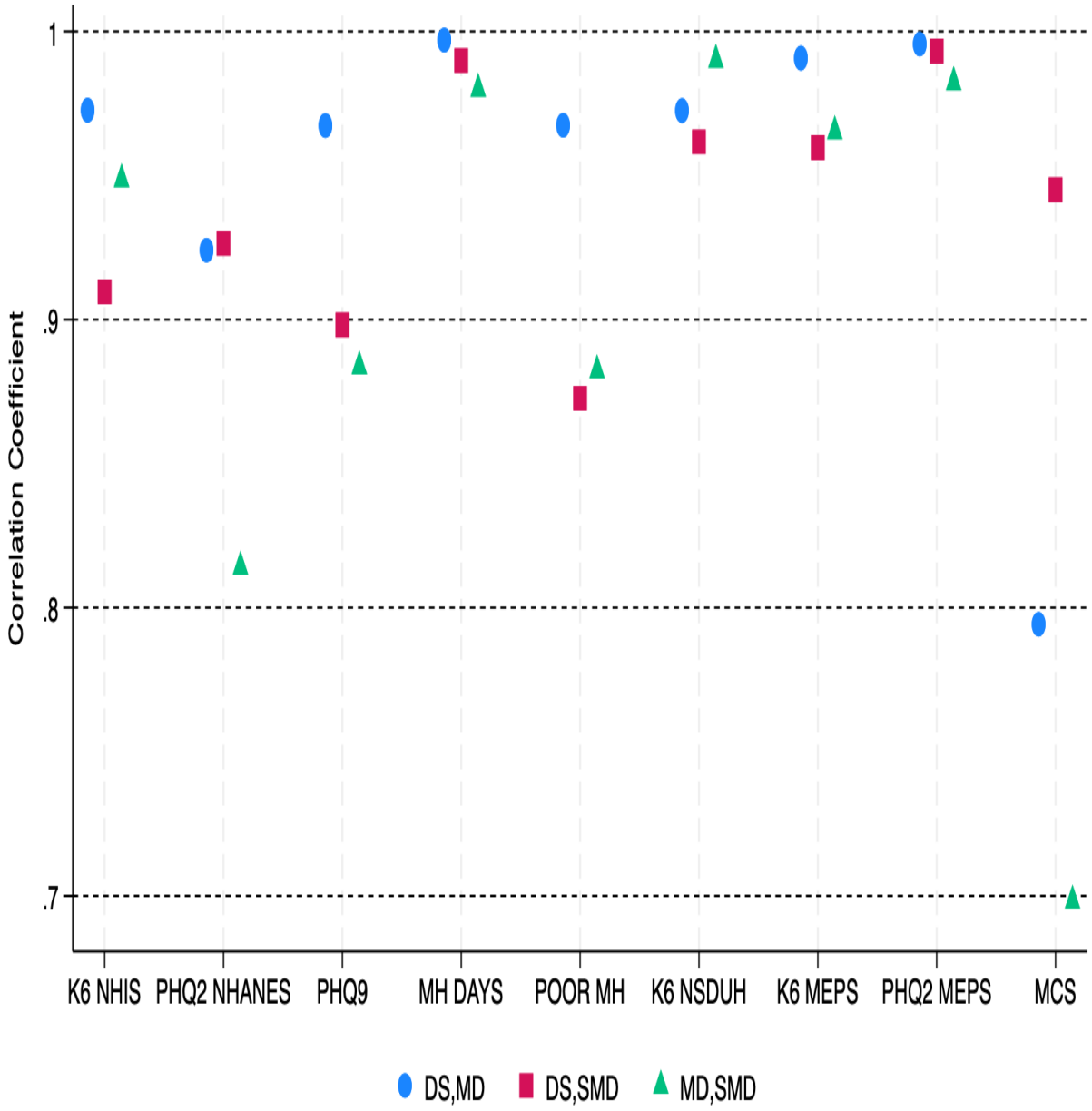
Note: Figure shows values of mental distress measures for 18-64 year olds from 1999-2019, or sub-periods with available data, as a percentage of 2008 values. The measures are obtained from the following data sets: Kessler K6 is from the National Health Interview Survey (NHIS), the National Survey on Drug Use and Health (NSDUH), and the Medical Expenditure Panel Survey (MEPS); Physician Health Questionnaire 2 and 9 scales (PHQ2 and PHQ9) are from the National Health and Nutrition Examination Survey (NHANES), with the PHQ2 also from the MEPS; poor mental health days in the last 30 (MH DAYS) are from the Behavior Risk Factor Surveillance System (BRFSS); self-rated poor/fair mental health (POOR MH) and the Mental Health Component Score (MCS) are from the Medical Expenditure Panel Survey (MEPS). All measures are coded so that higher values indicate greater mental distress. Additional details are provided in Table 1.

Figure 2. Regression Estimates of Trends in Normalized Measures of Mental distress from Individual Level Data



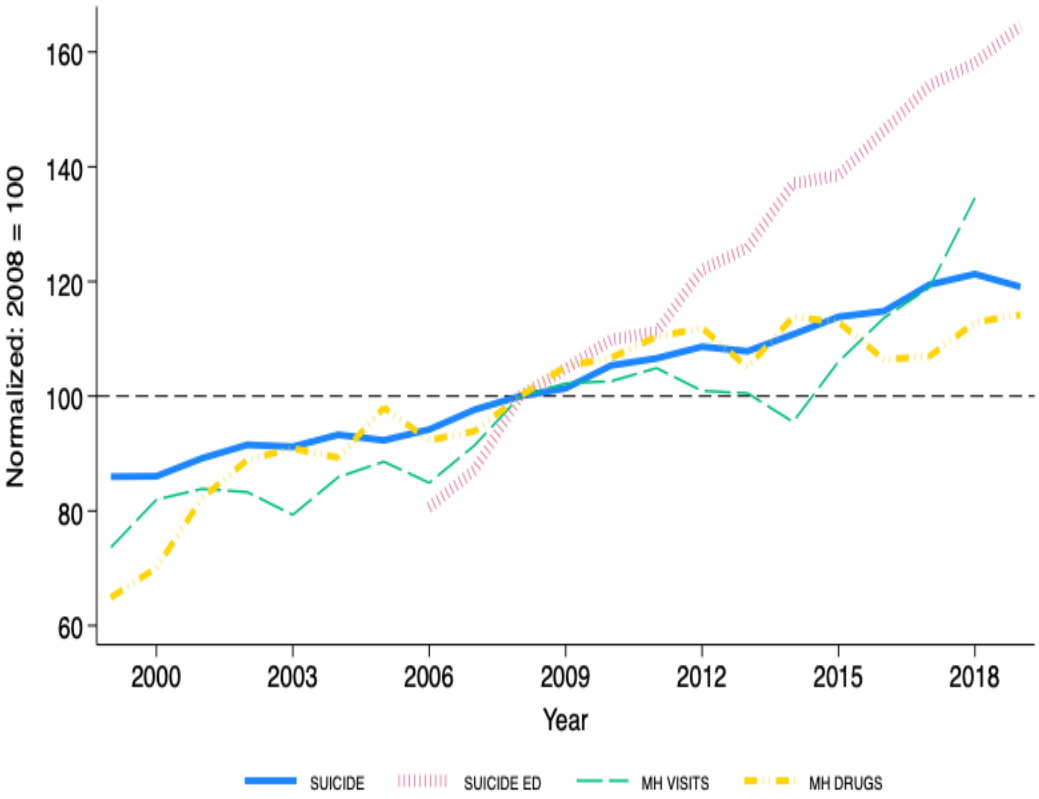
Note: Figure shows estimates of the coefficient from regressing the specified normalized mental distress measure as a percentage of 2008 values against a linear annual time trend for 18-64 year olds from 1999-2019, or sub-periods with available data, based on individual-level data. Error bars show 95% confidence intervals, adjusted for complex survey design. The measures and sources are described in Table 1. All scores are coded so that higher values indicate worse mental distress, as described in the text.

Figure 3. Correlation Coefficients Between Functional Forms of Mental Distress Indicators



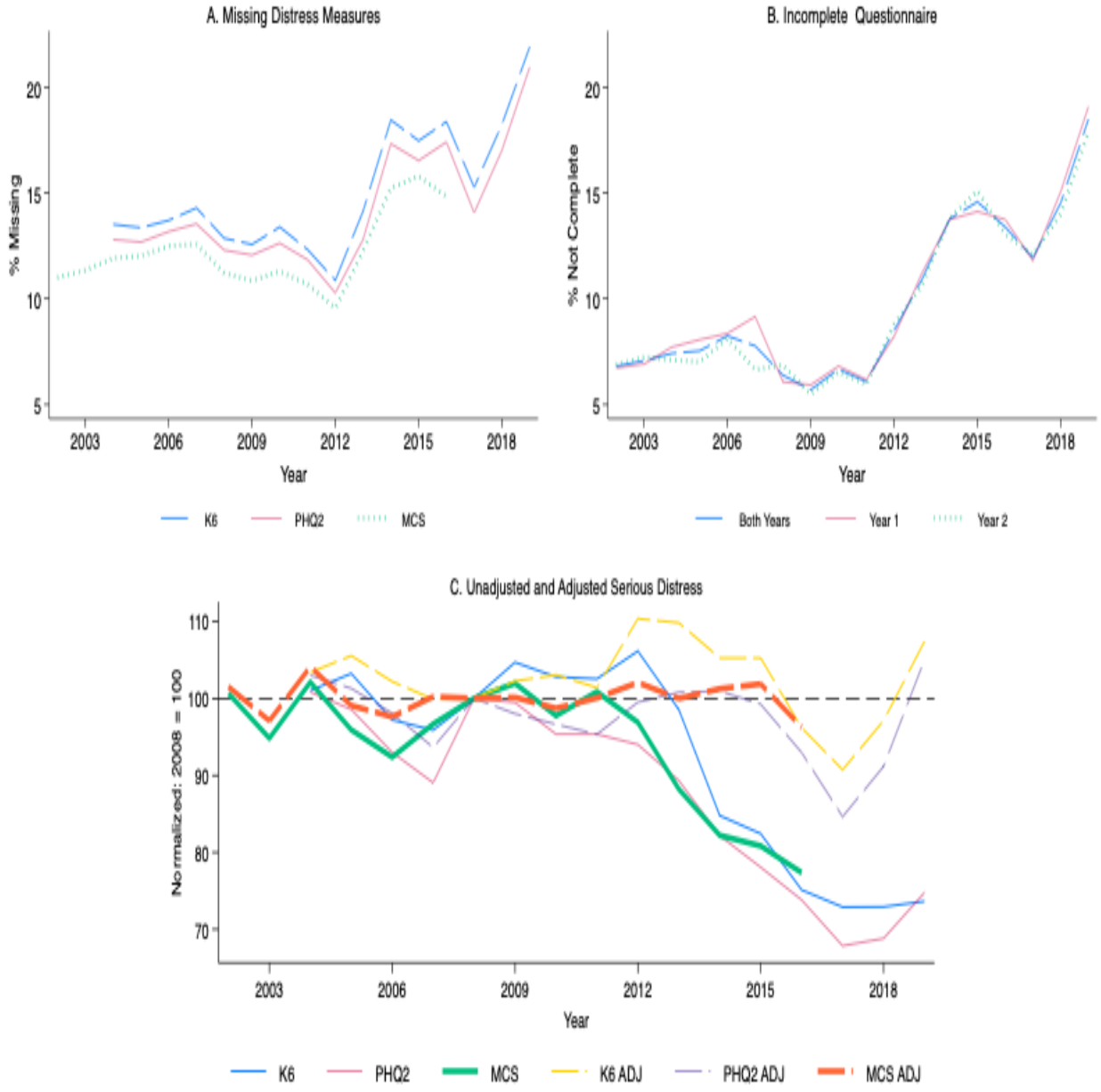
Note: Figure shows correlation coefficients between the average Distress Score (*DS*), Mental Distress (*MD*), and Serious Mental Distress (*SMD*) using the specified indicator for 18-64 year olds from 1999-2019, or sub-periods with available data. The measures and sources are described in Table 1. All scores are coded so that higher values indicate worse mental distress.

Figure 4. Trends in Normalized Potential Consequences of Mental Distress



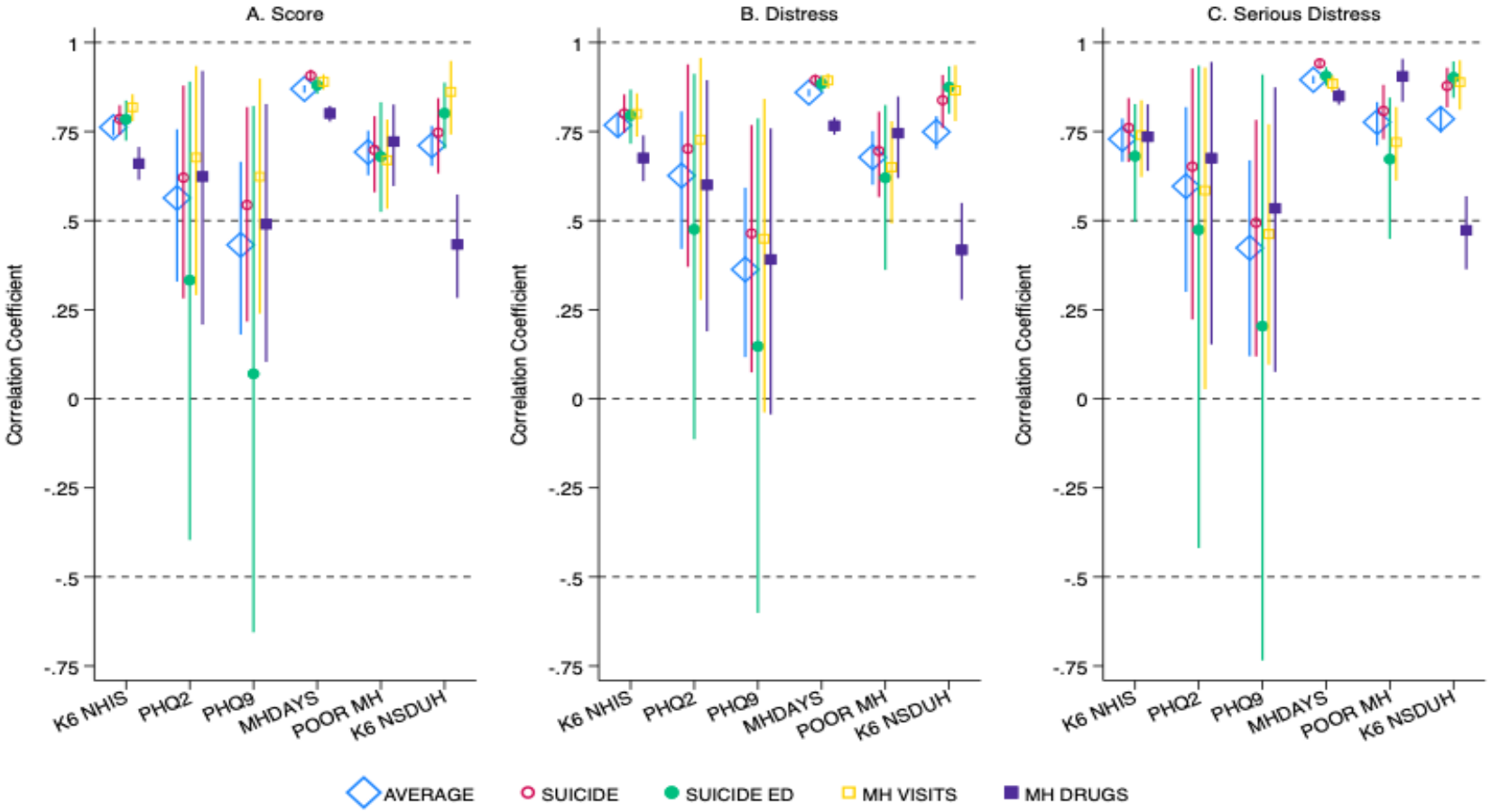
Note: Figure shows potential consequences of mental distress for 18-64 year olds from 1999-2019, or sub-periods with available data, as a percentage of 2008 values. The potential consequences are: suicide death rates (SUICIDE) obtained from the CDC Wonder Multiple Cause of Death (MCOD) data, emergency department visit rates with a diagnosis of suicidal ideation or attempts (SUICIDE ED) from the HCUP Nationwide Emergency Department Sample (NEDS), the share of persons reporting having seen or talked to a mental health professional in the last 12 months (MH VISITS) from the National Health Interview Survey (NHIS), and the share of individuals reporting use of prescribed psychotropic medications in the last year (MH DRUGS) using data from the Medical Expenditure Panel Survey (MEPS). ED visit rates in 2016 are interpolated as the average of 2015 and 2017 rates.

