Master Thesis Econometrics and Management Science
Erasmus University Rotterdam

Causal Effects of Social Media Usage on Mental Health

Elise van den Berge — 457994

Supervisor: Dinand Webbink

Second Assessor: Wendun Wang

We investigated the causal effect of social media usage on mental health. A distinction was made between total hours spent on social media and six different social networking activities: reading and viewing, blogging, posting, messaging, online dating and visiting forums. To account for the possible endogeneity of the treatment variable we corrected for the selection bias caused by unobserved confounders using methods provided by Altonji et al. (2005) and Oster (2019). We assumed equal selection of observed and unobserved variables to calculate bounds on the coefficients of social media usage. Online dating, blogging and visiting forums turned out to have negative causal effects on mental health scores. A sensitivity analysis showed that results were not robust against extreme values of social media use, as for a sub sample of moderate users all activities had negative causal effects except for messaging. The total amount of hours spent on social media had no significant causal effects for both the moderate users and the whole sample of both moderate and extreme social media users.

Date: 01-07-2021

Contents

1 Introduction

Almost every Dutch resident aged 12-55 uses social media, according to CBS (2020), we reported a prevalence of 95% or higher for different age categories. From 2014 to 2019 the percentage of social media users increased for all people from the age of 12. In 2014 the number of social media users is the highest among teenagers and adolescents (12-25 years), but that record is broken by residents 23-25 years old in 2018. Also, social media usage among elderly is increasing. Only 40% of people aged 65-75 used social media in 2014, which increased to 76% in 2019. Also among people that are older than 75 the percentage of social media users rose from 13% to 40%. The most used form of social media is exchanging text messages: 84% of the Dutch say they sent and received text messages in 2019. From the Dutch aged 12 years or older, 92% indicated they used a smartphone to access social networking sites, and 54% used a desktop.

From the past two decade virtual life is taking an increasingly important role. As an inevitable result physical contact with other people becomes more and more limited, thereby possibly affecting both mental and physical health of the social media user. Does a shift from the physical life to a virtual life cause a change in the mental health of individuals? And if so, would it improve or would it worsen our mental health?

Estimating the direct effects of social media usage on mental health is difficult because of the infinite regress of both variables. In order to estimate a true causal effect we must first solve the causation problem: which variable is the cause and which is the effect? Spending time online could cause social deprivation, leading to a worse mental health status. Moreover, using social media could lead to an increase in stress or anxiety levels because of the constant comparison of oneself to our virtual friends on the internet. On the other hand, people with problems that negatively affect mental health, such as loneliness or sleeping problems, might be spending more time on social media. Both the effect and the cause are then driven by a third variable, leading to endogeneity between the treatment variable and the error term in the standard regression models. Another example of this is that introvert and emotionally unstable individuals might both be spending more time on social media Liu and Campbell (2017) and also report lower scores on the quality of their mental health Diener et al. (2003).

Much research on this topic has already been done by various psychologists, psychiatrists and doctors, but also by statisticians and econometricians. Results by previous research seem to be inconsistent, as they vary from concluding that social media usage has either a positive, negative or zero effect on mental health.

Some research predicts no effects of social media use on mental health (Berryman et al., 2018). A longitudinal study over 8 years by Coyne et al. (2020) showed no individual changes in depression or anxiety that were caused by time spent using social media. Others say that social media might benefit social relationships and therewith increase the quality of mental health (Nowland et al., 2018). Of the adolescents that struggle with bad mental health, many (90%) use the internet for help, for example by researching mental health issues and by connecting with other people that have mental health problems (Rideout et al., 2018). Social media might offer support for those who suffer from depressive symptoms or anxiety. Many adolescents claim that social media helps them to find connection and support in times when mental health problems get worse: 30% says that

during heavy times, social media makes them feel less alone. A total of 30% of the adolescents says that social media makes them feel better during bad times, 22% claims it makes them feel worse and 47% claims that it does not affect them either way. Women are more likely than men to use social media as an instrument to help them deal with depressive symptoms or anxiety. A comparison by Facebook users and Facebook non-users by Brailovskaia and Margraf (2016) showed that Facebook users have significantly higher values of self-esteem, life satisfaction and subjective happiness. They also found that Facebook non-users showed significantly higher values of depressive symptoms than Facebook users.
Some inverse associations are established as well: Frost and Rickwood (2017) found that Facebook use was associated with higher levels of addiction, anxiety, depression, low body image and alcohol use. Young et al. (2020) also found a negative relationship between social media usage and both mental well being and life satisfaction. According to Hardy and Castonguay (2018) the relationship between social media use and mental health varies with age: they observed higher levels of anxiety for social media users that are older than 30. For adults aged 18-29 they found an opposite effect.

The aim of this study is to investigate the causal effect of social media usage on mental health. Previous research mainly investigate correlations, which are misinterpreted far too often as being causal relationships.

To interpret the coefficient of a treatment variable as being its causal effect, the treatment variable must be independent of any unobserved factors. Our treatment variable of interest is social media usage. Since the use of social media is not randomly assigned to the participants (that is, there might be factors that influence people's decision to engage in social media use) a selection bias occurs. If this selection into treatment is (partially) caused by unobservables we can not directly measure the influence of external factors on the causal coefficient value. Since instrumental variables - that explain the external variation in the treatment variable but do not have an association with mental health other than through social media usage - were hard to find for the currently used data set, we turned to the methods of Altonji et al. (2005) and Oster (2019). Both estimate the impact of the unobserved confounders on the treatment variable by making assumptions on the relative amount of selection of observables versus selection on unobservables. More specifically, they provide a lower bound and an upper bound for the causal coefficient using regular OLS and a restricted model. Altonji et al. (2005) establish bounds by assuming either that the selection on unobservables is zero (that is, zero correlation between the treatment variable and the error term) or that the selection on unobservables is equal to the selection of observables. Oster (2019) expands their research by providing a way to approximate the selection bias, by assuming that the selection on unobservables is proportional to selection on observables. To confirm their theories and results both introduce a coefficient to measure the relative amount of selection on unobservables that is necessary to explain away the treatment effect. In other words, they estimate for what amount of relative unobservable selection the effect of the treatment variable is equal to a certain number (e.g. zero).

We did not find significant results for the total time spent on social media, but we did find negative causal effects for various social media related activities: blogging, visiting forums and online dating. The negative effects of the first two variables can be explained by absence of communication that is needed to involve in both activities in comparison to the social activity that accompanies the other forms of social media use, and presumably keeps them from having negative causal effects.

Visiting dating websites was found to have the biggest negative causal effect on mental health, possibly because of the increasing dissatisfaction it can cause with someone's love life, leading to a decrease in perceived well-being of the individual. We also did an analysis on a subset of moderate users, excluding the 5% users that spent more than 12 hours a day on social media. For the subset of moderate social media users we found negative causal effects for almost all activities, except for messaging and total use. To rule out any possible reverse causality we did another analysis on a subset in which respondents rated their social media use before they responded to questions about mental health. We found similar results as we obtained from the complete data set.

In Section 2 we discuss our data source and explain our choice of variables. In Section 2.4 we introduce a sub sample of the data based on the mobile phone usage of the participants. Estimating the model on this sub sample enlarges the internal validity in our research, since the selection into treatment will intuitively be smaller for smartphone users. We prove this by evaluating the strength of the relationships between the treatment and the explanatory variables for the full sample and the sub sample. The theory is described in Section 3. To provide a clear interpretation of the bias estimator we first briefly explain the derivation of the bias estimator for the simple case of one observable variable. Thereafter, we describe the estimator for the general case (multiple observables). We then provide the formula for the proportionality coefficient, which substantiates previous methods. In Section 4 we describe the results, accompanied by a two sensitivity analyses. We conclude by summarizing our research goal and our findings, together with some suggestions for later research.

2 Data

The data on social media usage and mental health problems are retrieved from the Longitudinal Internet Studies for Social sciences (LISS) panel. The LISS panel collects yearly data among their participants on various topics with social relevance. Their data is freely available for research purposes.
The LISS panel consists of approximately 5000 households or 7500 participants. Participants are selected by CentERdata (institute for data collection and research) and the *Centraal Bureau voor de Statistiek* (CBS) and are a representative sample of the Dutch population. The participants get a financial compensation for each survey that they engage in.

The data used in this research are part of the so-called 'core studies' of the LISS panel, which are longitudinal studies that are carried out each year. Data on social media use among Dutch residents is collected from the Social Integration and Leisure study, wave 13 (2020). The data on mental health problems is retrieved from the Health core study, wave 13 (2020). We also included data on religion and personality from the Personality core study, which was conducted in the same year.Demographics on the participants are collected from Background Variables that were retrieved in October of 2020 and contain information on gender, age, income, education level and origin.

The Health survey was conducted from 02-11-2020 until 29-11-2020, and 6832 household members were selected, from which 5714 (83.6%) of all respondents fully completed the questionnaire. The Social Integration and Leisure study was conducted from 05-10-2020 to 24-11-2020, and 6680 household members were selected. Of all respondents 5883 (88.1%) fully completed the Social Integration and Leisure questionnaire. For a complete overview of the response information we refer to Table 1. In total, 5513 respondents participated in both studies, of which 4250 fully completed both questionnaires. The sociodemographics of the respondents were extracted in October 2020. For a total of 2059 respondents the complete data was available.

	Health	Social Integration and Leisure
Selected Participants	6832 (100%)	6680 (100%)
Respondents	5736 (84.0%)	5970 (89.4%)
Complete Questionaires	5714 (83.6%)	5883 (88.1%)

Table 1: Survey Response Information

The representativeness of the participants in terms of respondents versus non-respondents is investigated by van der Velden et al. (2016). They found no associations between response behavior and mental health.

2.1 Social Media Use

To measure the social media use of the panel members participants were asked to indicate how many hours per week, on average, they spend on the following online activities:

- Reading and viewing: Reading and viewing social media (e.g. Facebook, Instagram, Twitter, Youtube, LinkedIn, Pinterest, TikTok, etc.)

