**Title**: Data and Code for: Social Media and Mental Health

**Principal Investigator(s)**: Luca Braghieri, Bocconi University, CEPR, and CESifo; Ro’ee Levy, Tel Aviv University; Alexey Makarin, MIT Sloan School of Management, EIEF, and CEPR

**Version**: V1

**Summary**: The diffusion of social media coincided with a worsening of mental health conditions among adolescents and young adults in the United States, giving rise to speculation that social media might be detrimental to mental health. For this project, the analysis combines data on student mental health around the years of Facebook's expansion with a generalized difference-in-differences empirical strategy. It was found that the roll-out of Facebook at a college increased symptoms of poor mental health, especially depression. It was also found that, among students predicted to be most susceptible to mental illness, the introduction of Facebook led to increased utilization of mental healthcare services. Lastly, after the introduction of Facebook, students were more likely to report experiencing impairments to academic performance resulting from poor mental health.

**Citation**:

Braghieri, Luca, Levy, Ro’ee, and Makarin, Alexey. Data and Code for: Social Media and Mental Health. Nashville, TN: American Economic Association [publisher], 2022. Ann Arbor, MI: Inter-university Consortium for Political and Social Research [distributor], 2022-10-19. <https://doi.org/10.3886/E175582V1>

**Funding:** Deutsche Forschungsgemeinschaft, CRC TRR 190 (280092119); Foerder Institute for Economic Research

**Subject Terms**: Social media; mental health; Facebook; social comparisons

**Geographic Coverage**: USA

**Smallest Geographic Unit**: State

**Distributor(s)**: OPEN Inter-university Consortium for Political and Social Research (ICPSR)

**Curated by: Abigail Mor, Ishita Singh, Katy Lamb**

## Variables

*Birth Year*

Participants might have taken the survey at different times, spanning months. As a result, converting birth years to ages at a specific point in time (such as the survey start date) would inaccurately assume that all participants were surveyed simultaneously. This could lead to skewed age distributions if, for example, more younger individuals participated later in the survey period compared to older individuals.

*States*

The state names were originally collected by their full names, but were converted into their abbreviations for brevity purposes.

*General Health*

Participants were asked to report on how good they perceive their general health to be, considering their age. They could choose from five options: Poor, Fair, Good, Very Good, Excellent. In the datasheet, their responses were left in character form to help future users understand what the question measured, but these responses could always be converted to an ordinal scale in the future (e.g., “poor” = 1 and “excellent” = 5).

*PHQ-9 Quick Depression Assessment*

The PHQ-9 is a validated and reliable assessment of depression that is used by both researchers and healthcare professionals. It scores each of the nine DSM-5 criteria for depression on a scale from “not at all” to “nearly every day,” coded as on an ordinal scale of 1 to 4 respectively. Examples of these criteria include having little interest in doing things, feeling down, and having sleep issues. In the dataset, there are the original survey responses in character form as well as a copy of the responses in numerical form. The numerical form of the responses were later used to create a composite score, or a summed score, that allows for interpretation of depression severity. The severity of each individual's depression was based on the original scale interpretation (See Codebook).

*GAD-7 Anxiety Scale*

The GAD-7 is also a validated and reliable measure used in both the research and healthcare fields, but it is used as a screening tool to assess one’s symptoms of anxiety. It includes seven symptoms of anxiety, which are scored according to their frequency: “not at all” to “nearly every day,” coded on an ordinal scale from 1 to 4 respectively. Examples of these symptoms include feeling nervous/ on-edge, uncontrollable worrying, and restlessness. Also for these variables, there are the original survey responses in character form as well as a copy of the responses in numerical form. The numerical form of the responses were later used to create a composite score, or a summed score, that allows for interpretation of anxiety severity. The severity of each individual's anxiety was based on the original scale interpretation (See Codebook).

*ACHA 12-months Time*

This scale was based on the National College Health Assessment from the American College Health Association. Specifically, questions about college students' mental and physical health. The scale ranged from never (no instances) to 11 or more times for questions regarding mental health. There were also yes/no questions about access to services if a participant was diagnosed with depression. Finally, participants were asked to respond yes/no to 29 questions about if they have any medical conditions. We coded yes as “1” and no as “0”.

