# The Social Media Infodemic and Anxiety During the COVID-19 Pandemic

**Jack (Quan Cheng) Xie**

Student ID: 31091325

University of British Columbia

ECON 398: Introduction to Applied Economics

Professor Jonathan L. Graves

8 December 2020

# I. Introduction

In the outbreak of the novel coronavirus disease (COVID-19) and the resulting human and economic toll, news consumption has skyrocketed as the public seeks day-to-day information on the development of the disease during lockdown. The New York Times reports that time spent on news sites have increased 46 percent early in the 2020 pandemic compared to the previous year, while overall visits rose 57 percent.

The ubiquity of social media has made these platforms key channels for communication and dissemination of information, especially during the lockdown period. However, social media has come under increasing scrutiny for its ability to spread misinformation (see Cinelli, Quattrociocchi et al. 2020) and its potentially detrimental effects mental health (see Gao, Zheng et al. 2020). While some studies have linked social media use to negative mental health outcomes, others have found no evidence for harm, and even mental health benefits (Berryman et al 2018). The WHO has highlighted the perils of the "infodemic," the overabundance of information, as a concern in mitigating the physical and mental health during the COVID-19 pandemic (2020). Health officials are concerned that the prevalence of the misinformation threatens to cause social unrest and erode trust in institutions, undermining the response to mitigate the global public health crisis, and exasperating mental health challenges such as anxiety and depression during the pandemic lockdown.

This paper examines the effect of social media use on generalized anxiety symptoms in the Canadian population during the COVID-19 pandemic. We use data from Statistics Canada to examine the association of using social media as an information source on clinical symptoms of anxiety, and compare the its effect against alternative information sources and other behavioural and demographic factors. While we find evidence to support an association with anxiety outcomes with the interaction between social media use and exposure to misinformation, we find no conclusive evidence that social media contributes to anxiety caused by misinformation more than other prevalent sources of information.

# II. Data

We use microdata from the Canadian Perspectives Survey Series (SPSS) administered by Statistics Canada. The SPSS is a set of online surveys that began in March of 2020. Specifically, we use data from the fourth survey in the series, which was collected from July 20, 2020 to July 26, 2020. The survey included a wide range of questions related to the experiences and welfare of the Canadian population during the COVID-19 pandemic, including questions about the respondent's mental health, behaviour, changes in habits, and general sentiments and concerns during the pandemic.

The target population of the survey is residents of the ten Canadian provinces 15 years of age or older. The survey sample is a randomly selected subset of respondents to the Labor Force Survey. The sample represents groups that exclude less than 2% of Canadians aged 15 and over (Statistics Canada 2020).

Table 1: Categorization of GAD-7 score and cut-point

| | Severity of Generalized Anxiety | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | **No Symptoms** | | **Minimal Symptoms** | | **Mild Symptoms** | | **Moderate Symptoms** | | **Severe Symptoms** | | **Total** | |
| | Freq | Col % | Freq | Col % | Freq | Col % | Freq | Col % | Freq | Col % | Freq | Col % |
| **Generalized Anxiety Cut-point Classifications (Dependent Variables)** | | | | | | | | | | | | |
| NOGAD | 1,421 | 100% | | | | | | | | | 1,421 | 34.8% |
| MGAD | | | 1,360 | 100% | 872 | 100% | | | | | 2,232 | 54.6% |
| GAD | | | | | | | 285 | 100% | 149 | 100% | 434 | 10.6% |
| **Total** | 1,421 | 100% | 1,360 | 100% | 872 | 100% | 285 | 100% | 149 | 100% | 4,087 | 100% |
| **Generalized Anxiety Severity Score** | | | | | | | | | | | | |
| 0 | 1,421 | 100% | | | | | | | | | 1,421 | 34.8% |
| 1 | | | 409 | 30.1% | | | | | | | 409 | 10.0% |
| 2 | | | 364 | 26.8% | | | | | | | 364 | 8.9% |
| 3 | | | 310 | 22.8% | | | | | | | 310 | 7.6% |
| 4 | | | 277 | 20.4% | | | | | | | 277 | 6.8% |
| 5 | | | | | 235 | 26.9% | | | | | 235 | 5.7% |
| 6 | | | | | 224 | 25.7% | | | | | 224 | 5.5% |
| 7 | | | | | 229 | 26.3% | | | | | 229 | 5.6% |
| 8 | | | | | 92 | 10.6% | | | | | 92 | 2.3% |
| 9 | | | | | 92 | 10.6% | | | | | 92 | 2.3% |
| 10 | | | | | | | 79 | 27.7% | | | 79 | 1.9% |
| 11 | | | | | | | 56 | 19.6% | | | 56 | 1.4% |
| 12 | | | | | | | 51 | 17.9% | | | 51 | 1.2% |
| 13 | | | | | | | 56 | 19.6% | | | 56 | 1.4% |
| 14 | | | | | | | 43 | 15.1% | | | 43 | 1.1% |
| 15 | | | | | | | | | 27 | 18.1% | 27 | 0.7% |
| 16 | | | | | | | | | 23 | 15.4% | 23 | 0.6% |
| 17 | | | | | | | | | 20 | 13.4% | 20 | 0.5% |
| 18 | | | | | | | | | 18 | 12.1% | 18 | 0.4% |
| 19 | | | | | | | | | 14 | 9.4% | 14 | 0.3% |
| 20 | | | | | | | | | 8 | 5.4% | 8 | 0.2% |
| 21 | | | | | | | | | 39 | 26.2% | 39 | 1.0% |
| **Total** | 1,421 | 100% | 1,360 | 100% | 872 | 100% | 285 | 100% | 149 | 100% | 4,087 | 100% |