- Blogging: Reading and/or writing blogs
- Posting: Posting messages, photos and short films on social media yourself (e.g. Facebook, Instagram, Twitter, Youtube, LinkedIn, Pinterest, TikTok, etc.)
- Messaging: Chatting, video calling or sending messages (e.g. WhatsApp, Telegram, Signal, Snapchat, Skype, Facebook messenger, etc.)
- Online dating: Visiting dating websites (e.g. Relatieplanet, Lexa, Tinder, Grindr or similar services)
- Vistiting forums: Visiting (discussion)forums and internet communities

The various uses of social media are used because of different motivations and for various purposes, according to Khan (2017). He found that users *like* and *dislike* posts, messages, pictures or videos (activities performed while 'reading and viewing' social media) for relaxing entertainment. Another use of reading and viewing social media is to search for information. Posting (commenting and uploading messages, photos or films) is mainly done to socialize. Another use of social networking sites is blogging, which refers to self-writing of diary-styled media online and to reading self-published articles of other bloggers. Research has found that commenting on blogs is positively associated with perceived social support from both blog writers and blog readers (Rains and Keating, 2011). Another investigation looked into the psycho-social differences between bloggers and non-bloggers on MySpace over time. They found that social integration and friendship satisfaction increased significantly over time, suggesting that blogging has positive effects on well-being (Baker and Moore, 2008). However, whether the results in the latter study are reliable is debatable, since their investigated sample size was rather small ($N = 57$). As for online messaging we expect its effects to have different sides: on the one hand online messaging with peers might increase social activity and the satisfaction with one's social life, but it could also be that the online interaction reduces the real-life interaction of the individual, causing mental health problems like loneliness. These expectations are confirmed by Rosenbaum and Wong (2012), who found that online messaging is both helping and hindering societal mental health. Messaging with others can provide support in times of need, but excessive usage can lead to internet addiction, causing isolation and therefore worse mental health. We describe online dating as making online contact with new people for the purpose of developing a potentially romantic relationship. We hypothesize a negative correlation between online dating and mental health, as we assume that people that are on dating websites are in general not satisfied with their current love life. In earlier literature positive associations are found between life satisfaction and happy intimate relationships (Arrindell et al., 2001).

From Table 2 we observe that more than 92% of the respondents in our data set uses social media. The average weekly time spent on social media for the social media users is almost 10 hours. That time is mostly spent on reading and viewing social media, shortly followed by messaging via social media. The high corresponding user percentages indicate that those two are also the most popular forms of social media usage among the participants. Of all respondents 10% uses social media to visit forums or internet communities, which they do for on average 3 hours per week. Only 5.7% of the respondents use social media to visit dating websites. Blogging and posting takes the least amount of time when doing so.

	Average Time (hours)	Average Time of Users (hours)	Percentage Users
Total	8.969	9.737	0.921
Reading and viewing	4.033	5.321	0.758
Blogging	0.434	2.389	0.182
Posting	0.678	2.417	0.281
Messaging	3.339	4.064	0.822
Online dating	0.155	2.719	0.057
Visiting forums	0.331	3.033	0.109

Table 2: Statistics on Weekly Social Media Usage

To model the effect of social media we include the variable SM representing the hours spent on social media, distinguishing between total and separate use. We removed responses of participants that reported a social media usage of more than 168 hours a week since it is not feasible to spend more than 24 hours a day on social media.

2.2 Mental Health

The 'Mental Health Inventory 5', abbreviated as 'MHI-5', is an international standard to measure general mental health (Berwick et al., 1991). Participants are asked to give an indication of their personal prevalence over the last 4 weeks regarding the following five statements:

1. I felt very anxious;
2. I felt so down that nothing could cheer me up;
3. I felt calm and peaceful;
4. I felt depressed and gloomy;
5. I felt happy.

To each of the statements a level of applicability regarding the individual's personal situation must be assigned based on six answer categories: never, seldom, sometimes, often, mostly, continuously. The values of 5, 4, 3, 2, 1 and 0 are matched to the positively related questions (question 3 and 5). For the negatively phrased questions (question 1, 2 and 4) the matching is done in the reverse other. To calculate the individual scores the sum is taken over the matching values and multiplied by 4, such that the minimum score is equal to 0 (unhealthy) and the maximum score is equal to 100 (perfectly healthy). Individuals are qualified as having mental health disorders when they have a score of MHI-5 ≤ 70 (Rumpf et al., 2001).

The average mental health score among the LISS-panel members was 75, and 95% of the respondents had a mental health score between 38 and 100. Of all respondents 33% is qualified as having mental health disorders since they have a MHI-5 score below 70. The distribution of the mental health scores is graphically represented in Figure 1.

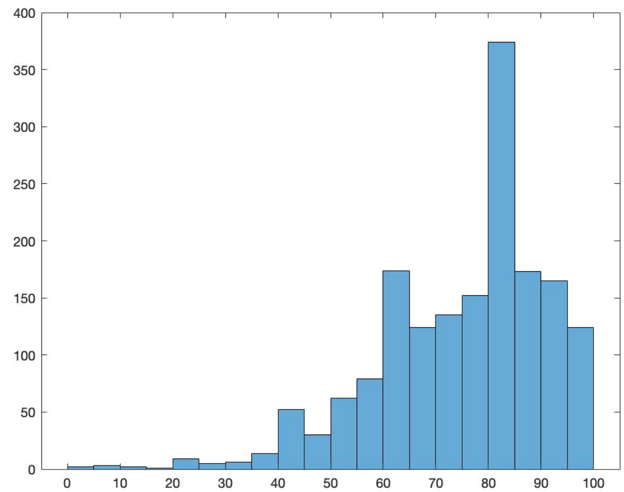


Figure 1: Distribution Mental Health Scores

2.3 Control Variables

Apart from the standard sociodemographics (gender, age, marital status, income, education level, origin, urban character of place of residence) to explain mental health we include religion, physical health, body size, personality and physical health. We removed any observations that reported a monthly gross income above €50000 per month and of €0 for employed participants. The summarizing statistics of the observed explanatory variables can be found in Table 8 in Appendix B.

Psychological literature reports a positive association between religion and mental health. For example, religiousness can aid in recovering from mental illnesses or other problems that affect mental health and can provide a protection from addiction and suicidal behavior (Unterrainer et al., 2014). To obtain the religiosity of the participants the panel members could rate how religious they found themselves to be on a 4-point Likert scale, where 1 = *certainly religious*, 2 = *somewhat religious*, 3 = *barely religious*, 4 = *certainly not religious.* Of all respondents more than 40% stated that they were certainly not religious.

Substantial evidence of a relationship between physical health and mental health is found by McCloughen et al. (2012), who found correlations in mental and physical health by reviewing 18 papers. Another research showed higher morbidity and mortality rates of chronic diseases for people with bad mental health, and showed that physical health can be enhanced by improving mental health (Robson and Gray, 2007). Individuals with bad mental health might be less able to take care of themselves, resulting in worse physical health (Kolappa et al., 2013). The effect goes both ways: bad physical health can lead to dissatisfaction with quality of life, which can negatively influence mental health. To measure physical health in the LISS-panel respondents were asked whether a physician has told them that they suffered from one of 18 diseases/problems in the past year. An overview of the diseases can be found in Table 9, Appendix C. We found that a person in our sample suffered from 1.3 diseases on average. We captured physical health by taking the sum of the amount of diseases/problems the individual reported he or she suffered from. We added a quadratic term to take into account the increasing effect of suffering from multiple diseases/problems at the same time.

Another health-related confounder is body size, which we will measure by BMI. The BMI of an individual can be calculated by dividing his or her body weight by the square of their length. People with higher values for BMI have a relatively high weight with respect to their body length and vice versa. An exceptionally high or low BMI can lead to a negative body image, leading to a worse mental health related quality of life. On the other hand, individuals with bad mental health might suffer from bad eating behaviors, which can lead to anorectic eating disorders on the one hand and to obesity on the other. Statistically significant associations between obesity and both depressive and anxiety disorders are found by Scott et al. (2008). We removed self reported lengths below 100 cm and observations corresponding to weights higher than 500 and lower than 20 kilograms. Since the effects for underweight and overweight people could be similar including BMI as a linear factor would not suffice. We therefore constructed two different variables for over- and underweight individuals:

$$BMI_{high} = \begin{cases} \frac{weight}{length \cdot length}, & \text{if } \frac{weight}{length \cdot length} \geq 30 \\ 0; & \text{otherwise} \end{cases}$$

$$BMI_{low} = \begin{cases} \frac{weight}{length \cdot length}, & \text{if } \frac{weight}{length \cdot length} \leq 18.5 \\ 0. & \text{otherwise} \end{cases}$$

Earlier research has found strong relationships between personality and mental health (Soldz and Vaillant, 1999; DeYoung et al., 2010). Different personalities explain different coping strategies, which in their turn affect how certain circumstances or events affect mental health. Gosling et al. (2003) developed a measure to capture personality by making a distinction in five different areas: openness, conscientiousness, extraversion, agreeableness and neuroticism. The different traits differ in the direction and size of the effect on mental health (Lamers et al., 2012). To recover the so-called big-five personality traits the respondents had to rate 50 statements (10 for each trait) from 1 = *very inaccurate* to 5 = *very accurate*. For each trait the scores are added up and divided by 10.

2.4 Smartphone Usage

Smartphones can be used for a variety of things, such as information seeking, productivity enhancement, relaxation and entertainment, or social media information and relationships (Van Deursen et al., 2015). Using a mobile phone has become a daily habitat for many people, and studies have shown that social media use is closely related to smartphone use (Abi-Jaoude et al., 2020; Greenwood et al., 2016). The found effects of smartphone usage on mental health vary between positive and negative throughout literature, depending on the ways the smartphone is used. Elhai et al. (2017) distinguish between process usage and social media usage. They found that the relationship between smartphone usage and mental health differ for the two categories. For example, they found that anxiety was related to process usage, but not to social usage. Furthermore, they found an inverse relationship between depression and social smartphone usage.