Figure 5. Missing MEPS Questionnaire Distress Measures, Self-Administered Questionnaires Not Completed, and Adjustments for Questionnaire Non-Completion



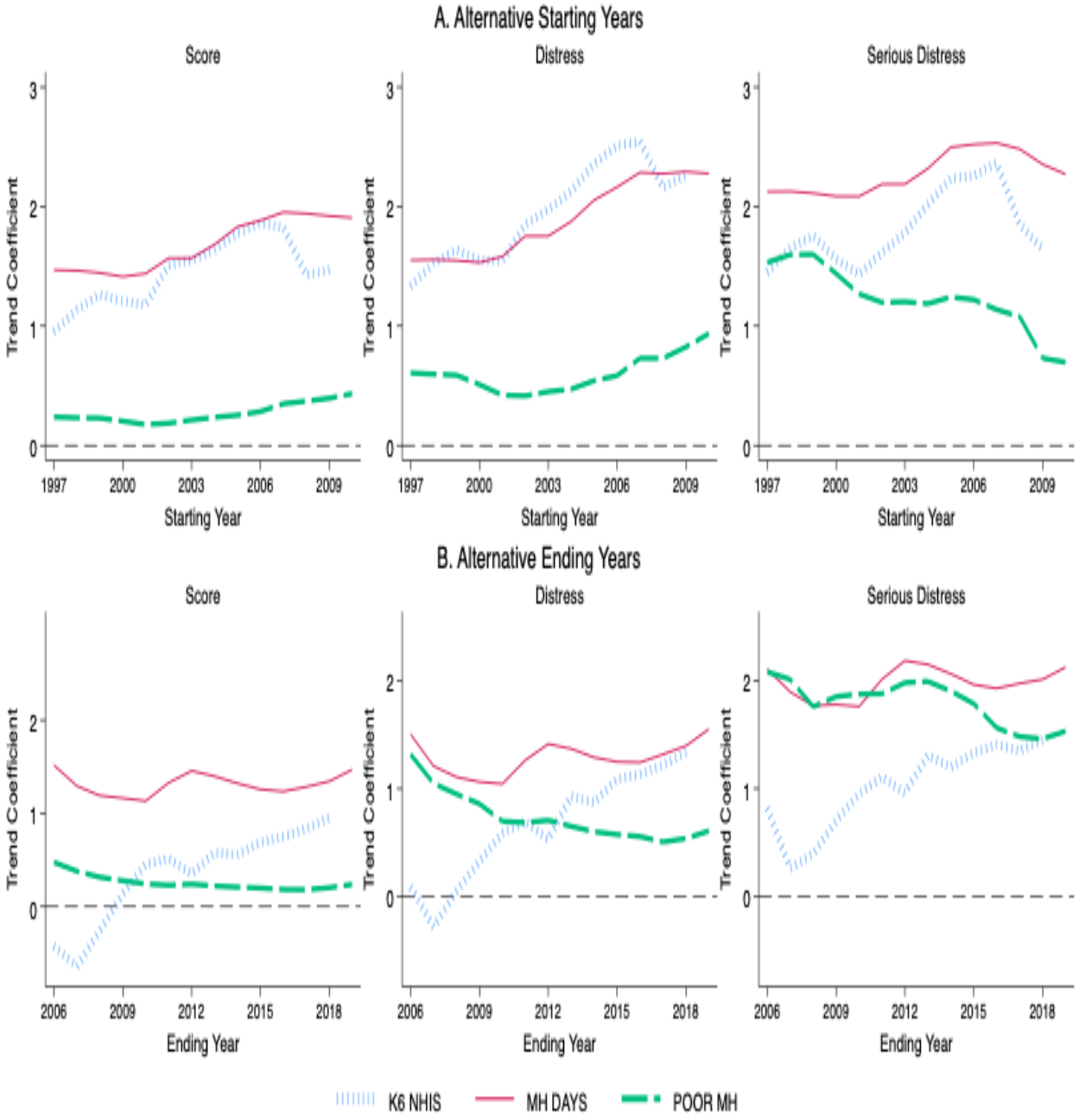
Note: Panel A. shows the percentage of eligible respondents with missing data for the MEPS K6, PHQ2, and MCS measures. Panel B. displays the percentage of eligible MEPS respondents not completing the self-administered questionnaire (SAQ), with Years 1 and 2 referring to the first and second year the respondent participates in the MEPS. Panel C. indicates values of the MEPS K6, PHQ2, and MCS serious mental distress for 18-64 year olds as a percentage of 2008 value; adjusted (ADJ) results refer to those predicted if all eligible respondents had returned the SAQ. Measures and sources are described in Table 1.

Figure 6. Correlation Coefficients Between Selected Measures and Potential Consequences of Poor Mental Health



Note: Figure shows correlation coefficients between mental distress measures, excluding those from the *MEPS* questionnaire data, and potential consequences of poor mental health for 18-64 year olds from 1999-2019, or sub-periods with available data. The distress measures are described in Table 1 and the potential consequences in Table 2. AVERAGE refers to the mean of the correlation coefficients over the four potential mental health consequences for the specified mental distress measure. Scores are coded so that higher values indicate worse mental health. Estimates are obtained through bootstrapping with 1,000 replications. Error bars show 95% confidence intervals.

Figure 7. Regression Estimates of Selected Trends in Mental Distress from Individual Level Data Using Alternative Analysis Periods



Note: Figure A. shows estimates of the coefficient from regressing the specified normalized measure of mental distress on a linear time trend for 18-64 year olds where the ending year is 2019 (2018 for the K6 NHIS) and the starting year is shown on the X-axis. Figure B. shows corresponding estimates where the starting year is 1997 and the ending year is shown on the X-axis. Shadings show 95% confidence intervals, adjusted for complex survey design. The measures and sources are described in Table 1.

Supplemental Appendices for "Estimating Trends in Mental Distress Using Survey Data"

Appendix A: Details on Mental Distress Measures and Potential Consequences of Poor Mental Health, Plus Additional Tables and Figures

Additional Details on Mental Distress Measures

K6

The Kessler K6 is based on responses to six questions asking whether, in the last 30 days, the respondent felt: "so sad nothing could cheer you up", "nervous", "restless or fidgety", "hopeless", "that everything was an effort", and "worthless". Each response is scored on a 0-4 scale corresponding to: "none of the time", "a little of the time", "some of the time", "most of the time", or "all of the time". Scores for the six responses are then summed to provide a total ranging from 0 to 24. This analysis used the conventional threshold defining *SMD* as scores of 13 or greater (Kessler et al. 2003; Mojtabai and Jorm 2015; Olfson et al. 2019; Tomitaka et al. 2019; Daly 2022; Jackson et al. 2023). When interested in more moderate levels of mental distress, researchers have used lower cutpoints such as K6 $\geq$5 (Prochaska et al. 2012; Wright et al. 2025) or 6-12 (Mojtabai and Jorm 2015). This analysis classified *MD* as scores of 8 and above.

PHQ2 and *PHQ9*

The PHQ2 is based on the responses to two questions asking how often in the last two weeks the individual was bothered by: "having little interest or pleasure in doing things", and "feeling down, depressed, or hopeless". The PHQ9 uses these questions plus seven others, also referring to the last two weeks, on: "trouble falling or staying asleep, or sleeping too much", "feeling tired or having little energy", "poor appetite or overeating", "feeling bad about yourself — or that you are a failure or have let yourself or your family down", "trouble concentrating on things, such as reading the newspaper or watching

television", "moving or speaking so slowly that other people could have noticed – or the opposite, being so fidgety or restless that you have been moving around a lot more than usual", "thoughts that you would be better off dead, or thoughts of hurting yourself in some way". Answers are scored from 0 to 3 for "not at all", "several days", "more than half the days", or "nearly every day". Aggregating the scores gives a range from 0 to 6 for the PHQ2 and 0 to 27 for the PHQ9. When using the PHQ2, scores of 3 and higher are often taken to indicate major depression,(Kroenke et al. 2003; Mitchell et al. 2016; Tápias et al. 2021) although ≥ 2 has also been used (Levis et al. 2020). For the PHQ9, scores of 10 or higher have been used to indicate serious depression, (Mitchell et al. 2016; Brody et al. 2018; Levis et al. 2020), with cutpoints of 5, 10, 15, and 20 sometimes indicating mild, moderate, moderately severe, and severe depression (Kroenke et al. 2001), and with other studies suggesting thresholds between 8 and 11 (Manea et al. 2012; Mitchell et al. 2016). For this analysis $\geq$2 and $\geq$3 and $\geq$6 and $\geq$10 respectively indicate *MD* and *SMD* for the PHQ2 and PHQ9.

MCS

The MCS is derived from the Medical Outcomes Study 12-Item Short Form Health Survey (SF12). The SF12 contains responses to 12 questions related to: 1) general health rating, 2) moderate activities limitations, 3) limitations climbing stairs, 4) pain interfering with normal work, 5) physical health interfering with work or daily activities, 6) physical pain limits kind of work/other activities, 7) felt calm/peaceful, 8) felt downhearted/blue, 9) accomplished less at work/other activities due to emotional problems, 10) did work/other activities less carefully due to emotional problems, 11) had lots of energy, 12) physical health/emotional problems interfered with social activities. Responses to questions 7-10 are particularly important for the MCS. The final scores are normed to have a mean of 50 and a standard deviation of 10 for a reference population (1990 for

2000-2003 and 1998 for later years), with higher scores indicating lower levels of depression. Data from the 2000-2002 *MEPS* use version 1 of the SF12 and those from 2003-2016 use version 2. To make these fully comparable, 0.019394 was subtracted from the untransformed version 1 (2000-2002) scores (Blewett et al. 2024). To reverse the scale, so that higher scores indicated greater mental distress, the scores were subtracted from 100. Vilagut et al. (2013) estimate an optimal cutpoint for detecting depressive disorders in the (untransformed) MCS of 45.6, although this was for six European (non-US) countries that had average scores of between 52 and 55, raising questions about applicability to the US. In this analysis the reverse-coded thresholds of $\geq$56.18 and $\geq$62.79 classify *MD* and *SMD*, which correspond to untransformed ranges of $\leq$43.82 and $\leq$37.2.

MH DAYS

The number of poor mental health days in the last 30 has been used, in various forms, to examine trends in mental distress (Blanchflower and Oswald 2020; Ruhm 2025) or changes during the COVID pandemic (Kessler et al. 2022). Blanchflower and Oswald (2020) utilize a dichotomous variable indicating poor mental health in all 30 days as the distress measure; Kessler et al. (2022) focuses on $\geq$15 days; Ruhm (2025) analyzes both of these, as well as the number of poor mental health days, and Wright et al. (2025) use $\geq$1 day. In this analysis, $\geq$6 and $\geq$16 days define *MD* and *SMD*.

POOR MH

There have been many examinations of self-rated overall health (Idler and Benyamini 1997; DeSalvo et al. 2006) but fewer of self-classified mental health. The available evidence indicates that this mental health variable correlates with previously validated measures such as K6 and PHQ scores, although it may capture not capture all aspects of mental distress (Fleishman and Zuvekas 2007; Ahmad et al. 2014). Some studies have also examined the determinants of self-rated mental health (Choi et al. 2022).

For this analysis "good" or better and "fair" or better indicate *MD* and *SMD*. Note that almost 30% of weighted observations are at or above the *MD* threshold, which is higher than for the other variables. This occurs because there are no responses between this and the 93rd percentile.

Additional Details on Potential Consequences of Poor Mental Health

SUICIDE

These were defined as deaths per 100,000 population with underlying ICD-10 cause of death codes *U03,X60-X84,Y87.0, obtained using vital statistics data from death certificates.

SUICIDE ED

These referred to emergency department admissions, per 100,000 persons, with a diagnosis of suicide ideation or attempt. Diagnoses are based on the ICD-9-CM coding system for 2006-2015 and the ICD-10-CM system for 2016-2019. Specifically, from 2006-2015, they referred to ED visits with a primary or secondary diagnosis of suicide or intentional self-inflicted injury, categorized using the Clinical Classification Software (CCS Code 662). For 2016-2019 they were for visits with a diagnosis of suicidal ideation/attempt/self-harm, categorized using Clinical Classification Software Refined (CCSR Code MBD012). ED visits in 2016 were likely to have been under-reported because hospitals needed time to transition to the new and more complicated coding system (Owens et al. 2020). Therefore, rates in that year were computed by averaging 2015 and 2017 values.

MH DRUGS

The prescribed psychotropic medications included in this variable referred to antidepressants, CNS stimulants, and anxiolytics/sedatives/hypnotics. These were

primarily prescribed to treat depression, ADD/ADHD, and anxiety/sleep disorders. Antipsychotics, and antimanic agents were not included because the types of mental illness they treat (e.g. schizophrenia, bipolar, and psychotic disorders) were likely to primarily have a genetic component, rather than being mainly caused by mental distress or depression.

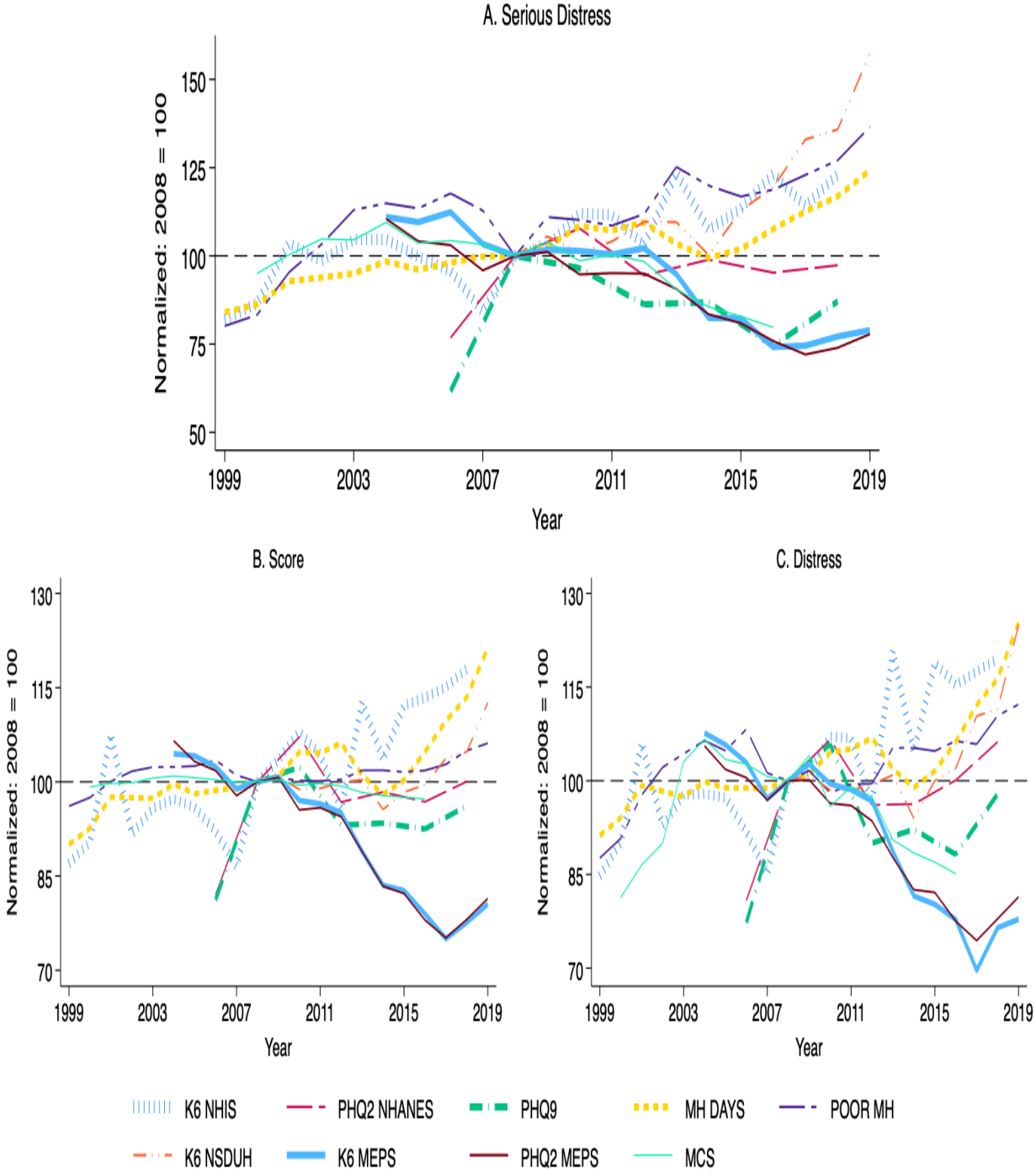
Additional Tables and Figures

Table A1. Correlation Coefficients Between Potential Consequences of Mental Distress

Outcome	SUICIDE	SUICIDE ED	MH VISITS	MH DRUGS
SUICIDE	1.000			
SUICIDE ED	0.985	1.000		
MH VISITS	0.932	0.820	1.000	
MH DRUGS	0.877	0.792	0.782	1.000

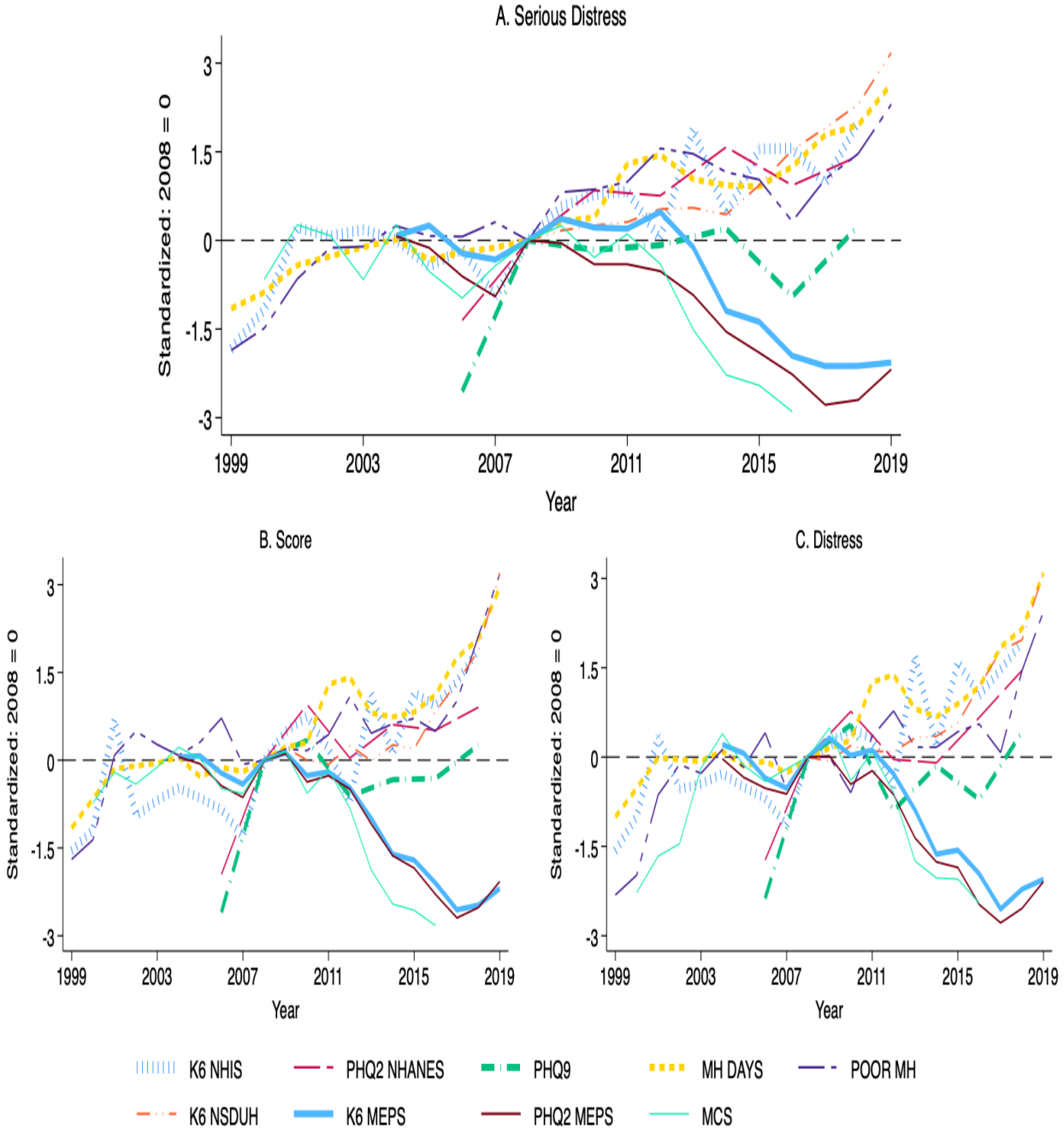
Note: Table shows correlation coefficients between alternative potential consequences of mental distress for 18-64 year olds from 1999-2019, or sub-periods with available data. The potential consequences and data sources are described in Table 2.

