ARTICLE

OPEN

Check for updates

Time windows of opportunities to fight earthquake under-insurance: evidence from Google Trends

Fabrizio Terenzio Gizzi [1]✉, Jonghun Kam[2,3] & Donatella Porrini [4]

Earthquake insurance can be a useful tool to build more sustainable societies and disaster-resilient communities. However, the coverage is not common in many countries. This article aims to contribute to the literature through an empirical analysis of the online interest in earthquake insurance through Google Trends. The proposed methodology implies to move from a top-down conceptual approach to a bottom-up/data-enabled one. It allows us to explore potential triggers and dynamic patterns of online interest in earthquake insurance at daily time-scale through the lens of Big Data. In order to validate the methodology, the article considers Italy as a test area. For this country, where the coverage rate is low, we fuse multiple databases to create 16-year daily time series of public search activities about the insurance in Italy and analyse it with other data sources. As a result, the peak analysis shows a connection with the occurrences of large domestic earthquakes, overseas earthquakes, and policy decisions, which create time windows of opportunities for insurers and policymakers to boost the public's motivation towards the coverages. The research outcomes suggest that the data-enabled approach can additionally be applied in other countries where the coverage rate is low and stakeholders are facing the challenge to strive against earthquake under-insurance.

---

[1] Institute of Heritage Science, National Research Council (ISPC-CNR), Area della Ricerca, C.da S. Loja, 85050 Tito (Potenza), Italy. [2] Division of Environmental Science and Engineering, Pohang University of Science and Technology, 202 Jigok Research Building, Pohang, South Korea. [3] Graduate School of Artificial Intelligence, Pohang University of Science and Technology, 202 Jigok Research Building, Pohang, South Korea. [4] Department of Management, Economics, Mathematics and Statistics, University of Salento, Centro Ecotekne - Via per Monteroni, 73100 Lecce, Italy. ✉email: fabrizioterenzio.gizzi@cnr.it

Introduction

Natural hazards can cause serious damages in communities so posing serious challenges in terms of policy choices. Among natural hazards, earthquakes are one of the main causes of disaster losses in the world (CRED, 2015). That imposes, jointly to the awareness of unpredictable features of earthquakes (Kagan, 2006), to focus particularly on preparedness, including mitigation actions and recovery tools to compensate for earthquake damages (Nigg, 2000; Dowrick, 2003; Spittal et al., 2008; Moon et al., 2019; Seko, 2019). One of these tools is the insurance and the literature on disasters highlights the role played by coverages in mitigating disaster risk (Kunreuther, 2000). In this perspective, homeowners earthquake insurance is often proposed as a useful mean for a post-disaster recovery plan, providing also incentives to preventive behaviours (Kunreuther et al., 1988; Lin, 2019a; Palm, 2020).

However, the literature often discusses the factors that determine the willingness to purchase disaster insurance that could vary substantially across developing and developed countries (Panda et al., 2020). In particular, the studies often report unwillingness to purchase policies (Wang et al., 2012; Zhu et al., 2012; Pothon et al., 2019) due to the high premiums, the socio-demographic and geographical features of the homeowners, and the low risk perception (Kunreuther et al., 1978; Slovic, 1987; Palm and Hodgson, 1992; Naoi et al., 2010). What's more the issue of under-insurance is also addressed by international institutions which at the same time underline the role of disaster insurance: "[…] Supporting comprehensive and affordable insurance coverage of disaster risks across economies with varying levels of insurance market development remain a challenge in many economies […]" (OECD, 2015, p. 11).

On an empirical point of view, previous studies mainly followed a top-down conceptualization approach based on the limited socio-economic data and developed conceptual (stochastic) models to simulate the interactions among the key local components of the affected community (e.g., premium amount, socio-economic structures, and risk perception) (Pothon et al., 2019). Often, some of the key components (e.g., perceived risk) have been simplified in the models for conceptualization. For example, perceived risk is often conceptualized by affecting local components, such as event occurrences and socio-economic structures. Accepting the intrinsic limitation from no existing data for validation, this top-down approach has been popular and has supplied lots of suggestions for the improvement of earthquake insurance purchase via scenario-based testing. However, the under-insurance phenomenon is still a challenge for earthquake disaster-prone communities.

Here, we propose a new bottom-up paradigm that can generate data-enabled science to analyse earthquake insurance interest. To test the methodology we consider Italy, one of the countries with the highest seismic hazard in the Mediterranean area, as test site. Despite of adverse effects of earthquakes, the insurance penetration in Italy is the lowest among the OECD countries (OECD, 2018), so that several attempts to introduce a compulsory or semi-mandatory insurance system have been made since 1998 (Amendola et al., 2000).

To evaluate the key factors that can motivate the purchase of an earthquake coverage, this study examines the spatio-temporal patterns of online search interest related to earthquake insurance on a national (Italian) scale. To monitor and model the dynamics of earthquake insurance interest in population, multiple data are integrated, ultimately having the 16 years daily data as target (2004–2019). The multiple data includes daily and monthly Google Trends (GT) statistics from various search options, GT GeoMaps (GMs), and annual percentages of internet users in Italy.

While Google does not explain in detail some aspects (e.g., the criteria followed to collect the raw data), the utility and reliability of GT has been proved across different disciplines (Carneiro and Mylonakis, 2009; Choi and Varian, 2012; Nuti et al., 2014). More recently, GT has been used as an indicator of public awareness/interest of disasters, such as earthquakes (Tan and Maharjan, 2018) and drought (Gonzales and Ajami, 2017; Kam et al., 2019; Kim et al., 2019) across the spatial scales at regional, national, and global level.