**Note:** This table tabulates the dependent variables in our models and scores on the GAD-7 questionnaire against categories of the severity of generalized anxiety disorder. Each classification in the general anxiety cut-point represents a dummy dependent variable we use to model mental health outcomes. In particular, the GAD variable is a dummy variable that indicates clinically significant generalized anxiety disorder symptoms. The reported percentages are column-wise unweighted frequency for each variable or category value.

## A. Dependent Variable: General Anxiety Disorder Symptoms

The survey incorporates the GAD-7, a standardized questionnaire designed to diagnose generalized anxiety disorder and screen for panic, social anxiety, and post-traumatic stress disorders. The efficacy of GAD-7 has been validated extensively in primary care settings, though it cannot act as a replacement for a clinical diagnoses (see PHQ and PRIME-MD 2010).

The results of the questionnaire are evaluated on a 21-point scale, which categorizes the severity of an respondent's anxiety symptoms. The dataset includes a dummy variable that distinguishes a cut-point for moderate and severe symptoms. The cut-point indicates that the symptoms constitute a "clinically significant condition" (PHQ and PRIME-MD 2010). Roughly 10.6% of our unweighted sample (13.14% in weighted sample) fall within the

cut-point group, with 434 observed cases. We use this variable (GAD) as the primary dependent variable in our models for anxiety outcomes. We also apply cut-points and mild or minimal symptoms (MGAD) and no symptoms (NOGAD) for comparison analysis. Table 1 shows the breakout of the categories and GAD variable cutoff.

## B. Treatment and Covariates

For the treatment and covariates, we use a set of behavioural and demographic survey responses as observables. Our main treatment is a categorical variable indicating the use of social media posts from news organizations or magazines for COVID-19 information since the beginning of the pandemic.

In the extension model, we also use categorical variables indicating the main source of information about COVID-19. The main information source variable is interacted with the frequency of seeing suspected misinformation about COVID-19 online. Table 2 shows the breakout of the treatment groups by whether they meet the cut-point for moderate or severe GAD.

To isolate the treatment effect of the social media use, we select behavioural covariates unrelated to social media use based on the statistical significance of their association with GAD outcomes. These covariates comprise mainly responses to questions regarding concerns about health and social or domestic situations. Table 2 lists these controls tabulated by the GAD cut-point group.