Figure A1. Normalized Mental Distress Trends Without Population Weights



Note: Figure shows values of mental distress measures for 18-64 year olds from 1999-2019, or sub-periods with available data, as a percentage of 2008 means, calculated using unweighted data. The measures and data sources are described in Figure 1. All measures are coded so that higher values indicate worse mental health, as described in the text.

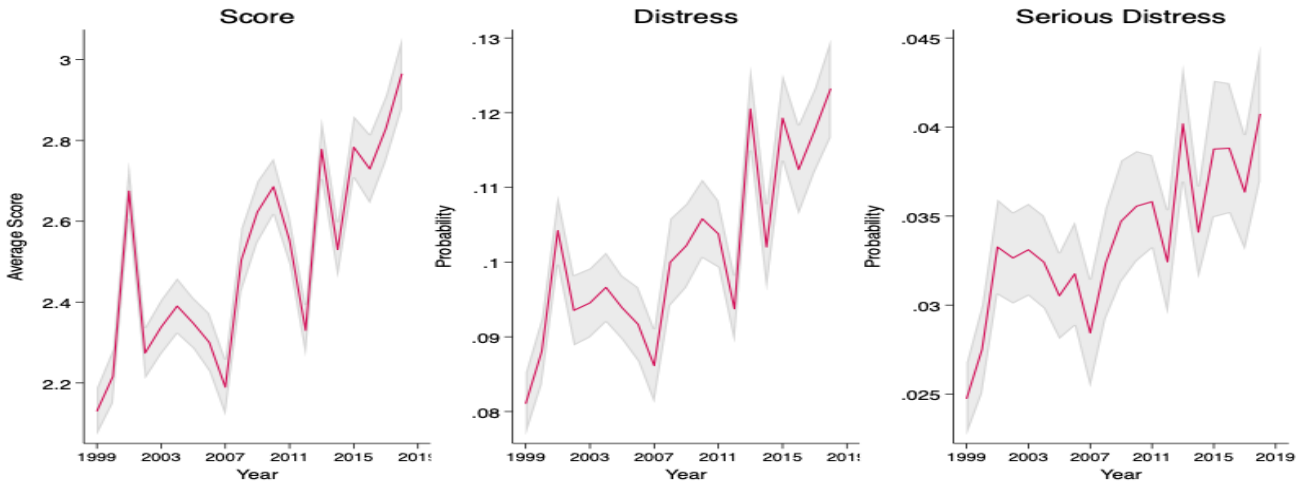
Figure A2. Standardized Mental Distress Trends



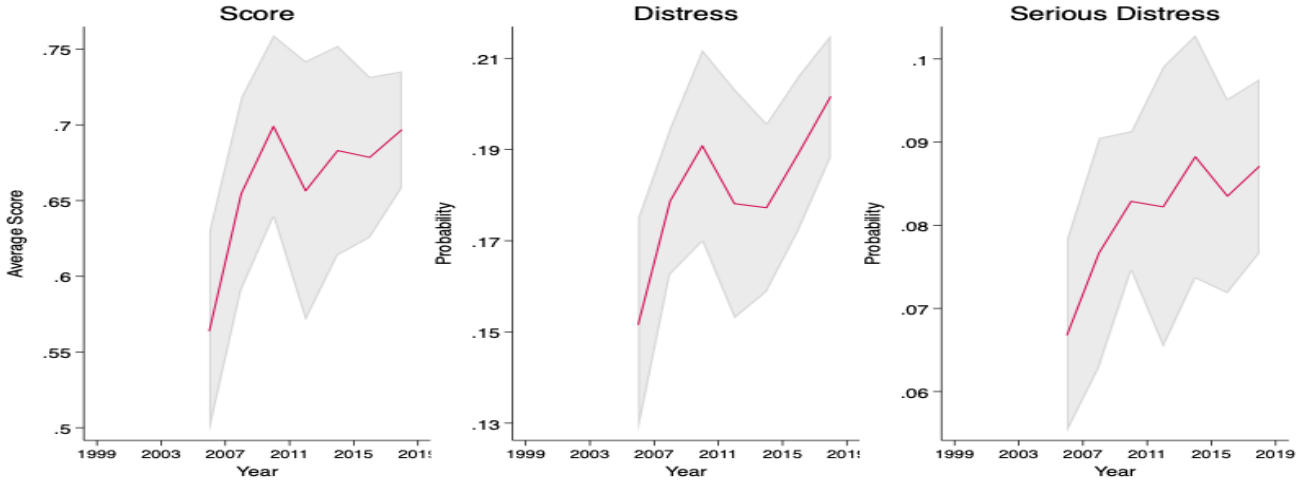
Note: Figure shows standard deviation differences of mental distress measures from 2008 means for 18-64 year olds from 1999-2019 or sub-periods with available data. The measures and data sources are described in Figure 1. All measures are coded so that higher values indicate worse mental health, as described in the text.