In the article, we analyse the GT public interest in earthquake insurance: (i) to link it to the possible increase of people's preparedness after the multi-facet public's hazard experiences gained as a results of seismic events occurred in Italy as well as overseas, (ii) to link it to the policy decisions, (iii) to argue possible relationship between information seeking and real actions (such as insurance purchase), (iv) to move the research outcomes towards insurance industry and policymakers, indicating time windows of opportunities to make people more confident with earthquake coverages so as to increase policy subscription rate.

Earthquake experience is recognized as influencing preparedness actions (Becker et al., 2017; Bronfman et al., 2019), among which we can include the insurance coverage (Moon et al., 2019). Recently, Becker et al. (2017) argued that the types of earthquake experiences swaying the household's behaviour and motivating preparedness are four: *direct experience* (i.e. direct involvement in injury or damage, physical feeling of the event), *indirect experience* (i.e. individuals directly exposed to the event without being personally affected such as participating as volunteer, having the emergency management role), *vicarious experience* (i.e. experience of disasters reported by relatives/others or media reports of national and international events), and *life experience*, not related to disaster experience (i.e. adverse life experiences, such as car accidents). Each type of these experiences can influence the preparedness actions differently, but all interact with each other in shaping the individual response to preparedness (Becker et al., 2017).

On the one hand direct experience is known as being the most effective in motivating preparation, on the other hand indirect and life experiences, even if less effective than direct one, also influence the preparedness (Lee, 1999; Becker et al., 2013). What's more, vicarious experience can be considered as an important predisposing factor (Becker et al., 2017). Bao et al. (2019) found that direct and/or vicarious experiences contribute to reduce the likelihood of unplanned responses to unexpected emergencies. Rogers and Sorensen (1991), analysing the adoption of emergency planning practices for chemical hazards in the, US found that vicarious experience contribute to variation of the community preparedness level. Moon et al. (2019), examining direct earthquake experience, identified the determinants motivating preparedness and insurance purchase in moderate-seismic zones such as Korea. Lin (2019a) showed that in case of earthquake insurance in California the "feeling" of an earthquake as a personal experience, even without having losses, can affect the insurance demand.

Methods

Search queries and time intervals. The Statista Research Department reported that Google is the most popular search engine in Italy as December of 2019 (93.18% share of the search engine market) (https://www.statista.com/statistics/623043/search-engines-ranked-by-market-share-in-italy/).

GT is a powerful tool to visualize Big Data from 3.5 billion searches every day (https://www.internetlivestats.com/). It provides relative search/query activity volumes of Google Products'

users with a search topic/term over an interest region (city-level, county-level, state-level, and nation-level) during a specific period. The GT activity volume data is available since 2004. The relative search activity volume is scaled from 0 to 100 based on the maximum of the absolute activity volume during the period that a user selects. The temporal scales (hourly, daily, weekly to monthly) of the interest search activity volumes are automatically determined based on the length of the selected period.

Recently, the GT platform has been updated with the two search options: "search topic" and "search term". The results from the "search topic" option are based on the search activities using the term across the languages while the results from the "search term" are based on the activities using the exact search word.

GT also provides interest search activities at different spatial scales. Users can select the spatial scales of the interest search activity volumes from city-level, state/province-level, national-level, and global-level. It hints which communities are a key contributor of the web-search activities volumes during the corresponding period.

In this study, first we retrieved monthly relative search volumes data in Italy over the period 2004–2019. Then, we retrieved the daily relative search activity volumes in Italy at the 6-month intervals (32 times from 01/2004 through 12/2019). These monthly and daily GT data were saved on January 28, 2020 through the two search terms and options: the term ("terremoto") and the term ("assicurazione terremoto"). It is worth noting that the raw daily data has at least 32 days with the relative search activity volume, 100 (every 6 months), over 2004–2019, which means that the post-processing is necessary to account for the relative search activity volumes across the 6-month intervals.

Furthermore, in order to have an overview of the Italian regions from which people performed the online searches, we built cumulative geographical maps (GMs) for significant dates. GMs were built according to the following steps: (1) selection of the days of interest, (2) identification of the Italian regions involved in the searches during the first day of interest by the "Interest by regions" field included in the GT results, (3) building of the first-day map, (4) expansion of the search period by one (or n-day) day of interest and identification of the key regions contributing to the total volume of relative search activities during the expanded search period, (5) building of the second (or n-day) map including in it all the regions involved in the searches from the first date through the 2nd (or nth) day. Data for the building of GMs were retrieved between February 9 and February 12, 2020.

Italy experienced a gradual growth of the percentage of internet users since 2004. According to the World Bank Open Data (https://data.worldbank.org/), 34% and 73% of the population in Italy had an internet access in 2004 and 2018, respectively. It means that absolute search activity volumes of Google Products' users in Italy during the first decade of the study period (2000s) are expected to be much smaller than those occurred during the later decade (2010s). In the next section, we develop a methodology for weighted daily GT data to account for the monthly weight for search volumes and the growth of internet user in Italy.