In addition to the behavioural covariates, we use variables to control for treatment effects conditioned on

Table 2: Summary of Behavioural Treatments and Covariates

| | **Generalized Anxiety Cut-point Classifications** | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| | **No GAD** | | **Mild GAD** | | **GAD** | | **Total** | |
| | Freq | Col % | Freq | Col % | Freq | Col % | Freq | Col % |
| **Use social media posts from news organizations for COVID-19 info** | | | | | | | | |
| From news orgs/magazines only | 220 | 16.2% | 433 | 19.9% | 99 | 23.4% | 752 | 19.0% |
| From other users/influencers only | 156 | 11.5% | 265 | 12.2% | 57 | 13.5% | 478 | 12.1% |
| From both sources | 136 | 10.0% | 366 | 16.8% | 88 | 20.8% | 590 | 14.9% |
| From Neither | 848 | 62.4% | 1,114 | 51.1% | 179 | 42.3% | 2,141 | 54.1% |
| **Total** | 1,360 | 100.0% | 2,178 | 100.0% | 423 | 100.0% | 3,961 | 100.0% |
| **Main source of information for COVID-19** | | | | | | | | |
| Social media | 99 | 7.0% | 186 | 8.3% | 54 | 12.4% | 339 | 8.3% |
| News outlets | 750 | 52.8% | 1,111 | 49.8% | 173 | 39.9% | 2,034 | 49.8% |
| Government sources | 384 | 27.0% | 673 | 30.2% | 155 | 35.7% | 1,212 | 29.7% |
| Personal sources | 122 | 8.6% | 184 | 8.2% | 39 | 9.0% | 345 | 8.4% |
| Other sources/None | 66 | 4.6% | 77 | 3.5% | 13 | 3.0% | 156 | 3.8% |
| **Total** | 1,421 | 100.0% | 2,231 | 100.0% | 434 | 100.0% | 4,086 | 100.0% |
| **Frequency of seeing suspected false COVID-19 info** | | | | | | | | |
| Multiple times a day | 220 | 19.5% | 500 | 25.2% | 135 | 34.9% | 855 | 24.4% |
| Once a day | 150 | 13.3% | 258 | 13.0% | 58 | 15.0% | 466 | 13.3% |
| At least once a week | 308 | 27.2% | 638 | 32.2% | 104 | 26.9% | 1,050 | 30.0% |
| Rarely | 372 | 32.9% | 514 | 25.9% | 78 | 20.2% | 964 | 27.5% |
| Never | 81 | 7.2% | 74 | 3.7% | 12 | 3.1% | 167 | 4.8% |
| **Total** | 1,131 | 100.0% | 1,984 | 100.0% | 387 | 100.0% | 3,502 | 100.0% |

(Continued)

Table 2: Summary of Behavioural Treatments and Covariates (Continued)

| | Generalized Anxiety Cut-point Classifications | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| | **No GAD** | | **Mild GAD** | | **GAD** | | **Total** | |
| | Freq | Col % | Freq | Col % | Freq | Col % | Freq | Col % |
| **Concern about my own health** | | | | | | | | |
| Not at all | 408 | 28.9% | 327 | 14.7% | 41 | 9.4% | 776 | 19.0% |
| Somewhat | 768 | 54.3% | 1,327 | 59.5% | 171 | 39.4% | 2,266 | 55.6% |
| Very | 179 | 12.7% | 448 | 20.1% | 143 | 32.9% | 770 | 18.9% |
| Extremely | 59 | 4.2% | 128 | 5.7% | 79 | 18.2% | 266 | 6.5% |
| **Total** | 1,414 | 100.0% | 2,230 | 100.0% | 434 | 100.0% | 4,078 | 100.0% |
| **Concern about family member's health** | | | | | | | | |
| Not at all | 396 | 28.8% | 396 | 18.0% | 59 | 13.9% | 851 | 21.3% |
| Somewhat | 648 | 47.1% | 951 | 43.3% | 114 | 26.9% | 1,713 | 42.9% |
| Very | 252 | 18.3% | 619 | 28.2% | 135 | 31.8% | 1,006 | 25.2% |
| Extremely | 80 | 5.8% | 230 | 10.5% | 116 | 27.4% | 426 | 10.7% |
| **Total** | 1,376 | 100.0% | 2,196 | 100.0% | 424 | 100.0% | 3,996 | 100.0% |
| **Concern about family stress from confinement** | | | | | | | | |
| Not at all | 821 | 58.7% | 729 | 32.8% | 80 | 18.5% | 1,630 | 40.2% |
| Somewhat | 466 | 33.3% | 1,047 | 47.1% | 137 | 31.6% | 1,650 | 40.7% |
| Very | 86 | 6.1% | 367 | 16.5% | 137 | 31.6% | 590 | 14.6% |
| Extremely | 26 | 1.9% | 79 | 3.6% | 79 | 18.2% | 184 | 4.5% |
| **Total** | 1,399 | 100.0% | 2,222 | 100.0% | 433 | 100.0% | 4,054 | 100.0% |
| **Concern about violence in the home** | | | | | | | | |
| Not at all | 1,344 | 96.3% | 2,133 | 96.0% | 390 | 90.5% | 3,867 | 95.5% |
| Somewhat | 29 | 2.1% | 60 | 2.7% | 27 | 6.3% | 116 | 2.9% |
| Very | 13 | 0.9% | 15 | 0.7% | 6 | 1.4% | 34 | 0.8% |
| Extremely | 9 | 0.6% | 15 | 0.7% | 8 | 1.9% | 32 | 0.8% |
| **Total** | 1,395 | 100.0% | 2,223 | 100.0% | 431 | 100.0% | 4,049 | 100.0% |
| **My general health** | | | | | | | | |
| Excellent | 547 | 38.6% | 367 | 16.5% | 26 | 6.0% | 940 | 23.1% |
| Very good | 610 | 43.1% | 970 | 43.5% | 94 | 21.8% | 1,674 | 41.1% |
| Good | 231 | 16.3% | 716 | 32.1% | 191 | 44.3% | 1,138 | 27.9% |
| Fair | 23 | 1.6% | 156 | 7.0% | 96 | 22.3% | 275 | 6.7% |
| Poor | 5 | 0.4% | 19 | 0.9% | 24 | 5.6% | 48 | 1.2% |
| **Total** | 1,416 | 100.0% | 2,228 | 100.0% | 431 | 100.0% | 4,075 | 100.0% |