Figure A3. Trends in Mental Distress Measures with Confidence Intervals

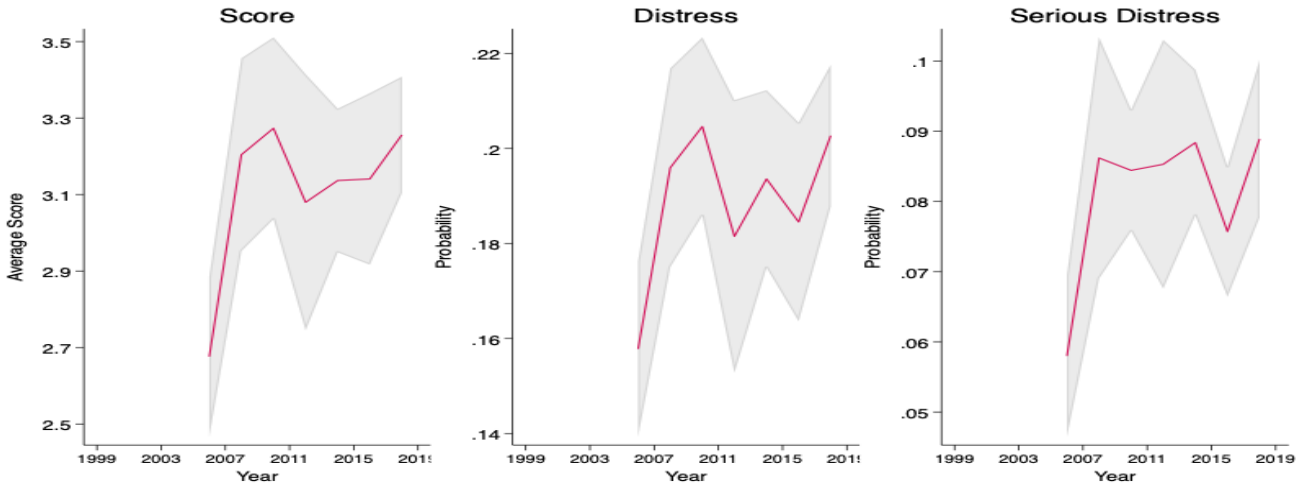
a. K6 NHIS



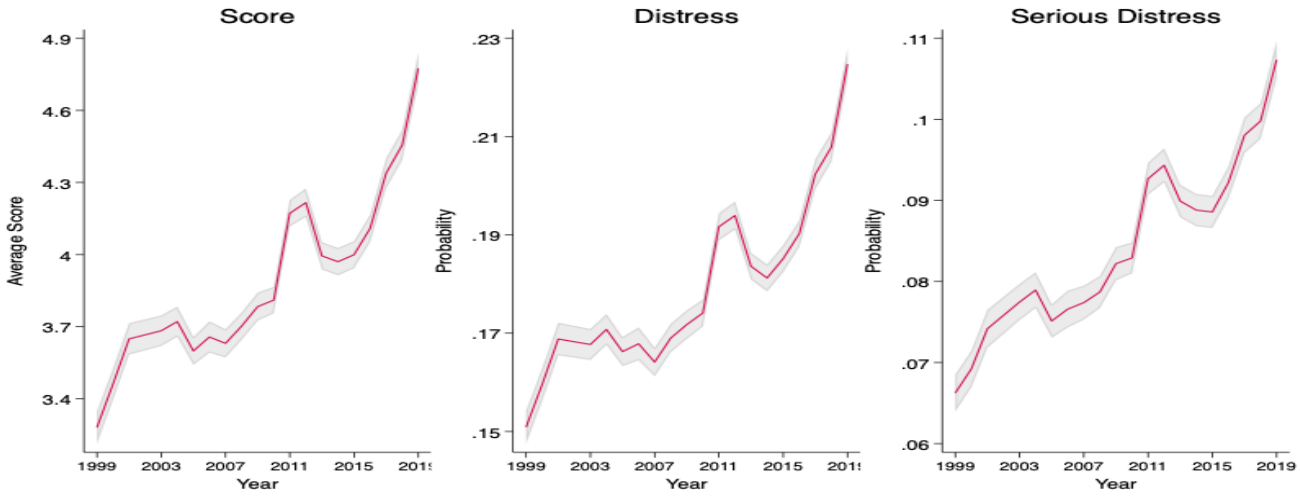
b. PHQ2 NHANES



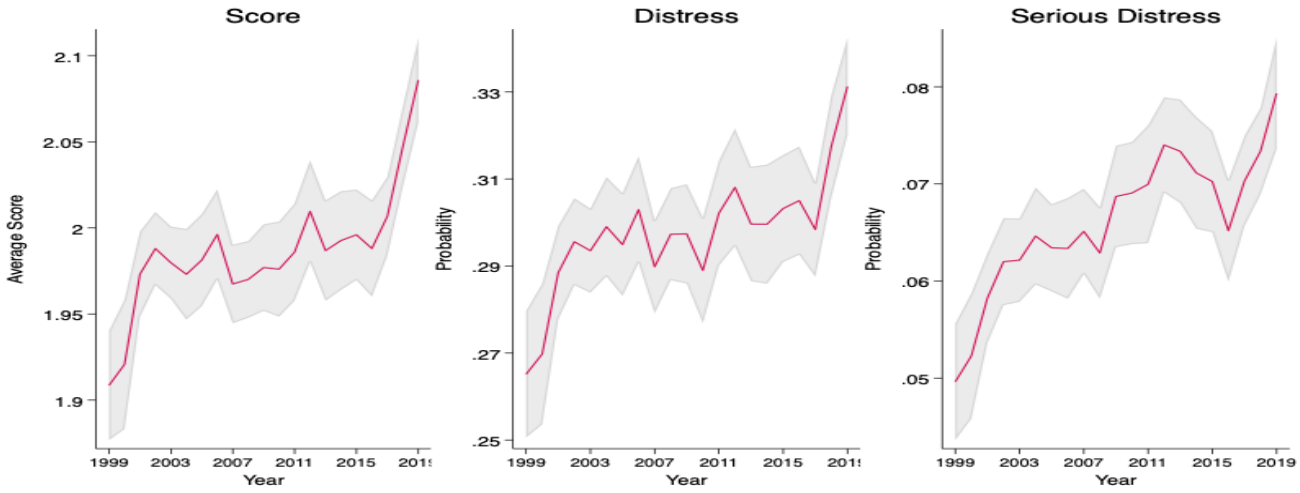
c. PHQ9



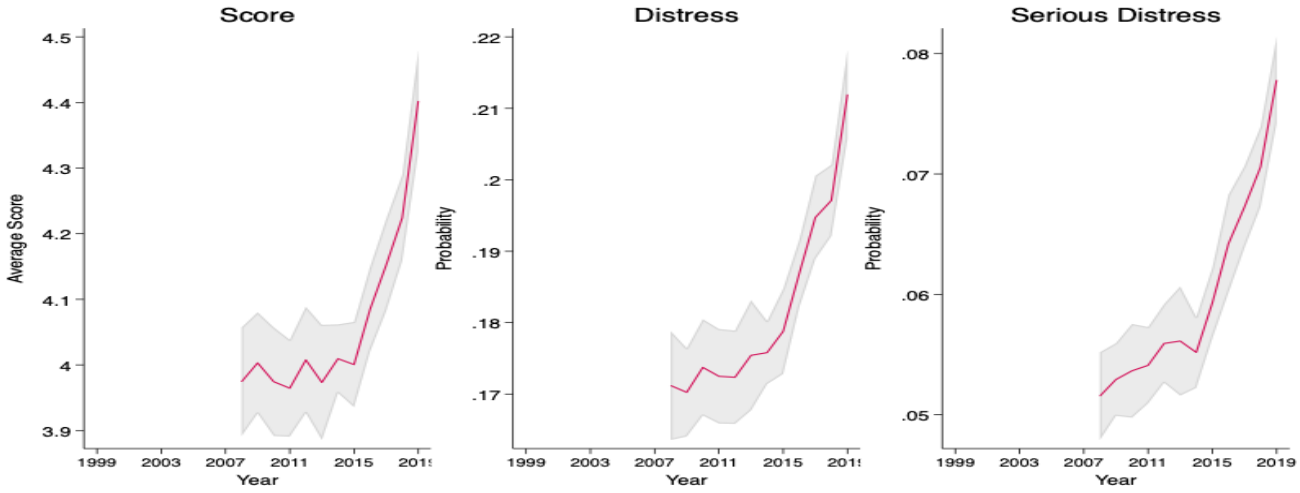
d. MH DAYS



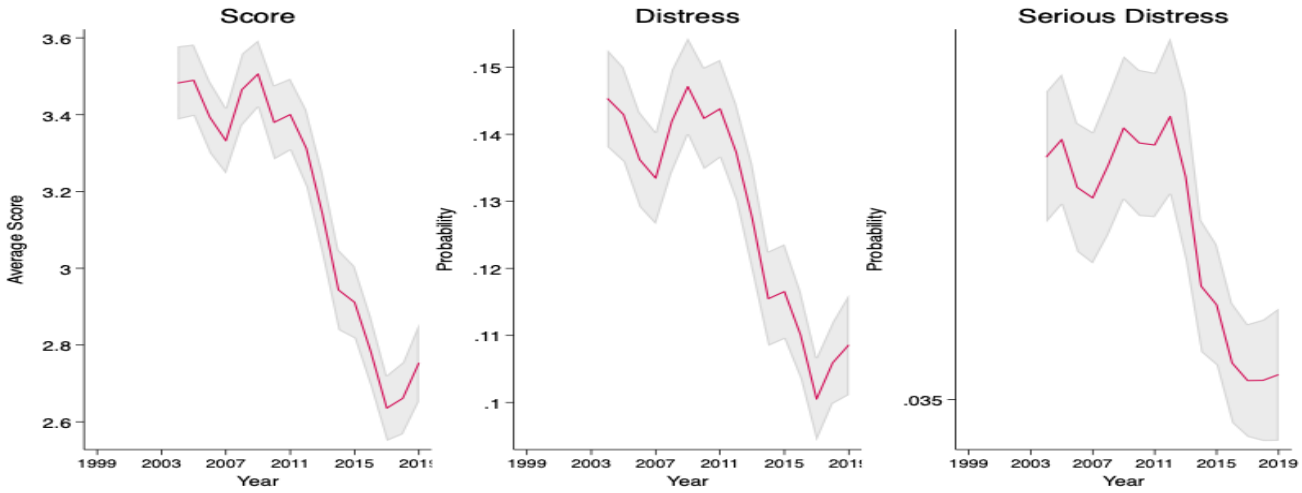
e. POOR MH



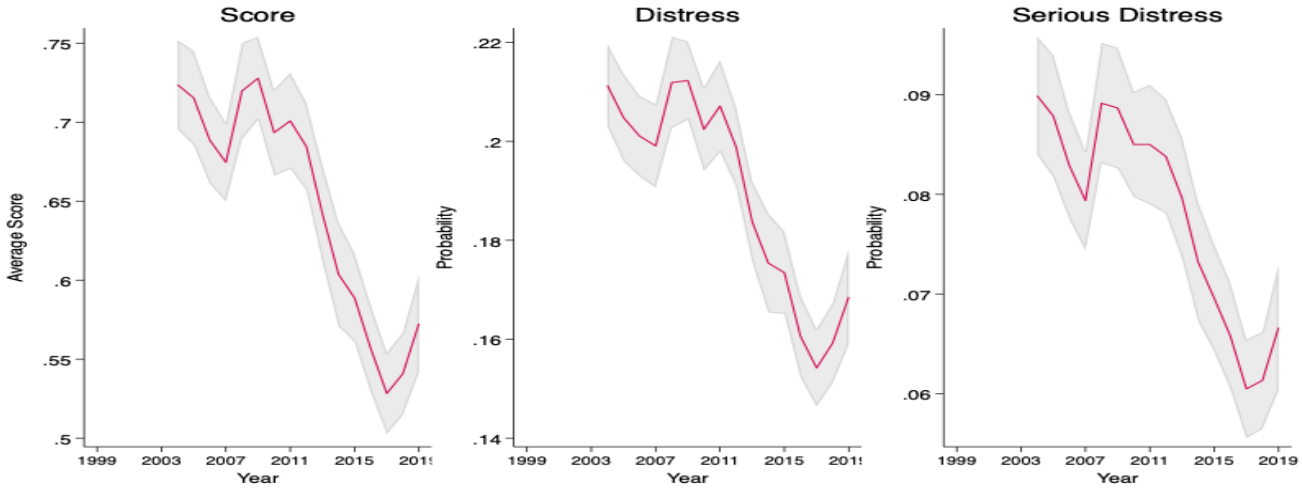
f. K6 NSDUH



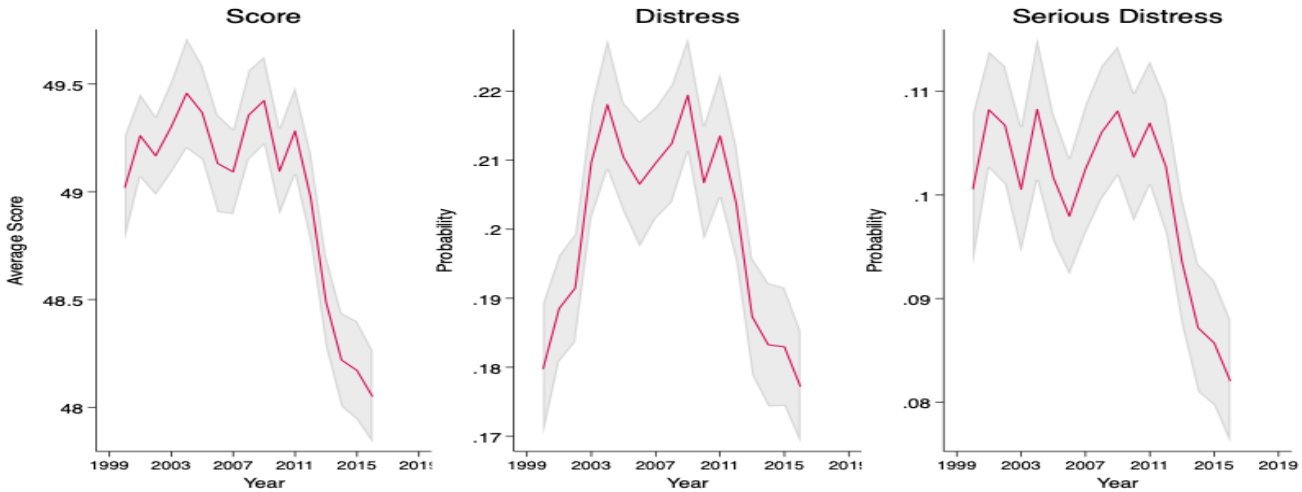
g. K6 MEPS



h. PHQ2 MEPS

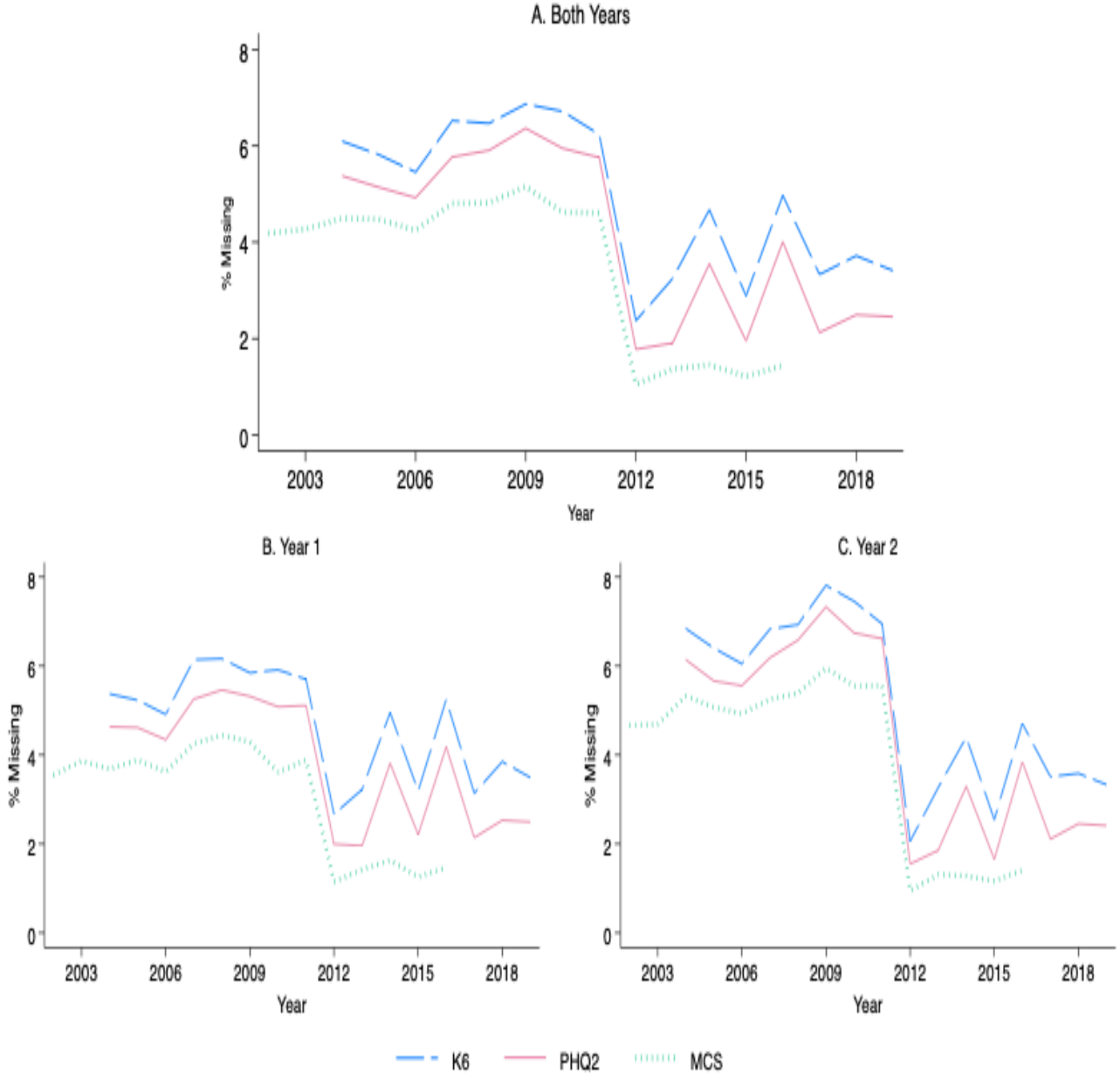


i. MCS



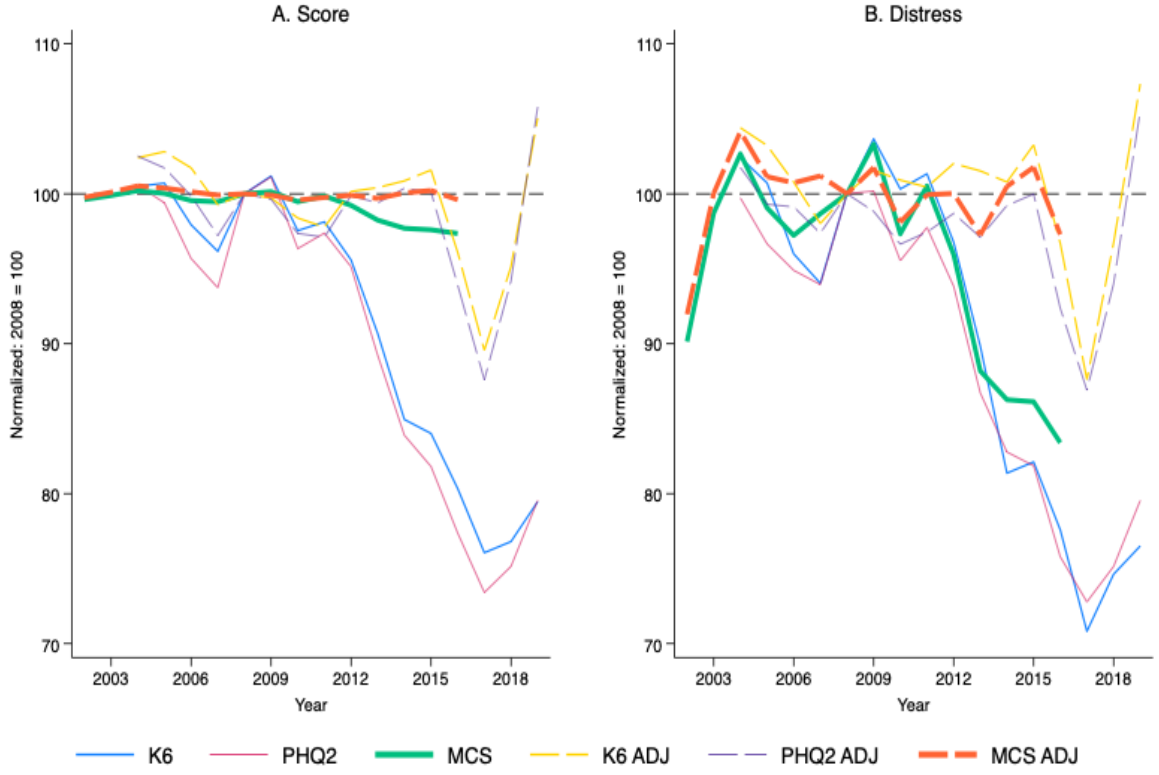
Notes: Figure displays average scores and probabilities of *DS*, *MD*, and *SMD* for 18-64 year olds from 1999-2019, or sub-periods with available data. The measures and sources are described in Table 1. Shaded areas show 95% confidence intervals based on standard errors adjusted for complex survey design.

Figure A4. Percent of Missing Observations for MEPS Questionnaire Measures Conditional on Completion of Self-Administered Questionnaire



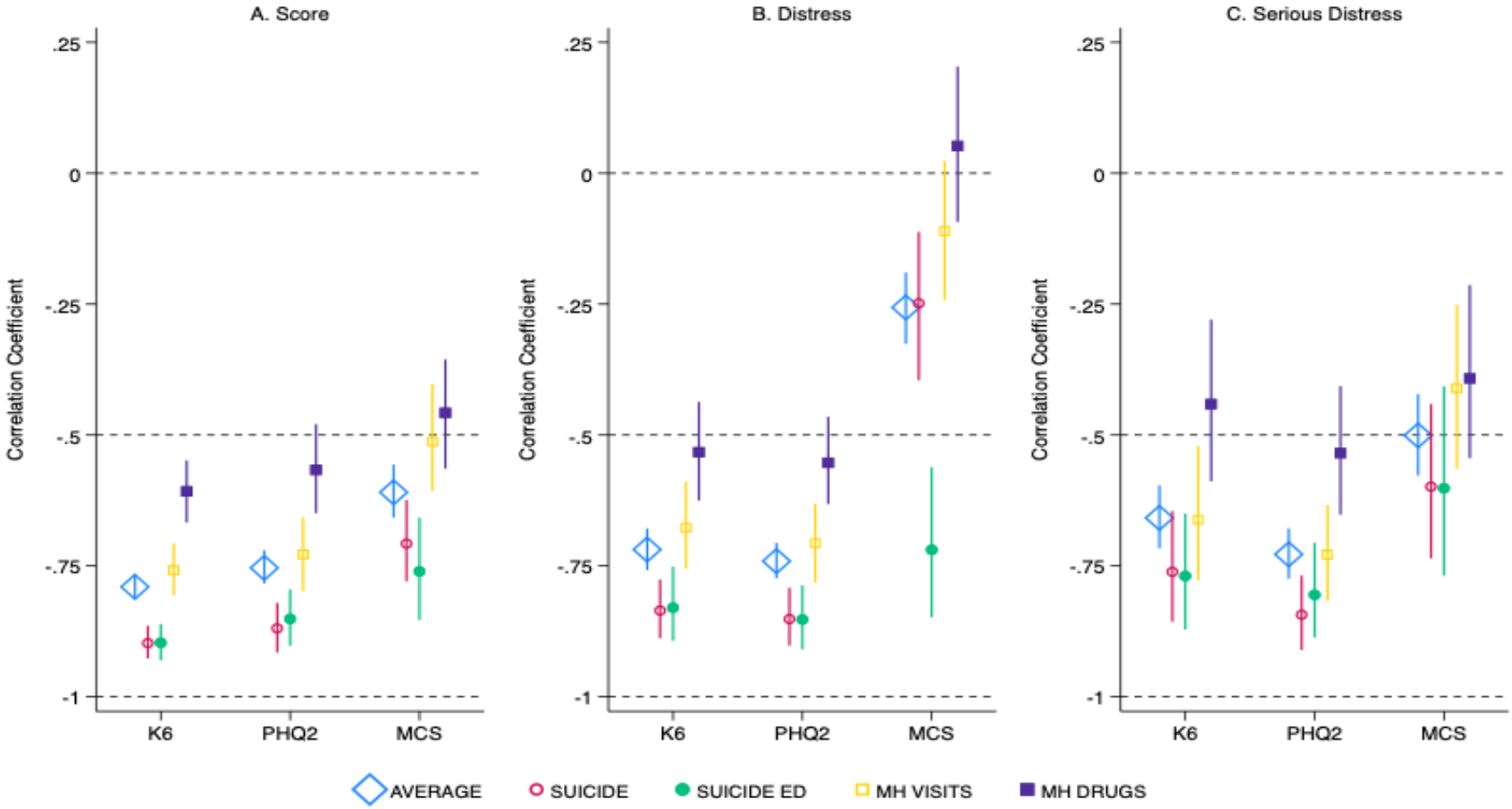
Note: Figure shows percent of observations with missing data for the MEPS Questionnaire mental distress measures conditional on respondents having completed and returned the self-administered questionnaire. Years 1 and 2 refer to the first and second year the respondent participates in the MEPS.

Figure A5. Normalized Mental Distress Measures With and Without Adjustment for MEPS Self-Administered Questionnaire Completion Status



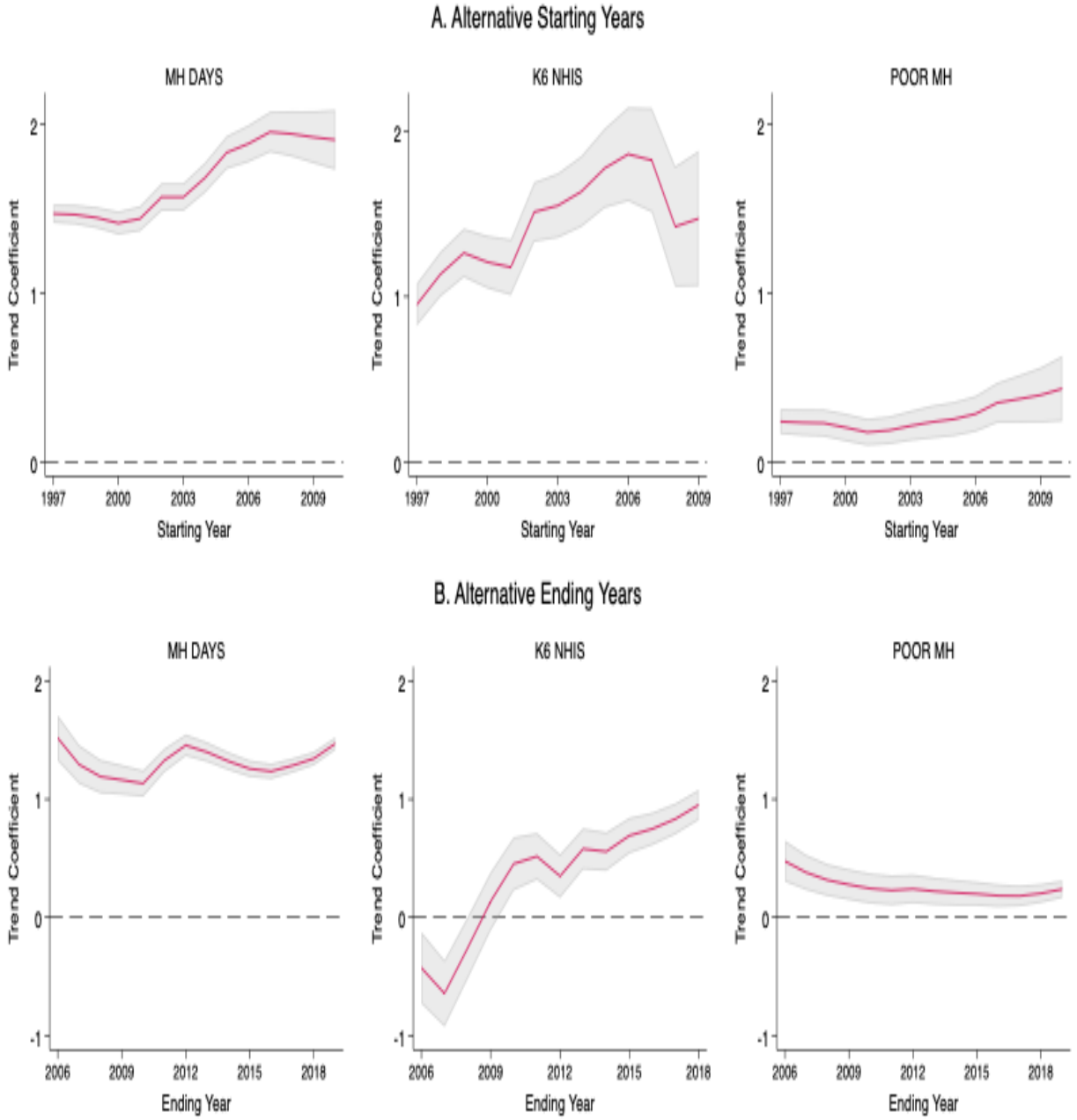
Note: Figure displays *DS* and *MD* values of MEPS questionnaire measures for 18-64 year olds, as a percentage of 2008 value, with the adjusted (ADJ) results referring to those predicted assuming that all eligible respondents had completed and returned the self-administered questionnaire. The measures and sources are described in Table 1. All scores are coded so that higher values indicate worse psychological distress.

