How's Life in the Village? Economic Resilience and Reaction During Pandemic Lockdowns

Ahmed Alsayed, Tiziana Balbi*, Giuseppe Gerardi*, Giancarlo Manzi*, Martina Viggiano**

Abstract

Although pandemics have been a recurring problem in history, the COVID-19 pandemic has some characteristics never experienced before. The human being has survived wars, nature catastrophes and economic shocks, always showing resilience in adapting to new situations. In this paper we want to check if this attitude is still strong. Our research question is: How do we react to emerging situations? In this paper we try to answer this question analyzing a dataset of Twitter messages collected through the second and third COVID-19 pandemic waves in Italy regarding everyday life during strict lockdowns and people's opinion on these situations. The small villages are our starting point, questioning first about the reactions in the population during severe restrictions and secondly looking at the responses to social and economic changes without any reference to the lockdown period; we focus then to the resilience behaviors considering the areas in the North, Center and South Italy thanks to the Twitter messages' geo-localization.

1. Introduction[1]

The impact of the COVID-19 pandemic over people's everyday life has been (and still is) devastating over many dimensions. It spread all over the globe on such a scale that, as of February 2022, it took almost 6 million lives from its beginning in late December 2019[2] (the so-called "Russian flu" in 1889-1990 and "Spanish

* University of Milan, Department of Economics, Management and Quantitative Methods, Milan, Italy, e-mail: ahmed.alsayed@unimi.it, tiziana.balbi@unimi.it, giu.gerardi@gmail.com, giancarlo.manzi@unimi.it (corresponding author), martina.viggiano2@studenti.unimi.it.

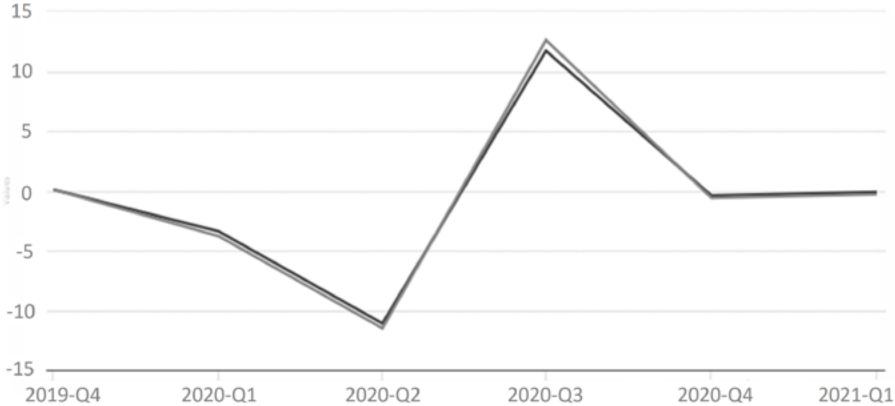
1. We would like to thank Guido Ravasi for his comments and suggestions on a first draft of this paper and for his precious help in developing the code for this work, and Flavio Verrecchia, National Italian Statistical Institute, for inviting us to present this work in a special session of the XLII Conference of the Italian Regional Science Association. We also would like to thank Emanuele Morales for his counselling on developing the Python code.

2. https://coronavirus.jhu.edu/map.html accessed on February 6th, 2022.

flu" in 1918-19 were also global pandemics but with a lower level of spread due to more limited possibilities for travel and movement at the time, even if the Spanish flu is considered the worst pandemic of history with up to 100 million deaths worldwide caused mainly by lack of possibilities for sufficient health countermeasures – see Aassve *et al*., 2021). Curbs to counter the spread of the virus have been adopted almost in all parts of the world: school closures, transport limitations, gathering restrictions are only a few examples of the many actions undertaken by the various national and regional authorities. Even in countries where at the beginning of the pandemic was decided to aim for herd immunity, such as Sweden, some containment measures were eventually taken. These measures have impacted on countries' economies on a large scale (Pak *et al*., 2020). For example, in the European Union from the fourth quarter 2019 to the second quarter 2021 there have been two strong contractions in the average EU GDP (Figure 1), with the Euro area countries slightly more affected. All world markets have been affected, more than in the financial crisis of 2008 and in the great depression of 1929 (Bagchi *et al*., 2020). All in all, the COVID-19 outbreak negatively affected the economy even in countries where no lockdowns have been introduced as in South Korea (Aum *et al*., 2021).

However, it is at the local level and for small businesses that the COVID-19 pandemic has had the most disastrous effects, as several studies have examined the impact of COVID-19 or lockdown on economic and social aspects at local level, counties or small towns.

Figure 1 – Quarterly GDP growth, European Union (27 countries, black line) and Euro area (19 countries, gray line) – Q4 2019/Q1 2021



Source: Eurotat

The literature on the economic impact of the COVID-19 pandemic has of course flourished in the last one year and a half. Ur Rahman, *et al*. (2021) investigated the economic impacts of COVID-19 on households based on differences in the socio-economic status (SES). Household-level effects were determined by using income sources, types of industries, communities' resilience, household susceptibility, and relevant policy measures. Data were collected using an online survey questionnaire from different villages located in Sichuan and other provinces in China (475 in Sichuan, 80 from other provinces). The statistical analysis was performed by applying stepwise binary logistic regression analysis. Findings suggested the significant use of SES to detect the impact of COVID-19 on different households. Households with low SES tend to depend more on farmland income and transfer payments from the government, while high SES households focus more on business and local employment as sources of income generation. Due to that, poor households or communities are less resilient and more likely tend to suffer poverty due to the COVID-19 crisis, whereas the reverse situation happens for households with high SES.

