Vax Populi: The social costs of online vaccine skepticism

Matilde Giaccherini[1,2] Joanna Kopinska[3] Gabriele Rovigatti[4]

[1] *"Tor Vergata" University of Rome - CEIS,*
[2] *CESifo ,*
[3] *Sapienza University of Rome,*
[4] *Bank of Italy*

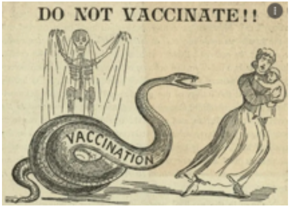
EEA-ESEM 2023 Conference - 28^{th} August 2023

Fake news and social media

Nothing spreads faster than a newly discovered airborne disease that could potentially kill millions...except rumors on social media.

- New tech internet & social media $\rightarrow$ **free access to news at lower quality** (not subject to fact-checking/editorial judgement)
- Result: **lower ability of consumers to distinguish authentic from fake news** $\rightarrow$ also due to the ideological echo chambers (Cinelli et al., 2021), increasing polarization (Flaxman et al., 2016, Sunstein, 2001, 2017, 2018), ideological self-segregation (Berinsky, 2017, Gentzkow and Shapiro, 2011) and misinformation spread (Allcott and Gentzkow, 2017)

Fake news and pediatric vaccines



Recent leading fake news: 1998 Wakefield's study **published and retracted** in The Lancet on the link between measles, mumps and rubella **(MMR) vaccine and autism**.

Italian novax movement exploded in 2012: the Court of Rimini recognizes the causal link between the MMR and autism $\rightarrow$ decrease in child immunization rates (Carrieri et al., 2019).

LANGUAGE

novax = anti-vax
activism = activity, movement, propaganda
hesitancy = opting-out, avoidance, skipping shots

Pediatric vaccines in Italy

National Plan of Vaccine Prevention (PNPV) establishes vaccine calendar and population eligible for shots free of charge at Local Health Authorities (LHAs)

MANDATORY

polio, diphtheria, tetanus, hepatitis B (combined with HIB and whooping cough as **hexavalent**, or 6-in-1 vaccine)

RECOMMENDED until 2017

MMR, chickenpox, meningo- and pneumococcal

- After 2010, coverage declined and several outbreaks of measles epidemics took place.
- In 2017 PNPV **enforced** and extended **mandatory shots to MMR and the recommended ones** stating that *"falling uptake was driven by* ***novax sentiment*** *and put in danger not only the eligible but also* ***the fragile****"* (Decree 73, 2017, "Lorenzin's Law").

RQ: Does online vaccine skepticism affect public health?

- As in Kim (2022), we use Twitter data to measure and track public attitudes $\longrightarrow$ vaccine-related tweets as a proxy for penetration of anti-vaccination movement (online and offline) in Italian municipalities .
- Pair geotagged tweets with disease-specific vaccine coverage rates, hospitalizations, and costs due to vaccine-preventable diseases for 2013-2018
- Formalize the sources of engodeneity that pervade the relationship between the spread of anti-vax opinions on media and vaccine hesitancy
- Use an IV strategy based on users' "followings-of-fofllowings" network to identify exogenous variation in anti-vax views
- In a Mixed 2SLS (Dhrymes and Lleras-Muney, 2005) estimate the effect of users' anti-vax stances on their municipality level vaccination rates and vaccine preventable hospitalizations
- Policy implications

Scraping data from `Twitter`

Through `Twitter` API for academia we collect all publicly available tweets on vaccines (Jan 1, 2013 to Dec 31, 2018) $\rightarrow \approx 2.04$ mln Query.

We map tweets with geolocation data ($\rightarrow$ 830,253 geotagged tweets for 80,471 unique users on 4220 municipalities) Descs. Maps User trend

Sentiment analysis: `VaxBERTo`

Train an anti-vax tweet classifier **`VaxBERTo`** on 2.04mln tweets, using the Italian version of the Natural Language Processing model `BERT` (as in Polignano et al 2019)

- capable of *understanding* the social media atypical language, with all contextual nuances (irony etc.)

For geeks, please see the Appendix

1 **manually** label (`anti-vax==0/1`) tweets created by media and renowned fake news accounts (48k tweets) (Pierri et al. 2020)

3 the **prediction phase**: the model is evaluated with the small test dataset and the remaining untagged textual data (2mln) are categorized: $l_\tau \in \{0, 1\}$

Identifying users' attitudes

`anti-vax == 1`

`anti-vax == 0`

Consider a user i producing a number of a_i of contents: $C_i = \{c_1, c_2, ...c_{a_i}\}$
The individual anti-vax stance is defined as the **share of anti-vax tweets in all their vax-related tweets** in year t:

$$s_{it} \equiv \frac{\sum_{\tau=1}^{a_{it}} c_\tau}{a_{it}} \times 100$$

Health data

Vaccination rates Descriptive statistics.