*Note:* This table tabulates the behavioural treatment and covariates in our model with the dependent dummy variables in our main regressions. The reported percentages are column-wise unweighted frequency for each variable or category value. Three categories for main source of information for COVID-19 are combined from several options in the available survey question responses. Other sources/None combines responses with sources not listed in the question and not looking for COVID-19 information. Government sources combines sources from federal, provincial, and municipal health agencies and daily announcements. Personal sources combine family/friends/colleagues, health professionals, and place of employment.

demographic groups, such as age, sex, education, and urban or rural location. Demographic covariates are also selected on statistical significance to GAD outcomes. Table 3 summarizes the demographic covariates.

Table 3: Summary of Demographic Covariates

| | Generalized Anxiety Cut-point Classifications | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| | **No GAD** | | **Mild GAD** | | **GAD** | | **Total** | |
| | Freq | Col % | Freq | Col % | Freq | Col % | Freq | Col % |
| **Age of respondent** | | | | | | | | |
| 15 to 24 years old | 26 | 1.8% | 97 | 4.3% | 45 | 10.4% | 168 | 4.1% |
| 25 to 34 years old | 100 | 7.0% | 322 | 14.4% | 93 | 21.4% | 515 | 12.6% |
| 35 to 44 years old | 171 | 12.0% | 413 | 18.5% | 87 | 20.0% | 671 | 16.4% |
| 45 to 54 years old | 225 | 15.8% | 366 | 16.4% | 78 | 18.0% | 669 | 16.4% |
| 55 to 64 years old | 362 | 25.5% | 492 | 22.0% | 74 | 17.1% | 928 | 22.7% |
| 65 to 74 years old | 384 | 27.0% | 413 | 18.5% | 41 | 9.4% | 838 | 20.5% |
| 75 years and older | 153 | 10.8% | 129 | 5.8% | 16 | 3.7% | 298 | 7.3% |
| **Total** | 1,421 | 100.0% | 2,232 | 100.0% | 434 | 100.0% | 4,087 | 100.0% |
| **Sex of respondent** | | | | | | | | |
| Male | 738 | 51.9% | 1,006 | 45.1% | 149 | 34.3% | 1,893 | 46.3% |
| Female | 683 | 48.1% | 1,226 | 54.9% | 285 | 65.7% | 2,194 | 53.7% |
| **Total** | 1,421 | 100.0% | 2,232 | 100.0% | 434 | 100.0% | 4,087 | 100.0% |
| **Highest level of education completed** | | | | | | | | |
| Less than highschool or equivalent | 82 | 5.8% | 92 | 4.1% | 33 | 7.6% | 207 | 5.1% |
| Highschool or equivalent | 262 | 18.4% | 405 | 18.1% | 106 | 24.4% | 773 | 18.9% |
| Trade school | 153 | 10.8% | 168 | 7.5% | 36 | 8.3% | 357 | 8.7% |
| College/non-university | 355 | 25.0% | 519 | 23.3% | 106 | 24.4% | 980 | 24.0% |
| University below bachelor's | 65 | 4.6% | 71 | 3.2% | 11 | 2.5% | 147 | 3.6% |
| Bachelor's degree | 320 | 22.5% | 612 | 27.4% | 99 | 22.8% | 1,031 | 25.2% |
| Degree above Bachelor's | 184 | 12.9% | 365 | 16.4% | 43 | 9.9% | 592 | 14.5% |
| **Total** | 1,421 | 100.0% | 2,232 | 100.0% | 434 | 100.0% | 4,087 | 100.0% |
| **Rural/Urban indicator** | | | | | | | | |
| Rural | 340 | 23.9% | 443 | 19.8% | 67 | 15.4% | 850 | 20.8% |
| Urban | 1,081 | 76.1% | 1,789 | 80.2% | 367 | 84.6% | 3,237 | 79.2% |
| **Total** | 1,421 | 100.0% | 2,232 | 100.0% | 434 | 100.0% | 4,087 | 100.0% |