Figure A6. Correlations Between MEPS Questionnaire Measures and Potential Consequences of Poor Mental Health



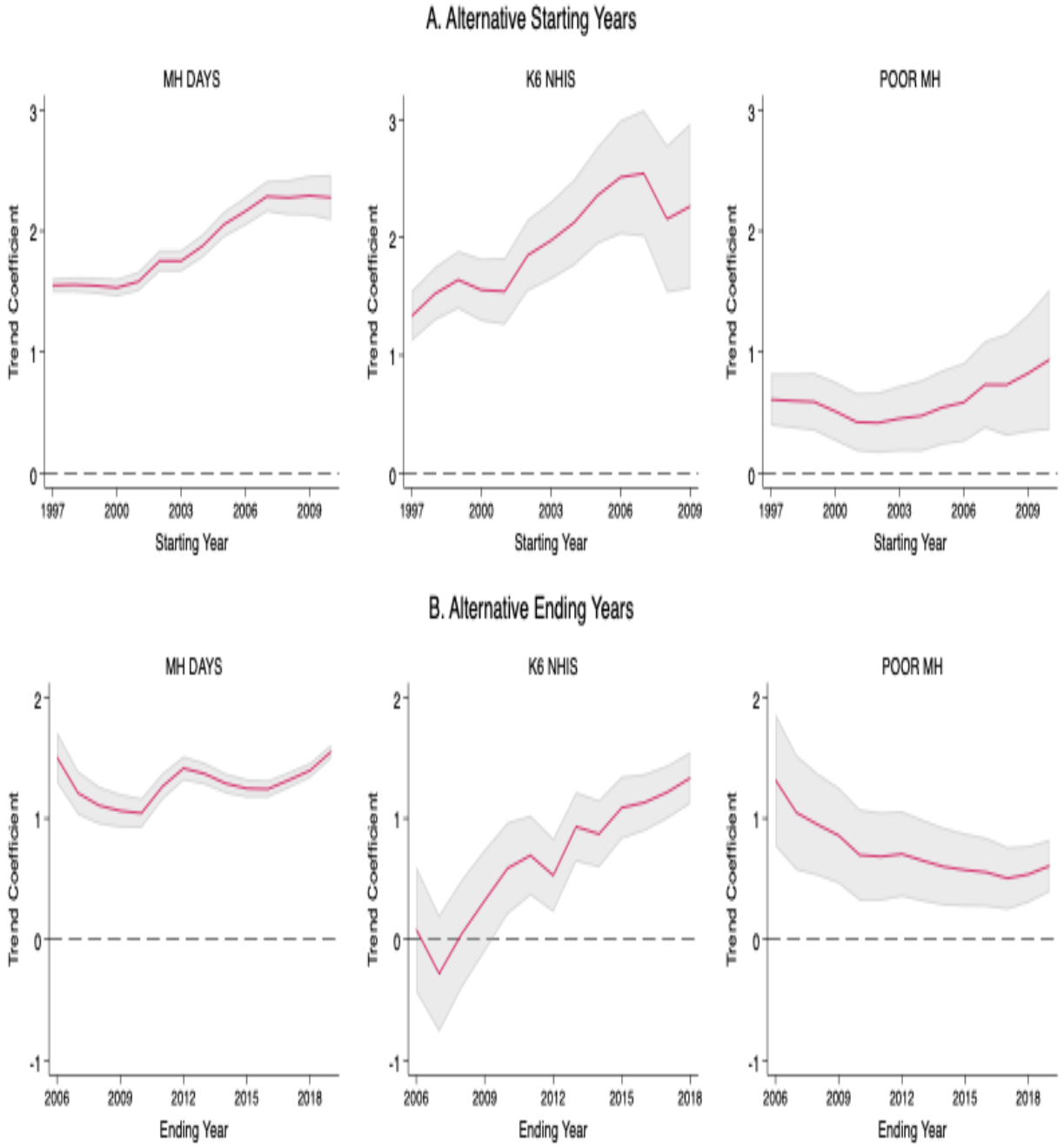
Note: Figure shows correlation coefficients between the mental distress measures from the MEPS Questionnaire data and potential consequences of poor mental health for 18-64 year olds from 1999-2019, or sub-periods with available data. The distress measures are described in Table 1 and the potential consequences in Table 2. AVERAGE refers to the mean of the correlation coefficients over the four potential consequences of poor mental health for the specified mental distress measure. Scores are coded so that higher values indicate worse mental distress. Estimates are obtained through bootstrapping with 1,000 replications. Error bars show 95% confidence intervals.

Figure A7. Regression Estimates of Trends in Distress Scores from Individual Level Data Using Alternative Analysis Periods



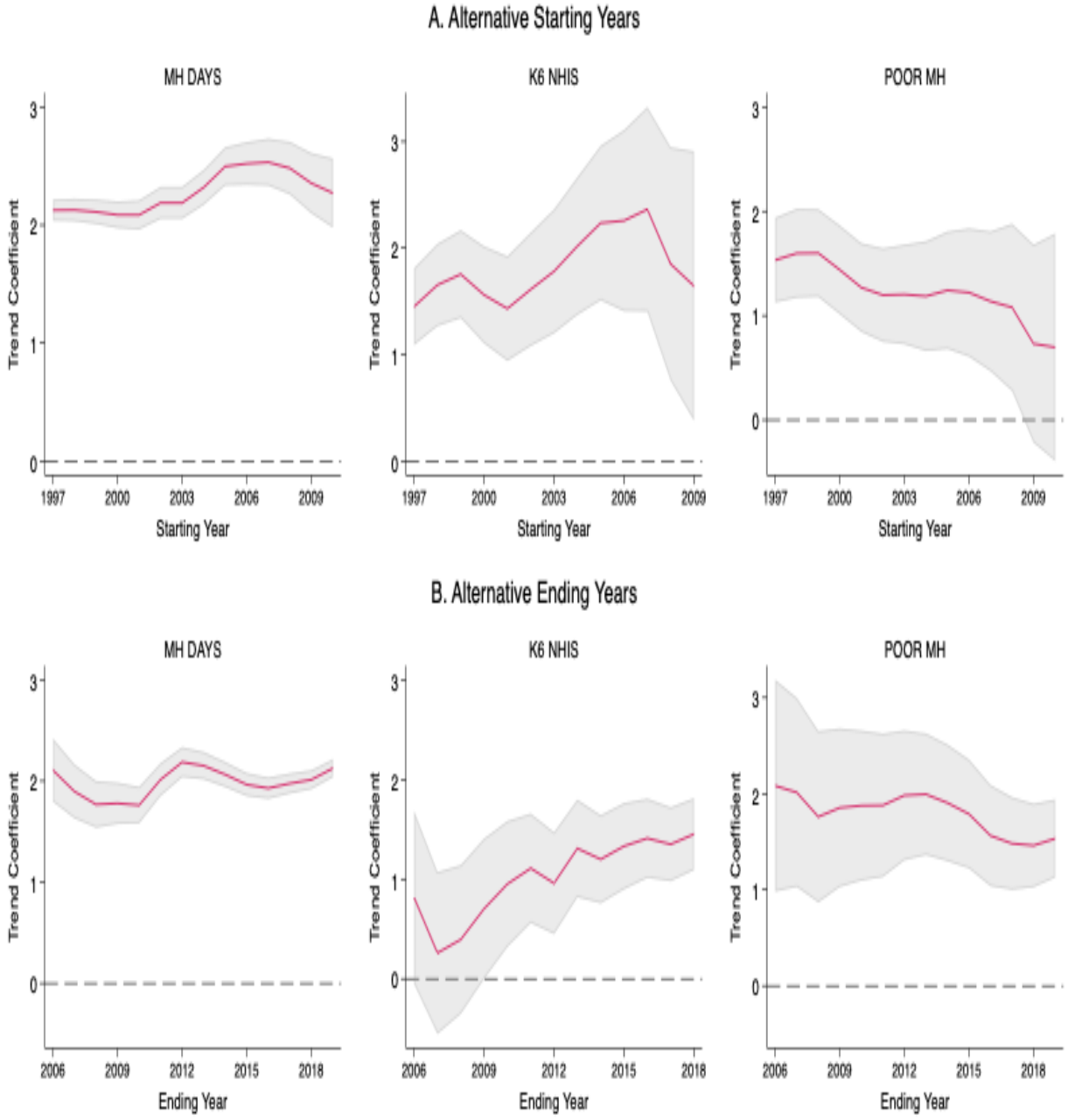
Note: Figure A. shows estimates of the coefficient from regressing the specified normalized *DS* measure on a linear time trend for 18-64 year olds where the ending year is 2019 (2018 for the K6 NHIS) and the starting year is shown on the X-axis. Figure B. shows corresponding estimates where the starting year is 1997 and the ending year is shown on the X-axis. Shadings show 95% confidence intervals, adjusted for complex survey design. The measures and sources are described in Table 1.

Figure A8. Regression Estimates of Trends in Mental Distress from Individual Level Data Using Alternative Analysis Periods



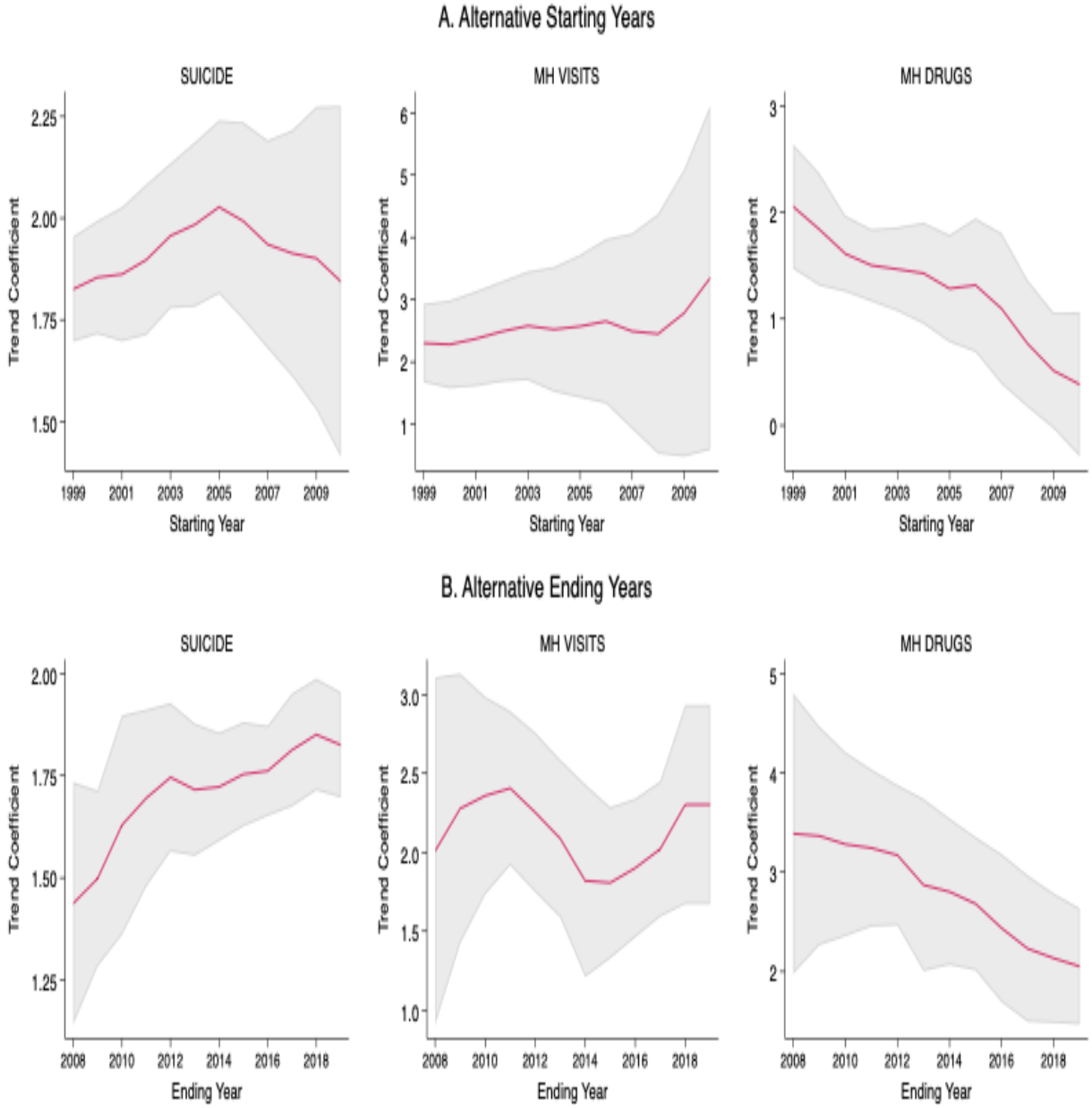
Note: Figure A. shows estimates of the coefficient from regressing the specified normalized *MD* measure on a linear time trend for 18-64 year olds where the ending year is 2019 (2018 for the K6 NHIS) and the starting year is shown on the X-axis. Figure B. shows corresponding estimates where the starting year is 1997 and the ending year is shown on the X-axis. Shadings show 95% confidence intervals, adjusted for complex survey design. The measures and sources are described in Table 1.

Figure A9. Regression Estimates of Trends in Serious Mental Distress from Individual Level Data Using Alternative Analysis Periods



Note: Figure A. shows estimates of the coefficient from regressing the specified normalized *SMD* measure on a linear time trend for 18-64 year olds where the ending year is 2019 (2018 for the K6 NHIS) and the starting year is shown on the X-axis. Figure B. shows corresponding estimates where the starting year is 1997 and the ending year is shown on the X-axis. Shadings show 95% confidence intervals, adjusted for complex survey design. The measures and sources are described in Table 1.

Figure A10. Regression Estimates of Trends in Potential Consequences of Mental Distress Using Aggregated Data and Alternative Analysis Periods



Note: Figure A. shows estimates of the coefficient from regressing the annual mean value of the specified normalized *CMH* measure on a linear time trend for 18-64 year olds where the ending year is 2019 and the starting year is shown on the X-axis. Figure B. shows corresponding estimates where the starting year is 1999 and the ending year is shown on the X-axis. Shadings show 95% confidence intervals, based on robust standard errors. The measures and sources are described in Table 2.

Appendix B: Additional Tables with Details on Figures

Table B1: Normalized Values of Mental Distress Measures, By Year

YEAR	K6 NHIS	PHQ2 NHANES	PHQ9	MH DAYS	POOR MH	K6 NSDUH	K6 MEPS	PHQ2 MEPS	MCS
Score									
1999	85.0			88.6	96.9				
2000	88.5			93.5	97.5				99.3
2001	106.8			98.6	100.2				99.8
2002	90.8				100.9				99.6
2003	93.4			99.5	100.5				99.9
2004	95.4			100.5	100.2		100.5	100.6	100.2
2005	93.7			97.2	100.6		100.7	99.4	100.0
2006	91.8	86.1	83.5	98.8	101.3		97.9	95.7	99.5
2007	87.4			98.1	99.9		96.2	93.7	99.5
2008	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0
2009	104.7			102.2	100.4	100.7	101.2	101.1	100.1
2010	107.2	106.8	102.2	102.9	100.3	100.0	97.5	96.3	99.5
2011	101.9			112.7	100.8	99.8	98.1	97.4	99.9
2012	93.0	100.3	96.1	113.9	102.0	100.8	95.6	95.1	99.2
2013	110.9			107.9	100.9	100.0	90.7	89.2	98.2
2014	101.0	104.3	97.9	107.3	101.1	100.9	84.9	83.9	97.7
2015	111.1			108.0	101.3	100.7	84.0	81.8	97.6
2016	109.0	103.6	98.0	111.0	100.9	102.8	80.3	77.4	97.4
2017	113.0			117.2	101.9	104.5	76.1	73.4	
2018	118.4	106.4	101.6	120.4	103.9	106.3	76.8	75.1	
2019				129.0	105.9	110.8	79.5	79.5	
Mental Distress									
1999	81.1			89.3	89.2				
2000	88.0			94.4	90.7				84.6
2001	104.2			99.9	97.0				88.7
2002	93.6				99.4				90.1
2003	94.6			99.2	98.7				98.7
2004	96.6			101.0	100.6		102.4	99.7	102.7
2005	93.9			98.4	99.2		100.7	96.7	99.1
2006	91.7	84.8	80.5	99.3	101.9		96.0	94.9	97.2
2007	86.2			97.1	97.5		94.0	93.9	98.7
2008	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0
2009	102.2			101.6	100.0	99.4	103.7	100.2	103.3
2010	105.8	106.8	104.4	103.0	97.2	101.5	100.3	95.5	97.3
2011	103.8			113.4	101.6	100.8	101.3	97.8	100.5
2012	93.8	99.7	92.6	114.8	103.6	100.7	96.8	93.8	95.9
2013	120.5			108.6	100.8	102.5	89.8	86.7	88.2
2014	102.0	99.2	98.8	107.2	100.8	102.7	81.4	82.8	86.3
2015	119.3			109.6	102.0	104.4	82.1	81.9	86.1

2016	112.4	105.9	94.2	112.6	102.6	109.2	77.6	75.8	83.4
2017	117.7			119.7	100.4	113.7	70.8	72.8	
2018	123.3	112.8	103.5	123.0	106.8	115.1	74.6	75.1	
2019				133.0	111.4	123.8	76.5	79.5	
Serious Mental Distress									
1999	76.5			84.2	78.9				
2000	85.1			87.9	83.1				94.8
2001	102.8			94.2	92.6				102.1
2002	100.9				98.6				100.6
2003	102.3			98.4	98.8				94.8
2004	100.2			100.3	102.8		101.0	100.8	102.1
2005	94.4			95.5	100.8		103.2	98.6	95.9
2006	98.2	87.1	67.3	97.3	100.7		97.2	93.0	92.4
2007	87.9			98.3	103.5		95.9	89.0	96.7
2008	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0
2009	107.3			104.4	109.3	102.7	104.7	99.5	102.0
2010	109.9	108.1	98.0	105.3	109.8	104.1	102.8	95.4	97.7
2011	110.7			117.8	111.3	105.0	102.5	95.4	100.9
2012	100.2	107.2	99.0	119.9	117.7	108.5	106.1	94.0	96.9
2013	124.2			114.2	116.7	108.9	98.5	89.3	88.2
2014	105.4	115.1	102.5	112.8	113.1	107.0	84.8	82.2	82.2
2015	119.8			112.5	111.7	115.0	82.4	78.1	80.8
2016	120.0	108.9	87.8	117.1	103.7	124.5	75.1	73.8	77.4
2017	112.3			124.6	111.8	130.5	72.9	67.9	
2018	125.9	113.6	103.2	126.8	116.7	136.9	72.9	68.8	
2019				136.4	126.1	150.9	73.6	74.8	

Note: Table shows mental distress measures for 18-64 year olds from 1999-2019, or sub-periods with available data, as a percentage of 2008 value. The measures and data sources are described in Table 1.