To evaluate smartphone usage in the Netherlands the respondents of the LISS panel were asked whether they use a smartphone for other things than completing the questionnaires of the panel, with answer categories 1 = *yes, often*, 2 = *yes, sometimes* and 3 = *no*.

From Table 3 it can be observed that the differences between the means of the observed confounders are smaller for a sub sample of smartphone users than for the full sample for 17 of the 27 variables. Consequently the relative strength of the relationship between the treatment variable and the observed explanatory variables is weaker for smartphone users. This suggests that we can improve the internal validity of our analysis of the effect of social media use on mental health by comparing the treatment and control group from the smartphone sample. Since participants in the smartphone sample have already chosen to use a smartphone the unobserved variables that affect this decision will not determine the selection into treatment, thereby decreasing a possible selection bias.

Using the smartphone users as our complete data sample we observe that people that spent relatively more time on social media are more likely to be female, they are younger and they are more often not married than people who spend less time on social media. Frequent social media users are on average less religious and have higher BMI-scores (see Appendix B for more details).

	Full Sample			Smartphone Users		
	SM_{low}	SM_{high}	Difference	SM_{low}	SM_{high}	Difference
Female	0.428	0.529	0.101	0.453	0.541	0.088*
Age	60.099	45.184	14.915	56.900	43.651	13.249*
Marital status	0.623	0.445	0.181	0.623	0.449	0.174*
Urbanicity						
Slightly Urban	0.229	0.177	0.112	0.226	0.184	0.042*
Moderately Urban	0.166	0.172	0.006*	0.158	0.173	0.015
Very Urban	0.212	0.243	0.021	0.225	0.243	0.018*
Extremely Urban	0.155	0.224	0.109	0.161	0.219	0.058*
Income	2885.0	2893.1	8.100	2973.8	2966.7	7.100*
Education						
VMBO	0.200	0.117	0.083	0.168	0.111	0.057*
HAVO/VWO	0.084	0.115	0.031*	0.088	0.120	0.032
MBO	0.227	0.247	0.020*	0.224	0.253	0.029
HBO	0.304	0.280	0.024*	0.318	0.288	0.030
WO	0.144	0.191	0.047	0.169	0.184	0.015*
Origin						
First Generation, Western	0.034	0.042	0.008	0.035	0.043	0.008
First Generation, Non-Western	0.039	0.086	0.047	0.041	0.080	0.039*
Second Generation, Western	0.049	0.052	0.003*	0.052	0.058	0.006
Second Generation, Non-Western	0.028	0.059	0.031	0.034	0.064	0.030*
Religiousness						
Certainly Religious	0.135	0.097	0.038	0.120	0.097	0.023*
Somewhat Religious	0.206	0.191	0.015	0.187	0.194	0.007*
Barely Religious	0.264	0.255	0.009*	0.281	0.241	0.040
BMI	25.830	25.762	0.068*	25.566	25.830	0.264
Total Sickness	1.408	1.217	0.191	1.359	1.200	0.159*
Personality						
Extraversion	21.504	22.826	1.322	21.775	22.904	1.129*
Agreeableness	28.183	28.374	0.191*	28.296	28.589	0.293
Conscientiousness	27.874	27.098	0.776	27.972	27.212	0.760*
Neuroticism	25.993	23.905	2.088*	25.903	23.691	2.212
Openness	24.639	25.733	1.094	25.148	25.861	0.071*
Number of observations	1113	947		947	740	

*We compared the characteristics of a group with the 50% smallest amount of hours spent on social media (SM_{low}) against a group with the 50% highest amount of hours spent on social media (SM_{high}). The smaller difference is indicated with an *.*

Table 3: Means of Observed Confounders

3 Methodology

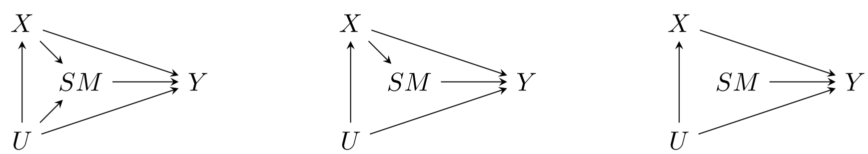
Let mental health be fully determined by the following variables:

$$Y = \beta SM + \gamma X + U \tag{1}$$

where SM is a dummy variable representing whether or not an respondent uses social media, and the outcome variable Y is a measure for mental health, specified by the MHI-5. The observed confounders are denoted by $X = \{x_1, \dots x_k\}$ with corresponding coefficients $\gamma = \{\gamma_1, \dots \gamma_k\}$ and $U = \sum_j^m \omega_j u_j$ is a linear combination of the unobserved confounders.

3.1 Exogeneity of the Treatment Variable

A commonly made mistake in interpreting regression models is to directly explain the correlation coefficient β as being the causal effect of the treatment on the outcome variable Y. In general, a correlation coefficient is not equal to the causal effect. In order to interpret the parameter β as being the causal effect of the treatment on Y, the treatment variable SM must not depend on both the measured explanatory variables X and the unmeasured explanatory variables U (Figure 2c). When the latter holds, the treatment variable is said to be causally exogenous. The treatment variable is said to be statistically exogenous if the probability of receiving the treatment is not correlated with U (Figure 2b). Note from the definitions above that a causal exogeneity implies statistical exogeneity. On the other hand, a statistically exogenous variable does not necessarily need to be causally exogenous.



(a) Endogeneity (b) Statistical Exogeneity (c) Causal Exogeneity

Figure 2: Correlation of the Treatment Variable

Only when all confounders are observed, that is $U = \emptyset$, statistical exogeneity implies causal exogeneity. If $U = \emptyset$, but $\text{corr}(X, SM) \neq 0$, the slope of the regression $\beta = \mathbb{E}[Y|SM = 1] - \mathbb{E}[Y|SM = 0]$ does not explain the causal effect. We can correct for the dependence of SM on X by taking conditional expectations on X, yielding the conditional treatment effect $\mathbb{E}[Y|SM = 1, X = x] - \mathbb{E}[Y|SM = 0, X = x]$. Taking the expectation over all $X = x$ will result in the average total treatment effect of SM on Y:

$$\mathbb{E}[Y|SM = 1, X] - \mathbb{E}[Y|SM = 0, X], \tag{2}$$

which then equals the desired causal effect.

3.2 Selection on Observed and Unobserved Variables

Since U is unobserved and in all probability it holds $U \neq \emptyset$ it is not possible to directly measure its effect on SM. To approximate the possible bias on the causal effect of SM that could arise because of U, Altonji et al. (2005) developed a new estimation method. They make assumptions on the effects that the unobserved variables have on the treatment variable. They state that the amount of selection on the observed explanatory variables in the model is representative for the amount of selection on the unobservables.

Consider the model in (1), and rewrite it as

$$Y = \tilde{\beta} SM + \tilde{\gamma} X + \epsilon, \tag{3}$$

where $\tilde{\gamma} = \{\tilde{\gamma}_1, \dots \tilde{\gamma}_k\}$ and ϵ are defined such that $\text{cov}(X, \epsilon) = 0$. Therefore, $\tilde{\gamma}$ does not only capture the direct effect γ from X on Y in (1), but also the relationship between X and U.

To make inference about the relationship with SM and the unobserved confounders U, Altonji et al. (2005) make use of the relationship between SM and the observed confounders X. Denote the index $\chi = \tilde{\gamma} X$. Imagine the linear projection of SM onto χ and ϵ:

$$\text{proj}(SM|\chi, \epsilon) = \phi_0 + \phi_\chi \chi + \phi_\epsilon \epsilon. \tag{4}$$

The correlation of the treatment variable with the index of observables and unobservables are denoted by ϕ_χ and ϕ_ϵ respectively. In their paper, Altonji et al. (2005) assume an equal amount of selection on observables and unobservables, implied by $\phi_\chi = \phi_\epsilon$. When alongside equal selection it also holds that $\phi_\epsilon = 0$, the treatment variable is said to be causally exogenous, since both observable and unobservable confounders that affect Y do not affect SM.

Let $\sigma_{SM,\chi}$ denote the covariance of SM and χ, and $\sigma_{SM,U}$ the covariance of SM and U. Denote the variances of χ and U as σ_χ and σ_U respectively. The equal selection relationship can be expressed as

$$\frac{\sigma_{SM,\chi}}{\sigma_\chi} = \frac{\sigma_{SM,U}}{\sigma_U}. \tag{5}$$

Note that the relationships are defined on the index $\tilde{\gamma} X$ and not directly on X.

Equal selection on observed and unobserved confounders seems however not realistic, since the factors that influence Y and or SM seem not to be randomly chosen from the total group that includes both observed and unobserved variables. Moreover, the *ex ante* selection on observables is likely to be stronger than selection on unobservables (Angrist and Pischke, 2010). Therefore, in a later research, Altonji et al. (2008) relax the condition of equal selection and suggest $0 \leq \phi_\epsilon \leq \phi_\chi$. If it holds that $\phi_\epsilon = 0$ and $\phi_\chi \neq 0$, the treatment variable is only statistically exogenous and we can calculate average treatment effects using (2). When implementing the latter, $\phi_\epsilon = 0$ and $\phi_\chi = \phi_\epsilon$ can be used to calculate either a lower and an upper bound to the causal effect β.

3.3 Selection Bias

A smaller bounding interval for β would intuitively imply a more stable coefficient. However, only the size of change in the causal coefficient is not sufficient to fully explain its quality (Imbens, 2003; Oster, 2019). In her research on unobservable selection and coefficient stability, Oster (2019) illustrates that coefficients can be relatively more stable for a decreasing precision in control variables because of the resulting extra noise. The severity of the magnitude of change in a coefficient when unobservables are added is determined by the corresponding movement in the determination coefficient R^2. An intuitive explanation for this is that whenever the outcome variable is already largely explained by the observed explanatory variables (that is: R^2 is close to 1), there is very little variation left that can lead to a bias in the coefficient values due to unobservables.