Weighted daily GT. Here, we propose a method to generate weighted daily GT data to account for the relative search volume across the 6-month search periods and the increasing trend of internet users in Italy. First, the raw monthly GT data (MGT(y, m); y, m, and d are the indicators of the yth year and mth month of the search period) are converted to the daily scale (MGT(y, m, d); d is the indicator of the dth day in the corresponding year and

month. Within a month, the daily values are equal to the value of the corresponding month and year (Eq. (1))

$$\text{MGT}(y, m, d) = \text{MGT}(y, m) \qquad (1)$$

Second, the annual percentages of relative internet users to the population (API(y)) are converted to the daily percentages of internet users (API(y,m,d)) by fixing the daily percentages at the annual percentage of the corresponding year that provided by the World Bank Open Database (Eq. (2)).

$$\text{API}(y, m, d) = \text{API}(y) \qquad (2)$$

Third, the daily weights are computed by multiply MGT(y,m,d)/100 by API(y,m,d)/100. These daily weights account for both the increasing trend of internet users in Italy and the relative internet activity volume over 2004–2019 (Eq. (3)).

$$\text{wgt}(y, m, d) = \frac{\text{MGT}(y, m, d)}{100} * \frac{\text{API}(y)}{100} \qquad (3)$$

Then, the weighted daily GT data (wDGT(y,m,d)) are generated by multiply the raw daily Google data (DGT(y,m,d)) by the daily weights (wgt(y,m,d)) and then dividing it by their maximum value (the denominator of Eq. (4)). Lastly, the weighted daily GT data is scaled from 0 to 100 by multiplying 100.

$$\text{wDGT}(y, m, d) = \frac{\text{wgt}(y,m,d)*\text{DGT}(y,m,d)}{\max_{2004 \leq y \leq 2019}\max_{1 \leq m \leq 12}\max_{1 \leq d \leq l}(\text{wgt}(y,m,d)*\text{DGT}(y,m,d))} * 100 \qquad (4)$$

where l is the last day of the corresponding year and month ($l = 28$, 29, 30, or 31, depending on the corresponding month of the year). Time series of wDGT are generated using the three search options. Results show similar dynamic patterns over 2004–2019 while the magnitudes of the peaks of the wDGT values, using the search term "assicurazione terremoto", are slightly different from that of the peaks of the wDGT values using the search term "terremoto" (Fig. 1).

For the diffusion pattern of search activities, GT GeoMap (GM (y, m, d)) data have been retrieved to generate the cumulative geographical distribution of search activities related to earthquakes. For example, first the GT GM data (GM(2016, 8, 24)) is retrieved on 24 August 2016 when a major earthquake occurred in Central Italy and then over August 24–25 (GM(2016, 8, 24–25)), 24–26 (GM(2016, 8, 24–26), and until the wDGT value is very low (for example, <5). The GT GM data provide the relative contribution of each state/province to the entire interest region (herein, Italy) during the chosen search period.

Results and discussion

From the wDGT data, firstly comparable tendency can be observed for "terremoto" and "assicurazione terremoto" so as one can infer that the peaks of interest for insurance are linked to interest in earthquake that in turn rise significantly when the territory is shaken by earthquakes. Then, the temporal correlations (Spearman rank order and Pearson linear correlation) are computed to quantify the degree of associations of the GT data with the number of daily earthquakes (Fig. 1d).

According to Table 1 and Fig. 1, "terremoto" (TR) and "assicurazione terremoto" (ATR) patterns can be divided into four main periods by the occurrences of major earthquakes. Moreover, their correlations indicate that the relationship between wGT and daily frequency of earthquakes in Italy becomes stronger in recent years than in early 2000s.