Table B2: Actual Values of Mental Distress Measures, By Year

YEAR	K6 NHIS	PHQ2 NHANES	PHQ9	MH DAYS	POOR MH	K6 NSDUH	K6 MEPS	PHQ2 MEPS	MCS
Score									
1999	2.13			3.28	1.91				
2000	2.22			3.46	1.92				49.02
2001	2.67			3.65	1.97				49.26
2002	2.27				1.99				49.17
2003	2.34			3.68	1.98				49.30
2004	2.39			3.72	1.97		3.48	0.72	49.46
2005	2.35			3.60	1.98		3.49	0.72	49.37
2006	2.30	0.56	2.68	3.66	2.00		3.39	0.69	49.13
2007	2.19			3.63	1.97		3.33	0.67	49.09
2008	2.50	0.65	3.20	3.70	1.97	3.97	3.47	0.72	49.36
2009	2.62			3.78	1.98	4.00	3.51	0.73	49.42
2010	2.69	0.70	3.27	3.81	1.98	3.97	3.38	0.69	49.10
2011	2.55			4.17	1.99	3.96	3.40	0.70	49.28
2012	2.33	0.66	3.08	4.22	2.01	4.01	3.31	0.68	48.98
2013	2.78			3.99	1.99	3.97	3.14	0.64	48.49
2014	2.53	0.68	3.14	3.97	1.99	4.01	2.94	0.60	48.22
2015	2.78			4.00	2.00	4.00	2.91	0.59	48.17
2016	2.73	0.68	3.14	4.11	1.99	4.08	2.78	0.56	48.05
2017	2.83			4.34	2.01	4.15	2.64	0.53	
2018	2.97	0.70	3.26	4.46	2.05	4.22	2.66	0.54	
2019				4.77	2.09	4.40	2.75	0.57	
Mental Distress									
1999	0.081			0.151	0.265				
2000	0.088			0.160	0.270				0.180
2001	0.104			0.169	0.288				0.189
2002	0.094				0.296				0.191
2003	0.095			0.168	0.294				0.210
2004	0.097			0.171	0.299		0.145	0.211	0.218
2005	0.094			0.166	0.295		0.143	0.205	0.210
2006	0.092	0.152	0.158	0.168	0.303		0.136	0.201	0.207
2007	0.086			0.164	0.290		0.133	0.199	0.210
2008	0.100	0.179	0.196	0.169	0.297	0.171	0.142	0.212	0.212
2009	0.102			0.172	0.297	0.170	0.147	0.212	0.219
2010	0.106	0.191	0.205	0.174	0.289	0.174	0.142	0.202	0.207
2011	0.104			0.192	0.302	0.172	0.144	0.207	0.214
2012	0.094	0.178	0.181	0.194	0.308	0.172	0.137	0.199	0.204
2013	0.121			0.184	0.300	0.175	0.128	0.184	0.187
2014	0.102	0.177	0.194	0.181	0.300	0.176	0.115	0.175	0.183
2015	0.119			0.185	0.303	0.179	0.117	0.173	0.183
2016	0.112	0.189	0.185	0.190	0.305	0.187	0.110	0.161	0.177
2017	0.118			0.202	0.298	0.195	0.101	0.154	

2018	0.123	0.202	0.203	0.208	0.318	0.197	0.106	0.159	
2019				0.225	0.331	0.212	0.109	0.169	
Serious Mental Distress									
1999	0.025			0.066	0.050				
2000	0.028			0.069	0.052				0.101
2001	0.033			0.074	0.058				0.108
2002	0.033				0.062				0.107
2003	0.033			0.077	0.062				0.101
2004	0.032			0.079	0.065		0.050	0.090	0.108
2005	0.031			0.075	0.063		0.051	0.088	0.102
2006	0.032	0.067	0.058	0.077	0.063		0.048	0.083	0.098
2007	0.028			0.077	0.065		0.048	0.079	0.102
2008	0.032	0.077	0.086	0.079	0.063	0.052	0.050	0.089	0.106
2009	0.035			0.082	0.069	0.053	0.052	0.089	0.108
2010	0.036	0.083	0.084	0.083	0.069	0.054	0.051	0.085	0.104
2011	0.036			0.093	0.070	0.054	0.051	0.085	0.107
2012	0.032	0.082	0.085	0.094	0.074	0.056	0.053	0.084	0.103
2013	0.040			0.090	0.073	0.056	0.049	0.080	0.094
2014	0.034	0.088	0.088	0.089	0.071	0.055	0.042	0.073	0.087
2015	0.039			0.089	0.070	0.059	0.041	0.070	0.086
2016	0.039	0.083	0.076	0.092	0.065	0.064	0.037	0.066	0.082
2017	0.036			0.098	0.070	0.067	0.036	0.060	
2018	0.041	0.087	0.089	0.100	0.073	0.071	0.036	0.061	
2019				0.107	0.079	0.078	0.037	0.067	

Note: Table shows raw (not normalized) values of mental distress measures for 18-64 year olds from 1999-2019, or sub-periods with available data. The measures and data sources are described in Table 1.

Table B3. Trends in Normalized Measures of Mental Distress From Individual Data

Mental Health Indicator	Score	Mental Distress	Serious Mental Distress
K6 NHIS	1.26 (1.12,1.41)	1.64 (1.39,1.89)	1.75 (1.34,2.17)
PHQ2 NHANES	1.18 (0.35,2.01)	1.58 (0.58,2.57)	1.86 (0.42,3.31)
PHQ9	0.82 (0.21,1.44)	0.92 (0.02,1.82)	1.57 (0.24,2.90)
MH DAYS	1.45 (1.38,1.51)	1.55 (1.48,1.62)	2.11 (2.01,2.22)
POOR MH	0.23 (0.15,0.31)	0.59 (0.35,0.83)	1.60 (1.18,2.02)
K6 NSDUH	0.75 (0.59,0.91)	1.90 (1.58,2.22)	4.09 (3.52,4.65)
K6 MEPS	-1.80 (-1.99,-1.61)	-2.10 (-2.42,-1.77)	-2.19 (-2.73,-1.65)
PHQ2 MEPS	-1.85 (-2.12,-1.58)	-1.81 (-2.08,-1.55)	-2.13 (-2.56,-1.71)
MCS	-0.13 (-0.17,-0.10)	-0.30 (-0.59,-0.01)	-1.03 (-1.43,-0.64)

Note: Table shows estimates of the coefficient from regressing the specified normalized mental distress measure as a percentage of 2008 values against a linear annual time trend for 18-64 year olds from 1999-2019, or sub-periods with available data. Estimates are based on individual-level (not aggregated) data; 95% confidence intervals, adjusted for complex survey design, are shown in parentheses. The measures and sources are described in Table 1. All scores are coded so that higher values indicate worse mental distress.

Table B4. Correlation Coefficients Between Functional Forms of Specified Mental Distress Measures

Mental Distress Indicator	DS,MD	DS,SMD	DS,SMD
K6 NHIS	0.973	0.910	0.949
PHQ2 NHANES	0.924	0.926	0.814
PHQ9	0.967	0.898	0.884
MH DAYS	0.997	0.990	0.980
POOR MH	0.959	0.855	0.869
K6 NSDUH	0.973	0.962	0.990
K6 MEPS	0.991	0.960	0.965
PHQ2 MEPS	0.996	0.993	0.982
MCS	0.794	0.945	0.698

Note: Table shows correlation coefficients between the average Distress Score (DS), mental distress (MD), and Serious Mental Distress (SMD) functional forms of the specified indicator for 18-64 year olds from 1999-2019, or sub-periods with available data. The measures and data sources are described in Table 1.

Table B5. Potential Consequences of Mental Distress With and Without Normalization

YEAR	SUICIDE	SUICIDE ED	MH VISITS	MH DRUGS
Normalized				
1999	86.0		73.6	64.9
2000	86.0		82.0	70.1
2001	89.2		83.8	82.2
2002	91.5		83.3	88.9
2003	91.2		79.3	90.9
2004	93.3		85.9	89.3
2005	92.3		88.6	98.0
2006	94.2	80.6	84.9	92.3
2007	97.7	87.2	91.4	93.9
2008	100.0	100.0	100.0	100.0
2009	101.4	104.8	102.2	105.2
2010	105.4	109.8	102.6	106.7
2011	106.6	111.2	104.9	110.4
2012	108.6	122.0	100.9	111.9
2013	107.8	125.9	100.5	105.0
2014	110.8	137.1	95.5	113.8
2015	113.8	138.4	106.0	112.9
2016	114.8		113.7	106.4
2017	119.5	154.1	118.9	106.9
2018	121.3	158.1	134.6	112.9
2019	119.1	164.6		114.2
Not Normalized				
1999	13.2		0.061	0.088
2000	13.2		0.068	0.095
2001	13.7		0.069	0.111
2002	14.0		0.069	0.121
2003	14.0		0.066	0.123
2004	14.3		0.071	0.121
2005	14.1		0.073	0.133
2006	14.4	374.0	0.070	0.125
2007	15.0	404.4	0.076	0.127
2008	15.3	464.0	0.083	0.136
2009	15.5	486.1	0.084	0.143
2010	16.1	509.6	0.085	0.145
2011	16.3	515.9	0.087	0.150

2012	16.6	566.2	0.083	0.152
2013	16.5	583.9	0.083	0.142
2014	17.0	636.0	0.079	0.154
2015	17.4	642.3	0.088	0.153
2016	17.6		0.094	0.144
2017	18.3	715.2	0.098	0.145
2018	18.6	733.5	0.111	0.153
2019	18.2	763.6		0.155

Note: Table shows potential consequences of mental distress for 18-64 year olds from 1999-2019, or sub-periods with available data. The variables and data sources are described in Table 2. Normalized measures indicate percentages of 2008 average values.

Table B6: Percentage of Incomplete MEPS Self-Administered Questionnaires and Missing Values for MEPS Questionnaire Variables

	SAQ			Both Years			Year 1			Year 2		
Year	Both Years	Year 1	Year 2	K6	PHQ2	MCS	K6	PHQ2	MCS	K6	PHQ2	MCS
2002	6.8	6.7	6.9			11.0			10.3			11.6
2003	7.1	6.9	7.2			11.3			10.8			11.9
2004	7.4	7.7	7.1	13.5	12.8	11.9	13.1	12.4	11.4	14.0	13.3	12.4
2005	7.5	8.1	7.0	13.4	12.7	12.0	13.3	12.7	12.0	13.4	12.7	12.1
2006	8.3	8.4	8.1	13.7	13.2	12.5	13.3	12.7	12.0	14.2	13.7	13.1
2007	7.8	9.2	6.7	14.3	13.5	12.6	15.3	14.4	13.4	13.5	12.8	11.9
2008	6.4	6.1	6.9	12.9	12.3	11.2	12.2	11.5	10.5	13.8	13.4	12.3
2009	5.7	5.9	5.5	12.6	12.1	10.9	11.8	11.3	10.2	13.3	12.8	11.4
2010	6.7	6.8	6.5	13.4	12.6	11.3	12.7	11.9	10.4	14.0	13.3	12.1
2011	6.1	6.2	6.0	12.3	11.8	10.7	11.9	11.3	10.1	12.9	12.6	11.5
2012	8.5	8.2	8.8	10.9	10.3	9.6	10.9	10.2	9.4	10.9	10.4	9.7
2013	10.9	11.2	10.6	14.1	12.8	12.3	14.4	13.2	12.6	13.9	12.5	11.9
2014	13.8	13.7	13.8	18.5	17.3	15.2	18.7	17.6	15.4	18.2	17.1	15.1
2015	14.6	14.1	15.1	17.5	16.5	15.8	17.3	16.3	15.4	17.7	16.8	16.3
2016	13.4	13.8	13.1	18.4	17.4	14.8	19.0	17.9	15.2	17.8	16.9	14.5
2017	11.9	11.8	12.1	15.3	14.1		14.9	13.9		15.6	14.2	
2018	14.5	15.1	14.0	18.2	17.0		18.9	17.6		17.6	16.4	
2019	18.5	19.1	17.9	21.9	21.0		22.6	21.6		21.2	20.3	

Note: First three data columns show the percentage of eligible MEPS respondents not completing the self-administered questionnaire (SAQ). The remainder of the table displays the percentage of eligible respondents with missing data for the MEPS questionnaire mental distress measures. Years 1 and 2 refer to the first and second year a respondent participates in the MEPS.

Table B7. Percent of Missing Observations for MEPS Questionnaire Measures Conditional on Completion of Self-Administered Questionnaire

	Both Years			Year 1			Year 2		
YEAR	K6	PHQ2	MCS	K6	PHQ2	MCS	K6	PHQ2	MCS
2002			4.2			3.5			4.7
2003			4.3			3.9			4.7
2004	6.1	5.4	4.5	5.4	4.6	3.7	6.8	6.1	5.3
2005	5.8	5.1	4.5	5.2	4.6	3.9	6.4	5.7	5.1
2006	5.5	4.9	4.2	4.9	4.3	3.6	6.1	5.6	4.9
2007	6.5	5.8	4.8	6.1	5.3	4.3	6.8	6.2	5.2
2008	6.5	5.9	4.8	6.2	5.5	4.4	6.9	6.6	5.4
2009	6.9	6.4	5.2	5.8	5.3	4.3	7.8	7.3	6.0
2010	6.7	6.0	4.6	5.9	5.1	3.6	7.5	6.7	5.6
2011	6.2	5.8	4.6	5.7	5.1	3.9	7.0	6.6	5.6
2012	2.4	1.8	1.0	2.7	2.0	1.1	2.1	1.6	0.9
2013	3.2	1.9	1.4	3.2	2.0	1.4	3.3	1.9	1.3
2014	4.7	3.6	1.5	5.0	3.8	1.6	4.4	3.3	1.3
2015	2.9	1.9	1.2	3.2	2.2	1.3	2.5	1.7	1.2
2016	5.0	4.0	1.4	5.2	4.2	1.5	4.7	3.8	1.4
2017	3.3	2.1		3.1	2.1		3.5	2.1	
2018	3.7	2.5		3.8	2.5		3.6	2.5	
2019	3.4	2.5		3.5	2.5		3.3	2.4	

Note: Table shows percent of observations with missing data for the MEPS Questionnaire mental distress measures conditional on respondents having completed and returned the self-administered questionnaire.

Table B8. Normalized Values of MEPS Questionnaire Measures, Adjusting for Partial Completion of Self-Administered Questionnaire

	Score			Distress			Serious Distress		
YEAR	K6	PHQ2	MCS	K6	PHQ2	MCS	K6	PHQ2	MCS
2002			99.8			92.0			101.5
2003			100.1			100.0			97.1
2004	102.4	102.5	100.5	104.4	101.8	104.1	103.4	103.0	104.1
2005	102.8	101.7	100.4	103.2	99.3	101.1	105.5	101.3	99.1
2006	101.7	99.9	100.1	100.8	99.1	100.7	102.2	98.1	97.6
2007	99.3	97.2	99.9	98.0	97.4	101.2	99.9	93.6	100.2
2008	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0
2009	99.7	99.6	99.9	101.6	98.8	101.8	102.3	98.0	100.1
2010	98.4	97.4	99.6	100.9	96.6	98.1	103.1	96.6	98.7
2011	97.8	97.1	99.8	100.4	97.4	99.9	101.4	95.3	100.1
2012	100.2	99.9	99.9	102.0	98.7	100.0	110.3	99.5	102.1
2013	100.4	99.4	99.7	101.5	97.1	97.3	109.8	100.8	100.0
2014	100.9	100.4	100.1	100.8	99.2	100.5	105.2	101.0	101.3
2015	101.6	100.1	100.2	103.3	100.0	101.8	105.2	99.2	101.9
2016	96.1	93.9	99.6	96.7	92.4	97.3	96.2	93.0	96.3
2017	89.6	87.6		87.6	86.9		90.7	84.6	
2018	95.1	94.2		96.7	94.0		97.1	91.2	
2019	105.1	105.8		107.3	105.5		107.4	105.0	

Note: Table displays values of MEPS questionnaire mental distress measures for 18-64 year olds as a percentage of 2008 value, predicted assuming that all eligible respondents completed and returned the self-administered questionnaire. The measures and sources are described in Table 1. All scores are coded so that higher values indicate worse mental distress.