3.3.1 Equal Selection, Single Observable

To provide intuition for the measure of selection bias and its calculation we first provide a derivation for the case of one observed confounder ($k = 1$), besides the treatment variable. Define the following auxiliary regressions:

$$Y = \mathring{\beta} SM + \xi, \tag{6}$$

$$Y = \tilde{\beta} SM + \tilde{\gamma} X + \epsilon, \tag{7}$$

with corresponding coefficients of determination, denoted by $\mathring{R}^2$ and $\tilde{R}^2$ for (6) and (7) respectively (note that (7) is the regression from (Altonji et al., 2005) in (3)). Denote R^2_{max} as the coefficient of determination for the regression in (1). When it is assumed that the model in (1) can be fully explained by the full set of observed and unobserved variables, the maximum coefficient of determination $R^2_{max} = 1$.

Furthermore, define the following regressions:

$$X = \lambda_{X|SM} SM + \zeta_1; \tag{8}$$

$$U = \lambda_{U|SM} SM + \zeta_2; \tag{9}$$

$$U = \lambda_{U|SM,X} SM + \kappa X + \zeta_3. \tag{10}$$

We can then assign the following probability limits to $\mathring{\beta}$ and $\tilde{\beta}$ to specify the convergence in probability:

$$\mathring{\beta} \xrightarrow{p} \beta + \gamma \lambda_{X|SM} + \lambda_{U|SM}; \tag{11}$$

$$\tilde{\beta} \xrightarrow{p} \beta + \lambda_{U|SM,X}. \tag{12}$$

Denote $\lambda_{U|SM,X}$ by Π. From (16) it can be seen that the true β converges in probability to $\tilde{\beta} - \Pi$. We can rewrite Π in terms of the auxiliary regressions (6) and (7) and the expressions for $\mathring{\beta}$ and $\tilde{\beta}$ in (15) and (16) respectively and define a new parameter β^*.

$$\beta^* = \tilde{\beta} - (\mathring{\beta} - \tilde{\beta}) \frac{R^2_{max} - \tilde{R}^2}{\tilde{R}^2 - \mathring{R}^2} \tag{13}$$

It can then be shown that $\beta^* \xrightarrow{p} \beta$. For a detailed derivation of (13) we refer to Appendix A. Note that for $R^2_{max} = 1$ the calculated value for β^* is equal to the resulting β obtained from the method of Altonji et al. (2008). However, in general $R^2_{max} < 1$ due to measurement errors, for example.

3.3.2 Proportional Selection, Single Observable

The proportional selection relationship can be defined as

$$\delta \frac{\sigma_{SM,X}}{\sigma_X} = \frac{\sigma_{SM,U}}{\sigma_U}, \tag{14}$$

where δ denotes the coefficient of proportionality. Equal selection corresponds to $\delta = 1$. Under the proportional selection relationship the coefficients $\lambda_{U|SM}$ and $\lambda_{U|SM,X}$ in (9) and (10) respectively should be multiplied by δ, such that $\mathring{\beta}$ and $\tilde{\beta}$ can be expressed as:

$$\mathring{\beta} \xrightarrow{p} \beta + \lambda_{X|SM} + \delta\lambda_{U|SM}; \tag{15}$$

$$\tilde{\beta} \xrightarrow{p} \beta + \delta\lambda_{U|SM,X}. \tag{16}$$

Subsequently, we can derive an approximation of the bias on β, yielding

$$\beta^* \approx \tilde{\beta} - \delta(\mathring{\beta} - \tilde{\beta}) \frac{R^2_{max} - \tilde{R}^2}{\tilde{R}^2 - \mathring{R}^2}. \tag{17}$$

3.3.3 Multiple Observables

In the general case of $k > 1$ we should rewrite the expression in (8) by

$$X_j = \lambda_{X_j|SM} SM + \zeta_{1,j} \quad \text{for } j = 1, \dots k. \tag{18}$$

Then $\mathring{\beta}$ is expressed as

$$\mathring{\beta} \xrightarrow{p} \beta + \sum_{j=1}^{k} \gamma_j \lambda_{X_j|SM} + \lambda_{U|SM}. \tag{19}$$

Let $(\mu_1, \dots \mu_k)$ be the coefficients of a regression from SM on X. Under the assumption of equal selection and if $\frac{\gamma_i}{\gamma_j} = \frac{\mu_i}{\mu_j} \forall i, j$ the average treatment effect β can be estimated by (17), using the expression for $\mathring{\beta}$ in (19). Since the latter assumption - that is, for SM and Y the effects of X are the same - is very restrictive another approach to calculate β is required.

Let σ_{SM} and σ_Y denote the variance of SM and Y respectively. Let τ_X be the estimated variance of the error term $\tilde{X}$ in the auxiliary regression $SM = \psi X + \tilde{X}$. Define the cubic function $f(v)$:

$$\begin{aligned} f(v) = {} & \delta(R^2_{max} - \tilde{R}^2)\sigma_Y(\mathring{\beta} - \tilde{\beta})\sigma^2_{SM} + v(\delta(R^2_{max} - \tilde{R}^2)\sigma^2_Y(\sigma^2_{SM} - \tau_X) - (\tilde{R}^2 - \mathring{R}^2)\sigma^2_Y \tau_X \\ & - \sigma^2_{SM}\tau_X(\mathring{\beta} - \tilde{\beta})^2) + v^2(\tau_X(\mathring{\beta} - \tilde{\beta})\sigma^2_{SM}(\delta - 2)) + v^3(\delta - 1)(\tau_X \sigma^2_{SM} - \tau^2_X). \end{aligned} \tag{20}$$

It can be shown that the roots of this equation are equal to Π. Since the function is a cubic function with real coefficient values it has either one or three roots. We can eliminate one of the roots by assuming that the bias resulting from the observables is not so large that it changes the sign of the true coefficient of the observables. This can be mathematically formulated as

$$\text{sign}(\text{cov}(X, \hat{\gamma}X)) = \text{sign}(\text{cov}(X, \gamma X)). \tag{21}$$

When we assume the above as well as equal selection ($\delta = 1$), $f(v)$ will have a unique solution and we can define $\beta^* = \tilde{\beta} - v$, such that $\beta^* \xrightarrow{p} \beta$. If we assume $\delta \neq 1$, $f(v)$ will have two solutions. Making an additional assumption on whether Π will be relatively large or small will make the researcher choose the solution leading to a smaller or a larger bias. The latter however requires strong assumptions on the size of Π, which is not desirable.

3.3.4 Required Selection to Explain the Causal Effect

The proportionality coefficient δ measures how large the ratio of selection on X relative to U must be in order to match a given treatment effect $\hat{\beta}$. Let $\hat{\delta}$ be the proportionality coefficient corresponding to $\hat{\beta}$. We can then determine how large the proportionality coefficient $\hat{\delta}$ must be in order to obtain an average treatment effect equal to $\hat{\beta}$:

$$\delta^* = \frac{\hat{\Pi}(\tilde{R}^2 - \mathring{R}^2)\sigma_Y^2\tau_X + \hat{\Pi}\sigma_{SM}^2\tau_X(\mathring{\beta} - \tilde{\beta})^2 + 2\hat{\Pi}^2\tau_X(\mathring{\beta} - \tilde{\beta})\sigma_{SM}^2 + \hat{\Pi}^3(\tau_X\sigma_{SM}^2 - \tau_X^2)}{(R_{max}^2 - \tilde{R}^2)\sigma_{SM}^2 + \hat{\Pi}(R_{max}^2 - \tilde{R}^2)\sigma_Y^2(\sigma_{SM}^2 - \tau_X) + \hat{\Pi}^2\tau_X(\mathring{\beta} - \tilde{\beta})\sigma_{SM}^2 + \hat{\Pi}^3(\tau_X\sigma_{SM}^2 - \tau_X^2)}. \tag{22}$$

where $\hat{\Pi} = \tilde{\beta} - \hat{\beta}$. It holds that $\delta^* \xrightarrow{p} \hat{\delta}$. The proof can be obtained by substituting $\Pi = \tilde{\beta} - \hat{\beta}$ in $f(v)$ in (20) and solving for δ (Oster, 2019).

3.3.5 Bounds on the Maximum Determination Coefficient

An important assumption that has to be made when implementing the method of Oster (2019) regards the explanatory power of the regressors in (1). It is evident that when assuming $R_{max}^2 = 1$ we assume exactly zero measurement error. In many cases, including this investigation, this conservative approach is not plausible. The length of interviews and a possible lack of clarity of the questions can easily cause measurement errors. Accuracy of internet panel data varies with the characteristics of the respondents (Schachter, 2015). For example, significantly less accuracy is observed for lower educated participants. Since many of the variables in the survey are subjective and therefore sensitive to the participant's own interpretation we reject the hypothesis of $R_{max}^2 = 1$. Since it is evident from (3) that $\tilde{R}^2$ is a lower bound on R_{max}^2, that is, $R_{max}^2 \in [\tilde{R}^2, 1]$, we must provide a reasonable upper bound.

Since the bias Π on $\tilde{\beta}$ and therefore the true coefficient β largely relies on the value of R_{max}^2 we evaluate two different bounds on the coefficient of determination. First, we discuss the method of Bellows and Miguel (2009), suggesting $R_{max}^2 = 2\tilde{R}^2 - \mathring{R}^2$. They imply that, when ignoring the treatment variable, the observables and unobservables both explain an equal part of the outcome variable. According to Oster (2019), this is not realistic, since R^2 values in regressions often do not move that much. She takes a similar approach by stating $R_{max}^2 = \min\{1.3\tilde{R}^2, 1\}$. Using this value she assumes that the explanation due to unobservables is relatively less than the explanation by observables. This substantiates Altonji et al. (2005) by assuming that the most important observables are chosen to predict the outcome variable.