Peters (2020) created a COVID-19 susceptibility scale at county level for the United States of America, considering 3,079 counties in the 48 conterminous states. Exploratory factor analysis (EFA) was used to construct this COVID-19 susceptibility scale. Also, the author assessed the health and socioeconomic resiliency of susceptible places across the rural-urban continuum, by applying multivariate general linear model (MANOVA) to estimate unconditional mean differences across several resiliency. Finding shows that 33% of rural counties are highly susceptible to COVID-19, driven by older and health-compromised populations, and care facilities for the elderly. Rural counties were more sensitive to COVID-19 critical situations as they lack social services which might to hinder local pandemic recovery.

Chirisa *et al*. (2020) examined the challenges experienced by poor urban communities during the lockdown in sub-Saharan Africa. This study is strongly focused on social and economic impacts of lockdowns on the poor and disadvantaged communities. Data was collected from secondary sources, i.e. from existing databases and published scientific works including Google Scholar, Bok.org and EBSCOhost. They used qualitative analysis tools such content analysis. The results showed that the COVID-19 scourge had a huge impact on the increase of urban poverty in sub-Saharan Africa. Many reach households continue earning an income similar to the pre-pandemic one working from home. Moreover, they save money by reducing commuting expenses for moving to workplaces, while households working in small business result more exposed to the risk of losing their jobs.

Karaye *et al*. (2020) examined the association between characteristics of infected COVID-19 people and social vulnerability in the U.S at global and local level using an ordinary least squares regression model at global level, and a geographically weighted model at county level. Independent variables were the social vulnerability index (SVI), the household composition and disability, minority status and language, housing and transportation. All independent variables were significant to predict new COVID-19 cases, and the SVI and minority status and language were associated with an increased number of new COVID-19 cases.

Our approach diverts from the literature above, as it is aimed at exploring resilience and reaction through social network text analysis. In particular, we analyze Twitter messages (in Italian) collected from October 28th, 2020, to March 19, 2021, i.e., the period between the second COVID-19 wave and the start of the first vaccine campaign in Italy. We proceed with four levels of analysis: (i) analysis of COVID-19 tweets directly related to the small villages put on strict local lockdown in the period considered (we refer to this analysis using the acronym ANA1 in the following); (ii) analysis of COVID-19-tweets with general comments on the resilience and reaction to restriction measures affecting the economy and social status of small municipalities (ANA2); (iii) a more general analysis of COVID-19 tweets concerning resilience and reaction in general in Italy with no particular reference to lockdown situations (ANA3). We perform some social network analysis, term frequency analysis, sentiment analysis and topic model analysis on these three levels. Finally, (iv) we conducted a non-automatic analysis on a tiny subset of tweets for which we were able to collect geographical information to see if there are difference between macro regions and between rural and urban areas (ANA4). The aim of our analysis is to check for economic and social behavior and feeling when facing shocks experienced like in this pandemic.

This article is organized as follows. Section 2 describes the data collection process. Section 3 briefly presents the natural language process tools we used in the analysis. Section 4 presents the results of the analysis and Section 5 concludes the paper

2. Data Collection

2.1. Small villages in "Red Zones"

In this study, we collect tweet data related to COVID-19 and the economic situation and health behavior during the first period when a system of 4-color alert scale has been in place in Italy. This system was introduced on November 6th, 2020. The upper level (called in Italian "zona rossa" or "red zone") of this scale is the strictest one and can be considered as a full lockdown. We selected Italy for our study as it is one of the countries firstly and more severely hit by the pandemic, especially the small towns and villages.

Normally, the areas put in a "red zone" corresponded to administrative regions (EU NUTS-02), but in some cases small areas were declared "red zone". We decided to consider areas smaller than NUTS-02 regions with a population under 100,000 inhabitants. In this way we ended up with 83 small villages experiencing at least one "red zones" period between October 28th, 2020, and March 19th, 2021. We obtained the list of these small municipalities by scraping the websites of municipal and regional authorities and local newspapers and media. We refer to this list for the ANA1 analysis.

Figure 2 shows the distribution of small "red zones", i.e., small municipalities put under the highest restriction level. These red areas have been mostly concentrated in the two most southern Italian regions, i.e. Calabria and Sicily. The power to put local municipalities in this "red zone" status was in the hands of regional authorities which, in doing this, had to consider economic and social implications for the small communities of these villages. Probably the red zones were established

Figure 2 – Spatial distribution of small "red zones" in Italy – From October 28th, 2020, to March 19th, 2021

Source: authors' computations

easier in Southern Italy because the impact of these restrictions was considered less invasive than the average economic condition there. In the rich North, a lockdown action was considered more problematic because it more probably could disrupt the production process there. Moreover, as in Italy there exist a decentralized region-based health system and the regions of southern Italy have a less efficient and more critical health system, these extreme measures have been taken more easily than in other regions in order not to overwhelm the regional health system.

2.2. Tweet collection

To collect the tweets concerning COVID-19 we used the standard API that Twitter makes available to users provided that they submit a project in which these tweets are used. There are several problems with this API because it does not allow to exceed the limit of 500 thousand tweets downloaded for each project and does not give you the possibility to specify the date. However, one can indicate the date until he or she wants to go back in the retrieving the tweets, with a maximum limit of 7 days before the Twitter queries are submitted[3].

We started collecting Twitter data from the very beginning of the pandemic outbreak in Italy daily, in so doing avoiding the issue of the time limit. Once the APIs have been queried tweets were returned in JSON format, the results were first loaded into a non-relational DB (MongoDB) and then into a relational DB (PostgreSQL), taking care not to duplicate the records, which were uniquely recognized by the field 'tweet_id', and filtering only the fields that were strictly necessary for our research, in order to avoid storing limit problems. This methodology, together with the elimination from the non-relational DB of the tweets that are transferred to the relational DB, and the use of a server deployed in a firewall protected network, allowed us to contain the possibility that personal data attributable to the tweet's author could be breach. API's queries took place daily thanks to the use of a scheduler that launched the scripts automatically. The query strings used to collect these tweets were "COVID-19" and "Coronavirus" only, and we focused on tweets written in Italian. Tweets were collected from March 1st, 2020, to March 19th, 2021, but we focus on the period October 28th, 2020, to March 19th, 2021, corresponding to the first months in which the "color alert system" was implemented in Italy.