- Disease specific vaccination rates in the target pediatric population at municipality yearly level provided by LHAs, for the period 2013-2018

Hospitalization data Descriptive statistics.

- Hospital Discharge Data (SDO) on the **universe** of Italian hospital admissions for the period 2013-2016.
- focus on the diagnosis of vaccine-preventable diseases in:
  - vaccine-target population (children aged between 1 and 10 y.o.)
  - fragile population not targeted by the vaccines: newborns, pregnant women, and patients with immunosuppressing conditions (based on ICD-09)
- Construct hospitalization **rates and costs** per 100k residents at yearly municipality level.

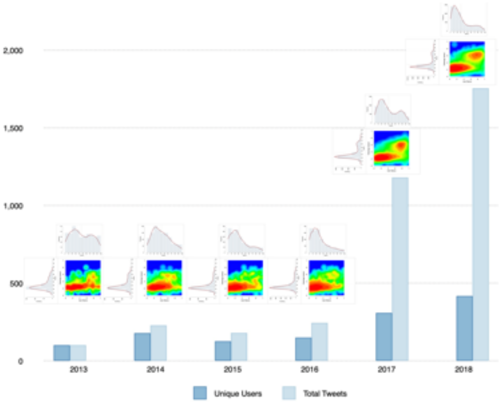
trend

The model of Opinion Dynamics and Network formation

Rationalize the evolution of social media anti-vax stances in Italy based on a model of social networks opinion dynamics proposed by Baumann et al, 2020: Details. Simulations.

- *exposure effect*: exposure to extreme-stances influences users' stance
- *link formation effect*: the controversialness of a vaccine-related topic endogenously exacerbates polarization by influencing the network formation process

Figure 1: Dynamics of Twitter activity on vaccination (2013-2018)



IV strategy

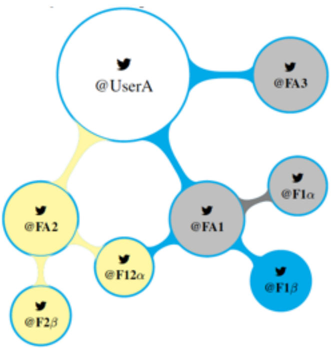
Exogenous variation in novax views of `Twitter` $\rightarrow$ **intransitivity** in network connections (Bramoullé et al., 2009)

- when **user's followings-of-followings** are not direct followings of the user, they have an impact on user's outcomes only through their effect on direct followings

LANGUAGE

friend = a user I follow
friend of friend = a user that my friend follows
follower = a user that follows me

followings-of-followings network



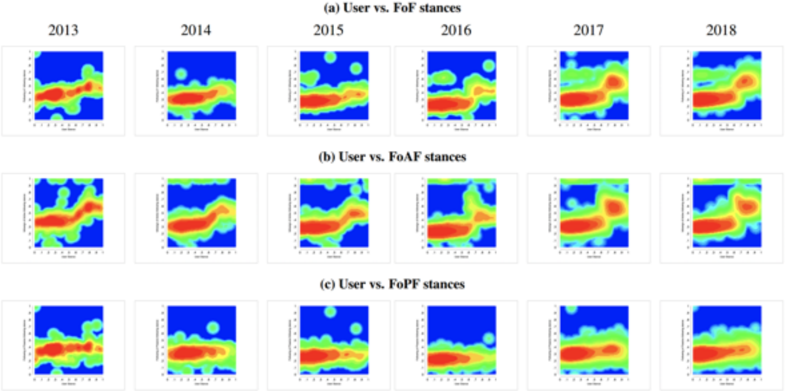
For each geotagged $\boxed{\text{initial user}}$ we define a 2-step neighborhood:

1. *followings*
   - ~~*active*~~
   - *passive* (∼ 48,2mln nodes for ∼ 8mln unique followings)
2. *followings-of-followings*
   - *active* (∼ 103,3mln nodes for ∼ 222k unique Followings of Passive followings - FoPf)
   - ~~*passive*~~

IV variable: each $\boxed{\text{initial user's}}$ i in time t features an indirect exposure to their *N* *followings- of -followings* novax stances as the **group-specific average anti-vax stances**:

$$ffs_{it} \equiv \frac{\sum_{\tau=1}^{N} s_\tau}{N_{it}} \times 100$$

Dynamics of Twitter activity on vaccination



Mixed 2SLS

We adopt M2SLS for estimation with grouped data (Dhrymes and Lleras-Muney, 2005):

- endogenous regressor s_{it} (*novax stance*) at the individual level i
- dependent variables $\overline{V}_{mt}$ (*vaccination rates/hospitalization rates/hospitalization costs*) at the municipality level m

First stage - (individual/year level)

$$s_{it} = \alpha + \beta f\bar{f} s_{it}^{ind} + \mathbf{T}'_{mt}\zeta + \mathbf{C}'_{mt}\phi + \gamma_m + \rho_r \times t + \theta_t + \varepsilon_{it} \tag{1}$$