From the earthquake records and wDGT data, we identified three types of trends regarding insurance interest: (1) peaks related to the immediate occurrence of earthquakes; (2) peaks not connected to the immediate occurrence of earthquakes; and
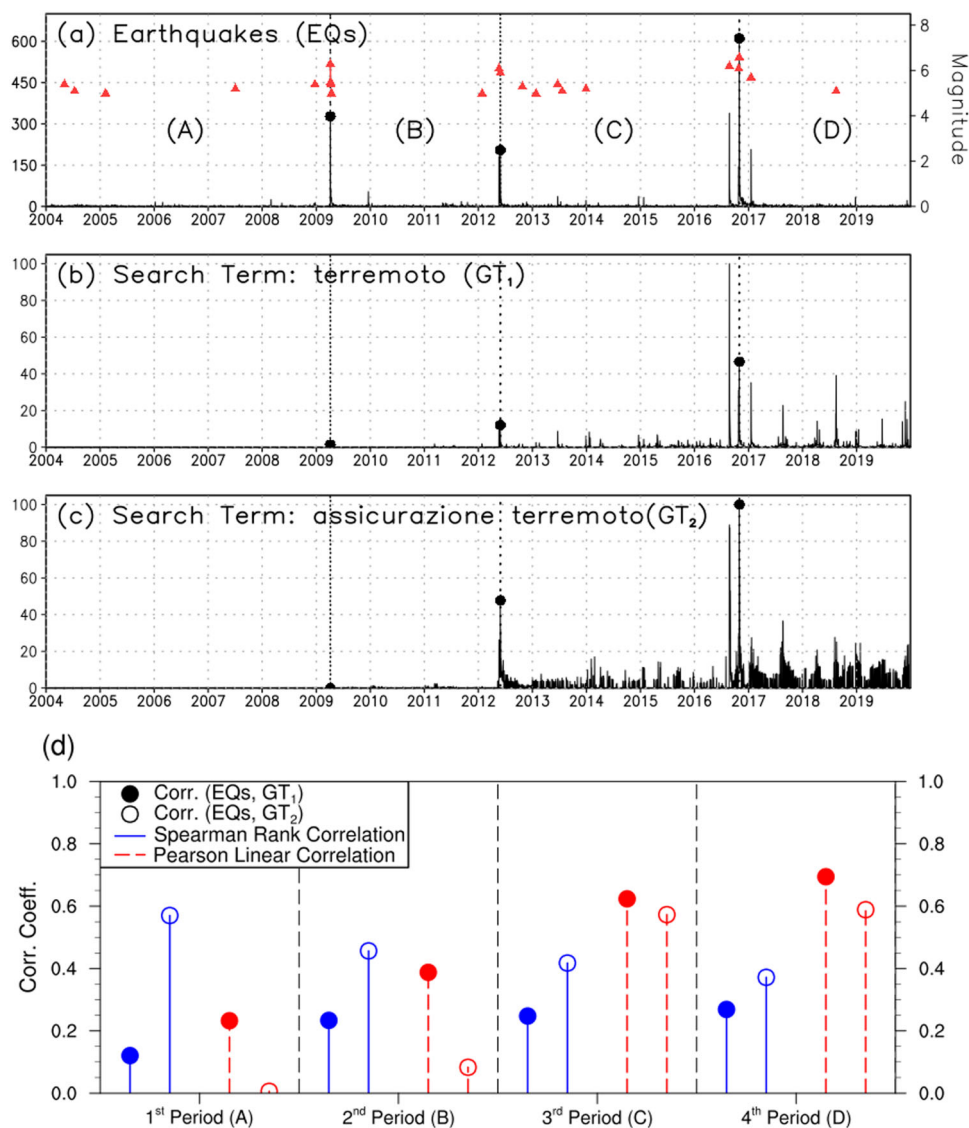
Fig. 1 Earthquakes and Google Trends data considered for the period (2004–2019) under investigation in the article. a Triangles: daily earthquakes ($M \geq 5.0$) (data from Rovida et al. (2019) for 2004–2017, from http://terremoti.ingv.it/ for years 2018–2019, in case of multiple events the same days plots refer to the maximum magnitude of that day), bars: daily number of earthquakes; **b**, **c** weighted daily time series of Google Trends (DGT) data using the two search options; **d** correlation between DGT data and daily earthquakes.

(3) no peaks in the occurrence of earthquakes. In the following section, we analyse in details these types of data and characterize four different periods, such as a first period of no interest, a second period when the interest begins, a third period when the interest gets a foothold, and a fourth period when the interest grows and consolidates.

At the beginning of the interest. After the previous non-interest period, then the following period ("B" in Fig. 1a) showed TR and ATR interests increased slightly even with only very low volumes (Fig. 1b, c). The most significant earthquake of this period occurred on April 6, 2009. The main shock (Mw 6.3, Imax 9–10 MCS scale), followed by numerous aftershocks (Rovida et al., 2019), struck the Abruzzo region with 308 fatalities, 1600 hurt and more than 65,000 people had to leave their homes. In the bordering regions of Lazio, Molise and Marche and subordinately in Umbria, Campania, and Apulia, people felt the earthquake (Tosi et al., 2007).

GT data show that TR search volumes peaked on the same day ($V = 1.4$) of the major shock, but volumes quickly dropped below 1 as early as the next day. Interest in insurance started on April 6, 2009 and decreased gradually over about the following 20 days with an interest that was, however, always low. Higher values were recorded within the first week after the main shock (maximum $V_{ATR} = 2.4$).

Looking at the cumulative geographical distribution (Fig. 2), 50% of regions (10) were involved altogether in the internet activity. In particular, on April 6, only one region was in GM (Lazio) while on April 9, the regions were five. In the following days there was a slowdown in geographical diffusion: it is necessary to wait for April 14 to count seven regions among those overall involved in the online searches. The region directly hit by the earthquake (Abruzzo) was involved in the searches only after 3 weeks (April 28) by the main shock, date on which not all the regions bordering Abruzzo were implicated in online search. This user's behaviour can be linked to the circumstance that the residents of Abruzzo were focused on the emergence phase of the first 3 weeks.

Table 1 Main statistical data about the period analysed.