2.3. Keywords for tweet selection

From the huge twitter database created as described in the previous section, subsamples of tweets for ANA1, ANA2 and ANA3 analyses were created in the following way. For ANA1 we simply used the names of the villages in the queries and merged the resulting tweets. So, for example, if the village was "San

3. https://developer.twitter.com/en/docs/twitter-api/v1/tweets/search/api-reference/get-search-tweets.

Miniato”, the resulting query were “(‘COVID-19’ AND ‘San Miniato’) OR “(‘Coronavirus’ AND ‘San Miniato’)”. This was done for each of the 83 selected villages. The resulting dataset was formed by 1,116 records.

As for the ANA2 analysis, the “COVID-19” and “Coronavirus” keywords were added in an AND clause to keywords regarding more general comments about small villages or small cities and resilience/reaction/recovery. Table 1 contains all the query strings used in this analysis together with the resulting number of tweets retrieved. In the query we also included terms for indirectly detect comments on small environments. For example, “family shops” is more probable to be a key business activity in small villages rather than in big cities. Other possible queries were dismissed as they did not retrieve any tweets. Table 2 shows the query strings used in ANA3 analysis together with the resulting number of tweets retrieved. As these queries were more general on reaction/resilience/recovery to/from the COVID-19 pandemic, the number of retrieved tweets is by far larger.

3. Method

3.1. Pre-processing treatment

We use standard pre-processing techniques to clean, filter, stemming and tokenize the Twitter messages. In particular:

- We removed non-Italian words and retained foreign words which are currently in the standard Italian vocabulary (“lockdown” is an example). For this, we exploited the current Italian corpora embedded in some Python libraries.
- We performed text cleaning by removing punctuation, double spaces, hyperlinks, numbers, special symbols, etc. For this, we used the standard regular expression treatment in Python using the re library.
- We performed lemmatization (Prabhakaran, 2018), converting each word in its basic root (“ross” is an example of a lemmatization of “rosse”, by removing the last letter “-e” stating in this case the gender and number agreement of words in Italian; in English this corresponds to getting “red” from “reds”). For this, we used the WordNetLemmatizer lemmatizer from the nltk Python library.
- We removed so-called stop words, i.e. articles, prepositions, and other function words which are not essential in this kind of analysis (Malik, 2020). For this, we used the list of the stop words in the Python nltk corpus.

3.2. Sentiment analysis

We used the Python TextBlob library used in multiple contexts (Loria, 2014; Schumacher, 2015; Hasan *et al.*, 2018; Morales, 2021). TextBlob is a a Python

Table 1 – Query strings used in the ANA2 analysis

Query (in Italian)	*English translation*	*No. of tweets retrieved*
"(piccoli comuni) AND (rossa)"	"(small municipalities) AND (red)"	36
"(piccoli comuni) AND (lockdown)"	"(small municipalities) AND (lockdown)"	17
"(resilienza) AND (rossa)"	"(resilience) AND (red)"	10
"(resilienza) AND (lockdown)"	"(resilience) AND (lockdown)"	49
"(reazione) AND (rossa)"	"(reaction) AND (red)"	25
"(reazione) AND (lockdown)"	"(reaction) AND (lockdown)"	170
"(ripresa) AND (rossa)"	"(recovery) AND (red)"	33
"(ripresa) AND (lockdown)"	"(recovery) AND (lockdown)"	221
"(ristoranti) AND (rossa)"	"(restaurants) AND (red)"	516
"(ristoranti) AND (lockdown)"	"(restaurants) AND (lockdown)"	1,176
"(piccoli negozi) AND (rossa)"	"(family shops) AND (red)"	6
"(piccoli negozi) AND (lockdown)"	"(family shops) AND (lockdown)"	8
"(piccoli comuni) AND (economia)"	"(small municipalities) AND (economy)"	3
"(economia locale)"	"(local economy)"	26
Total		2,196

Source: authors' computation

Table 2 – Query strings used in the ANA3 analysis

Query (in Italian)	*English translation*	*No. of tweets retrieved*
"disoccupat*"	"unemployed"	430
disoccupazione"	unemployment"	466
"lavoro"	"job/labor"	16,088
"finanza"	"finance"	1,178
"reazione"	"reaction"	2,027
"ristori"	"compensation"	4,330
"piccole imprese"	"small enterprises"	405
"economia"	"economy"	7,077
"ripresa"	"recovery"	2,282
"resilienza"	"(restaurants) AND (lockdown)"	487
Total		34,770

Source: authors' computation

library for Natural Language Processing that is built on top of NLTK to achieve the "polarization" goal. It takes a corpus as input and takes into consideration the order of the words, being important in this context. The case of the words "great" and "not great" is an example where for "good" the polarity of the word is positive, but in the second case, with the negation, it becomes negative. The polarity will be given a value between -1 and 1 with negative values signifying negative polarity and positive values signifying positive polarity. According to the TextBlob help, "great" will receive a polarity of 0.8, whereas "not great" will receive a polarity of -0.4. Scores are not symmetric around zero to consider irony and other linguistic figures like the litotes (the use of "she is not so beautiful" to say in a politer way "she is ugly").