Second stage - (municipality/year level)

$$V_{mt} = \alpha + \lambda\bar{\hat{s}}_{mt} + \overline{\mathbf{T}}'_{mt}\xi + \mathbf{C}'_{mt}\phi + \gamma_m + \rho_r \times t + \theta_t + \eta_{mt} \tag{2}$$

Second stage - vaccine coverage

Table 1: Results of the OLS and the Second stage of the M2SLS - Vaccination rates

	(1) OLS V_{mt}	(2) M2SLS V_{mt}
Panel a: Hexavalent (94.06)		
s_{mt}	-0.001	-0.023
	[0.002]	[0.015]
N	7,239	7,239
Panel b: MMR (89.53)		
s_{mt}	-0.005	-0.043**
	[0.003]	[0.021]
N	7,238	7,238
Panel c: Menigococcal (81.32)		
s_{mt}	-0.002	-0.040
	[0.008]	[0.054]
N	7,061	7,061
Panel d: Pneumococcal (82.64)		
s_{mt}	-0.0001	-0.029
	[0.008]	[0.052]
N	7,066	7,066
Controls (Twitter)	✓	✓
Controls (socioeconomics)	✓	✓
City and year FE	✓	✓
Reg × year	✓	✓

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.
$Notes$: All estimates include city and year fixed effects as well as region specific time trends. Standard errors (in brackets) are clustered at the municipality level and have been corrected in the second stage. Estimates as well as averages of V_{mt} are weighted by the municipality population size.

First stage.

Second stage - hospitalizations and costs

Table 2: Results of the OLS and the Second stage of the M2SLS - Hospitalizations .

	(1) OLS V_{mt} non-target pop.	(2) M2SLS V_{mt} non-target pop.	(3) OLS V_{mt} non-target pop.(MMR)	(4) M2SLS V_{mt} non-target pop.(MMR)	(5) OLS V_{mt} Children age 1-10 (MMR)	(6) M2SLS V_{mt} Children age 1-10 (MMR)
Panel a: Hopitalizations						
s_{mt}	0.0211	0.213*	0.018**	0.234***	0.007	0.145**
	[0.0159]	[0.113]	[0.00841]	[0.0601]	[0.008]	[0.065]
Panel b: Healthcare costs						
s_{mt}	129.8*	731.1**	71.96**	722.1***	47.13*	366.9**
	[66.39]	[353.8]	[30.92]	[243.1]	[25.95]	[161.1]
N	3,331	3,331	3,331	3,331	3,331	3,331
Controls (Twitter)	✓	✓	✓	✓	✓	✓
Controls (socioec.)	✓	✓	✓	✓	✓	✓
City and year FE	✓	✓	✓	✓	✓	✓
Reg × year	✓	✓	✓	✓	✓	✓

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.
Notes: All estimates include city and year fixed effects as well as region specific time trends. Standard errors (in brackets) are clustered at the municipality level and have been corrected in the second stage. Estimates are weighted by the municipality population size.

A 10pp increase in the municipality level novax stance → 2 additional hospitalization every 100k residents, and 7311 euro additional expenditure, which is a + 11% increase. Second stage (Mandatory).

Robustness checks

Table 3: M2SLS Individual - First stage.

	(1) Main	(2) Twitter algorithm	(3) Emilia Romagna Law	(4) Populist party	(5) Network distance	(6) Excluding FoF geolocated in user's municipality	(7) 2013-2016
	s_{it} (30.33)	s_{it} (30.33)	s_{it} (30.33)	s_{it} (30.33)	s_{it} (30.33)	s_{it} (30.33)	s_{it} (30.33)
ffs_{it}	0.704*** [0.017]	0.528*** [0.035]	0.706*** [0.017]	0.691*** [0.022]	0.611*** [0.021]	0.731*** [0.016]	0.512*** [0.031]
$ffs_{it} \times$ TWalg		0.251*** [0.039]					
$ffs_{it} \times$ ER			0.005 [0.0742]				
$ffs_{it} \times$ PP				0.048 [0.043]			
N	127,754	127,754	127,754	127,754	127,754	127,746	48,180
Controls (Twitter)	✓	✓	✓	✓	✓	✓	✓
Controls (socioec.)	✓	✓	✓	✓	✓	✓	✓
City and year FE	✓	✓	✓	✓	✓	✓	✓
Reg × year	✓	✓	✓	✓	✓	✓	✓
F-stat	1,757.86	998.690	870.815	943.98	875.82	2102.95	266.18

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.
$Notes$: The numbers refer to an initial sample of 830,253 tweets to a population of 80,471 unique users across 4220 municipalities. All estimates include city, region and year fixed effects and region-specific time trends fixed effects. Standard errors (in brackets) are clustered at the municipality level. Averages of s_{it} in parentheses is weighted by population size.

second stage checks Hospitalizations   second stage vaccination rate

The role of online debates' topics

Table 4: User exposure to friends-of-friends stances and the role of online debates' topics.