Period (no.)	Period	Total days (no.)	Days volume > 0 no. (%)	Volume range ATR (min-max)	Total volumes (ATR)	Interest
1st (A)	January 1, 2004– April 5, 2009	1922	10 (0.5)	0.0–0.1	1.0	Absent
2nd (B)	April 6, 2009–May 19, 2012	1140	112 (9.8)	00–3.8	75.0	Very low
3rd (C)	May 20, 2012–August 23, 2016	1557	291 (18.7)	0.0–47.8	1568.1	Low–moderate
4th (D)	August 24, 2016–December 31, 2019	1225	359 (29.3)	0.0–100.0	4223.8	Moderate–high
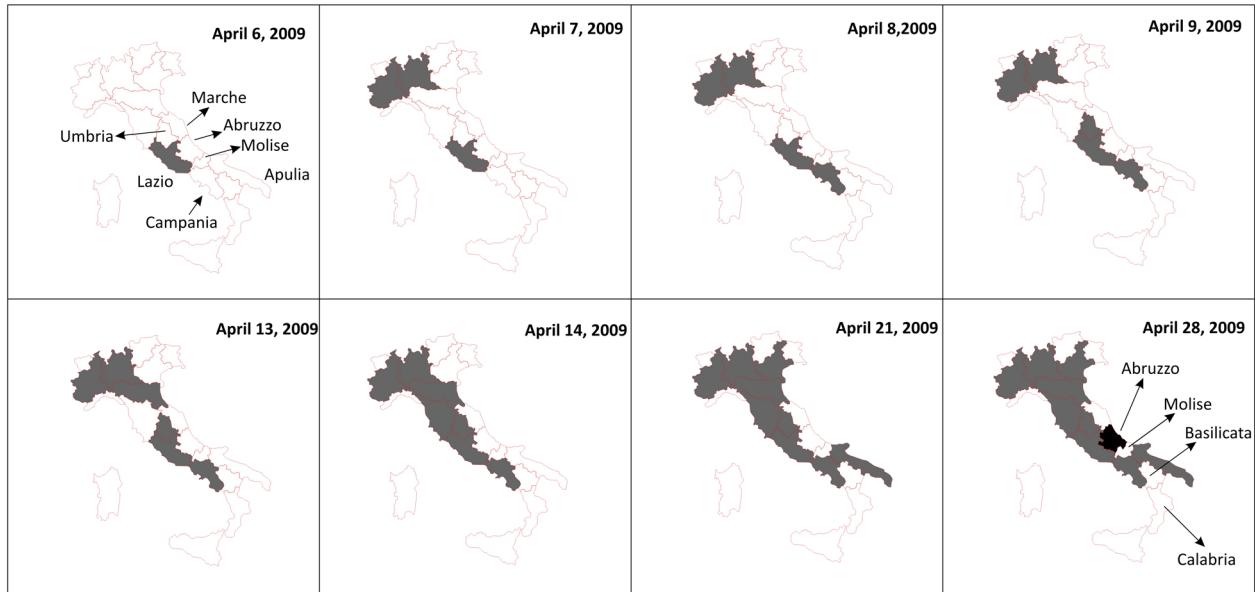
Fig. 2 2009 Abruzzo earthquake. Spread of internet searches (GMs) across Italy for ATR. Each GM represents the overview of the Italian regions involved in the searches from April 6, 2009 to the indicated date.

From these GMs data, one can argue that in the immediacy of earthquake, the ATR search regarded people who were only feeling the shake. During the following almost 3 weeks the ATR online activity regarded users which were involved in direct and/ or vicarious experiences. Vicarious experience takes place, as we defined above, when a person is exposed to secondary sources of information about the disaster (Breakwell, 2007; Becker et al., 2017). In particular, the public interest in the regions of Lazio, Campania, and Apulia could be also increased by the vicarious experiences of others (relatives/friends). Indeed, the search activities were probably motivated by the concerns related to two circumstances. The first is that the Universities of Abruzzo hosted about 15,000 students resident in those three regions in 2009 (https://anagrafe.miur.it/), the second concerns the partial collapse of the "Casa dello Studente" (Student House) at L'Aquila with eight students killed and three suffering serious injury (Mulas et al., 2013). This event had also a wide echo on the media.

Finally, from GMs one can also note that from some regions of Southern Italy (Molise, Basilicata and Calabria) characterized by a medium-high seismic hazard, there was no ATR online activity. However, GMs might be influenced by two aspects: (1) Molise and Basilicata are among the regions with the lowest population in Italy; (2) Internet access (families with home Internet access) by the three regions is below the national average (http://dati.istat. it/, data referred to 2016).

Other earthquakes of moderate magnitude that occurred in this year aroused no immediate interest in searches.

In contrast, ATR search activities on March 2011 was not linked to immediate occurrence of earthquakes in Italy,

considering that seismic activity was not mentionable in that period. Anyway, moving the attention to the worldwide scale, from March 9 to 19, over 20 earthquakes with $M > 6.0$ and maximum magnitude M 9.0 shaken Japan. The occurrences of extremely high-magnitude earthquake and associate aftershocks were reflected in the (low) searches whose values so deviated from the "zero" baseline of previous weeks. In fact, the search for "terremoto Giappone" (earthquake Japan) in the period 2004–2009 gives its maximum value just in March 2011, thus stimulating a low interest for earthquake insurance, due to the media vicarious experience (Becker et al., 2017) (see Fig. SI, Supplementary Information, SI).

The highest of the entire second period (3.8) was recorded on May 15, 2012 and it was reached with no significant earthquake activity (magnitude <2, http://terremoti.ingv.it/), confirming that even non-seismic factors could have driven the ATR online activity. This slight peak could be connected to the enactment of a Decree concerning provisions for the reorganization of Civil Protection. Even if an article of the Decree tried to introduce a voluntary hazard insurance coverage, the lack of an agreement between the policymakers on what insurance scheme to be adopted undermined the efforts once again, and the article was suppressed.