*Sex*

Participants were asked to report whether they were a male or female. When asked to provide feedback in the original survey, a couple participants voiced that they would’ve preferred a more inclusive question related to gender-identity instead of this question. This may indicate that not all participants’ answer to this question is the most accurate representation of them.

*Full-time Student*

Participants were asked to report whether they were a full-time status student or part-time status student, and it was a yes/no question. This question could be indicative of hours spent on schoolwork.

*International Student*

Participants were asked to report whether they were an international student, or a domestic student. This was also a yes/no question.

*Race*

Participants were asked to report their race in a “select all that apply” question. Options included “White - not Hispanic (includes Middle Eastern),” “Black - not Hispanic,” “Hispanic or Latino,” “Asian or Pacific Islander,” “American Indian or Alaskan Native,” and/ or “Other.”

## Potential Analyses to Run

First, users of this dataset should calculate summary/descriptive statistics to get a general overview of the representation within their sample, which will indicate how generalizable their analysis findings will be to their target population. Examples of summary/descriptive statistics to calculate would be the ratio of races and/or sexes, number of participants by state, or looking at the ratio between full-time and/or international students. Additionally, understanding the range, central measures of tendency, and standard deviations for the various variables will provide further insight about the dataset and responses received; for example, finding the mean, median, and modal responses for gratitude scores..

Furthermore, there are a wide variety of potential analyses to run with this dataset given its range of variables. Since the data was collected via an online survey, causality cannot be concluded, so correlations would be the most appropriate type of analysis to conduct. The following are a series of possible research questions that correlation tests can explore:

1. What is the relationship between participants who may be at risk for depression (based on their cumulative score) versus participants with an actual diagnosis of depression?
   1. The same analysis could be conducted which instead explores participants who show symptoms of anxiety vs participants with an anxiety diagnosis.
   2. Does the number of participants with an official diagnosis (of anxiety or depression) appear to be similar to the number of participants whose cumulative scores indicate they’re at-risk?
2. What is the relationship between participants’ identified race and their cumulative mental health scores?
   1. Do certain races appear to be more/ less at risk for higher cumulative depression or anxiety scores?
3. What is the relationship between participants’ identified sex and their cumulative mental health scores?
   1. Does one sex appear to be more/ less at risk for higher cumulative depression or anxiety scores?
   2. With this analysis, it is important to recognize that the survey question only included the option to identify as male or female. Those who do not identify with either sex were excluded by this question. An inclusive question about gender identification would’ve been more appropriate.
4. What is the relationship between states and participants’ cumulative mental health scores?
   1. Do certain states appear to have higher/ lower cumulative depression or anxiety scores?
      1. If certain states do appear to have higher/ lower scores on average, this would provide an opportunity to explore other social/ economic/ political information about the state that could be associated with those scores.
   2. How do the participant’s cumulative scores compare to the average scores for their state?
      1. To complete this analysis, additional research into the average depression and anxiety scores for each state would have to be conducted – this information is not provided in this dataset.
5. What is the relationship between full-time student status and their cumulative mental health scores?
   1. The same analysis could be conducted which instead explores the correlation between international students and their cumulative mental health scores.
   2. Do full-time/ international students appear to have higher/ lower anxiety and depression scores?
      1. If they do appear to have higher/ lower scores on average, this would provide an opportunity to explore external factors (not included in this dataset) that could be associated with those scores.
   3. Investigate the relationship between mental health outcomes and academic performance indicators (e.x. GPA) among college students.
6. What is the relationship between anxiety and depression cumulative scores?
   1. Do depressive or anxious symptoms appear to be higher, on average, across all participants?
   2. Do high depression and high anxiety cumulative scores appear to be related?

**Codebook**

### **Year\_1 - Birth Year**

Range: 1971-2003

### **State\_1 - State (abbreviations)**

Students from 48 different states within the United States of America

### **General\_Health**

Considering your age, how would you describe your general health?