	(1) s_{it} (30.31)	(2) Pro_{it} (0.495)	(2) $Anti_{it}$ (0.204)
$f\bar{f}s^{ind}_{it}$	0.2884***	-0.3309***	0.2295***
	[0.0693]	[0.0757]	[0.0728]
$f\bar{f}s^{ind}_{it} \times Efficacy$	-0.3425	0.3765	-0.3548
	[0.2724]	[0.2754]	[0.2961]
$f\bar{f}s^{ind}_{it} \times TrustfulSource$	-0.3136***	0.2656**	-0.3805***
	[0.0992]	[0.1127]	[0.1057]
$f\bar{f}s^{ind}_{it} \times PoliticsandMandate$	-0.1749***	0.0660	-0.3899***
	[0.0530]	[0.0408]	[0.0589]
$f\bar{f}s^{ind}_{it} \times VaccinesUnsafe$	-0.0697	0.1369	-0.0387
	[0.2292]	[0.2442]	[0.2495]
N	531352	531352	531352
User FE	✓	✓	✓
Daily date FE	✓	✓	✓

$^{*}\ p < 0.10$, $^{**}\ p < 0.05$, $^{***}\ p < 0.01$.

$Notes$: The numbers refer to an initial sample of 830,253 tweets to a population of 80,471 unique users across 4220 municipalities. All estimates include individual and daily date fixed effects. Standard errors (in brackets) are clustered at the individual. Mean values of s_{it}, $ffsPro_{it}$, and $ffsAnti_{it}$ in parentheses are weighted by population size.

Non-linear effects

Conclusions

- Novax propaganda in social media is **contagious** among users
- In the absence of vaccination mandate, local exposure to novax propaganda causes a reduction in vaccination rates
- Novax propaganda has economically relevant negative spillovers, where hospitalizations of patients non-targeted by the vaccines for vaccine-preventable disease are more frequent and impose extra costs on society.
- Controversial vaccination mandates (e.g. enforced on school enrollment) have the potential to **backfire**
- Policy makers should invest in raising awareness, especially using trustful sources in order to mitigate the impact of vaccine skeptic social media campaigns

Thank you!

matildegiaccherini@uniroma2.it

APPENDIX

Scraping data from `Twitter`

We run the query based on very general keywords related to vaccines - more specifically, we focus on all tweets in Italian which include the translation of “vaccine(s)”,“vaccination”, “vaccinating”, “novax”, “vax”, but for those (mainly ads) referring to mozzarella or cow milk (“latte vaccino” in Italian). The current version of the dataset was downloaded on April 23^{rd}, 2021.

```
query = "(vaccino OR vaccini OR vaccinazione OR vaccinazioni
OR vaccinarsi OR vaccinato OR vaccinata OR novax OR vax
-latte vaccino) lang:it"
start_date = "01-01-2013T00:00"
end_date = "01-01-2019T00:00"
```

Scraping.

Descriptive statistics of `Twitter` data

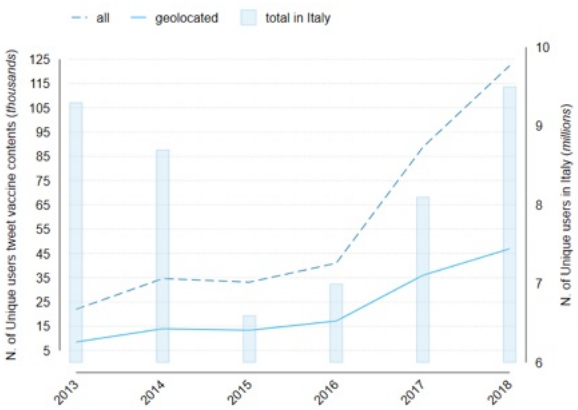
	median	mean	sd	min	max
(a) User characteristics					
Tweets about vaccine	1.00	6.24	32.82	1.00	3,720
Total *tweets*	5,586.00	19,793.54	50,699.13	1.00	1,825,203
Total *followers*	335.00	3,692.14	51,951.40	0.00	3,262,940
Total *friends*	462.00	970.31	2,759.93	0.00	189,582
Account's date of creation		2012	2.49	2006	2018
Verified accounts		0.007	0.084	0	1
(b) Tweets' characteristics					
Length of the tweet (number of characters)		102.42	42.05	0	306
Number of words		16.13	6.96	0	62
Retweets (%)		0.60	0.49	0	1
Replies (%)		0.10	0.30	0	1
(c) Tweets' popularity					
Retweet count		2.59	35.85	0.00	6696
Reply count		0.73	7.10	0.00	1106
Quote count		0.06	1.31	0.00	341
Like count		5.71	90.44	0.00	14188

$Notes$: *(a)*: summary statistics of 80,471 geotagged unique users tweeting on vaccines (2013-2018); *(b)*: summary statistics of 830,253 geotagged tweets cleaned by hashtag, "RT @", "@", url and emoji; *(c)*: Tweet-related popularity metrics of 328,879 original tweets.

Scraping data.