Interest gets a foothold. The third period begins on May 20, 2012 with the damaging Emilia earthquake (Mw 6.1, Imax 7 EMS-98 scale, Rovida et al., 2019). The event, followed by numerous aftershocks (Mw 4.0+) the same day, was felt by people particularly in the bordering regions of Veneto, Lombardy, Tuscany, and Marche (Tosi et al., 2007).
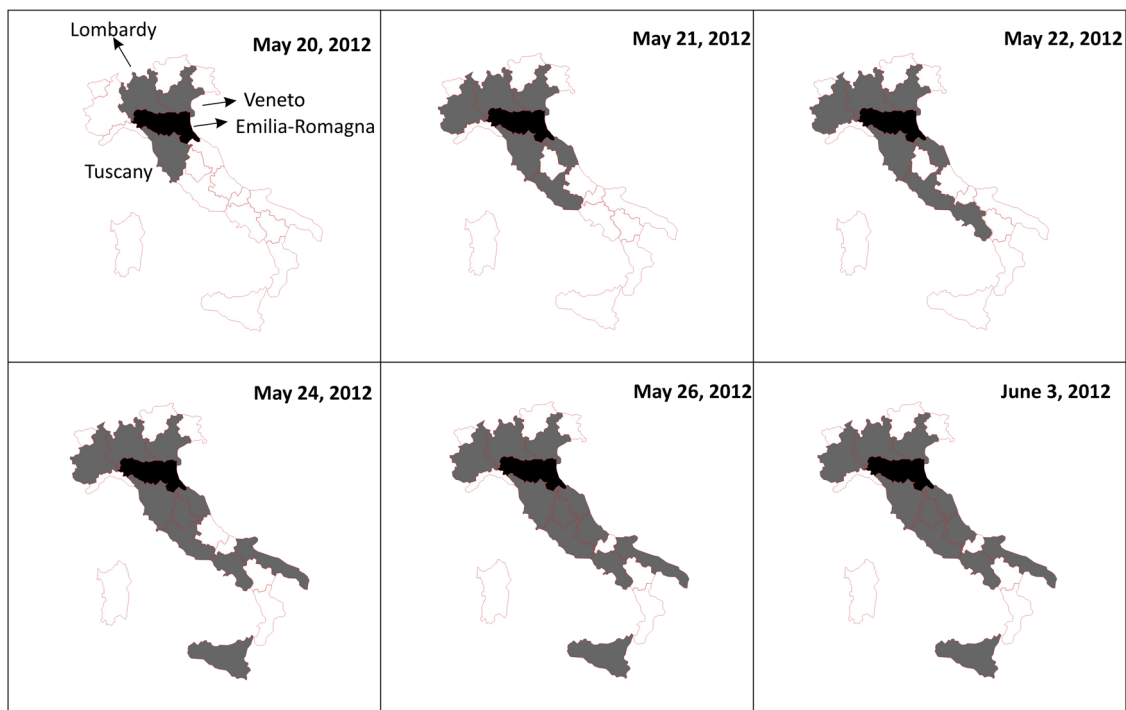
Fig. 3 2012 Emilia earthquake. Spread of ATR internet searches over the Italian regions from May 20 to June 3, 2020.

The ATR interest increased on the same day, reaching immediately volumes (20.6) significantly higher than the previous periods. The volumes still increased the following day (26.3) and then decreased at least until May 28 probably as it was not "sustained" by further significant ground shakings. On May 29, the volumes went up and reached the maximum peak (47.8) of the period, likely due to the occurrence of numerous shakes among which two of medium–high magnitude (Mw 5.9 and 5.5). However, the interest was quite relevant (ATR > 10.0) until June 5. Conversely to what ascertained for ATR, interest in TR peaked just in 2 days, May 20 and 29.

On May 20, the ATR searches concentrated in Emilia and bordering regions (Fig. 3), where the residents felt the earthquake (s) (Tosi et al., 2007). These findings do not seem random because these regions are among the most sensitive to the issue of hazard insurance, with the highest percentages of policies subscribed (ANIA, 2017). Consequently, the online search activities of the people in this period might be also addressed to get information on how to prepare the insurance claims. In the next days, searches expanded gradually also towards Central and Southern Italy involving 60% of regions, but again excluding three of them (Molise, Basilicata and Calabria) characterized by a medium-high seismic hazard, so that our online data are in line with the very low natural hazards coverage rate of these three regions (ANIA, 2017). However, it is noteworthy that information seeking by the residents in Southern Italy could be linked to both media and (relates/friends) vicarious experiences, considering that economic differences among Italian regions have forced many people of South Italy to emigrate towards North and in particular in Emilia-Romagna, which is the most attractive region for job seekers (ISTAT, 2018). Another interesting observation is related to the comparison between our findings related to 2009 Abruzzo earthquake and those concerning the Emilia event. Indeed, for the Abruzzo event the ATR searches involved the region directly hit by the earthquake at the last time, conversely what happened after the Emilia earthquake. This behaviour can be linked to different peoples' awareness about natural hazards insurance.

For about a year since June 2012 the ATR volumes, while remaining low, were spread over many days (108). However, in most of the period, online activity was not related to the immediate occurrence of earthquakes, as already observed in the previous (second) period. Indeed, single shock with moderate magnitude and involving online TR search increase, usually did not cause variations in ATR, probably due to the circumstance that earthquakes did not cause damage and the residents in those areas had no experiences of significant recent earthquakes. Therefore, the interest might be considered mainly a long-lasting effect of the recent Emilia earthquake. It is worth noting that, in this case, the GT data cannot be useful to identify the regions from which the search started, as the data volumes were not sufficient.

The earthquakes that affected Tuscany in June 2013 had impact on insurance search. On June 21 an event (Mw 5.4) took place in the northwestern Tuscany. Other three events with magnitude between 4.3 and 4.8 were recorded in the month in the area: one on day 21, one on day 23 and one on day 30 (Rovida et al., 2019). In the first 4 days, there was an increase both in TR and ATR volumes.