Table B9. Correlation Coefficients Between Measures and Potential Consequences of Poor Mental Health

Distress Measure	SUICIDE	SUICIDE ED	MH VISITS	MH DRUGS	AVERAGE
Score					
K6 NHIS	0.79 (0.74,0.82)	0.78 (0.72,0.84)	0.82 (0.78,0.85)	0.66 (0.61,0.71)	0.76 (0.74,0.78)
PHQ2 NHANES	0.62 (0.28,0.88)	0.33 (-0.40,0.89)	0.68 (0.29,0.93)	0.62 (0.21,0.92)	0.56 (0.33,0.76)
PHQ9	0.54 (0.22,0.82)	0.07 (-0.66,0.82)	0.62 (0.24,0.90)	0.49 (0.10,0.83)	0.43 (0.18,0.67)
MH DAYS	0.91 (0.89,0.92)	0.88 (0.86,0.90)	0.89 (0.87,0.91)	0.80 (0.78,0.82)	0.87 (0.86,0.88)
POOR MH	0.70 (0.58,0.79)	0.68 (0.52,0.83)	0.67 (0.53,0.78)	0.72 (0.60,0.83)	0.69 (0.63,0.75)
K6 NSDUH	0.75 (0.63,0.84)	0.80 (0.70,0.89)	0.86 (0.74,0.95)	0.43 (0.28,0.57)	0.71 (0.65,0.77)
K6 MEPS	-0.90 (-0.93,-0.86)	-0.90 (-0.93,-0.86)	-0.76 (-0.81,-0.71)	-0.61 (-0.67,-0.55)	-0.79 (-0.81,-0.77)
PHQ2 MEPS	-0.87 (-0.92,-0.82)	-0.85 (-0.90,-0.79)	-0.73 (-0.80,-0.66)	-0.57 (-0.65,-0.48)	-0.75 (-0.78,-0.72)
MCS	-0.71 (-0.78,-0.62)	-0.76 (-0.85,-0.66)	-0.51 (-0.61,-0.40)	-0.46 (-0.56,-0.36)	-0.61 (-0.66,-0.56)
Mental Distress					
K6 NHIS	0.80 (0.75,0.85)	0.80 (0.72,0.87)	0.80 (0.74,0.86)	0.68 (0.61,0.74)	0.77 (0.74,0.80)
PHQ2 NHANES	0.70 (0.37,0.94)	0.48 (-0.11,0.91)	0.73 (0.28,0.96)	0.60 (0.19,0.89)	0.63 (0.42,0.81)
PHQ9	0.46 (0.07,0.77)	0.15 (-0.60,0.79)	0.45 (-0.04,0.84)	0.39 (-0.04,0.76)	0.36 (0.12,0.59)
MH DAYS	0.89 (0.88,0.91)	0.88 (0.86,0.91)	0.89 (0.87,0.91)	0.77 (0.74,0.79)	0.86 (0.85,0.87)
POOR MH	0.70 (0.57,0.81)	0.62 (0.36,0.82)	0.65 (0.49,0.78)	0.75 (0.62,0.85)	0.68 (0.60,0.75)
K6 NSDUH	0.84 (0.76,0.91)	0.87 (0.80,0.93)	0.87 (0.78,0.94)	0.42 (0.28,0.55)	0.75 (0.70,0.79)
K6 MEPS	-0.84 (-0.89,-0.78)	-0.83 (-0.89,-0.75)	-0.68 (-0.76,-0.59)	-0.53 (-0.63,-0.44)	-0.72 (-0.76,-0.68)
PHQ2 MEPS	-0.85 (-0.90,-0.79)	-0.85 (-0.91,-0.79)	-0.71 (-0.78,-0.63)	-0.55 (-0.63,-0.47)	-0.74 (-0.77,-0.71)
MCS	-0.25 (-0.40,-0.11)	-0.72 (-0.85,-0.56)	-0.11 (-0.24,0.02)	0.05 (-0.09,0.20)	-0.26 (-0.33,-0.19)
Serious Mental Distress					
K6 NHIS	0.76 (0.66,0.84)	0.68 (0.50,0.83)	0.74 (0.62,0.84)	0.74 (0.64,0.83)	0.73 (0.67,0.79)
PHQ2 NHANES	0.65 (0.22,0.93)	0.47 (-0.42,0.94)	0.58 (0.03,0.93)	0.68 (0.15,0.95)	0.60 (0.30,0.82)
PHQ9	0.49 (0.12,0.78)	0.20 (-0.73,0.91)	0.46 (0.10,0.77)	0.53 (0.08,0.87)	0.42 (0.12,0.67)
MH DAYS	0.94 (0.93,0.96)	0.91 (0.88,0.93)	0.88 (0.86,0.91)	0.85 (0.83,0.87)	0.90 (0.88,0.91)
POOR MH	0.81 (0.73,0.88)	0.67 (0.45,0.85)	0.72 (0.61,0.82)	0.90 (0.83,0.95)	0.78 (0.71,0.83)
K6 NSDUH	0.88 (0.82,0.93)	0.90 (0.84,0.95)	0.89 (0.81,0.95)	0.47 (0.36,0.57)	0.79 (0.75,0.82)
K6 MEPS	-0.76 (-0.86,-0.65)	-0.77 (-0.87,-0.65)	-0.66 (-0.78,-0.52)	-0.44 (-0.59,-0.28)	-0.66 (-0.72,-0.60)
PHQ2 MEPS	-0.84 (-0.91,-0.77)	-0.81 (-0.89,-0.71)	-0.73 (-0.82,-0.63)	-0.53 (-0.65,-0.41)	-0.73 (-0.77,-0.68)
MCS	-0.60 (-0.74,-0.44)	-0.60 (-0.77,-0.41)	-0.41 (-0.57,-0.25)	-0.39 (-0.54,-0.21)	-0.50 (-0.58,-0.42)

Note: Table shows correlation coefficients between mental distress measures and potential consequences of poor mental health for 18-64 year olds from 1999-2019, or sub-periods with available data, with 95% confidence intervals shown in parentheses. The mental health measures and sources are described in Table 1 and the potential consequences in Table 2. Estimates are obtained through bootstrapping with 1,000 repetitions. AVERAGE refers to the mean of the correlation coefficients over the four potential mental health consequences for the specified mental distress measure.

Table B10. Regression Estimates of Trends in Selected Distress Measures from Individual Data Using Different Starting Analysis Years

Starting	K6 NHIS			MH DAYS			POOR MH		
Year	DS	MD	SMD	DS	MD	SMD	DS	MD	SMD
1997	0.95	1.33	1.45	1.47	1.55	2.13	0.24	0.61	1.53
	(0.82,1.08)	(1.12,1.55)	(1.09,1.81)	(1.41,1.52)	(1.49,1.61)	(2.04,2.22)	(0.16,0.31)	(0.39,0.82)	(1.12,1.94)
1998	1.13	1.52	1.65	1.46	1.56	2.13	0.23	0.60	1.60
	(1.00,1.27)	(1.29,1.75)	(1.27,2.04)	(1.40,1.52)	(1.49,1.62)	(2.03,2.22)	(0.15,0.31)	(0.37,0.82)	(1.17,2.02)
1999	1.26	1.64	1.75	1.45	1.55	2.11	0.23	0.59	1.60
	(1.12,1.41)	(1.39,1.89)	(1.34,2.17)	(1.38,1.51)	(1.48,1.62)	(2.01,2.22)	(0.15,0.31)	(0.35,0.83)	(1.18,2.02)
2000	1.21	1.55	1.56	1.41	1.53	2.09	0.20	0.51	1.44
	(1.05,1.37)	(1.29,1.82)	(1.10,2.02)	(1.34,1.48)	(1.46,1.61)	(1.97,2.20)	(0.12,0.29)	(0.26,0.76)	(1.01,1.86)
2001	1.18	1.54	1.43	1.44	1.58	2.09	0.18	0.42	1.27
	(1.01,1.35)	(1.26,1.83)	(0.94,1.92)	(1.36,1.51)	(1.50,1.66)	(1.96,2.21)	(0.10,0.26)	(0.18,0.66)	(0.84,1.70)
2002	1.51	1.85	1.61	1.57	1.75	2.19	0.19	0.42	1.20
	(1.33,1.69)	(1.55,2.15)	(1.08,2.15)	(1.48,1.65)	(1.66,1.84)	(2.05,2.32)	(0.10,0.27)	(0.16,0.67)	(0.75,1.65)
2003	1.55	1.97	1.78	1.57	1.75	2.19	0.21	0.45	1.20
	(1.35,1.75)	(1.64,2.31)	(1.20,2.37)	(1.48,1.65)	(1.66,1.84)	(2.05,2.32)	(0.12,0.30)	(0.18,0.72)	(0.72,1.68)
2004	1.64	2.13	2.02	1.68	1.88	2.32	0.24	0.47	1.19
	(1.42,1.85)	(1.76,2.49)	(1.37,2.67)	(1.59,1.77)	(1.78,1.97)	(2.17,2.47)	(0.14,0.33)	(0.18,0.77)	(0.66,1.71)
2005	1.78	2.36	2.24	1.83	2.05	2.50	0.25	0.54	1.24
	(1.54,2.02)	(1.94,2.77)	(1.51,2.96)	(1.73,1.93)	(1.95,2.16)	(2.33,2.66)	(0.15,0.35)	(0.23,0.85)	(0.68,1.81)
2006	1.86	2.51	2.26	1.88	2.16	2.52	0.28	0.58	1.22
	(1.58,2.15)	(2.03,3.00)	(1.41,3.11)	(1.77,1.99)	(2.04,2.28)	(2.34,2.70)	(0.18,0.39)	(0.26,0.91)	(0.61,1.84)
2007	1.83	2.54	2.36	1.95	2.29	2.53	0.35	0.73	1.14
	(1.51,2.14)	(2.00,3.08)	(1.40,3.33)	(1.83,2.07)	(2.16,2.41)	(2.33,2.73)	(0.23,0.47)	(0.37,1.09)	(0.47,1.81)
2008	1.42	2.16	1.85	1.94	2.28	2.48	0.37	0.73	1.08
	(1.06,1.79)	(1.53,2.78)	(0.75,2.95)	(1.81,2.08)	(2.13,2.42)	(2.26,2.71)	(0.23,0.51)	(0.31,1.15)	(0.28,1.88)
2009	1.47	2.26	1.64	1.92	2.29	2.35	0.40	0.82	0.73 (-
	(1.06,1.88)	(1.56,2.97)	(0.38,2.91)	(1.77,2.08)	(2.13,2.46)	(2.10,2.61)	(0.23,0.56)	(0.34,1.31)	0.22,1.68)
2010				1.91	2.28	2.27	0.43	0.94	0.70 (-
				(1.73,2.08)	(2.09,2.47)	(1.98,2.57)	(0.24,0.63)	(0.36,1.52)	0.39,1.79)

Note: Table shows the coefficient from regressing the mental distress measure, as a percentage of 2008 average, on a linear time trend for 18-64 year olds where the ending year is 2019 (2018 for the K6 NHIS) and the starting year ranges from 1997 through 2010. 95% confidence intervals, adjusted for complex survey design are in parentheses. The mental distress measures and sources are described in Table 1. *DS*, *MD*, and *SMD* refer to Distress Score, Mental Distress, and Serious Mental Distress.

Table B11. Regression Estimates of Trends in Selected Distress Measures from Individual Data Using Different Ending Analysis Years

Ending	K6 NHIS			MH DAYS			POOR MH		
Year	DS	MD	SMD	DS	MD	SMD	DS	MD	SMD
2006	-0.43	0.08	0.82	1.52	1.50	2.11	0.48	1.32	2.08
	(-0.73,-0.12)	(-0.44,0.60)	(-0.05,1.68)	(1.32,1.71)	(1.29,1.72)	(1.79,2.42)	(0.30,0.65)	(0.77,1.87)	(0.98,3.18)
2007	-0.64	-0.28	0.26	1.29	1.21	1.90	0.38	1.05	2.01
	(-0.93,-0.36)	(-0.77,0.20)	(-0.55,1.07)	(1.13,1.46)	(1.03,1.39)	(1.63,2.16)	(0.23,0.53)	(0.57,1.52)	(1.03,3.00)
2008	-0.26	0.05	0.39	1.19	1.11	1.77	0.32	0.95	1.76
	(-0.52,0.00)	(-0.40,0.50)	(-0.35,1.14)	(1.05,1.33)	(0.95,1.26)	(1.54,2.00)	(0.18,0.45)	(0.53,1.37)	(0.87,2.65)
2009	0.14	0.32	0.70	1.16	1.06	1.78	0.28	0.86	1.85
	(-0.11,0.38)	(-0.10,0.74)	(0.00,1.41)	(1.04,1.29)	(0.92,1.20)	(1.57,1.98)	(0.15,0.41)	(0.46,1.26)	(1.03,2.68)
2010	0.45	0.59	0.95	1.13	1.04	1.76	0.24	0.70	1.88
	(0.23,0.68)	(0.21,0.97)	(0.32,1.58)	(1.02,1.25)	(0.92,1.17)	(1.58,1.94)	(0.12,0.37)	(0.32,1.08)	(1.09,2.66)
2011	0.52	0.69	1.11	1.33	1.26	2.01	0.23	0.69	1.88
	(0.32,0.71)	(0.36,1.03)	(0.56,1.66)	(1.23,1.43)	(1.15,1.37)	(1.85,2.18)	(0.10,0.36)	(0.32,1.05)	(1.13,2.63)
2012	0.35	0.53	0.96	1.46	1.41	2.19	0.24	0.71	1.98
	(0.16,0.54)	(0.22,0.83)	(0.45,1.47)	(1.36,1.55)	(1.31,1.51)	(2.04,2.34)	(0.12,0.36)	(0.35,1.06)	(1.31,2.66)
2013	0.58	0.93	1.31	1.40	1.37	2.15	0.22	0.65	1.99
	(0.41,0.75)	(0.64,1.22)	(0.82,1.80)	(1.31,1.48)	(1.28,1.46)	(2.01,2.29)	(0.10,0.34)	(0.31,0.99)	(1.36,2.63)
2014	0.56	0.87	1.20	1.32	1.29	2.06	0.21	0.60	1.90
	(0.39,0.72)	(0.59,1.15)	(0.76,1.64)	(1.24,1.40)	(1.21,1.37)	(1.94,2.19)	(0.10,0.32)	(0.28,0.92)	(1.30,2.51)
2015	0.69	1.09	1.33	1.26	1.25	1.96	0.20	0.57	1.79
	(0.54,0.84)	(0.83,1.35)	(0.90,1.77)	(1.19,1.33)	(1.17,1.33)	(1.85,2.08)	(0.10,0.30)	(0.27,0.88)	(1.22,2.36)
2016	0.75	1.13	1.41	1.24	1.24	1.93	0.18	0.56	1.56
	(0.61,0.89)	(0.90,1.37)	(1.01,1.81)	(1.17,1.30)	(1.17,1.31)	(1.82,2.04)	(0.09,0.28)	(0.27,0.84)	(1.04,2.09)
2017	0.83	1.22	1.35	1.28	1.32	1.98	0.18	0.51	1.48
	(0.70,0.97)	(1.00,1.44)	(0.98,1.72)	(1.22,1.35)	(1.25,1.38)	(1.87,2.08)	(0.09,0.27)	(0.25,0.77)	(1.00,1.96)
2018	0.95	1.33	1.45	1.34	1.40	2.01	0.20	0.54	1.46
	(0.82,1.08)	(1.12,1.55)	(1.09,1.81)	(1.29,1.40)	(1.33,1.46)	(1.92,2.11)	(0.12,0.28)	(0.30,0.77)	(1.03,1.90)
2019				1.47	1.55	2.13	0.24	0.61	1.53
				(1.41,1.52)	(1.49,1.61)	(2.04,2.22)	(0.16,0.31)	(0.39,0.82)	(1.12,1.94)

Note: Table shows the coefficient from regressing the mental distress measure, as a percentage of 2008 average, on a linear time trend for 18-64 year olds where the starting year is 1997 and the ending year ranges from 2006 through 2019. 95% confidence intervals, calculated using standard errors adjusted for complex survey design are in parentheses. The mental distress measures and sources are described in Table 1. *DS*, *MD*, and *SMD* refer to Distress Scores, Mental Distress, and Serious Mental Distress.

Table B12. Regression Estimates of Trends in Potential Consequences of Mental Distress Using Alternative Starting and Ending Analysis Dates

	SUICIDE	SAWMH	MH DRUG
Starting Year			
1999	1.83 (1.70,1.95)	2.30 (1.67,2.93)	2.05 (1.47,2.64)
2000	1.85 (1.71,1.99)	2.28 (1.58,2.99)	1.84 (1.31,2.37)
2001	1.86 (1.70,2.03)	2.37 (1.61,3.14)	1.61 (1.25,1.97)
2002	1.90 (1.71,2.08)	2.49 (1.68,3.30)	1.50 (1.16,1.84)
2003	1.96 (1.78,2.13)	2.58 (1.71,3.45)	1.47 (1.07,1.86)
2004	1.98 (1.78,2.19)	2.52 (1.52,3.53)	1.43 (0.95,1.90)
2005	2.03 (1.81,2.24)	2.57 (1.42,3.72)	1.28 (0.78,1.79)
2006	1.99 (1.75,2.24)	2.65 (1.33,3.97)	1.32 (0.69,1.94)
2007	1.94 (1.68,2.19)	2.49 (0.92,4.06)	1.10 (0.39,1.80)
2008	1.91 (1.61,2.22)	2.45 (0.53,4.38)	0.77 (0.17,1.36)
2009	1.90 (1.53,2.27)	2.79 (0.48,5.09)	0.51 (-0.03,1.06)
2010	1.84 (1.41,2.28)	3.35 (0.59,6.11)	0.38 (-0.29,1.06)
Ending Year			
2008	1.44 (1.14,1.73)	2.01 (0.91,3.11)	3.39 (1.98,4.80)
2009	1.50 (1.28,1.71)	2.28 (1.42,3.13)	3.37 (2.26,4.47)
2010	1.63 (1.36,1.90)	2.36 (1.73,2.99)	3.28 (2.35,4.21)
2011	1.69 (1.48,1.91)	2.41 (1.92,2.89)	3.24 (2.45,4.03)
2012	1.75 (1.56,1.93)	2.25 (1.75,2.76)	3.17 (2.46,3.88)
2013	1.72 (1.55,1.88)	2.09 (1.59,2.59)	2.87 (2.00,3.74)
2014	1.72 (1.59,1.86)	1.82 (1.21,2.43)	2.80 (2.07,3.54)
2015	1.75 (1.63,1.88)	1.81 (1.33,2.29)	2.68 (2.01,3.35)
2016	1.76 (1.65,1.87)	1.90 (1.46,2.34)	2.44 (1.69,3.18)
2017	1.81 (1.67,1.95)	2.02 (1.59,2.45)	2.23 (1.49,2.97)
2018	1.85 (1.71,1.99)	2.30 (1.67,2.93)	2.13 (1.48,2.79)
2019	1.83 (1.70,1.95)	2.30 (1.67,2.93)	2.05 (1.47,2.64)

Note: Table shows estimates of the coefficient from regressing the annual average potential consequence of mental distress, as a percentage of the 2008 mean, on a linear time trend for 18-64 year olds where the ending year is 2019 and the starting year ranges from 1997 through 2010 (upper panel) and where the starting year is 1999 and the ending year ranges from 2008 through 2019 (lower panel). 95% confidence intervals, based on robust standard errors, are shown in parentheses. The potential consequences and data sources are described in Table 2.

Appendix C: Potential Sources of Trends in MEPS Questionnaire Measures

The *MEPS* questionnaire measures (K6, PHQ2, and MCS) results suggesting that mental health improved in the second half of the analysis period are almost certainly anomalous since the same measures (K6 and PHQ2) from other surveys indicate rising mental distress, as do the potential consequences of poor mental health. The roles of changing rates of self-administered questionnaire (SAQ) completion and item non-response conditional on such completion were provided in main analysis. This appendix explores other possibilities.