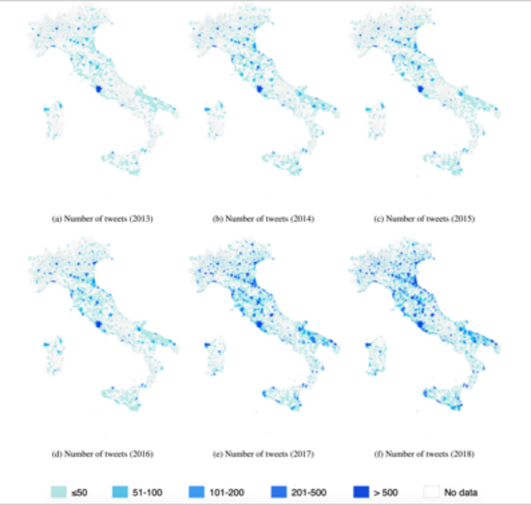
Number of unique users

Figure 2: Number of unique users



Scraping data.

Tweets mapping



Scraping data.

Descriptive statistics of vaccination rates (2013-2018

		Median	Mean	SD	Min	Max	N
Hexavalent	Diphteria*	94.97	94.29	3.15	54.69	100.00	44,750
	Hephatitis B*	94.80	94.15	3.19	54.69	100.00	44,750
	Polio*	95.00	94.31	3.14	54.69	100.00	44,750
	Tetanus*	95.00	94.38	3.13	54.69	100.00	44,777
	Pertussis**	94.94	94.29	3.14	54.69	100.00	44,750
	HIB**	94.64	94.04	3.17	54.69	100.00	44,749
Hexavalent		94.88	94.24	3.14	54.69	100.00	44,779
MMR	Measles**	91.05	89.52	5.97	10.72	100.00	44,750
	Rubella**	91.00	89.50	5.97	10.72	100.00	44,750
	Mumps**	91.00	89.48	5.96	10.72	100.00	44,750
MMR		91.02	89.50	5.97	10.72	100.00	44,752
Meningococcus		87.32	81.22	15.86	0.17	99.61	43,219
Pneumococcus		91.46	87.26	11.94	.17	100	43,167

Notes: exavalent and MMR vaccination rates across 7,929 Italian municipalities for the period 2013-2018. Average values are weighted by the municipality population size. * marks 2013-2017 set of compulsory vaccinations, ** indicates additional mandatory shots introduced by the 2017 Law Decree 73.

Health data.

Descriptive statistics of hospitalization due to vaccine-preventable diseases (2013-2016)

	Median	Mean	sd	Min	Max	N
Panel a: Hopitalizations						
non-target population	14.71	22.21	30.95	0.00	3,202.85	31,760
non-target population (MMR)	0.00	4.99	17.58	0.00	2,846.98	31,760
non-target population (Hexav.)	10.40	16.99	22.02	0.00	355.87	31,760
non-target population (Meningo.)	0.00	0.02	0.26	0.00	29.02	31,760
non-target population (Pneumo.)	0.00	0.88	2.25	0.00	155.04	31,760
Children age 1-10 (MMR)	0.00	2.96	6.87	0.00	1,617.25	31,760
Children age 1-10 (Hexav.)	0.00	1.27	2.70	0.00	152.44	31,760
Children age 1-10 (Meningo.)	0.00	0.04	0.41	0.00	26.21	31,760
Children age 1-10 (Pneumo.)	0.00	0.50	1.76	0.00	132.04	31,760
Panel b: Healthcare costs						
non-target population	38,581.69	66,477.60	116,320.65	0.00	59,880,842.11	31,760
non-target population (MMR)	0.00	15,381.55	96,931.58	0.00	59,880,842.11	31,760
non-target population (Hexav.)	46,275.59	83,151.57	119,925.38	0.00	14,819,697.72	31,760
non-target population (Meningo.)	0.00	150.92	3,976.38	0.00	411,341.22	31,760
non-target population (Pneumo.)	0.00	2,332.30	9,004.03	0.00	1,941,927.83	31,760
Children age 1-10 (MMR)	0.00	4,749.99	25,506.58	0.00	2,274,286.39	31,760
Children age 1-10 (Hexav.)	0.00	2,545.85	9,407.74	0.00	759,286.31	31,760
Children age 1-10 (Meningo.)	0.00	190.58	3,185.72	0.00	409,748.10	31,760
Children age 1-10 (Pneumo.)	0.00	1,255.36	5,365.51	0.00	259,504.65	31,760

Notes: The statistics refer to 7,940 municipalities for the time period between 2013-2016 and are weighted by the municipality population size.

Health data.