3.3. Topic modelling

We performed a topic model analysis on the three corpora resulting from ANA1, ANA2 and ANA3 tweet selection. We used the Latent Dirichelet Allocation (LDA) method (Blei *et al*., 2003) which is one of the most popular topic modelling methods. The LDA topic generative process works by assigning a score to a given topic within each document in a corpus, building on the concept that each document can be described by a distribution of topics and each topic can be described by a distribution of words. The number of topics to be chosen is given by the so-called coherence score which assign a level of coherence of words used in each topic. For example, a topic music has more coherence if it has "sound", "scores", "guitar", etc. rather than "sound", "noise", "vibration", etc. For implementing LDA we used the *gensim* package in Python 3.8.

4. Results

4.1. ANA1 analysis

In the ANA1 analysis we extracted the relative frequency of words from tweets related to resilience for each village which had classified as red-zone area during the lockdown period by the decision-maker.

Figure 3 displays the word clouds of the corpus formed by the negative and positive unigrams in the ANA1 dataset. In the negative cloud among verbs those expressing worry ("to worry", ["preoccupa"]), are the most frequent. Among nouns, those revealing the stress of the health system ("nurse", ["infermiere"]), economic crisis ("crisi" ["crisis"]) and those used in pandemic texts ("English variant", ["variante inglese"]) are among the most frequent. In the positive cloud among verbs those expressing resilience ("to face", ["contrastare"]), hope ("to hope", ["speriamo"]) and unity ("unity", ["unione"]) are the most frequent.

Figure 3 – Word clouds from the ANA1 corpus

*(a) Negative tweets** *(b) Positive tweets*

Note: * "Crocera", "stadium" and "crociera" in the world clouds refer to the social and sporting garrison located in Sampierdarena (GE), called "Crocera Stadium di Sampierdarena": see, for example https://www.genovatoday.it/attualita/coronavirus/piscina-crocera-stadium-sampierdarena-chiude.html (accessed March 21st, 2022). Being the world clouds constructed with unigrams (and containing sometimes also spelling errors as for "crociera" instead of "crocera"), only part of the garrison name is reported in them.

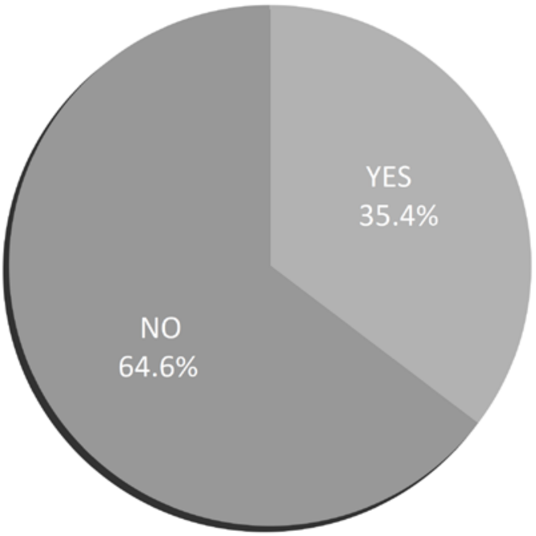
Source: authors' computation

Among nouns, those revealing the health "solution" for the pandemic ("vaccine", ["vaccino"]), responsibility ("responsibility" ["responsibilità"]) are among the most frequent. Adverbs revealing the need immediate action ("now!", ["subito!"]) are also a sign of positive attitude.

Figure 4 shows results from a sentiment analysis on the ANA1 corpus, after excluding the neutral tweets. We performed this analysis completely manually as the number of tweets in this corpus was relatively low and to check for the "average" vocabulary used in the tweets of this type. About two third of the tweets showed a rather negative sentiment and an attitude toward a gloomy mood with regard to the evolution of the pandemic and about the possibility of any recovery/reaction.

After using coherence analysis for determining the number of the most important topics, we ended up choosing 2 topics which, according to the most frequent words in them, we labeled "Political imposition" (as words like "government", "pd" – i.e, the Italian Acronym of the Democratic party, one of the government coalition party at the time – "regione", etc., were more frequent) and "Health system", as words related to healthcare and used to describe the coping of the epidemic were mostly used.

Figure 4 – Polarity of non-neutral tweets in the ANA1 analysis ("no" = negative; "yes" = positive)



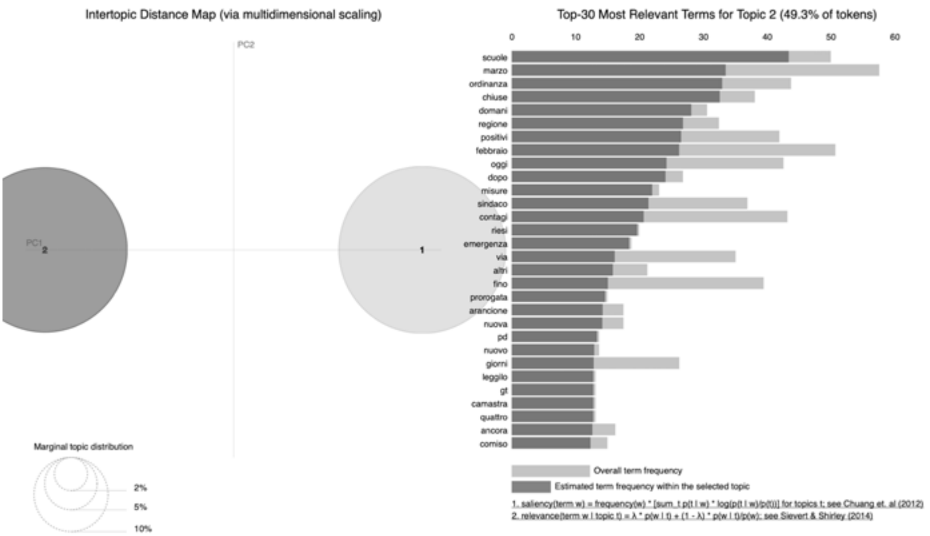
Source: authors' computations

In Figure 5 the inter-topic distance map of the two topics is showed together with the most frequent words in the topic "Health systems". The distribution of topics among the tweets was almost uniform resulting in 50% in the "Political imposition" topic and the same in the "Health system topic" (Figure 6).