*Note:* This table tabulates the demographic covariates in our model with the dependent dummy variables in our main regressions. The reported percentages are column-wise unweighted frequency for each category value. Category values are reported as on the survey.

# III. Methodology

To estimate the effect of social media use on likelihood of developing significant symptoms of generalized anxiety disorder, we specify the following regression model:

$$\text{GAD}_i = \beta_0 + \mathbf{SM_i'}\beta_1 + \mathbf{M_i'}\beta_2 + \mathbf{F_i'}\beta_3\mathbf{X_{Bi}'}\beta_4 + \mathbf{X_{Di}'}\beta_5 + \varepsilon_i, \tag{1}$$

where $\text{GAD}_i$ is the probability of individual $i$ having clinically significant symptoms of general anxiety disorder (moderate or severe). $\mathbf{SM_i}$ is a category vector for social media use as a COVID-19 information source. The categories are the source of the social media posts used (from media organizations, other users, both, or neither). $\beta_1$ is a estimator vector for the increased probability of developing significant symptoms of generalized anxiety disorder associated each category of social media use.

$\mathbf{M_i}$ is a categorical vector indicating the main source of information about COVID-19 (social media, news

outlets, government, personal, or other/none). $\mathbf{F_i}$ is a vector categorizing the frequency of seeing suspected false or misinformation online. $\mathbf{X_{Bi}}$ and $\mathbf{X_{Di}}$ are the behaviour observables unrelated to social media use and demographics associated with anxiety listed in Table 2 and Table 3. $\varepsilon_i$ is the error term.

The key causal assumptions required for this model are that observables in the model sufficiently explains anxiety outcomes, and that the treatment effects are independent to the treatment group conditioned on the model's observable variables (selection on observables). In our data, this is a particularly strong assumption. The Mayo Clinic explains that the causes of mental health outcomes, such as anxiety, arise "a complex interaction of biological and environmental factors" (n.d.), including personality, genetics, and the experiences of individuals. Given the broad intent of the survey design, there are likely significant explanatory variables for anxiety outcomes not captured in the data. Additionally, with the sample size of 434 observations in the GAD cut-point group and the administering of the survey, there is likely to be measurement errors in regards to the outcomes.

We use a second model to focus on the interaction of information source and perceived veracity of COVID-19 information online on generalized anxiety.

$$\text{GAD}_i = \theta_0 + \mathbf{M_i'}\theta_1 + \theta_2 G_i + \left(G_i \times \mathbf{M_i'}\right)\theta_3 + \mathbf{X_{Bi}'}\theta_4 + \mathbf{X_{Di}'}\theta_5 + \varepsilon_i, \tag{2}$$

where $G_i$ is a dummy variable for frequent (multiple times a day) suspicion of false or misinformation about COVID-19. This model allows us to better compare across the effect of information sources in infodemic mental health outcomes.