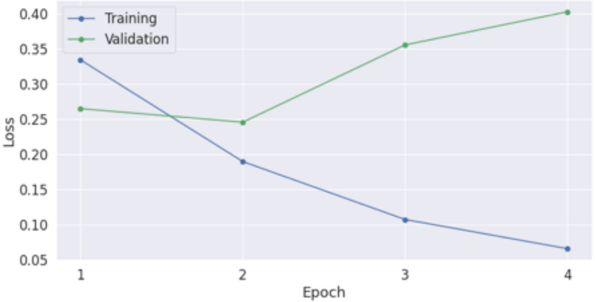
Training phase

Table 5: vaxBERTo last layer training

epoch	Training Loss	Valid. Loss	Valid. Accur.	Training Time	Validation Time
1	0.3342	0.2650	0.8885	0:05:50	0:00:13
2	0.1897	0.2456	0.9072	0:05:47	0:00:13
3	0.1074	0.3554	0.9023	0:05:47	0:00:13
4	0.0660	0.4025	0.9055	0:05:46	0:00:13

Notes: training and validation losses (columns 2 and 3), accuracy (4) and computing time (5 and 6) for each vaxBERTo training epoch.

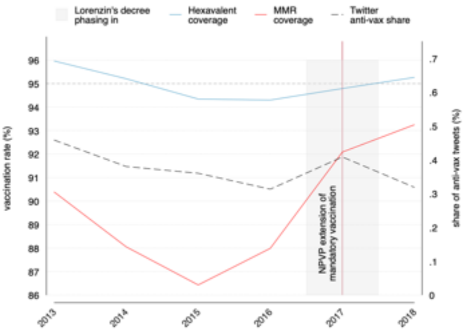
Figure 3: Training and validation loss



How it works

Twitter stances and vaccination rate

Figure 4: (b) Vaccination rate and share of tweets anti-vax geolocated



- Progressive decline in coverage until 2015, when the Rimini Court sentence was reversed by the Bologna Appeal Court
- Coverage rates (MMR in particular) started to rise from 2016, when extension and enforcement of mandatory vaccines was debated, and reinforced by National Law 117,2017.

Conceptual Framework

$$\dot{s}_i = -s_i + \mathbb{I} \sum_{j=1}^{N} W_{ij}(t) tanh(\alpha s_j) \quad (3)$$

- $\mathbb{I}$ measures the strength of the interaction among users of the platform
- $W(t)$ is a time-varying spatial contiguity matrix, whose i^{th}, j^{th} elements represent every link between individuals in the network
- $tanh(\cdot)$ is the hyperbolic tangent function, which provides a sigmoidal influence function of peers on individuals' stances.
- α captures the degree of *controversialness* of the topic

The contiguity matrix $W(t)$ evolves according to an activity-driven (AD) temporal network (Perra et al 2012), where each agent is characterized by the propensity to interact with a share $\omega_i \in [\epsilon, 1]$ of other agents, and the probability of an interaction is driven by homophily (Bessi et al 2016) $\longrightarrow$individuals are more likely to interact with like-minded peers, and we model it as a decreasing function of the (absolute) distance between i and j's opinions, $p_{ij}(t) = \frac{|s_i(t) - s_j|^{-\beta}}{\sum_j |x_i - x_j|^{-\beta}}$.

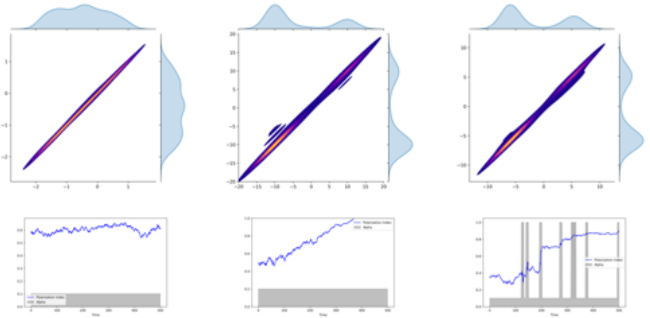
Conceptual framerwork

Simulated distribution of stances

Simulations →Micro-interactions of users on **controversial** topics give rise to transitions from a relative consensus to polarization

Figure 5: Simulated distribution of stances

(a) Low controversialness (b) High controversialness (c) Time-varying controversialness



Notes: user (x-axis) and average friends' (y-axis) distribution of stances in a simulated model with low - (a), $\alpha = .1$ - high - (b), $\alpha = .2$ - and low with exogenous, short-term outbursts controversialness - (c). In all models, the number of individuals is $N = 500$ and the periods are $T = 5$ - divided in 100 subperiods. We also set $\beta = 2$, $K = 3$ and $a = .2$. Initial values (s_0) are randomly drawn from a gaussian distribution with $\mu = -0.2$ and $\sigma = 0.5$ to match the asymmetry of the initial opinions in the data.

Conceptual framework.

Descriptive statistics of ego network

	Median	Mean	sd	Min	Max
Friends	469	973.46	2,717.55	1.00	189,433
Friends of friends (ff)	7,687	12,556.24	14,078.73	1.00	139,508
Total ff tweets with vaccine contents	59,535.50	142,261.09	186,460.83	1.00	1,685,355

Notes: The statistics refer to 80,471 geotagged unique users tweeting on vaccines (2013-2018) for 132,190 observations.