4.2. ANA2 analysis

In the ANA2 analysis we extracted the relative frequency of words from tweets related to resilience and recovery in small areas without direct reference to the villages in red zone. Figure 7 displays the word clouds of the corpus formed by the negative and positive unigrams in the ANA2 dataset. Negative tweets were much more related to the lockdown condition and its time span ("due" is the Italian word for "two" which is the number of weeks or months regions or other territorial authorities have been locked down). The "two" word ("due" in Italian) assumes an important weight in this context, since, as a matter of fact, it is about fourteen days (in the case of *two* weeks) or sixty days (in the case of *two* months), a time in which people felt to be deprived of autonomy in everyday life.

Figure 5 – "Political imposition": the two most important topics for the ANA1 corpus. Areas of the circles are proportional to the word frequencies in the topics



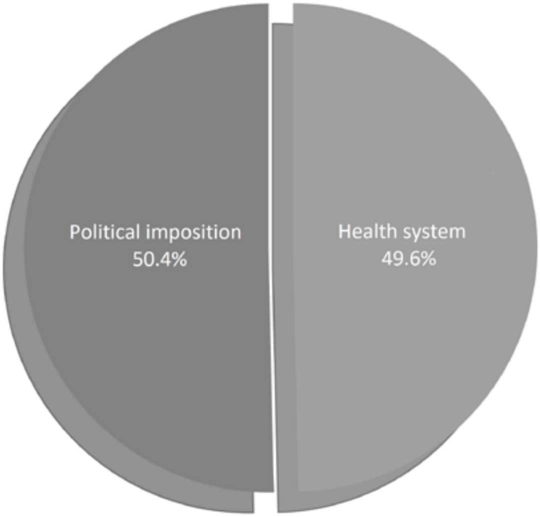
Source: authors' computations

"dpcm" stands for "Decree of the Presidency of the Council of Ministers" and is the apex institutional body in charge of the decisions for implementing restrictions. In positive weeks the word "chiusi" ["locked up"] is quite frequent like in the negative tweets but this time it is present in conjuction with the word "fine" ["end"], meaning a more optimistic attitude for the future. The world "ripresa" ["recovery"] means that this subgroup of tweets more prone toward seeing a light at the end of the tunnel.

Sentiment analysis showed again a more polarized attitude in the positive perspectives than in the ANA1 corpus, as about two third of the tweets showed a rather positive sentiment and an attitude toward a better mood with regard to the evolution of the pandemic and the possibility of any recovery/reaction (Figure 8). This analysis and the ANA3 analysis were performed automatically using the TextBlob technique.

Coherence analysis led us to choose two topics which, according to the most frequent words in them, we labeled "Political imposition" similarly to the ANA1 corpus and "Suggestion for recovery", as words related to healthcare and used to describe the coping of the epidemic were mostly used.

Figure 6 – Distribution of topics in the ANA1 corpus



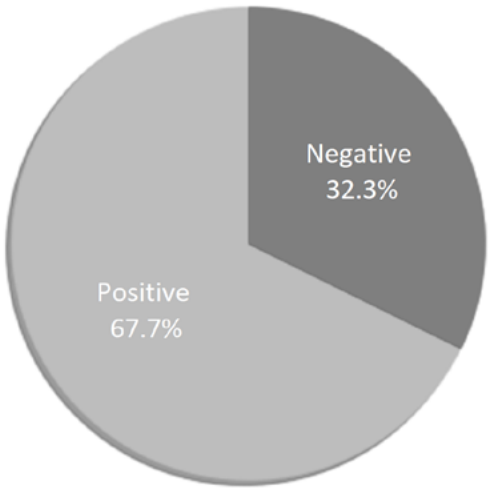
Source: authors' computations

Figure 7 – Word clouds from the ANA2 corpus – (a) Negative tweets; (b) Positive tweets

(a) *(b)*

Source: authors' computations

Figure 8 – Polarity of non-neutral tweets in the ANA2 analysis



Source: authors ‘calculations

In Figure 9 the inter-topic distance map of the two topics is showed together with the most frequent words in the topic “Suggestion for recovery”. This time the distribution of topics among the tweets was quite unbalanced with almost 90% in the “Political imposition” (Figure 10).

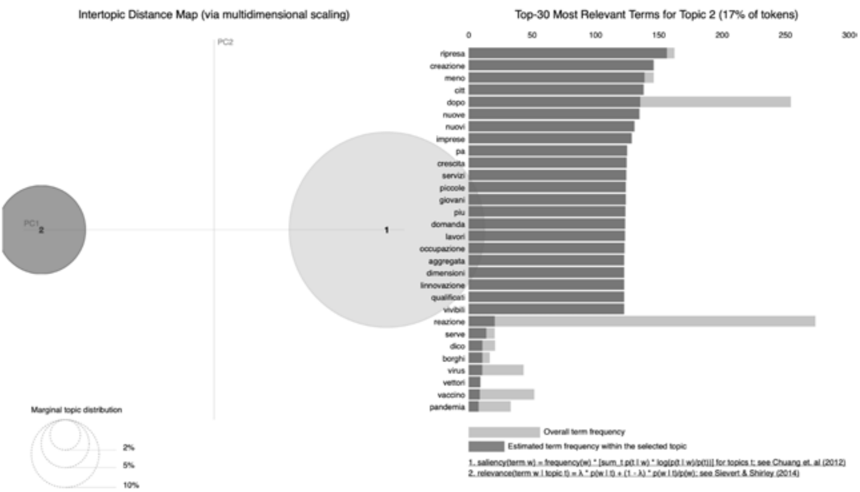
4.3. ANA3 analysis

In the ANA3 analysis we extracted the relative frequency of words from tweets related to resilience and recovery in general with reference neither to small areas nor to villages in red zone.