4 Results

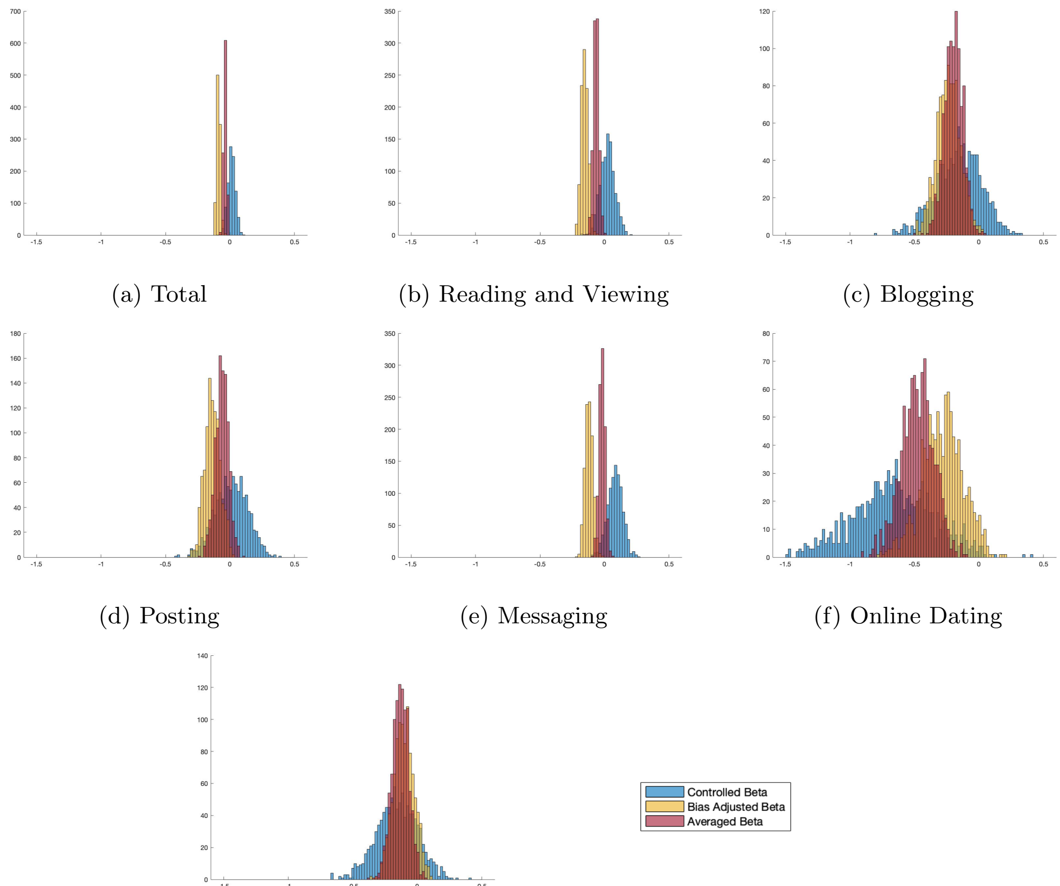
We calculate the results in two different ways that yield the same conclusions. First we provide bounds on β using $\Delta_\beta = [\tilde{\beta}, \beta^*(R^2_{max}, \delta)]$ under the assumption that $\delta = 1$. For the cubic function in (20) to yield to a unique solution Oster (2019) assumes equal signs for the controlled effect $\tilde{\beta}$ and the true effect β. Since the controlled coefficients in our analysis are relatively close to zero we expect changes in the sign of $\tilde{\beta}$ might occur. Therefore, instead of using $\text{sign}(\text{cov}(SM, \hat{\gamma}X)) = \text{sign}(\text{cov}(SM, \gamma X))$ we advocate assuming that the bias in $\tilde{\beta}$ is small. That is, if (20) yields two solutions one should choose the value closest to $\tilde{\beta}$ to provide the bound on β. However, in this particular investigation we could dispense with this assumption as it turns out that (20) yields a unique solution when only using $\delta = 1$. To substantiate these results we also provide values of δ that describe the required selection to explain the causal effect. The results are presented in Table 4. Note that the results apply to the sub sample of the LISS participants that indicated to use a smartphone.

	Controlled effect $\tilde{\beta}$ (std. error) $[\tilde{R}^2]$	Bias adjusted effect β^* (std. error) $[R^2_{max}]$	Proportionality coefficient δ such that $\beta = 0$
Total	0.011 (0.028) [0.455]	-0.089 (0.014) [0.591]	0.099
Reading and viewing	0.023 (0.054) [0.455]	-0.154 (0.026) [0.591]	0.129
Blogging	-0.161 (0.183) [0.455]	-0.235 (0.092) [0.591]	0.671
Posting	0.011 (0.126) [0.455]	-0.136 (0.060) [0.591]	0.079
Messaging	0.081 (0.060) [0.455]	-0.120 (0.030) [0.592]	0.665
Online dating	-0.676 (0.317) [0.456]	-0.284 (0.161) [0.593]	2.244
Visiting forums	-0.149 (0.164) [0.455]	-0.105 (0.079) [0.591]	1.360

Table 4: Bounds on the causal effect of social media usage on mental health

The relationship between accumulated time spent on social media and mental health is positive: respondents who reported spending more time on social media had a higher mental health score. The signs of the relationships between the various forms of social media usage differ, however. Reading, viewing and posting messages or other content of yourself on social media is on average positively correlated with mental health, while reading and/or writing blogs and visiting dating websites or visiting forums and internet communities in general are negatively associated with mental well-being. The coefficients of the negatively associated uses of social media have a larger size. The largest (absolute) association between social media use and mental health is indicated by the time spent on dating websites: for one extra hour per week spent on a dating website the corresponding mental health of the respondent is on average more than 0.6 point lower. Note that the latter relationship is also the most stable one among the different ways to use social media, according to the relatively small standard error of its coefficient. If we assume normality of the controlled coefficients, this is also the only significant coefficient at the 5%-level in our analysis (see Figure 5 for a graphical representation of the bootstrapped coefficients).

The bias adjusted effect β^* is negative for both the cumulative time spent on social media and time spent in the individual manners separately. The coefficient values differ in the way in which they bound the true causal effect β: some amplify the effect of $\tilde{\beta}$ (blogging), others weaken the effect (online dating, visiting forums) and for some (reading and viewing, posting, messaging and the accumulated time spent on social media) even the sign of the effect changes (see Figure 5 for a graphical overview of the effect of β^* on β, where we used $\beta = \Delta_\beta/2$). We assume that calculated coefficients in Table 4 suggest causal impact of social media usage on mental health if the identified set $\Delta_\beta = [\tilde{\beta}, \beta^*(R^2_{max}, \delta)]$ does not contain zero, which is in line with the proposed method of Oster (2019).



(a) Total

(b) Reading and Viewing

(c) Blogging

(d) Posting

(e) Messaging

(f) Online Dating

(g) Visiting Forums

Figure 3: Beta Coefficients ($\tilde{\beta}$, β^*, β)

Only blogging, online dating and visiting forums seem to have causal effect on mental health. The impact of all three is negative and is the largest for online dating, with an effect of almost twice as large as the second most impactful way of social media usage, which is blogging. The negative causal effect of visiting dating websites can be caused by a possible aggravation of the dissatisfaction the participant has with his or her love life, causing mental health scores to drop significantly. The direct and conscious search for a partner might be exhausting, and can affect the mental state of the individual. This is also confirmed by literature (see Section 2.1). The joint negative effect of blogging and visiting forums might be explained by the similarities of the two: they mostly involve looking for information/readable entertainment and do not necessarily involve communication. Besides, they are one of the least used forms of social media usage, see Section 2.1 among the participants. An explanation for their negative causal effects on mental health could be the lack of direct communication with others. The last thing to notice is the non-significant β for the total time spent on social media: this indicates no significant causal effect for the accumulated time spent on social networking sites. The non-significance of the variable can be explained by the fact that it is the sum of all significant and non-significant sub-uses of social media. Since the variables that account for most of the time spent on social media (reading and viewing and messaging) are not significant it is explicable why the accumulated time of the sub-uses is not significant as well.

The results described above are confirmed by the reported values of the proportionality coefficient δ. For the non-significant effects the value of δ is relatively close to zero, implying that an inclusion of only very few information on unobservables can distort the estimated relationship between social media use and mental health to be zero. The value of δ corresponding to online dating is the largest, implying that the selection on unobservables must be more than twice as large as the selection on unobservables to explain away the causal effect of online dating on mental health.

4.1 Sensitivity to Extreme Values of Social Media Usage

Since respondents had to rate their own weekly social media usage by themselves some noise might occur in the data leading to unrealistic values. To verify the robustness of the results we perform the same analysis as above on a subset of the data, where we removed any observations that contain a time spent on social media higher than on average 12 hours per day (we excluded $SM > 84$). Only 3.6% of the data was eliminated. For a graphical interpretation of the original set and the sub sample we refer to Figure 4.

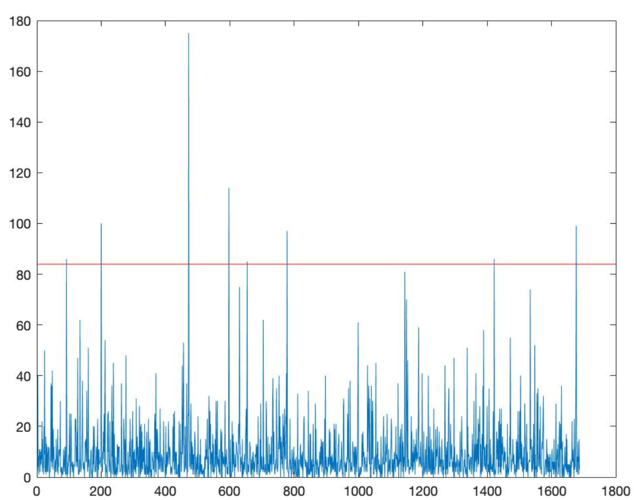


Figure 4: Outliers Social Media Usage ($SM > 84$)

	Controlled effect $\tilde{\beta}$ (std. error) $[\tilde{R}]$	Bias adjusted effect β^* (std. error) $[R_{max}]$	Proportionality coefficient δ such that $\beta = 0$
Total	0.003 (0.033) [0.454]	-0.111 (0.016) [0.591]	0.014
Reading and viewing	-0.004 (0.060) [0.454]	-0.174 (0.030) [0.591]	0.043
Blogging	-0.116 (0.193) [0.455]	-0.239 (0.095) [0.591]	0.477
Posting	-0.023 (0.147) [0.454]	-0.192 (0.071) [0.591]	0.120
Messaging	0.095 (0.069) [0.455]	-0.147 (0.036) [0.592]	0.634
Online dating	-0.924 (0.371) [0.456]	-0.312 (0.193) [0.593]	2.787
Visiting forums	-0.157 (0.164) [0.455]	-0.097 (0.084) [0.591]	1.154