Friend-of-friends network

First stage

Table 6: M2SLS Individual - First stage.

	(1)	(2)	(3)	(4)	(5)	(6)
	s_{it}	s_{it}	s_{it}	s_{it}	s_{it}	s_{it}
	(30.31)	(30.31)	(30.31)	(30.31)	(30.31)	(30.31)
ffs_{it} (28.77)	0.799***	0.751***	0.703***	0.703***	0.704***	0.704***
	[0.021]	[0.021]	[0.017]	[0.017]	[0.017]	[0.017]
N	127,754	127,754	127,754	127,754	127,754	127,754
CONTROL (Twitter)				✓		✓
CONTROL (socioeconomics)					✓	✓
YEAR FE	✓	✓	✓	✓	✓	✓
CITY FE		✓	✓	✓	✓	✓
Reg Year			✓	✓	✓	✓
F-stat	1,501.16	1,288.96	1,765.22	1,763.52	1,755.84	1,757.86

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

$Notes$: The numbers refer to the sample of 830,253 tweets and to a population of 80,471 unique users across 4,220 municipalities. All estimates include municipal and year fixed effects as well as region specific time trends. Standard errors (in brackets) are clustered at the municipality level. The average values of s_{it} and ffs_{it} in parentheses are weighted by population size.

In a set of balance tests we rule out potential non-random assignment of our IV with respect to contextual features of municipalities where the users reside.

Second stage. Balance test.

Balance Test

Is our IV randomly assigned with respect to contextual features of municipalities where the users reside?

Variation in novax stances of friends of friends should be unrelated to predetermined characteristics of the municipalities after controlling for municipality and year fixed effects First stage.

Table 7: Balance test

	(1) Health public cost per capita	(2) Income per capita	(3) Lower secondary school att. (%)	(4) Avg. mother's age at birth	(5) Birth rate	(6) Populist party
Panel a: geolocated in the same user's municipality						
$f\bar{f}s^{ind}_{it}$	-0.0211	-0.403	0.0001	0.0001	-0.0002	0.0002
	[0.0246]	[0.442]	[0.0002]	[0.0001]	[0.0002]	[0.0002]
	110639	110639	110639	110589	110639	110639
Panel b: geolocated in municipalities different from the user's municipality						
$f\bar{f}s^{ind}_{it}$	-0.0001	-0.447	-0.0001	-0.0001	-0.00002	0.0001
	[0.0126]	[0.337]	[0.0004]	[0.0001]	[0.0001]	[0.0001]
	131003	131003	131003	130817	131003	131003
Panel c: not geolocated						
$f\bar{f}s^{ind}_{it}$	0.0037	1.001	-0.00004	-0.00001	0.0001	0.0002
	[0.0121]	[0.912]	[0.0002]	[0.00003]	[0.0001]	[0.0002]
	130977	130977	130977	130791	130977	130977
CITY and YEAR FE	✓	✓	✓	✓	✓	✓

$^{*}\ p < 0.10,\ ^{**}\ p < 0.05,\ ^{***}\ p < 0.01.$

$Notes$: Figures in parentheses are standard errors robust to clustering at the municipality level.

Second stage - vaccine coverage (MANDATORY)

Table 8: Results of the OLS and the Second stage of the Mixed 2SLS - Hospitalizations.

	(1)	(2)	(3)	(4)	(5)	(6)
	OLS	Mixed 2SLS	OLS	Mixed 2SLS	OLS	Mixed 2SLS
	V_{mt} (Hexav.)	V_{mt} (Hexav.)	V_{mt} (Meningo.)	V_{mt} (Meningo.)	V_{mt} (Pneumo.)	V_{mt} (Pneumo.)
			Non-target population			
Panel a: Hopitalizations						
s_{mt}	0.009	0.025	-0.0001	-0.0003	-0.0006	-0.021
	[0.012]	[0.092]	[0.0002]	[0.0009]	[0.002]	[0.015]
Panel b: Healthcare costs						
s_{mt}	102.0	-628.4	-4.756	-20.81	-10.53*	-46.519
	[100.6]	[700.3]	[3.976]	[16.46]	[6.103]	[37.26]
			Children age 1-10			
Panel a: Hopitalizations						
s_{mt}	-0.00007	0.002	0.00005	0.0003	-0.002	0.009
	[0.003]	[0.016]	[0.0006]	[0.004]	[0.002]	[0.011]
Panel b: Healthcare costs						
s_{mt}	12.74	-66.18	-0.528	10.36	-3.788	-37.99
	[18.45]	[49.21]	[2.887]	[14.90]	[6.229]	[42.28]
N	3331	3331	3331	3331	3331	3331
CONTROL (Twitter)	✓	✓	✓	✓	✓	✓
CONTROL (socioeconomics)	✓	✓	✓	✓	✓	✓
CITY and YEAR FE	✓	✓	✓	✓	✓	✓
Reg Year	✓	✓	✓	✓	✓	✓

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.
$Notes$: All estimates include city and year fixed effects and region-specific time trends. Standard errors (in brackets) are clustered at the municipality level. Estimates are weighted by the municipality population size.