Figure 11 displays the word clouds of the corpus formed by the negative and positive unigrams in the ANA3 dataset. In this case one word was present both in negative and in positive tweets: “vaccino” [“vaccine”]. The vaccination campaign started in Italy in January 2021, more or less in the middle of the considered period. The presence of this word in the two groups reveals how this word was considered important by both pessimistic and optimistic people. We found some examples like “il vaccino ci salverà” [“the vaccine will save us all”] and “il vaccino non risolverà i nostri problemi” [“the vaccine won’t solve our problems”] in the positive and in the negative tweets, respectively, in any case revealing how the vaccine was a strongly debated topic all over the corpus.

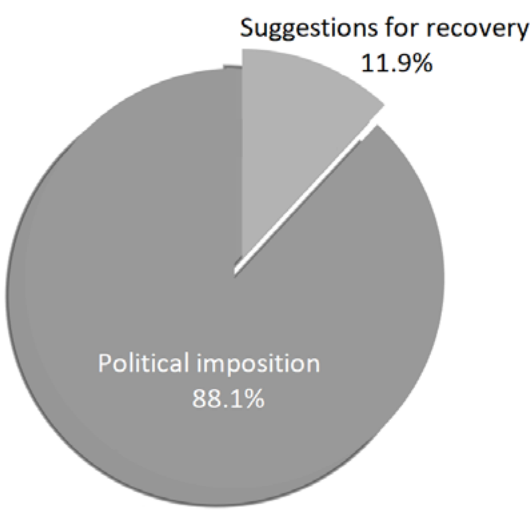
Sentiment analysis was again in favor of positive tweets as almost 60% of the tweets were positive (Figure 12).

Figure 9 – "Political imposition": the two most important topics for the ANA2 corpus. Areas of the circles are proportional to the word frequencies in the topics



Source: authors' computations

Figure 10 – Distribution of topics in the ANA2 corpus



Source: authors' computations

Figure 11 – Word clouds from the ANA3 corpus

(a) Negative tweets *(b) Positive tweets*

Source: authors' computations

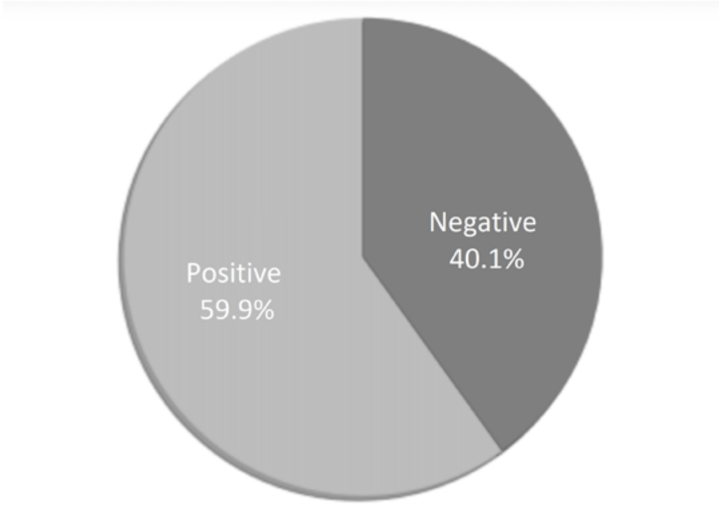
Coherence analysis this time led us to choose three topics which, according to the most frequent words in them, we labeled "Suggestions for recovery" as in the ANA2 corpus, "Take action", and "Uncertainty for the future".

In Figure 13 the inter-topic distance map of the two topics is showed together with the most frequent words in the topic "Take action". In Figure 14 the distributions of tweets in these three topics is showed. "Take action" is the minority topic, whereas the "Suggestions for recovery" is the most frequent.

4.3. ANA4 analysis

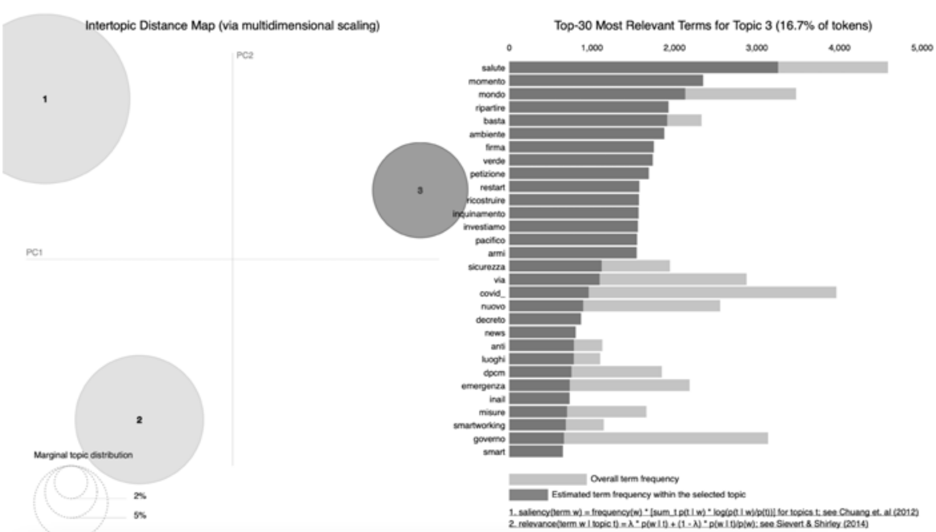
The identification of the true geographic location where tweets are launched is one of the challenging tasks for researchers aiming at retrieving location space information, as the tweet location information fields, namely "user location" and "place name", are not reliable for many reasons, among which the following three are the most important: 1) they are set when the account has been opened and most probably they do not correspond to the true locations where the tweets have been launched; 2) users can give fake addresses on purpose; 3) users can give any other information not related to location (Kumar, Singh, 2019; Liu *et al*., 2020). It is possible to process the text string contained in these fields with machine learning tools (for example, see Indira *et al*., 2019), but still results on inferring the true tweet location applying this methodology are controversial.