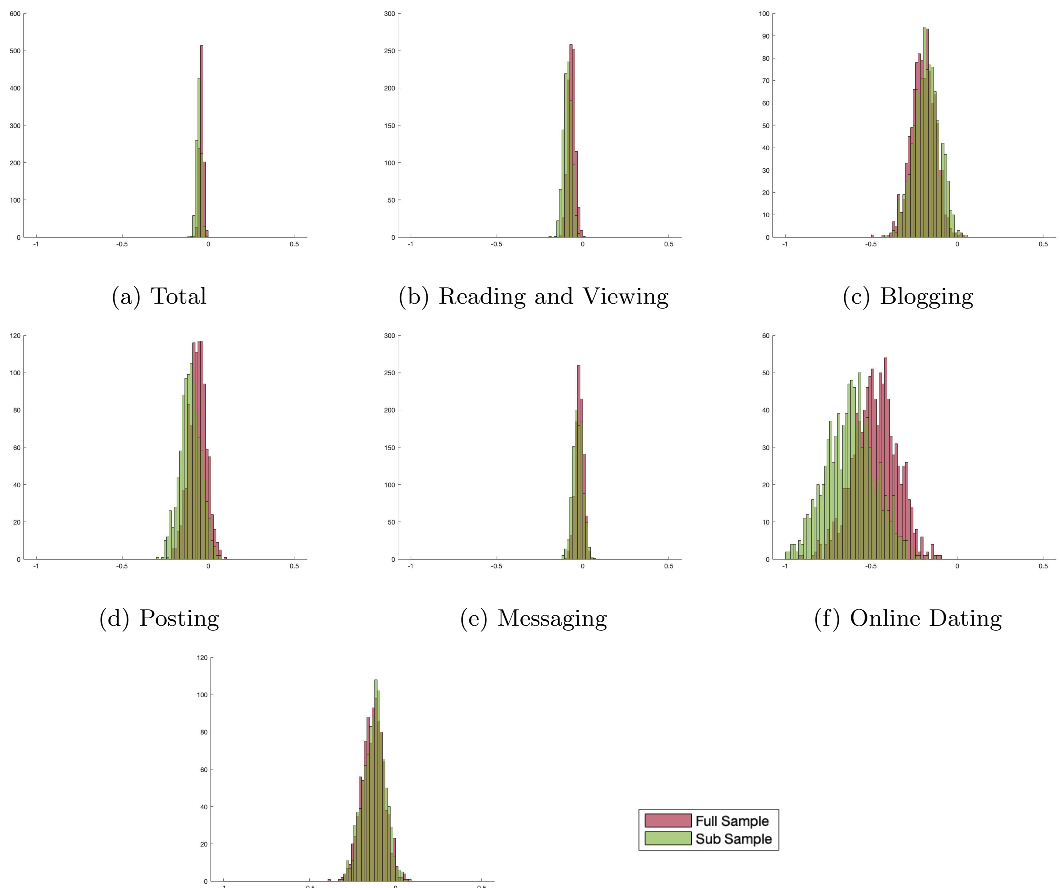
Table 5: Bounds on the causal effect of social media usage on mental health on data without outliers

When comparing the results of the regression on the full sample against the sample without the extreme social media users we notice a few differences. First of all, the relationships in the controlled regressions deviate: the average correlation between total hours spent on social media usage and mental health is smaller for the sub sample, and the relationship between between reading and viewing social media and posting becomes negative for mental health. The correlation between online dating and mental health is even more negative in the sub sample. The values of $\tilde{R}$ and corresponding values of R_{max} are almost the same for the model with moderate users compared to the model on the full sample, indicating similar explanatory power of the models.

The bias adjusted effects β^* stay negative and even increase in magnitude for all social media uses but visiting forums. In combination with the sign changes in $\tilde{\beta}$ this consequently means that more variables have a significantly negative causal effect on mental health, namely reading and viewing, blogging, posting, online dating and visiting forums. The effect of online dating stays the largest, which is confirmed by its large corresponding value of δ.

We can thus conclude that for the sampling population the effects of the different ways to use social media deviate when we filter out the extreme social media users. Where in the full sample the effects were only significant for blogging, online dating and visiting forums we now also observe significantly lower mental health scores for reading and viewing social media and for posting messages on social media platforms. The largest shift is however visible for online dating, see Figure 5f, suggesting enlarged effects for moderate social media users. The positive effect of extreme social media users is clearly visible for posting as well (Figure 5d). The only non-negative relationships are for messaging and the accumulated time spent on social media. Moderate social media users that use social media to directly communicate with others on average have higher mental health scores. Since messaging accounts for almost 50% of the accumulated time spent on social media (see Table 2) this can also be the cause of the positive $\tilde{\beta}$ for the correlation between total hours spent on social media and mental health.

The sensitivities of β^* correspond to the sensitivities of δ: the proportionality coefficient decreases for all variables except for posting and online dating. These are also the variables for which the estimated effect of β becomes more severe. This corresponds to theory behind the proportionality coefficient, resulting in a larger amount of selection on unobservables to be needed in order to obtain $\beta = 0$.



(a) Total (b) Reading and Viewing (c) Blogging

(d) Posting (e) Messaging (f) Online Dating

(g) Visiting Forums

Figure 5: Sensitivity of β to extreme values of social media use

4.2 Sensitivity to Reverse Causality

As described earlier this research can possibly be subject to reverse causality: the mental health of respondents can influence their social media behavior. Due to reverse causation the causal impact of social media use can be larger: if worse mental health causes an increase in social media use we expect larger negative coefficient values β. We can rule out some of this reverse causation by excluding observations from our set where the mental health interview is responded to earlier than the interview containing the questions on social media use. By construction of the questionnaires it is likely that only few of the observations are answered in the reverse order, as the Social Integration and Leisure questionnaire (containing information on social media use) was conducted earlier (05-10-2020 to 24-11-2020) than the Health questionnaire (02-11-2020 to 29-11-2020). Since the dates overlap some reverse causation could still be present, therefore we create a sub sample of observations that satisfy the correct chronological order of response. The results can be found in Table 6.

	Controlled effect $\tilde{\beta}$ (std. error) $[\tilde{R}]$	Bias adjusted effect β^* (std. error) $[R_{max}]$	Proportionality coefficient δ such that $\beta = 0$
Total	0.012 (0.029) [0.459]	-0.089 (0.014) [0.597]	0.108
Reading and viewing	0.020 (0.054) [0.459]	-0.155 (0.026) [0.597]	0.103
Blogging	-0.123 (0.185) [0.459]	-0.235 (0.093) [0.597]	0.510
Posting	0.011 (0.126) [0.459]	-0.130 (0.064) [0.597]	0.082
Messaging	0.074 (0.061) [0.4560]	-0.121 (0.031) [0.598]	0.604
Online dating	-0.657 (0.318) [0.461]	-0.273 (0.161) [0.600]	2.231
Visiting forums	-0.078 (0.171) [0.597]	-0.095 (0.090) [0.591]	0.803

Table 6: Bounds on the causal effect of social media usage on mental health excluding possible reverse causality

To rule out the reverse causation 58 observations (3.4%) were excluded from the data sample. In terms of significance we observe no differences: blogging, online dating and visiting forums are the only variables that have a significant causal effect on mental health, which is the same as for the whole sample. When we look at the mean values for the significant $\tilde{\beta}$ and β^* we notice that they are less in magnitude, as we expected. These results are substantiated by the lower values of the proportionality coefficient. The value of the determination coefficient R^2 is higher for all regressions. An explanation of the better fit of the model is that the time in which the explanatory variables X are observed is closer to the time in which SM is observed (the demographics were observed in October 2020).

4.3 A Review on van der Velden et al. (2019)

We performed the analysis of Oster (2019) on another similar investigation by van der Velden et al. (2019). In their research they investigate whether social media usage can predict mental health problems and sleeping problems using data on the Dutch population provided by LISS. They use a logistic regression model to predict the dependent variable, which is a binary transformation of the MHI-5 for mental health ($Y = 1$ if MHI-5 ≤ 60). Their model may however be subject to endogeneity as few of the regressors (employment status, loneliness, sleeping problems) are likely to be correlated with the error term. They removed observations that reported a social media usage of more than 10 hours a day for each of the specific uses (the cut-off was chosen in an arbitrary manner). A good feature of their research is that they investigate the effect of social media usage on both the short term (in a same fashion as we do in our research) and the long term (where they estimate the effect of social media usage in the previous year on mental health in the current year). We do however question the validity of the latter, as the estimated coefficient of SM_{t-1} does not only capture its own effect, but also the effect of SM_t and Y_{t-1} when the latter two are not included in the model. Furthermore, we think that analyzing the sample over one period only gives better results as the questionnaires on social media usage are held earlier than the questionnaires on mental health (October/November versus November/December), which logically already excludes part of the possible reverse causality. Lastly, we think that transforming the dependent variable to a binary variable (as they do) decreases the validity of the model. The larger variance in continuous variables provides more valuable information and thus more precise and valid coefficient estimates. The paper of van der Velden et al. (2019) only uses three measures of social media use, namely reading and viewing, messaging and posting. The data they used was provided by the LISS panel and contained information on demographics (gender, age, education), social media usage and mental health in 2016 and 2017. The way in which the questions were asked to the respondents was the same as in the data set we used for our own investigation. Their research found that social media use had no significant effect on mental health for both the short term and the long term.