*Poor, Fair, Good, Very Good, Excellent*

| Response | # |
| --- | --- |
| Poor | 29 |
| Fair | 103 |
| Good | 227 |
| Very Good | 178 |
| Excellent | 42 |

### **(**[**phq9**](https://med.stanford.edu/fastlab/research/imapp/msrs/_jcr_content/main/accordion/accordion_content3/download_256324296/file.res/PHQ9%20id%20date%2008.03.pdf)**\_*X*/phq9\_XNUM) Over the last two weeks, how often have you been bothered by the following problems?**

1. Little interest or pleasure in doing things?

| Response | | PUNCH | ALL |
| --- | --- | --- | --- |
| Not at all…………………………….. | | 0 | 182 |
| Several days……………………….. | | 1 | 248 |
| More than half of the days………... | | 2 | 102 |
| Nearly every day…………………… | | 3 | 47 |

1. Feeling down, depressed, or hopeless?

| Response | | PUNCH | ALL |
| --- | --- | --- | --- |
| Not at all…………………………….. | | 0 | 176 |
| Several days……………………….. | | 1 | 248 |
| More than half of the days………... | | 2 | 96 |
| Nearly every day…………………… | | 3 | 59 |

1. Trouble falling or staying asleep, or sleeping too much?

| Response | | PUNCH | ALL |
| --- | --- | --- | --- |
| Not at all…………………………….. | | 0 | 130 |
| Several days……………………….. | | 1 | 213 |
| More than half of the days………... | | 2 | 143 |
| Nearly every day…………………… | | 3 | 93 |

1. Feeling tired or having little energy?

| Response | | PUNCH | ALL |
| --- | --- | --- | --- |
| Not at all…………………………….. | | 0 | 76 |
| Several days……………………….. | | 1 | 222 |
| More than half of the days………... | | 2 | 157 |
| Nearly every day…………………… | | 3 | 124 |

1. Poor appetite or overeating?

| Response | | PUNCH | ALL |
| --- | --- | --- | --- |
| Not at all…………………………….. | | 0 | 175 |
| Several days……………………….. | | 1 | 187 |
| More than half of the days………... | | 2 | 128 |
| Nearly every day…………………… | | 3 | 89 |

1. Feeling bad about yourself â€” or that you are a failure or have let yourself or your family down?

| Response | | PUNCH | ALL |
| --- | --- | --- | --- |
| Not at all…………………………….. | | 0 | 208 |
| Several days……………………….. | | 1 | 167 |
| More than half of the days………... | | 2 | 129 |
| Nearly every day…………………… | | 3 | 75 |

1. Trouble concentrating on things, such as reading the newspaper or watching television?

| Response | | PUNCH | ALL |
| --- | --- | --- | --- |
| Not at all…………………………….. | | 0 | 174 |
| Several days……………………….. | | 1 | 203 |
| More than half of the days………... | | 2 | 121 |
| Nearly every day…………………… | | 4 | 81 |

1. Moving or speaking so slowly that other people could have noticed? Or the opposite - being so fidgety or restless that you have been moving a lot more than usual?

| Response | | PUNCH | ALL |
| --- | --- | --- | --- |
| Not at all…………………………….. | | 1 | 403 |
| Several days……………………….. | | 2 | 118 |
| More than half of the days………... | | 2 | 38 |
| Nearly every day…………………… | | 3 | 20 |

1. Thoughts that you would be better off dead, or thoughts of hurting yourself in some way?

| Response | | PUNCH | ALL |
| --- | --- | --- | --- |
| Not at all…………………………….. | | 0 | 429 |
| Several days……………………….. | | 1 | 93 |
| More than half of the days………... | | 2 | 35 |
| Nearly every day…………………… | | 3 | 22 |

**(phq9\_COMP/phq9\_severity)** Based on previous phq9\_XNUM responses, individuals’ composite scores for depression severity were calculated.