Second stage.

Table 9: M2SLS Municipal - Second stage (Vaccination rate, hospitalizations and healthcare costs).

	(1) Main	(2) Twitter algorithm	(3) Emilia Romagna Law	(4) Populist Party Law	(5) Network distance	(6) Excluding FoF geolocated in user's municipality	(7) 2013-2016
Panel a: MMR vaccination rate (89.53)							
s_{mt}	-0.043**	-0.047**	-0.048**	-0.055**	-0.050**	-0.042**	-0.087**
	[0.021]	[0.022]	[0.021]	[0.026]	[0.023]	[0.021]	[0.044]
N	7,238	7,238	7,238	7,238	7,238	7,238	3,137
Controls (Twitter)	✓	✓	✓	✓	✓	✓	✓
Controls (socioec.)	✓	✓	✓	✓	✓	✓	✓
City and year FE	✓	✓	✓	✓	✓	✓	✓
Reg × year	✓	✓	✓	✓	✓	✓	✓

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.
$Notes$: All estimates include city and year fixed effects and region-specific time trends. Standard errors (in brackets) are clustered at the municipality level and have been corrected in the second stage. Estimates, as well as averages of V_{mt}, are weighted by the municipality population size.

Robustness checks - Second stage - hospitalizations

Robustness checks

Table 10: M2SLS Municipal - Second stage (Vaccination rate, hospitalizations and healthcare costs).

	(1) Main	(2) Twitter algorithm	(3) Emilia Romagna Law	(4) Populist Party Law	(5) Network distance	(6) Excluding FoF geolocated in user's municipality	(7) 2013-2016
Panel b: Non-target population							
Hopitalizations							
s_{mt}	0.213*	0.231*	0.204*	0.215*	0.220*	0.205*	0.319*
	[0.113]	[0.121]	[0.112]	[0.112]	[0.115]	[0.108]	[0.167]
Healthcare costs							
s_{mt}	731.1**	821.3**	712.8**	746.5*	794.0**	909.9**	-162.2
	[409.8]	[434.7]	[406.6]	[412.2]	[411.0]	[402.0]	[952.1]
Panel c: Non-target population (MMR)							
Hopitalizations							
s_{mt}	0.234***	0.256***	0.233***	0.231***	0.242***	0.211***	0.320**
	[0.0601]	[0.0675]	[0.0596]	[0.0603]	[0.0621]	[0.0578]	[0.128]
Healthcare costs							
s_{mt}	722.1***	716.7***	725.1***	734.0***	743.7***	713.8***	422.6*
	[243.1]	[250.6]	[242.8]	[247.7]	[247.1]	[235.6]	[214.3]
Panel d: Children age 1-10 (MMR)							
Hopitalizations							
s_{mt}	0.145**	0.150**	0.145**	0.146**	0.142**	0.115*	0.184*
	[0.0650]	[0.0664]	[0.0651]	[0.0653]	[0.0659]	[0.0619]	[0.096]
Healthcare costs							
s_{mt}	366.9**	428.7**	366.5**	363.6**	390.2**	375.5**	233.8*
	[161.1]	[171.8]	[160.9]	[163.9]	[163.7]	[162.3]	[117.1]
N	3,331	3,331	3,331	3,331	3,331	3,331	3,331
Controls (Twitter)	✓	✓	✓	✓	✓	✓	✓
Controls (socioec.)	✓	✓	✓	✓	✓	✓	✓
City and year FE	✓	✓	✓	✓	✓	✓	✓
Reg $\times$ year	✓	✓	✓	✓	✓	✓	✓

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. $Notes$: All estimates include city and year fixed effects and region-specific time trends. Standa errors (in brackets) are clustered at the municipality level and have been corrected in the second stage. Estimates, as well as averages of V_{mt}, a weighted by the municipality population size.

Non-linear effects and policy implications

Table 11: Mixed 2SLS for pro-vax vs. anti-vax users - First stage.

	(1)	(2)
	Pro_{it}	$Anti_{it}$
	(0.495)	(0.204)
$f\bar{f}s_{it}^{ind}$ (28.77)	-0.0076 ***	0 .0046***
	[0 .0003]	[0.0001]
N	127754	127754
CONTROLs	✓	✓
CITY and YEAR FE	✓	✓
Reg Year	✓	✓
F-stat	1765.22	1763.52

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table 12: Mixed 2SLS for pro-vax vs. anti-vax users - Second stage Vaccination rates

	(1)	(2)
	Pro_{mt}	$Anti_{mt}$
	V_{mt}	V_{mt}
Panel b: MMR (89.53)		
	3.9086*	-6.6162*
	[2.1978]	[3.5315]
	7238	7238
CONTROLs	✓	✓
CITY and YEAR FE	✓	✓
Reg Year	✓	✓

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.