So far, we have analyzed the methods of van der Velden et al. (2019) and came to the conclusion that in order to translate their model to be giving unbiased results we need to exclude any endogenous variables and restrict the analysis to only one time period. Since after this removal the model can still be subject to selection bias due to reasons similar to the original problem that is investigated in this paper we use the method of Oster (2019) to estimate true causal effects. Since van der Velden et al. (2019) combined 2016 demographics with a 2017 dependent variable and a 2017 treatment variable we did an analysis on two models:

$$Y_{2017} = \tilde{\beta}_{2016} SM_{2017} + \tilde{\gamma}_{2016} X_{2016} + \epsilon_{2016} \tag{23}$$

$$Y_{2017} = \tilde{\beta}_{2017} SM_{2017} + \tilde{\gamma}_{2017} X_{2017} + \epsilon_{2017} \tag{24}$$

We included the first one for a better comparison with the results of van der Velden et al. (2019) and the latter since we think this gives more reliable results. Furthermore we decided not to restrict the originally continuous variable Y to a binary variable to prevent a loss of information and to fit the model of Oster (2019) better. The results can be found in Table 7.
If we assume the causal effect of SM is significant if and only if the identified set $\Delta_\beta = [\tilde{\beta}, \beta^*(R^2_{max}, \delta)]$ does not contain zero we see that for the 2016 demographics only posting is significant. This is not in line with the results of van der Velden et al. (2019), who state zero significance for all three variables. When we look at the regression model with the 2017 demographics we observe the same results. This was expected as the X-variables (gender, age and education level) do not change or only change little in the period of one year. We do however observe a higher value of R^2 for the regression models using 2017 demographics in comparison to the 2016 demographics, although the value of R^2 is generally low for all regressions.

The significance of the coefficients using the data of van der Velden et al. (2019) is different from our data, as in our analysis all three variables are not significant. A possible cause of this is a likely omitted variable bias in the model of van der Velden et al. (2019) or the bad fit of the model,

	Controlled effect $\tilde{\beta}$ (std. error) $[\tilde{R}^2]$	Bias adjusted effect β^* (std. error) $[R^2_{max}]$	Significance stated in reviewing paper
Reading and viewing^{2016}	-0.180 (0.077) [0.051]	0.304 (0.675) [0.067]	*Not significant*
Posting^{2016}	-0.314 (0.129) [0.052]	-0.029 (0.058) [0.067]	*Not significant*
Messaging^{2016}	-0.046 (0.063) [0.048]	0.537 (0.748) [0.062]	*Not significant*
Reading and viewing^{2017}	-0.153 (0.058) [0.053]	0.456 (0.745) [0.069]	-
Posting^{2017}	-0.391 (0.114) [0.055]	-0.022 (0.053) [0.072]	-
Messaging^{2017}	-0.097 (0.052) [0.051]	0.638 (0.779) [0.067]	-

Table 7: Bounds on the causal effect of social media usage on mental health

referring to the low values of R^2. Another possibility is a change in behavior and/or content on the social media platforms over a four year time period.

Due to the possible misspecification of the model of van der Velden et al. (2019) and the conflicting results resulting from the methods of Oster (2019) we think the conclusions drawn by van der Velden et al. (2019) should be revised. More variables should be added to the model to substantiate the results that we obtained using their data.

5 Conclusion

We investigated the effect of social media usage on mental health using data from the LISS panel in October, November and December 2020. We divided social media use into 7 different categories: accumulated time spent on social media, reading and viewing, blogging, posting, messaging, online dating and visiting forums. We found that participants differ in usage of the different sub-uses, and that the average user spends 9.7 hours a week on social media. Of all participants in the LISS panel 92.1% used social media at least once per week, from which most people used the online platforms for reading and viewing (75.8%) and messaging (82.2%). Visiting dating websites was the least popular among the participants, as only 5.7% engaged in online dating. To measure the mental health of the participants we used the so-called 'MHI-5' index, an instrument to measure mental health on a score of 1-100. The calculations of the score are based on 5 different aspects: feeling anxious, feeling down, feeling calm and peaceful, feeling depressed and gloomy, feeling happy. Apart from the standard sociodemographics (gender, age, marital status, income, education level, origin, urban character of place of residence) we used religion, physical health, BMI and personality as other confounders to explain the mental health of the participants. We then re-sampled the research population into a sample of only smartphone users, since this made the sample of social media users and non-social media users more homogeneous, leading to an increase in internal validity of the investigation.

Using the method of Oster (2019) to account for influence of unobservables we found significant negative effects for blogging, online dating and visiting forums on mental health. Online dating had the largest negative effect, shortly followed by blogging and thereafter visiting forums. The results of the former could be due to an increasing dissatisfaction with ones love life, leading to lower perceived well-being. The negative effect of the latter two can be caused by the lack of social communication that may be associated with both blogging and visiting forums, since all of the other variables involve social contact with other social media users.

Since the data contained values for social media use that exceeded an average use of 12 hours a day we calculated the results for a sub sample of non-extreme ('moderate') social media users. In this sub sample reading and viewing and posting turned out to have negative causal effects on mental health, in addition to reading and viewing, blogging and visiting forums. This not only told us that the effects of the various ways to use social media is sensitive to extreme uses thereof, but also that for moderate users the effects are different than for the whole set of users. More specifically, the effects of reading and viewing and posting became negative instead of zero, the effect of online dating became even more negative and the effect of blogging became less severe. The effect of visiting forums was roughly the same for moderate users when compared to the whole set. There was no significant effect for total hours spent on social media in the whole sample and the sub sample, which can be explained by the addition of positive and negative effects that the various uses of social media have.

We found that our main results were not sensitive to reverse causality: we excluded observations from the sample such that the sample only contained observations in which the responses were given in the required chronological order. The same variables were significant as in our main analysis, only their magnitude decreased.

All in all we can conclude that social media as a whole has no significant causal effects on mental health, but various different uses of social media can make mental health worse.

5.1 Discussion

As for every research there are some limitations to the results presented in this study. First of all it is important to note that the data from 2020 might not be fully representative for the behavior of the Dutch population over time, since COVID-19 might have had some influence on both social media behavior as (self-perceived) mental health. Therefore results could be biased. However, performing the same analysis on earlier years might not be representative for the future as well, since social media is rapidly developing and the number of its users is increasing.

For further research we are interested in investigating the external validity and possible bias caused by the latter by new methods provided in Andrews and Oster (2019). In their research they compare means of data samples against population means including information on movements in the coefficients of determination. Another extension would be to further investigate on the method of Oster (2019). In her research she makes many assumptions, such as a linear relationship between the dependent and the independent variable and normality of the error terms. It would be very interesting to investigate whether one of these assumptions could be changed into less strict conditions. One could for example do research on the effect on unobservables on a logit or probit model, or a model with dynamics. Furthermore, Oster (2019) does not provide information on what to do when the coefficients of the treatment variable in the OLS-regression are not significant, as it is for many of the different treatment variables in our case. In this research we assume that insignificant $\tilde{\beta}$ values are not statistically different from 0, and therefore we used $\Delta_\beta = [0, \beta^*(R_{max}, \delta)]$ in those cases. We presume this is correct, but further investigation on this can give more clarity on the topic.

References

Abi-Jaoude, E., Naylor, K. T., and Pignatiello, A. (2020). Smartphones, social media use and youth mental health. *Cmaj*, 192(6):E136–E141.

Altonji, J. G., Elder, T. E., and Taber, C. R. (2005). Selection on observed and unobserved variables: Assessing the effectiveness of catholic schools. *Journal of political economy*, 113(1):151–184.

Altonji, J. G., Elder, T. E., and Taber, C. R. (2008). Using selection on observed variables to assess bias from unobservables when evaluating swan-ganz catheterization. *American Economic Review*, 98(2):345–50.

Andrews, I. and Oster, E. (2019). A simple approximation for evaluating external validity bias. *Economics Letters*, 178:58–62.

Angrist, J. D. and Pischke, J.-S. (2010). The credibility revolution in empirical economics: How better research design is taking the con out of econometrics. *Journal of economic perspectives*, 24(2):3–30.

Arrindell, W., van Nieuwenhuizen, C., and Luteijn, F. (2001). Chronic psychiatric status and satisfaction with life. *Personality and individual differences*, 31(2):145–155.

Baker, J. R. and Moore, S. M. (2008). Blogging as a social tool: A psychosocial examination of the effects of blogging. *CyberPsychology & Behavior*, 11(6):747–749.

Bellows, J. and Miguel, E. (2009). War and local collective action in sierra leone. *Journal of public Economics*, 93(11-12):1144–1157.

Berryman, C., Ferguson, C. J., and Negy, C. (2018). Social media use and mental health among young adults. *Psychiatric quarterly*, 89(2):307–314.

Berwick, D. M., Murphy, J. M., Goldman, P. A., Ware Jr, J. E., Barsky, A. J., and Weinstein, M. C. (1991). Performance of a five-item mental health screening test. *Medical care*, pages 169–176.

Brailovskaia, J. and Margraf, J. (2016). Comparing facebook users and facebook non-users: Relationship between personality traits and mental health variables–an exploratory study. *PloS one*, 11(12):e0166999.

CBS (2020). More elderly active on social media. `https://www.cbs.nl/en-gb/news/2020/04/more-elderly-active-on-social-media`.

Coyne, S. M., Rogers, A. A., Zurcher, J. D., Stockdale, L., and Booth, M. (2020). Does time spent using social media impact mental health?: An eight year longitudinal study. *Computers in Human Behavior*, 104:106160.

DeYoung, C. G., Hirsh, J. B., Shane, M. S., Papademetris, X., Rajeevan, N., and Gray, J. R. (2010). Testing predictions from personality neuroscience: Brain structure and the big five. *Psychological science*, 21(6):820–828.

Diener, E., Oishi, S., and Lucas, R. E. (2003). Personality, culture, and subjective well-being: Emotional and cognitive evaluations of life. *Annual review of psychology*, 54(1):403–425.

Elhai, J. D., Levine, J. C., Dvorak, R. D., and Hall, B. J. (2017). Non-social features of smartphone use are most related to depression, anxiety and problematic smartphone use. *Computers in Human Behavior*, 69:75–82.

Frost, R. L. and Rickwood, D. J. (2017). A systematic review of the mental health outcomes associated with facebook use. *Computers in Human Behavior*, 76:576–600.