First, I examined whether declines over time in overall non-response rates to household surveys (briefly summarized in the main test) could also be important. Survey non-response rates were computed using information from the following sources: *NHIS* (National Center for Health Statistics 2019); *NHANES* (Centers for Disease Control and Prevention 2021); *MEPS* (Agency for Healthcare Research and Quality 2025); *BRFSS* (Centers for Disease Control and Prevention 2020); *NSDUH*: 1999-2014 (Czajka and Beyler 2016). In some cases, specific decisions had to be made when computing response rates. For the *NHANES*, these were based on persons completing the medical examinations. The *BRFSS* provided a variety of methods of computing the rates and did not report national average rates in most year. The rates used here referred to median state response rates computed as CARSO rates through 2010 and AAPOR RR4 rates subsequently. *NHIS* response rates were computed for sample adults. Since age-specific response rates were generally not available, the rates refer to all ages for the relevant group (e.g. 18 and older for adults), rather than to 18-64 year olds.

Figure C1 summarizes the annual response rates for the five surveys, with additional details in Table C1. The decreasing response was fairly smooth for the *MEPS* and *NHIS* but the decline in *NHANES* response rates was much sharper, particularly after 2009-

2010, falling from 77% in 2005-2006 to 47% in 2018-2018. By contrast, *NSDUH* response rates were fairly flat from 2000-2009 and then declined. The *BRFSS* presents a contrast, with fairly stable response rates, generally in the 50%-55% range from 1999-2010, followed by a reduction to 45% to 50% range starting in 2011, when a cell phone sample was added to the survey.

While it is difficult to know exactly how changes in response rates will have affected the main estimates of trends in the mental distress variables, it seems unlikely that these accounted for the large declines in measured mental distress observed after 2012 for the *MEPS* questionnaire variables. Specifically, as just noted, response rates to this survey where declining fairly steadily over time, but with no sharp discontinuities. In addition, there was so corresponding decrease in POOR MH, which also came from the *MEPS*, but from the core survey and which did not experience corresponding item non-response (i.e. question-specific missing data).[1]

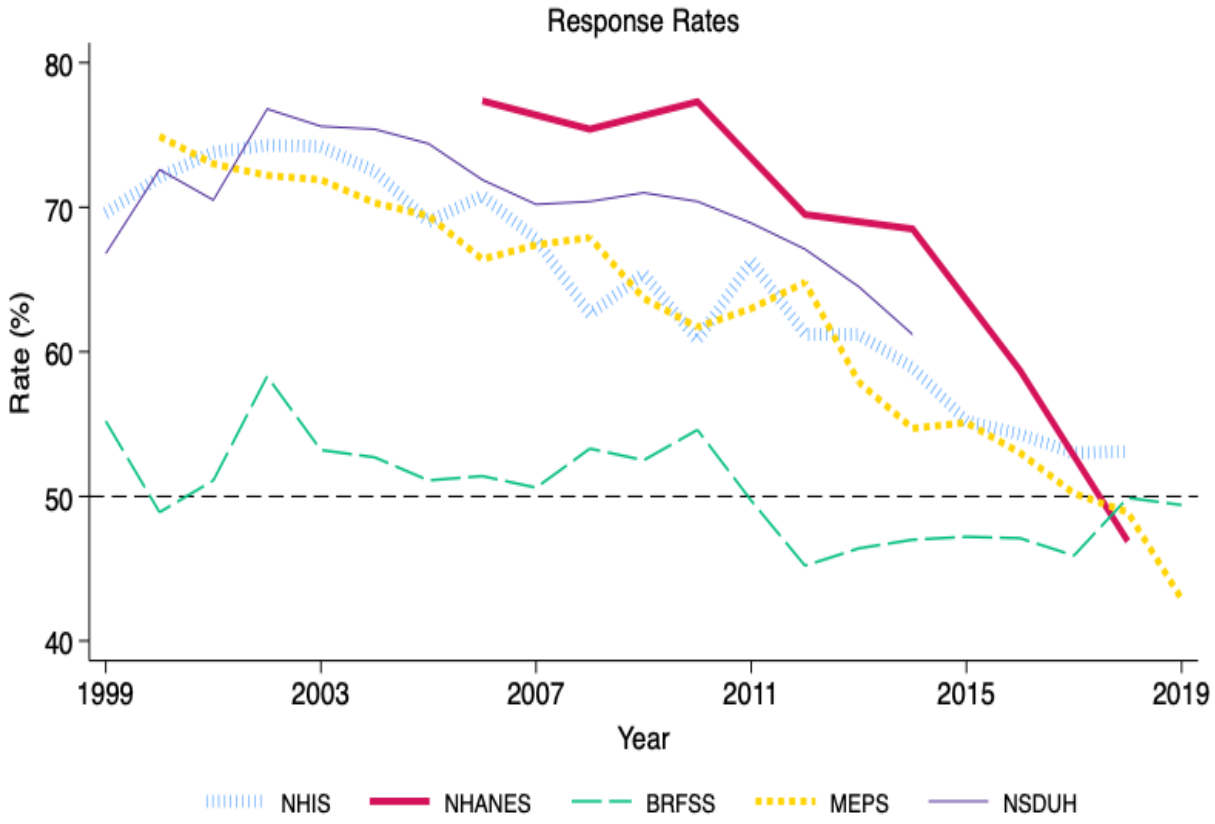
Changes in the design of the *MEPS* survey provide another possible source of the unexpected decline in mental distress measured by the K6, PHQ2, and MCS. Table C2 provides information on this by showing how the structure of the self-administered questionnaire changed during the analysis period – specifically indicating the total number of questions (or parts of questions) included as well as the location of the first question used to construct the three measures.[2] The table confirms that there were some changes in the instrument design but it seems doubtful that these played much of a role. Specifically, the key questions were moved slightly later in the survey and the total number of

[1] Prior research also indicates that changes in survey non-response rates do not necessarily cause significant bias, particularly when the overall response propensity is not strongly related to the variables of key interest (Groves 2006; Groves and Peytcheva 2008).

[2] This information was obtained through my review of the survey instruments which were available at Agency for Health Care Research and Quality (2025).

questions marginally increased starting in 2011, and they were moved earlier and the total length of the instrument was reduced beginning in 2017. Response rates and quality tend to decline with the length of the survey and for questions asked later,(Jeong et al., 2023; Sahlqvist et al., 2011) so both of these changes could be relevant. Also, adult respondents were asked to complete additional self-assessed questionnaires in later years, including instruments indicating cancer experiences in 2011, 2016, and 2017, as well as preventive care behaviors in 2014, and 2018 (Agency for Health Care Research and Quality, 2025), adding to possible perceived burden and survey fatigue in these years.

Figure C1. Response Rates by Survey and Year



Note: Figure shows overall response rates by year for each survey.

Table C1. Response Rates by Survey and Year

Year	NHIS	NHANES	BRFSS	MEPS	NSDUH
1999	69.6		55.2		66.8
2000	72.1		48.9	74.9	72.6
2001	73.8		51.1	73.0	70.5
2002	74.3		58.3	72.2	76.8
2003	74.2		53.2	71.9	75.6
2004	72.5		52.7	70.3	75.4
2005	69.0		51.1	69.4	74.4
2006	70.8	77.4	51.4	66.4	71.9
2007	67.8		50.6	67.4	70.2
2008	62.6	75.4	53.3	67.9	70.4
2009	65.4		52.5	63.7	71.0
2010	60.8	77.3	54.6	61.7	70.4
2011	66.3		49.7	63.0	68.9
2012	61.2	69.5	45.2	64.8	67.1
2013	61.2		46.4	57.9	64.5
2014	58.9	68.5	47.0	54.7	61.2
2015	55.2		47.2	55.1	
2016	54.3	58.7	47.1	53.0	
2017	53.0		45.9	50.2	
2018	53.1	46.9	49.9	48.9	
2019			49.4	42.9	

Note: Table shows overall response rates by year and survey. Response rates for the examination portion of NHANES are shown, where the years refer to the second of a two-year interval (e.g. 2005-2006 for the number reported in 2006), except that the 2018 percentage refers to 2017 through March of 2020.

Table C2: Location of Mental Distress Questions in MEPS Adult Self-Administered Questionnaire

Year	First Question # SF12	K6	PHQ2	Total # of Questions
2000	17			27
2001	17			38
2002	18			39
2003	18			35
2004-2010	18	30	36	41
2011-2016	23	35	41	46
2017, 2019	1	13	19	42

Note: Table shows the number of questions (or parts of multi-part questions) until the first one administered for computing the mental distress measure, as well as the total number of questions. The SF12 is used to compute the MCS score. No questionnaire was available online for 2018.

Appendix References

Agency for Health Care Research and Quality. 2025. "Medical Expenditure Panel Survey Self-Administered Questionnaires." https://meps.ahrq.gov/mepsweb/survey_comp/survey_results_saq_ques.jsp?utm_source.

Agency for Healthcare Research and Quality. 2025. "MEPS-HC Response Rates by Panel." https://meps.ahrq.gov/survey_comp/hc_response_rate.jsp.

Ahmad, Farah, Anuroop K. Jhajj, Donna E. Stewart, Madeline Burghardt, and Arlene S. Bierman. 2014. "Single Item Measures of Self-Rated Mental Health: A Scoping Review." *BMC Health Services Research* 14 (September): 398. https://doi.org/10.1186/1472-6963-14-398.

Blanchflower, David G., and Andrew J. Oswald. 2020. "Trends in Extreme Distress in the United States, 1993–2019." *American Journal of Public Health* 110 (10): 1538–44. https://doi.org/10.2105/AJPH.2020.305811.

Blewett, Lynn A., Julia A. Rivera Drew, Daniel Backman, et al. 2024. "IPUMS Health Surveys: Medical Expenditure Panel Survey: Version 2.4." Version 2.4. Minneapolis, MN: IPUMS. https://doi.org/10.18128/D071.V2.4.

Centers for Disease Control and Prevention. 2020. *Behavioral Risk Factor Surveillance System (BRFSS) Summary Data Quality Reports, 1999–2019*. Technical report. U.S. Department of Health and Human Services, Centers for Disease Control and Prevention.

Centers for Disease Control and Prevention. 2021. *National Health and Nutrition Examination Survey (NHANES) Response Rate Reports, 2005–2020*. Technical report. U.S. Department of Health and Human Services, Centers for Disease Control and Prevention, National Center for Health Statistics. https://wwwn.cdc.gov/Nchs/data/ResponseRates/.

Choi, Briana M., Mavis Obeng-Kusi, and David R. Axon. 2022. "Association between Patient-Provider Communication and Self-Perceived Mental Health in US Adults with Cancer: Real-World Evidence through Medical Expenditure Panel Survey." *Diseases (Basel, Switzerland)* 10 (4): 88. https://doi.org/10.3390/diseases10040088.

Czajka, John L., and Amy Beyler. 2016. *Background Paper Declining Response Rates in Federal Surveys: Trends and Implications*. Mathematica Policy Research. https://aspe.hhs.gov/sites/default/files/private/pdf/255531/Decliningresponserates.pdf.

DeSalvo, Karen B., Nicole Bloser, Kristi Reynolds, Jiang He, and Paul Muntner. 2006. "Mortality Prediction with a Single General Self-Rated Health Question. A Meta-Analysis." *Journal of General Internal Medicine* 21 (3): 267–75. https://doi.org/10.1111/j.1525-1497.2005.00291.x.

Fleishman, John A., and Samuel H. Zuvekas. 2007. "Global Self-Rated Mental Health: Associations with Other Mental Health Measures and with Role Functioning." *Medical Care* 45 (7): 602–9. https://doi.org/10.1097/MLR.0b013e31803bb4b0.

Groves, Robert M. 2006. "Nonresponse Rates and Nonresponse Bias in Household Surveys." *International Journal of Public Opinion Quarterly* 70 (5): 646–75.

Groves, Robert M., and Emilia Peytcheva. 2008. "The Impact of Nonresponse Rates on Nonresponse Bias: A Meta-Analysis." *Public Opinion Quarterly* 72 (2): 167–89.

Idler, E. L., and Y. Benyamini. 1997. "Self-Rated Health and Mortality: A Review of Twenty-Seven Community Studies." *Journal of Health and Social Behavior* 38 (1): 21–37.

Jeong, Dahyeon, Shilpa Aggarwal, Jonathan Robinson, Naresh Kumar, Alan Spearot, and David Sungho Park. 2023. "Exhaustive or Exhausting? Evidence on Respondent Fatigue in Long Surveys." *Journal of Development Economics* 161 (March): 102992. https://doi.org/10.1016/j.jdeveco.2022.102992.

Kessler, Ronald C., Jennifer Greif Green, Michael J. Gruber, et al. 2010. "Screening for Serious Mental Illness in the General Population with the K6 Screening Scale: Results from the WHO World Mental Health (WMH) Survey Initiative." *International Journal of Methods in Psychiatric Research* 19 Suppl 1 (Suppl 1): 4–22. https://doi.org/10.1002/mpr.310.

Kessler, Ronald C., Christopher J. Ruhm, Victor Puac-Polanco, et al. 2022. "Estimated Prevalence of and Factors Associated With Clinically Significant Anxiety and Depression Among US Adults During the First Year of the COVID-19 Pandemic." *JAMA Network Open* 5 (6): e2217223. https://doi.org/10.1001/jamanetworkopen.2022.17223.

Kroenke, K., R. L. Spitzer, and J. B. Williams. 2001. "The PHQ-9: Validity of a Brief Depression Severity Measure." *Journal of General Internal Medicine* 16 (9): 606–13. https://doi.org/10.1046/j.1525-1497.2001.016009606.x.

Kroenke, Kurt, Robert L. Spitzer, and Janet B. W. Williams. 2003. "The Patient Health Questionnaire-2: Validity of a Two-Item Depression Screener." *Medical Care* 41 (11): 1284–92. https://doi.org/10.1097/01.MLR.0000093487.78664.3C.

Levis, Brooke, Ying Sun, Chen He, et al. 2020. "Accuracy of the PHQ-2 Alone and in Combination With the PHQ-9 for Screening to Detect Major Depression: Systematic Review and Meta-Analysis." *JAMA* 323 (22): 2290–300. https://doi.org/10.1001/jama.2020.6504.

Manea, Laura, Simon Gilbody, and Dean McMillan. 2012. "Optimal Cut-off Score for Diagnosing Depression with the Patient Health Questionnaire (PHQ-9): A Meta-Analysis." *CMAJ: Canadian Medical Association Journal = Journal de l'Association Medicale Canadienne* 184 (3): E191-196. https://doi.org/10.1503/cmaj.110829.

Meyer, Bruce D., Wallace KC Mok, and James X. Sullivan. 2015. "Household Surveys in Crisis." *Journal of Economic Perspectives* 29 (4): 199–226.

Mitchell, Alex J., Motahare Yadegarfar, John Gill, and Brendon Stubbs. 2016. "Case Finding and Screening Clinical Utility of the Patient Health Questionnaire (PHQ-9 and PHQ-2) for Depression in Primary Care: A Diagnostic Meta-Analysis of 40 Studies." *BJPsych Open* 2 (2): 127–38. https://doi.org/10.1192/bjpo.bp.115.001685.

Mojtabai, Ramin, and Anthony F. Jorm. 2015. "Trends in Mental Distress, Depressive Episodes and Mental Health Treatment-Seeking in the United States: 2001-2012." *Journal of Affective Disorders* 174 (March): 556–61. https://doi.org/10.1016/j.jad.2014.12.039.

National Center for Health Statistics. 2019. *Survey Description, National Health Interview Survey, 2018*. Technical report. U.S. Department of Health and Human Services, Centers for Disease Control and Prevention, National Center for Health Statistics. https://nhis.ipums.org/nhis/resources/srvydesc2018.pdf.

Owens, Pamela L., Kimberly W. McDermott, Rachel N. Lipari, and Megan M. Hambrick. 2020. "Emergency Department Visits Related to Suicidal Ideation or Suicide

Attempt, 2008–2017." *Healthcare Cost and Utilization Project (HCUP) Statistical Brief #263*, September. http://www.ncbi.nlm.nih.gov/books/NBK563021/.

Prochaska, Judith J., Hai-Yen Sung, Wendy Max, Yanling Shi, and Michael Ong. 2012. "Validity Study of the K6 Scale as a Measure of Moderate Mental Distress Based on Mental Health Treatment Need and Utilization." *International Journal of Methods in Psychiatric Research* 21 (2): 88–97. https://doi.org/10.1002/mpr.1349.

Ruhm, Christopher J. 2025. "Mental Health and Mortality Trends in the United States." *Journal of Health Economics* 102 (August): 103015. https://doi.org/10.1016/j.jhealeco.2025.103015.

Sahlqvist, Shannon, Yena Song, Fiona Bull, et al. 2011. "Effect of Questionnaire Length, Personalisation and Reminder Type on Response Rate to a Complex Postal Survey: Randomised Controlled Trial." *BMC Medical Research Methodology* 11 (May): 62. https://doi.org/10.1186/1471-2288-11-62.

Vilagut, Gemma, Carlos G. Forero, Alejandra Pinto-Meza, et al. 2013. "The Mental Component of the Short-Form 12 Health Survey (SF-12) as a Measure of Depressive Disorders in the General Population: Results with Three Alternative Scoring Methods." *Value in Health: The Journal of the International Society for Pharmacoeconomics and Outcomes Research* 16 (4): 564–73. https://doi.org/10.1016/j.jval.2013.01.006.

Williams, Douglas, and J. Michael Brick. 2018. "Trends in US Face-to-Face Household Survey Nonresponse and Level of Effort." *Journal of Survey Statistics and Methodology* 6 (2): 186–211.

Wright, Emily, Emily C. Dore, Karestan C. Koenen, Christina Mangurian, David R. Williams, and Rita Hamad. 2025. "Prevalence of and Inequities in Poor Mental Health Across 3 US Surveys, 2011 to 2022." *JAMA Network Open* 8 (1): e2454718. https://doi.org/10.1001/jamanetworkopen.2024.54718.