Figure 12 – Polarity of non-neutral tweets in the ANA3 analysis



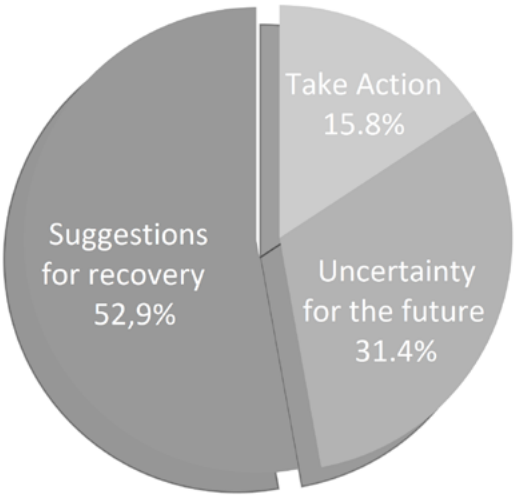
Source: authors' computations

Figure 13 – "Take action": the two most important topics for the ANA3 corpus. Areas of the circles are proportional to the word frequencies in the topics



Source: authors' computations

Figure 14 – Distribution of topics in the ANA3 corpus



Source: authors' computations

The tweet field really trustable is only the "geo" field because it is filled with the geographical coordinates taken from the device used for tweeting.

Unfortunately, in our case, among the total 38,072 collected tweets we were able to retrieve information about true geotagging on only 74 tweets (rate: 0.2%). This low rate is in line with the most recent outcomes from the literature on social network analysis. For example, in a cross-country analysis about the use of geo-services and geotagging on Twitter, Sloan and Morgan (2015) found that only around 3% of the tweeters enables Tweets to geotag their tweets. However, Sloan and Morgan's work dates back to 2015 and since then Twitter have deeply reduced the possibility to geotag tweets. In particular, in a 2019 tweet, Twitter announced that they were reducing the geotagging option (Figure 15). As a result, the rate of geotagging tweets must have reduced further as a consequence of this decision.

Out of these 74 tweets, 32 were from Northern Italy regions (i.e., regions north the Apennine mountains, Liguria included) and 42 from Center-Southern Italy regions (rest of regions). Moreover, 30 were from rural areas, and 44 from urban areas if we consider as urban areas those corresponding to provincial capitals.

We performed a "manual" sentiment analysis by reading the tweets' texts. We also manually extracted the most important topics and the most common way to show disappointment, negativity and mistrust in the way problems are faced on one side, and, on the other side, trust, positivity and confidence.

Figure 15 – Twitter support tweets on reducing geotagging

Table 3 summarizes our manual analysis for the two variables "Area type" and "Geography".

Polarity seems more positive in rural areas rather than in urban areas (53.6% positive tweets in rural areas vs. 41% in urban areas), whereas geography does not seem to be decisively discriminant for a positive attitude (46.7% positive tweets in the North against 45.9% in the Center-South).

As for the topic analysis, positive messages in rural tweets are more in the sense of an economic outlook based on resilience and proacting (especially with reference to some economic categories to rely on: "We want to be part of these numbers or better react by relying on professionals who do not make companies fail", referring to unemployment dramatic numbers at the time). Negative outlook in rural tweets come mainly from no-vax, no-green pass categories of tweeters, right-wing oriented ("The left seriously harms health, produces poverty. They say data on millions in poverty and absolute hunger… thousands of unemployed…companies… bankrupt… despair… divorces…. suicides… barbarians didn't do that", referring to the left-wing government at the time"). Tweets from urban areas were more neutral (probably because of some local news media reporting about COVID-19); positive tweet were focused in trusting the government ("only taxes would attract huge investments from abroad, would favor growth zero unemployment, debt reduction; the only political way is an adequate preparation") and recovery funds from the European Union and the government ("pills on government trust – recovery fund and Conte… no need of a task force; EU OK!... loans to Italy… USA requests for unemployment benefits"). Negative tweets from urban areas were

Table 3 – ANA 4 results on area type and geography

Variable	*Category*	*Sentiment*	*Percentage*	*Percentage (only positive and negative)*
Area type	Urban	Positive	36.4%	41.0%
		Negative	52.3%	59.0%
		Neutral	11.3%	
	Rural	Positive	50.0%	53.6%
		Negative	43.3%	46.4%
		Neutral	6.7%	
Geography	North	Positive	43.8%	46.7%
		Negative	50.0%	53.3%
		Neutral	6.2%	
	Center-South	Positive	40.5%	45.9%
		Negative	47.6%	54.1%
		Neutral	11.9%	

Source: authors' computation

more about inadequate school closures, conspiracy theories and a more general discussion about reasons causing the pandemic ("delay payments and repayments perpetually overlapping one another…. The big secret…. modern capitalist regime values finance ethics… debt…. capitalism crisis").

As for geography, difference are minimal and focused on economic arguments, more on development and recovery in the North and more on unemployment worries in the Center-South ("financial sector in a 'red place'; business development largely affected by the rest of the market; news related to circulation; banks, finance and insurance helping recovery…"; "'Unimpresa', Confindustria newspaper raises the alarm about loss of thousands of jobs; tourism sector highly affected…increase in unemployed people…"; "Recovery in the hands of artisans and retail sectors…").