* 0-4: Minimal depression
* 5-9: Mild depression
* 10-14: Moderate depression
* 15-19: Moderately severe depression
* 20-27: Severe depression

### **(**[**gad7**](https://adaa.org/sites/default/files/GAD-7_Anxiety-updated_0.pdf)**\_*X*/gad7\_XNUM)Over the last two weeks, how often have you been bothered by the following problems?**

1. Feeling nervous, anxious, or on edge

| Response | | PUNCH | ALL |
| --- | --- | --- | --- |
| Not at all…………………………….. | | 0 | 83 |
| Several days……………………….. | | 1 | 256 |
| More than half of the days………... | | 2 | 129 |
| Nearly every day…………………… | | 3 | 111 |

1. Not being able to stop or control worrying

| Response | | PUNCH | ALL |
| --- | --- | --- | --- |
| Not at all…………………………….. | | 0 | 180 |
| Several days……………………….. | | 1 | 195 |
| More than half of the days………... | | 2 | 114 |
| Nearly every day…………………… | | 3 | 90 |

1. Worrying too much about different things

| Response | | PUNCH | ALL |
| --- | --- | --- | --- |
| Not at all…………………………….. | | 0 | 131 |
| Several days……………………….. | | 1 | 205 |
| More than half of the days………... | | 2 | 134 |
| Nearly every day…………………… | | 3 | 109 |

1. Trouble relaxing

| Response | | PUNCH | ALL |
| --- | --- | --- | --- |
| Not at all…………………………….. | | 0 | 137 |
| Several days……………………….. | | 1 | 233 |
| More than half of the days………... | | 2 | 118 |
| Nearly every day…………………… | | 3 | 91 |

1. Being so restless that it's hard to sit still

| Response | | PUNCH | ALL |
| --- | --- | --- | --- |
| Not at all…………………………….. | | 0 | 249 |
| Several days……………………….. | | 1 | 193 |
| More than half of the days………... | | 2 | 88 |
| Nearly every day…………………… | | 3 | 49 |

1. Becoming easily annoyed or irritable

| Response | | PUNCH | ALL |
| --- | --- | --- | --- |
| Not at all…………………………….. | | 0 | 146 |
| Several days……………………….. | | 1 | 229 |
| More than half of the days………... | | 2 | 138 |
| Nearly every day…………………… | | 3 | 66 |

1. Feeling afraid as if something awful might happen

| Response | | PUNCH | ALL |
| --- | --- | --- | --- |
| Not at all…………………………….. | | 0 | 254 |
| Several days……………………….. | | 1 | 190 |
| More than half of the days………... | | 2 | 76 |
| Nearly every day…………………… | | 3 | 59 |

**(gad7\_COMP/gad7\_severity)** Based on previous gad7\_XNUM responses, individuals’ composite scores for anxiety severity were calculated:

* 0–4: minimal anxiety
* 5–9: mild anxiety
* 10–14: moderate anxiety
* 15–21: severe anxiety

### **(**[**acha\_12months\_times**](https://www.acha.org/documents/ncha/SampleSurvey_ACHA-NCHA_Fa00-FA07.pdf)**\_*X*) Within the last 12 months how many times have you:**

1. Felt things were hopeless

| Response | # |
| --- | --- |
| Never | 109 |
| 1-2 times | 108 |
| 3-4 times | 77 |
| 5-6 times | 52 |
| 7-8 times | 60 |
| 9-10 times | 43 |
| 11 or more | 131 |

1. Felt overwhelmed by all you had to do

| Response | # |
| --- | --- |
| Never | 30 |
| 1-2 times | 41 |
| 3-4 times | 52 |
| 5-6 times | 55 |
| 7-8 times | 60 |
| 9-10 times | 72 |
| 11 or more | 270 |

1. Felt exhausted (not from physical activity)

| Response | # |
| --- | --- |
| Never | 37 |
| 1-2 times | 48 |
| 3-4 times | 49 |
| 5-6 times | 60 |
| 7-8 times | 60 |
| 9-10 times | 73 |
| 11 or more | 253 |

1. Felt very sad

| Response | # |
| --- | --- |
| Never | 49 |
| 1-2 times | 97 |
| 3-4 times | 80 |
| 5-6 times | 61 |
| 7-8 times | 60 |
| 9-10 times | 48 |
| 11 or more | 185 |

1. Felt so depressed that it was difficult to function

| Response | # |
| --- | --- |
| Never | 155 |
| 1-2 times | 114 |
| 3-4 times | 63 |
| 5-6 times | 50 |
| 7-8 times | 46 |
| 9-10 times | 38 |
| 11 or more | 114 |