# IV. Results

## A. Baseline Results

Table 4 reports the regression estimates of model (1). Column 1 shows the results of a simplified treatment model without covariates. The treatments are the categories of social media use for COVID-19 information, if used at all, based on source of posts. The result indicates that use of social media posts from news organizations is associated with a 7% increase in probability of clinically significant symptoms of generalized anxiety disorder (GAD). The result shows no statistically significant effect from using social media posts by other users or influencers only, and a 7.8% probability increase in GAD outcomes from using both social media sources.

Column 2 extends the baseline model by adding two categorical covariates, the primary source of information (social media, news outlets, etc.) and frequency of seeing suspected misinformation about COVID-19 online. The effects of the social media use estimators in the baseline model are persistent in this specification (6.9% and 8.4% for news org posts and both sources).

The covariate estimators show significant effects on GAD outcomes. Using social media as the main information source corresponds to a 14% probability increase in GAD outcomes. This effect is compared to a baseline

Table 4: Social Media Use for COVID-19 Information and Generalized Anxiety Disorder Symptoms

| | Moderate or Severe Symptoms | | | Mild or Minimal Symptoms | No Symptoms |
|---|---|---|---|---|---|
| | $GAD_i$ (1) | $GAD_i$ (2) | $GAD_i$ (3) | $MGAD_i$ (4) | $NOGAD_i$ (5) |
| ***Use COVID-19 information from social media*** | | | | | |
| Posts from news organizations or magazines only | 0.070** (0.030) | 0.069** (0.028) | 0.025 (0.023) | 0.0014 (0.035) | -0.027 (0.028) |
| Posts from other users only | 0.019 (0.028) | 0.019 (0.027) | -0.033 (0.030) | 0.065 (0.045) | -0.032 (0.035) |
| Posts from both sources | 0.078** (0.031) | 0.084*** (0.031) | 0.023 (0.031) | 0.021 (0.041) | -0.044 (0.031) |
| ***Main source of information for COVID-19 (Base level: Other sources/None)*** | | | | | |
| Social media | | 0.14*** (0.040) | 0.11*** (0.041) | -0.24*** (0.090) | 0.13 (0.085) |
| News outlets | | 0.086*** (0.031) | 0.058 (0.036) | -0.25*** (0.084) | 0.19** (0.082) |
| Government sources | | 0.15*** (0.036) | 0.11*** (0.038) | -0.31*** (0.087) | 0.21** (0.084) |
| Personal sources | | 0.13*** (0.048) | 0.095* (0.049) | -0.30*** (0.097) | 0.20** (0.094) |
| ***Frequency of suspecting false COVID-19 info (Base level: Never)*** | | | | | |
| Multiple times a day | | 0.14*** (0.034) | 0.081** (0.035) | -0.042 (0.068) | -0.039 (0.063) |
| Once a day | | 0.077** (0.035) | 0.056 (0.038) | -0.022 (0.071) | -0.033 (0.065) |
| At least once a week | | 0.042 (0.028) | 0.026 (0.030) | 0.013 (0.066) | -0.039 (0.062) |
| Rarely | | 0.028 (0.028) | -0.0046 (0.032) | 0.015 (0.066) | -0.010 (0.060) |
| ***Controls*** | | | | | |
| Behavioural Controls | No | No | Yes | Yes | Yes |
| Demographic Controls | No | No | Yes | Yes | Yes |
| Observations | 3,961 | 3,495 | 3,397 | 3,397 | 3,397 |
| R-squared | 0.010 | 0.039 | 0.243 | 0.095 | 0.230 |

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

*Note:* Columns 1 to 3 report the estimated regression coefficient of the probability of having clinically significant symptoms of generalized anxiety disorder (GAD) on social media use for information about COVID-19. The dependent and treatments are categories of social media use, with a baseline of not using social media for information. Columns 2 and 3 report regression results that incorporate observables based on survey responses to behaviour and demographic questions. Columns 4 and 5 report regression results with the control model on having mild or minimal GAD symptoms (column 4) and no symptoms (column 5). Standard errors for estimator coefficients are displayed in brackets. Sampling weights and robust standard errors are used.

of using other sources (not included in survey options) or not looking for COVID-19 information. Other categories of information sources also have strong effects, particularly government sources (15%). Estimators for

higher frequencies of suspecting misinformation have significant effects on GAD outcomes. Multiple instances of suspicions a day corresponds to a 14% risk increase, and once a day to a 7.7% increase.