Gosling, S. D., Rentfrow, P. J., and Swann Jr, W. B. (2003). A very brief measure of the big-five personality domains. *Journal of Research in personality*, 37(6):504–528.

Greenwood, S., Perrin, A., and Duggan, M. (2016). Social media update 2016. *Pew Research Center*, 11(2):1–18.

Hardy, B. W. and Castonguay, J. (2018). The moderating role of age in the relationship between social media use and mental well-being: An analysis of the 2016 general social survey. *Computers in Human Behavior*, 85:282–290.

Imbens, G. W. (2003). Sensitivity to exogeneity assumptions in program evaluation. *American Economic Review*, 93(2):126–132.

Khan, M. L. (2017). Social media engagement: What motivates user participation and consumption on youtube? *Computers in human behavior*, 66:236–247.

Kolappa, K., Henderson, D. C., and Kishore, S. P. (2013). No physical health without mental health: lessons unlearned?

Lamers, S. M., Westerhof, G. J., Kovács, V., and Bohlmeijer, E. T. (2012). Differential relationships in the association of the big five personality traits with positive mental health and psychopathology. *Journal of Research in Personality*, 46(5):517–524.

Liu, D. and Campbell, W. K. (2017). The big five personality traits, big two metatraits and social media: A meta-analysis. *Journal of Research in Personality*, 70:229–240.

McCloughen, A., Foster, K., Huws-Thomas, M., and Delgado, C. (2012). Physical health and wellbeing of emerging and young adults with mental illness: An integrative review of international literature. *International journal of mental health nursing*, 21(3):274–288.

Nowland, R., Necka, E. A., and Cacioppo, J. T. (2018). Loneliness and social internet use: pathways to reconnection in a digital world? *Perspectives on Psychological Science*, 13(1):70–87.

Oster, E. (2019). Unobservable selection and coefficient stability: Theory and evidence. *Journal of Business & Economic Statistics*, 37(2):187–204.

Rains, S. A. and Keating, D. M. (2011). The social dimension of blogging about health: Health blogging, social support, and well-being. *Communication Monographs*, 78(4):511–534.

Rideout, V., Fox, S., et al. (2018). Digital health practices, social media use, and mental well-being among teens and young adults in the us.

Robson, D. and Gray, R. (2007). Serious mental illness and physical health problems: a discussion paper. *International journal of nursing studies*, 44(3):457–466.

Rosenbaum, M. S. and Wong, I. A. (2012). The effect of instant messaging services on society's mental health. *Journal of Services Marketing*.

Rumpf, H.-J., Meyer, C., Hapke, U., and John, U. (2001). Screening for mental health: validity of the mhi-5 using dsm-iv axis i psychiatric disorders as gold standard. *Psychiatry research*, 105(3):243–253.

Schachter, A. (2015). Measurement error in panel data: A comparison of face-to-face and internet survey samples.

Scott, K. M., Bruffaerts, R., Simon, G. E., Alonso, J., Angermeyer, M., De Girolamo, G., Demyttenaere, K., Gasquet, I., Haro, J. M., Karam, E., et al. (2008). Obesity and mental disorders in the general population: results from the world mental health surveys. *International journal of obesity*, 32(1):192–200.

Soldz, S. and Vaillant, G. E. (1999). The big five personality traits and the life course: A 45-year longitudinal study. *Journal of Research in Personality*, 33(2):208–232.

Unterrainer, H.-F., Lewis, A. J., and Fink, A. (2014). Religious/spiritual well-being, personality and mental health: a review of results and conceptual issues. *Journal of religion and health*, 53(2):382–392.

van der Velden, P. G., Bosmans, M. W., van der Meulen, E., and Vermunt, J. K. (2016). Pre-event trajectories of mental health and health-related disabilities, and post-event traumatic stress symptoms and health: A 7-wave population-based study. *Psychiatry research*, 246:466–473.

van der Velden, P. G., Setti, I., van der Meulen, E., and Das, M. (2019). Does social networking sites use predict mental health and sleep problems when prior problems and loneliness are taken into account? a population-based prospective study. *Computers in Human Behavior*, 93:200–209.

Van Deursen, A. J., Bolle, C. L., Hegner, S. M., and Kommers, P. A. (2015). Modeling habitual and addictive smartphone behavior: The role of smartphone usage types, emotional intelligence, social stress, self-regulation, age, and gender. *Computers in human behavior*, 45:411–420.

Young, L., Kolubinski, D. C., and Frings, D. (2020). Attachment style moderates the relationship between social media use and user mental health and wellbeing. *Heliyon*, 6(6):e04056.

A Alternative Proof Equal Selection

Below we present an alternative derivation of the selection bias on the $\hat{\beta}$ coefficient. In our opinion, this derivation gives clearer insights into the proof of Oster (2019) as we managed to exclude one of the regressions necessary in Oster (2019). The proof is for the case of one observable under the assumption of equal selection.

Denote σ_{SM} as the variance of SM. Using the mathematical definitions of the OLS regression coefficients and the equal selection rule, we obtain the following expressions for $\lambda_{X|SM}$, $\lambda_{U|SM}$ and $\lambda_{U|SM,X}$.

$$\begin{aligned}
\lambda_{X|SM} &\xrightarrow{p} \frac{\sigma_{SM,\chi}}{\gamma\sigma_{SM}} \\
\lambda_{U|SM} &\xrightarrow{p} \frac{\sigma_{SM,U}}{\sigma_U} = \frac{\sigma_{SM,\chi}\sigma_U}{\sigma_\chi\sigma_{SM}} \\
\lambda_{U|SM,X} &\xrightarrow{p} \frac{\sigma_\chi\sigma_{SM,U}}{\sigma_{SM}\sigma_\chi - \sigma^2_{SM,\chi}} = \frac{\sigma_U\sigma_{SM,\chi}}{\sigma_{SM}\sigma_\chi - \sigma^2_{SM,\chi}}
\end{aligned}$$

Denote $\lambda_{U|SM,X} = \Pi$. Then the regression coefficients of the auxiliary regressions in (6) and (7) can be expressed in terms of covariances and Π.

$$\begin{aligned}
\mathring{\beta} &\xrightarrow{p} \beta + \gamma\frac{\sigma_{SM,\chi}}{\gamma\sigma_{SM}} + \frac{\sigma_{SM,\chi}\sigma_U}{\sigma_\chi\sigma_{SM}} \\
&= \beta + \frac{\sigma_{SM,\chi}}{\sigma_{SM}} + \Pi\frac{\sigma_{SM} - \frac{\sigma^2_{SM,\chi}}{\sigma_\chi}}{\sigma_{SM}} \\
&= \beta + \frac{\sigma_{SM,\chi}}{\sigma_{SM}} + \Pi\left(1 - \frac{\sigma^2_{SM,\chi}}{\sigma_\chi\sigma_{SM}}\right)
\end{aligned}$$

$$\tilde{\beta} \xrightarrow{p} \beta + \Pi$$

Subtracting yields

$$\mathring{\beta} - \tilde{\beta} \xrightarrow{p} \frac{\sigma_{SM,\chi}}{\sigma_{SM}} + \Pi\left(1 - \frac{\sigma^2_{SM,\chi}}{\sigma_\chi\sigma_{SM}}\right) - \Pi \tag{25}$$

$$\xrightarrow{p} \frac{\sigma_{SM,\chi}}{\sigma_{SM}}\left(1 - \frac{\sigma_{SM,\chi}}{\sigma_\chi}\Pi\right). \tag{26}$$

Substituting the expressions above and using the definitions of $\tilde{R}^2$, $\mathring{R}^2$ and R^2_{max} yields

$$\beta^* = \tilde{\beta} - (\mathring{\beta} - \tilde{\beta})\frac{R^2_{max} - \tilde{R}^2}{\tilde{R}^2 - \mathring{R}^2}. \tag{27}$$

For a derivation of (27) we refer to Oster (2019).

B Summary Statistics

	Full Sample	Smartphone Users
Female	0.474	0.492
Age	53.242	51.088
Marital status	0.541	0.547
Urbanicity		
Slightly Urban	0.205	0.208
Moderately Urban	0.169	0.165
Very Urban	0.226	0.233
Extremely Urban	0.187	0.186
Income	2888.7	2970.7
Education		
VMBO	0.162	0.143
HAVO/VWO	0.098	0.102
MBO	0.236	0.237
HBO	0.293	0.305
WO	0.166	0.176
Origin		
First Generation, Western	0.038	0.039
First Generation, Non-Western	0.061	0.058
Second Generation, Western	0.050	0.055
Second Generation, Non-Western	0.042	0.047
Religion		
Certainly Religious	0.118	0.110
Somewhat Religious	0.199	0.190
Barely Religious	0.260	0.263
BMI	25.799	25.682
Total Sickness	1.320	1.289
Personality		
Extraversion	22.112	22.270
Agreeableness	28.271	28.425
Conscientiousness	27.517	27.639
Neuroticism	25.033	24.933
Openness	25.142	25.461
Number of Observations	2060	1687

Table 8: Summary Statistics

C List of Diseases

1	angina, pain in the chest
2	a heart attack including infarction or coronary thrombosis or another heart problem including heart failure
3	high blood pressure or hypertension
4	high cholesterol content in blood
5	a stroke or brain infarction or a disease affecting the blood vessels in the brain
6	diabetes or a too high blood sugar level
7	chronic lung disease such as chronic bronchitis or emphysema
8	asthma
9	arthritis, including osteoarthritis, or rheumatism, bone decalcification or osteoporosis
10	cancer or malignant tumor, including leukemia or lymphoma, but excluding less serious forms of skin cancer
11	a gastric ulcer or duodenal ulcer, peptic ulcer
12	Parkinson's disease
13	cataract
14	a broken hip or thigh bone
15	another fracture
16	Alzheimer, dementia, organic brain syndrome, senility, or another serious memory problem
17	benign tumor (skin tumor, polyps, angioma)
18	COVID-19 (new corona virus)
19	other afflictions not yet mentioned

Table 9: Physical Health: List of Diseases