1. Seriously considered attempting suicide

| Response | # |
| --- | --- |
| Never | 417 |
| 1-2 times | 85 |
| 3-4 times | 25 |
| 5-6 times | 15 |
| 7-8 times | 14 |
| 9-10 times | 9 |
| 11 or more | 15 |

1. Attempted suicide

| Response | # |
| --- | --- |
| Never | 561 |
| 1-2 times | 16 |
| 3-4 times | 0 |
| 5-6 times | 1 |
| 7-8 times | 1 |
| 9-10 times | 0 |
| 11 or more | 1 |

(**acha\_depression**)Have you ever been diagnosed with depression?

| Response | # |
| --- | --- |
| No | 379 |
| Yes | 201 |

**(acha\_services\_1)**If your answer to the previous question (acha\_depression) is yes, then: Have you been diagnosed with depression within the last 12 months?

| Response | # |
| --- | --- |
| No | 62 |
| Yes | 336 |
| No Response (N/A) | 181 |

**(acha\_services\_2)**If your answer to the previous question is yes, then: Are you currently in therapy for depression?

| Response | # |
| --- | --- |
| No | 83 |
| Yes | 311 |
| No Response (N/A) | 185 |

**(acha\_services\_3)**If your answer to the previous question is yes, then: Are you currently taking medication for depression?

| Response | # |
| --- | --- |
| No | 273 |
| Yes | 123 |
| No Response (N/A) | 183 |

### 

### **(acha\_12months\_any\_X/acha\_12months\_any\_XNUM)Within the last 12 months, have you had any of the following?**

| X | Medical Condition | **Yes (1)** | **No (0)** |
| --- | --- | --- | --- |
| 1 | Allergy problems | 224 | 355 |
| 2 | Anorexia | 43 | 536 |
| 3 | Anxiety Disorder | 280 | 299 |
| 4 | Asthma | 76 | 503 |
| 5 | Bulimia | 19 | 560 |
| 6 | Chronic Fatigue Syndrome | 44 | 535 |
| 7 | Depression | 247 | 332 |
| 8 | Diabetes | 10 | 569 |
| 9 | Endometriosis | 13 | 566 |
| 10 | Genital Herpes | 8 | 571 |
| 11 | Genital warts/HPV | 5 | 574 |
| 12 | Hepatitis B or C | 3 | 576 |
| 13 | High blood pressure | 39 | 540 |
| 14 | High cholesterol | 35 | 544 |
| 15 | HIV infection | 1 | 578 |
| 16 | Repetitive stress injury (e.g. carpal tunnel syndrome | 36 | 543 |
| 17 | Seasonal Affect Disorder | 93 | 486 |
| 18 | Substance abuse problem | 42 | 537 |
| 19 | Back pain | 246 | 333 |
| 20 | Broken bone/fracture | 12 | 567 |
| 21 | Bronchitis | 24 | 555 |
| 22 | Chlamydia | 10 | 569 |
| 23 | Ear Infection | 37 | 542 |
| 24 | Gonorrhea | 3 | 576 |
| 25 | Mononucleosis | 7 | 572 |
| 26 | Pelvic Inflammatory Disease | 3 | 576 |
| 27 | Sinus Infection | 112 | 467 |
| 28 | Strep throat | 49 | 530 |
| 29 | Tuberculosis | 2 | 577 |

**(acha\_12months\_any\_comp)** Based on acha\_12months\_any\_X responses, individuals’ total health issues scores were calculated:

* 0 - no health issues
* 30 - extreme health issues

### **(sex) Sex**

| Response | # |
| --- | --- |
| Male | 186 |
| Female | 393 |

### **(fulltime) full time student status**

| Response | # |
| --- | --- |
| No | 520 |
| Yes | 59 |

### **(international\_student\_status) International Student**

| Response | # |
| --- | --- |
| No | 563 |
| Yes | 15 |
| No response | 1 |

### **(Race\_X) Race**

How do you usually describe yourself? (Mark all that apply)

1. White - not Hispanic (includes Middle Eastern): 368 participants
2. Black - not Hispanic: 55 participants
3. Hispanic or Latino: 72 participants
4. Asian or Pacific Islander: 121 participants
5. American Indian or Alaskan Native: 9 participants
6. Other: 4 participants