Column 3 reports results of the fully-specified model, which incorporates additional controls for behavioural and demographic observables. The behavioural controls are responses to questions about health and domestic concerns during COVID-19. Demographic controls are age, sex, education level, and urban or rural residence. The specifications of the control variables are shown in Table 2 and Table 3. With the controls, the statistical significance of the treatment estimators disappear. This suggests that the behavioural or demographic covariates are more relevant factors of GAD outcomes. The estimators of main information sources and suspicion frequency largely persist with the controls. We focus on the effects of these covariates in the extension model on Table 5.

Columns 4 and 5 report regression results from the fully-specified model with two dummy outcomes for less severe GAD symptoms. The outcome groups are mild or minimal GAD symptoms (MGAD) in column 4 and no GAD symptoms (NOGAD) in column 5. The results show no significant effects for less severe GAD outcomes as a result of social media use.

However, strong negative effects for MGAD are prevalent in each of the main information categories. This indicates that the base level (using other sources or not seeking COVID-19 information) has strong effects with MGAD outcomes. For NOGAD outcomes, except for social media, the main information source categories have stronger effects than with GAD outcomes. This suggests that the main information sources effects are prevalent in the extremities of symptoms (either severe/moderate or no symptoms). Respondents who keep up with COVID-19 information tend to be either very anxious or not anxious at all. Though the magnitude of the other information source estimators are large, the effect of using social media as the main source is more inconclusive due to high standard error. The statistical power of the estimator may be improved with a larger sample size or more effective sampling methods.

Frequency of suspecting misinformation is insignificant for MGAD and NOGAD, as may be expected as a causal factor for GAD outcomes.

## B. Interaction Effects

Table 5 shows results of model (2), which focuses on the interaction between main information source and frequency of suspecting misinformation.

Column 1 reports results for GAD outcomes without interaction. Except for two differences, the specifications are similar to the control model in Table 4. First, we omit the previous social media use treatments, as their estimators are insignificant with the controls. Second, we reduce the categories of frequency of suspected misinformation to a dummy that indicates only multiple instances of suspicion a day. This specification captures the most significant category while reducing the dimensionality of our interaction models.

The results are also similar to the baseline. Social media corresponds to a 11% (0% change from baseline model) increase in risk of GAD outcomes. Government sources is also 11% (0% change from baseline model), and

Table 5: Interaction Effects of Information Sources With Frequency of Suspecting Misinformation and Generalized Anxiety Disorder Symptoms

| | Moderate or Severe Symptoms | | Mild or Minimal Symptoms | No Symptoms |
|---|---|---|---|---|
| | $GAD_i$ (1) | $GAD_i$ (2) | $MGAD_i$ (3) | $NOGAD_i$ (4) |
| ***Main source of information for COVID-19 (Base level: Other sources/none)*** | | | | |
| Social media | 0.11*** | 0.058 | -0.17 | 0.11 |
| | (0.041) | (0.047) | (0.11) | (0.11) |
| News outlets | 0.058 | 0.016 | -0.21* | 0.19* |
| | (0.036) | (0.040) | (0.11) | (0.11) |
| Government sources | 0.11*** | 0.039 | -0.27** | 0.23** |
| | (0.038) | (0.044) | (0.11) | (0.11) |
| Personal sources | 0.089* | 0.043 | -0.26** | 0.22* |
| | (0.048) | (0.052) | (0.12) | (0.12) |
| ***Frequency of suspecting false COVID-19 info (Base level: Once a day or less)*** | | | | |
| Multiple times a day | 0.063** | -0.085* | 0.067 | 0.018 |
| | (0.025) | (0.050) | (0.11) | (0.12) |
| ***Main info × Freq suspecting false COVID-19 info*** | | | | |
| Social media × Multiple times a day | | 0.17** | -0.20 | 0.025 |
| | | (0.077) | (0.14) | (0.13) |
| News outlets × Multiple times a day | | 0.11* | -0.099 | -0.013 |
| | | (0.060) | (0.12) | (0.12) |
| Government sources × Multiple times a day | | 0.22*** | -0.14 | -0.083 |
| | | (0.066) | (0.13) | (0.13) |
| Personal sources × Multiple times a day | | 0.13 | -0.066 | -0.067 |
| | | (0.092) | (0.15) | (0.15) |
| ***Controls*** | | | | |
| Behavioural Controls | Yes | Yes | Yes | Yes |
| Demographic Controls | Yes | Yes | Yes | Yes |
| Observations | 3,402 | 3,402 | 3,402 | 3,402 |
| R-squared | 0.237 | 0.241 | 0.094 | 0.230 |

