Title: Tables appendix for “How COVID drove the evolution of fact-checking”
Authors: Samikshya Siwakoti (1), Kamya Yadav (1), Nicola Bariletto (2), Luca Zanotti (2), Ulaş Erdoğdu (3), Jacob N. Shapiro (4)
Date: May 6th, 2021
Note: The material contained herein is supplementary to the article named in the title and published in the Harvard Kennedy School (HKS) Misinformation Review.

Appendix A: Tables

Table A1. *List of coronavirus-related terms used in analysis.*

coronavirus, corona, koronavirus, wuhancoronavirus, kungflu, n95, covid-19, corona virus, covid19, sars-cov-2, covd, pandemic, coronapocalypse, chinese virus, chinavirus, cronvirus, virus, ١٩-كوفيد, وباء, لقاح, حظر تجوال, حظر تجول, اصاب, كوفيد, فيروس كورونا, وباء كورونا, كورونا, Korona, koronavirüs, aşı, mutasyon, virusi vya corona, कोविड, कोरोना , कोरोना वाइरस, कोभिड, चीनी वाइरस , चिन भाइरस , कोभिड-19

Table A2. *Difference-in-differences analyses – Engagement.*

The estimating equation for the model is:

$$Engagement_{j,i,t} = \alpha + \beta * did_{j,i,t} + \delta * pandemic\_dummy_{j,t} + \gamma * 2020\_dummy_{j,i} + \mathbf{X}'_j + \epsilon_{j,i,t}$$

Here, $Engagement_{j,i,t}$ refers to the total engagement metric (sum of likes, quote tweets, replies, retweets and no. of tweets by users) for website j in year i and month t; $did_{j,i,t}$ is the difference-in-differences indicator for website j, year i, month t (obtained as $2020\_dummy_{j,i} * pandemic\_dummy_{j,t}$); $pandemic\_dummy_{j,t}$ is an indicator equal to 1 for observations on website j in pandemic months (variable) t, 0 for pre-pandemic months; $2020\_dummy_{j,i}$ is an indicator equal to 1 for year 2020, 0 for 2019 observations for website j and year i; $\mathbf{X}'_j$ are website fixed effects and $\epsilon_{j,i,t}$ is the error term.

Table A2: Difference-in-differences, total engagement

	First pandemic month: March		First pandemic month: April	
	All year	Jan-Jun	All year	Jan-Jun
Diff-in-diff	-13252.2	3203.8	-10733.0	10259.4
	(-0.80)	(0.16)	(-0.81)	(0.50)
Pandemic months dummy	4660.1	8471.8	1698.5	6076.7
	(0.61)	(0.65)	(0.26)	(0.50)
2020 dummy	31211.1*	31211.1*	28217.3**	28217.3*
	(2.09)	(2.14)	(2.65)	(2.54)
Observations	360	180	360	180
Website FE	YES	YES	YES	YES

t statistics in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Notes: Difference-in-differences on total engagement (sum of likes, quotes, retweets, replies to, and number of, tweets containing a link to fact-checking organizations' websites). While we detect a significant increase in engagement in 2020 compared to 2019, there is no specific association with the pandemic period.

Table A3. *Difference-in-differences analyses – Posts.*

The estimating equation for the supply-side model is:

$$Posts_{j,i,t} = \alpha + \beta * did_{j,i,t} + \delta * pandemic\_dummy_{j,t} + \gamma * 2020\_dummy_{j,i} + \mathbf{X}'_j + \epsilon_{j,i,t}$$

Here, $Posts_{j,i,t}$ is the total number of posts by website j in year i and month t; $did_{j,i,t}$ is the difference-in-differences indicator for website j, year i, month t (obtained as $2020\_dummy_{j,i} * pandemic\_dummy_{j,t}$); $pandemic\_dummy_{j,t}$ is an indicator equal to 1 for observations on website j in pandemic months (variable) t, 0 for pre-pandemic months; $2020\_dummy_{j,i}$ is an indicator equal to 1 for year 2020, 0 for 2019 observations for website j and year i; $\mathbf{X}'_j$ are website fixed effects and $\epsilon_{j,i,t}$ is the error term.

Table A3: Difference-in-differences, number of posts

	First pandemic month: March		First pandemic month: April	
	All year	Jan-Jun	All year	Jan-Jun
Diff-in-diff	-50.13	-16.72	-64.91	-35.82
	(-0.83)	(-0.27)	(-1.34)	(-0.71)
Pandemic months dummy	83.94	77.58	67.58	64.91
	(1.62)	(1.40)	(1.77)	(1.60)
2020 dummy	115.7*	115.7*	122.5**	122.5**
	(2.07)	(2.00)	(2.96)	(2.89)
Observations	356	180	356	180
Website FE	YES	YES	YES	YES

t statistics in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Notes: Difference-in-differences analysis on the number of tweets by fact-checking organizations' handles. While we detect a significant increase in the number of posts in 2020 compared to 2019, there is no specific association with the pandemic period.

Table A4. *Difference-in-differences analyses - Per-post engagement.*

The estimating equation for the model is:

$$Per\_post\_engagement_{j,i,t} = \alpha + \beta * did_{j,i,t} + \delta * pandemic\_dummy_{j,t} + \gamma * 2020\_dummy_{j,i} + \mathbf{X}'_j + \epsilon_{j,i,t}$$

Here, $Per\_post\_engagement_{j,i,t}$ refers to the per-post engagement metric (sum of likes, quote tweets, replies, retweets and no. of tweets by users, divided by the total no. of tweets by **both users and fact-checking websites' handles**) for website j in year i and month t; $did_{j,i,t}$ is the difference-in-differences indicator for website j, year i, month t (obtained as $2020\_dummy_{j,i} * pandemic\_dummy_{j,t}$); $pandemic\_dummy_{j,t}$ is an indicator equal to 1 for observations on website j in pandemic months (variable) t, 0 for pre-pandemic months; $2020\_dummy_{j,i}$ is an indicator equal to 1 for year 2020, 0 for 2019 observations for website j and year i; $\mathbf{X}'_j$ are website fixed effects and $\epsilon_{j,i,t}$ is the error term.

Table A4: Difference-in-differences, per-post engagement

	First pandemic month: March		First pandemic month: April	
	All year	Jan-Jun	All year	Jan-Jun
Diff-in-diff	-6.811*	-9.872**	-7.944**	-13.69***
	(-2.42)	(-2.80)	(-3.18)	(-3.82)
Pandemic months dummy	2.128	6.383*	2.415	8.586**
	(1.06)	(2.27)	(1.31)	(2.80)
2020 dummy	4.392	4.392	4.658*	4.658*
	(1.81)	(1.72)	(2.36)	(2.31)
Observations	356	180	356	180
Website FE	YES	YES	YES	YES

t statistics in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Notes: Difference-in-differences on per-post engagement (sum of likes, quotes, retweets, replies to, and number of, tweets containing a link to fact-checking organizations' websites, divided by the total no. of tweets by both users and fact-checking organizations' websites' handles containing a link to fact-checking organizations' websites). In all specifications, we detect a statistically significant reduction in average engagement associated with the pandemic months and as compared to the control year (2019).