5. Conclusions

Our analysis shows that the biggest concerns are expressed when it comes to areas affected by tight lockdowns. When the discussion becomes broader, embracing the whole country and without reference to the small areas placed in the red zone, the attitude is more positive. From the topic analysis it turns out that some tweets are against the government taking such drastic measures; they concern problems related to the national health service and more generally there are references to uncertainty in the future and the need to take urgent action to put an end to the problems generated from the pandemic. In the georeferenced analysis a positive attitude is more spread in rural areas and slightly more in the North than in the Center-South. In the future a more semantic-based textual analysis will be implemented to better capture people's real feeling.

References

Aassve A., Alfani G., Gandolfi F., Le Moglie M. (2021), Epidemics and trust: The case of the Spanish Flu. *Health Economics*, 30, 1: 840-857. Doi: 10.1002/hec.4218.

Aum S., Lee S.Y., Shin Y. (2021), COVID-19 doesn't need lockdowns to destroy jobs: The effect of local outbreaks in Korea. *Labour Economics*, 70: 101993. Doi: 10.1016/j.labeco.2021.101993.

Bagchi B., Chatteriee S., Gosh R., Dandapat D. (2020), *Coronavirus Outbreak and the Great Lockdown*. Berlin: Springer. Doi: 10.1007/978-981-15-7782-6.

Blei D.M., Ng A.Y., Jordan M.I. (2003), Latent Dirichelet Allocation. *Journal of Machine Learning Research*, 3: 993-1002.

Chirisa I., Mutambisi T., Chivenge M., Mabaso E., Matamanda A.R., Ncube R. (2020), The urban penalty of COVID-19 lockdowns across the globe: manifestations and lessons for Anglophone sub-Saharan Africa. *GeoJournal*, 87: 815-828. Doi: 10.1007/s10708-020-10281-6.

Hasan A., Moin S., Karim A., Shamshirband S. (2018), Machine Learning-Based Sentiment Analysis for Twitter Accounts. *Mathematical and Computational Applications*, 23: 49-54. Doi: 10.3390/mca23010011.

Indira K., Brumancia E., Siva Kumar P. (2019), Location prediction on Twitter using machine learning Techniques. *IEEE – Proceedings of the Third International Conference on Trends in Electronics and Informatics (ICOEI 2019)*, 700-703. Doi: 10.1109/ICOEI.2019.8862768.

Karaye I.M., Horney J.A. (2020), The impact of social vulnerability on COVID-19 in the US: an analysis of spatially varying relationships. *American Journal of Preventive Medicine*, 59, 3: 317-325. Doi: 10.1016/j.amepre.2020.06.006.

Kumar A., Singh J.P. (2019), Location reference identification from tweets during emergencies: A deep learning approach. *International Journal of Disaster Risk Reduction*, 33: 365-375. Doi: 10.1016/j.ijdrr.2018.10.021.

Loria S. (2014), TextBlob repository – https://github.com/sloria/TextBlob/blob/dev/textblob/en/en-sentiment.xml – last access May 2022.

Malik U. (2020), Removing Stop Words from Strings in Python – Stackabuse.com – last access May 2022.

Morales E. (2021), *BikeMi: Data analysis, optimization and monitoring of the service quality for Bike sharing in Milan*. Master's thesis. University of Milan, Italy.

Pak A., Adegboye O.A., Adekunle A.I., Rahman K.M., McBryde E.S., Eisen D.P. (2020), Economic Consequences of the COVID-19 Outbreak: The Need for Epidemic Preparedness. *Frontiers in Public Health*, 8: 241. Doi: 10.3389/fpubh.2020.00241.

Peters D.J. (2020), Community susceptibility and resiliency to COVID-19 across the rural-urban continuum in the United States. *The Journal of Rural Health*, 36, 3: 446-456. Doi: 10.1111/jrh.12477.

Prabhakaran S. (2018), Lemmatization Approaches with Examples in Python. Url: www.machinelearningplus.com – last access July 2021.

Schumacher A. (2015), TextBlob Sentiment: Calculating Polarity and Subjectivity. https://planspace.org/20150607-textblob_sentiment/ – last access May 2022.

Sloan L., Morgan J. (2015), Who Tweets with Their Location? Understanding the Relationship between Demographic Characteristics and the Use of Geoservices and Geotagging on Twitter. *Plos One*, 10, 11: e0142209. Doi: 10.1371/journal.pone.0142209.

Ur Rahman I., Jian D., Junrong L., Shafi M. (2021), Socio-economic status, resilience, and vulnerability of households under COVID-19: Case of village-level data in Sichuan province. *Plos One*, 16, 4: e0249270. Doi: 10.1371/journal.pone.0249270.

Come va la vita in paese? Resilienza economica e reazione durante i *lockdown* pandemici

Sommario

Sebbene le pandemie rappresentino un problema ricorrente nella storia, la pandemia di COVID-19 ha alcune caratteristiche mai viste prima. L'essere umano è sopravvissuto a guerre, catastrofi naturali e shock economici, dimostrando sempre resilienza nell'adattarsi a nuove situazioni. In questo lavoro vogliamo verificare se questo atteggiamento è ancora forte. La nostra domanda di ricerca è: come reagiamo alle situazioni emergenziali? In questo articolo cerchiamo di rispondere a questa domanda analizzando un database di messaggi Twitter raccolti durante la seconda e la terza ondata di COVID-19 in Italia riguardanti la vita quotidiana durante i rigidi lockdown e le opinioni su queste situazioni. Le piccole città e i paesi sono il nostro punto di partenza, interrogando dapprima le reazioni della popolazione nel periodo di forte restrizione e secondariamente osservando le risposte ai cambiamenti sociali ed economici senza particolare riferimento al periodo di lockdown; sposteremo poi la lente sul territorio italiano a livello macro indagando i comportamenti di resilienza, ed infine daremo uno sguardo alle aree di nord, centro e sud grazie alla geo-localizzazione dei messaggi Twitter.