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

*Note:* Columns 1 and 2 report the regression estimators of the probability of having moderate or severe symptoms of generalized anxiety disorder. Column 2 features the interaction of different information sources and the frequency of suspecting misinformation. Column 3 show the results of the same model estimating outcomes with mild or minimal symptoms, and column 4 estimates outcomes with no symptoms. Standard errors are shown in brackets. Sampling weights and robust standard errors are used.

personal sources (information from family, friends, health professionals, workplace) around 9% (-0.5% change). Frequent suspicion of misinformation corresponds to a 6.3% risk increase in GAD outcomes (-2% change).

Column 2 adds interactions term between main information sources and suspicion frequency. The stand-alone estimators for the main information sources become insignificant, while the effect of suspicion frequency increases to 8.5%. The interaction effects are significant. Social media and high suspicion frequency (multiple times a day) corresponds to a 17% risk increase in GAD outcomes. However, the highest interaction effect is government

sources, with 22% risk increase. Perhaps unsurprisingly, generalized anxiety is associated strongly with distrust of information and following the information closely, though not only from following social media in particular.

Column 3 and 4 repeat the same comparison analysis of MGAD outcomes (mild or minimal GAD symptoms) and NOGAD outcomes (no symptoms) in Table 4 with the extension model. The results indicate no significant interaction effects between main information source and suspicion frequency. The stand-alone effects of information sources are similar to the baseline. This would be expected if misinformation or distrust in information is a major contributor to anxiety, as it would not appear to affect the probability of less severe anxiety outcomes.

# V. Conclusion

In this paper, we examine the effect of social media use and misinformation on clinical anxiety symptoms in the Canadian population during the COVID-19 pandemic. In our survey data, we find initially significant correlations between social media use and generalized anxiety when confounded, but the correlation does not exists as an unconfounded factor.

Rather, We find a significant effects on GAD outcomes in frequent perceived encounters of misinformation and generalized anxiety. Though dominant informational use of social media together with a high level distrust in information leads to greater risk of clinical anxiety, the effects with other information sources are comparable. While misinformation is a potentially significant generalized anxiety factor in the COVID-19 crisis, we find inconclusive evidence that social media in particular contributes to the problem more than other channels of information.

# References


- **Berryman, C., Ferguson, C.J. & Negy, C.** 2018. "Social Media Use and Mental Health among Young Adults." Psychiatr Q 89, 307–314. https://doi.org/10.1007/s11126-017-9535-6
- **Cinelli, M., Quattrociocchi, W., Galeazzi, A. et al.** 2020. "The COVID-19 social media infodemic." Sci Rep 10, 16598. https://doi.org/10.1038/s41598-020-73510-5
- **Gao J, Zheng P, Jia Y, Chen H, Mao Y, Chen S, et al.** 2020. "Mental health problems and social media exposure during COVID-19 outbreak." PLoS ONE 15(4): e0231924. https://doi.org/10.1371/journal.pone.0231924
- **Mayo Clinic.** n.d. "Generalized anxiety disorder." *Mayo Foundation for Medical Education and Research (MFMER),* 1998-2020. https://www.mayoclinic.org/diseases-conditions/generalized-anxiety-disorder/symptoms-causes/syc-20360803.
- **New York Times.** 2020. "Coronavirus Brings a Surge to News Sites" March 20, 2020. https://nyti.ms/2J1AuY7

▷ **PHQ and PRIME-MD.** 2010. "Instructions for Patient Health Questionnaire (PHQ) and GAD-7 Measures." *PHQscreeners,* 2010. https://www.phqscreeners.com/images/sites/g/files/g10016261/f/201412/instructions.pdf.

▷ **Statistics Canada.** 2020. "Canadian Perspectives Survey Series 4: Information Sources Consulted During the Pandemic Public Use Microdata File." *Abacus Data Network,* 2020. https://hdl.handle.net/11272.1/AB2/IIIOGC.

▷ **World Health Organization.** 2020. "Let's flatten the infodemic curve." https://www.who.int/news-room/spotlight/let-s-flatten-the-infodemic-curve