Another earthquake, whose epicentre was at the border between Italy and France, happened on April 7, 2014 (Mw 4.7) determining increase in interest. The most involved Italian area was Piedmont, a region with low-seismic hazard, where the earthquake caused only slight damage in very few sites but people felt the shock in many localities (Tosi et al., 2007). The high level of interest in Piedmont can be connected with the fact that it is among the regions with the highest natural hazards coverages in Italy. Therefore, awareness about coverages was already developed to a certain extent. This finding indicates a potential remedy to adverse selection, which is often considered as a fundamental limit for the spread of seismic insurance (Lin, 2019b). Indeed, our results support the possibility to offer the insurance policies in areas where the insolvency risk for companies is low due to low seismic hazard.

Another seismic event (Mw 4.1) had repercussions on online searches. It hit Tuscany again and particularly the Florence
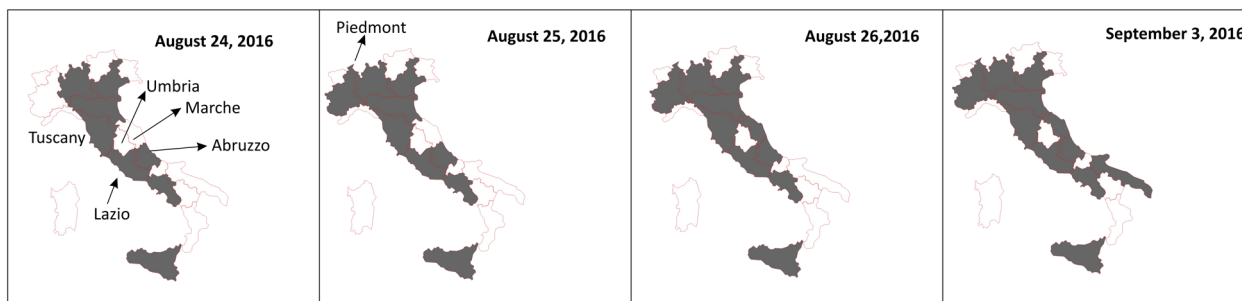
Fig. 4 2016 Central Italy earthquake. Spread of ATR internet searches over the Italian regions from August 24 to September 3, 2016.

province on December 19, 2014, without damage in the involved localities. The event was preceded and followed by some foreshocks and many aftershocks on the same day (http://terremoti.ingv.it/). The event(s) had upshots on insurance volume searches (4.2) only on day 19 and had no impact on the online activity of the following days. The low search activity volumes is probably due to the circumstance that shakings caused no damage and the shocks were clustered only in one day.

The earthquake on January 23, 2015 (Castiglione dei Pepoli (BO), Mw 4.3) occurred in the same region affected by the Emilia 2012 earthquake, even if in a different area. The main shock was between foreshocks and aftershocks (3.0+) occurred on the same day, when the TR and ATR searches peaked. Noteworthy, search volumes in insurance (11.1) can be due to both aftershocks and a high level of insurance awareness of the residents who were also sensitive to the topic by the discussions open in the media in the aftermath of the 2012 Emilia earthquake.

We found sporadic (about 40 days) and low search volumes in insurance over the next 20 months. However, most of these peaks were not connected to the direct occurrence of earthquakes in Italy. In particular, a peak of interest in insurance was recorded on August 1, 2016. This peak, the highest over the 20 months ($V_{ATR} = 17.1$), can be considered the late reflection of the issue of an awaited Resolution of the Italian Council of Ministers of July 28, 2016 relating to damages compensation after disasters. The provision includes references to insurance coverage for natural hazards, stating how damages had to be compensated for. The relationship can be confirmed by the correlation between the insurance searches with the online activity related to the Resolution, looking at the days at the turn of August 1, 2016. These data indicate that just on August 1 there was the second most important peak in online searches about the Resolution, thus suggesting a probable correlation with the insurance online activity peak.

Interest grows and consolidates. The fourth period is characterized by the seismic sequence of 2016–2017 in Central Italy. The sequence started on August 24, 2016, causing about 300 casualties, damage, and failures in parts of four regions of Central-Southern Italy (Lazio, Umbria, Marche, and Abruzzo).

The shake on August 24 (Mw 6.2, Imax 10 EMS-98 scale) was followed by about 10 events of magnitude between 4.0 and 5.5 approximately within 22 h and about 10 events of magnitude between 4.0 and 4.9 within the following 9 days (Rovida et al., 2019). The ATR interest grew rapidly and reached very significant volumes (88.9) on the same day when the main shock occurred. Furthermore, the ATR volumes remained higher than 10 until September 3. Averaged search values remained quite notable until August 31 (51.7), reaching a relative high peak on August 27 (87.3). The persistent prominent search activities can be connected to both the occurrence of an earthquake of Mw 4.3

on August 27 and the "renewed" interest triggered by some strong M4+ events that occurred in the previous days.

Since August 24 the people showed interest in the regions affected directly by the Mw 6.2 shock or where the seismic event(s) caused damage and the geographical spread involved many of the regions where the earthquake was only felt, according to the data of Tosi et al. (2007) (Fig. 4). The spatial pattern of searches was certainly also influenced by vicarious experiences considering that the earthquake attracted high media attention. In detail, 40% of the Italian regions (8) were involved in internet activity on August 24, when the first strong earthquake hit the Lazio region directly (epicentre within the region). On August 25, aftershocks (Mw 4.0+) occurred only in the Lazio region, close to the boundary with Abruzzo. Therefore, the geographic picture was quite similar to that of the previous day, adding only one region (Piedmont, NW of Italy). On day 26, the Marche region, which suffered damage by the earthquake of day 24, "enters" in GM, while Umbria was not mapped probably due to the limited damage in the region (Fig. 4). Again, the three regions of Southern Italy with a medium-high seismic hazard were not involved in the internet activity.

From September 4 to October 25, 2016 the interest in insurance was relatively low, but it was not guided by the immediate earthquake occurrence. Conversely, the severe earthquake of October 26 (Mw 6.1) and in particular that of October 30 (Mw 6.6), the strongest event in the sequence, aroused great interest pushing the search volume to the maximum (100) of the entire period analysed in this article (2004–2019). In the 30 days following the October 30 earthquake, many M4+ shocks occurred. Throughout this period, online attention related to insurance was always present.

In the following month and a half the attention towards ATR was lower. However, it went up promptly with the occurrence of four moderate–high (Mw 5.0+) earthquakes in Abruzzo on January 18, 2017. Nevertheless, the interest in ATR, even if modest in amount, lasted for almost a week beyond the "triggers" caused by the earthquakes. This might be due to great attention developed during the long seismic sequence of Central Italy which motivated the preparation intentions more.

In the 7 months following, the insurance interest had ups and downs. Again, even if it was rather low, the interest can only partially be related to the occurrence of significant earthquakes. The attention was likely due to the interest accumulated previous events and "released" only weeks and months later.

After these months, the Casamicciola earthquake (August 21, 2017; Mw 3.9) occurred on the island of Ischia (Campania region, Southern Italy). The event had a very superficial hypocentre thus causing intense local damage especially in Casamicciola and triggering a clear volume peak in TR the same day (17.8). However, the highest peak (22.8) was recorded the day after, probably due to the evening hours not conducive to a complete understanding of the earthquake effects. The importance of the
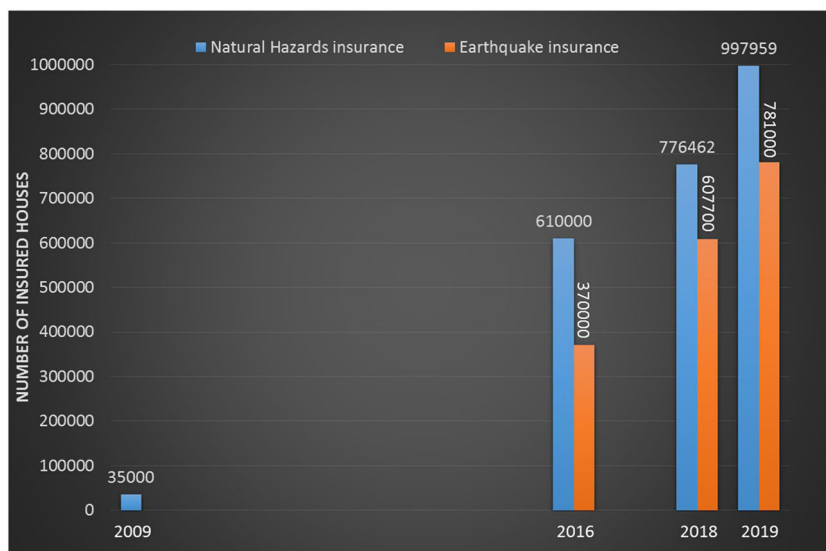
Fig. 5 Number of insured houses for natural hazards (floods and earthquakes) and earthquakes in Italy since 2009. Overall data about natural hazards coverage rise from ANIA (2017) for 2009 and 2016, from ANIA (personal communication[1]) for 2018 and 2019; earthquake insurance data for 2016 rise from ANIA (2017), for 2019 from ANIA (2019), for 2018 were estimated by the authors based on 2019 data.

occurring time on information search activities was evident on the following day, with the resulting echo in the media. The highest ATR peak (36.6) was on the same day of the main shock, but the next day the search volume persisted at a high level and close to the highest (35.8). The highest interest came from some regions where people (internet users) had the experience from the previous 2012 and 2016–2017 earthquakes, so that the media vicarious experience of August 2017 awakened that past event in the memory (Fig. S2, SI). These data might confirm that having frequent thoughts about earthquakes induces more preparation and that also distant but time-closed events can have positive effects on judgments about preparedness needs (McClure et al., 2016). We found similar patterns of information seeking over the period from September 1, 2017 to December 30, 2019 (Fig. S3, SI).

On June 23, 2019, a modest earthquake of magnitude Mw 3.6 hit a site located about 20 km southeast of Rome, in the Lazio region. Despite the low magnitude and the limited effects not exceeding the instrumental intensity of VI (http://shakemap.rm.ingv.it), the interest for coverage grew rapidly and lasted, with an up-and-down trend, at least for a week, involving users only by the Lazio region. This could be linked to the experiences of two recent earthquakes. Indeed, a local earthquake occurred in July 2012 and another main event shaken Lazio directly during the 2016–2017 Central Italy sequence. These two events probably motivated the residents/homeowners to think about the needs of "protecting" their properties from earthquake hazards. In particular, the 2019 earthquake would have reawaken the sensitivity of the population on the issue of insurance, especially after the controversy that broke out in the media about the slowness of reconstruction in the regions affected by the 2016–2017 Central Italy earthquake. These disputes occurred in the weeks immediately preceding the earthquake of June 23, 2019 (this can be made evident examining the results of the search for "ricostruzione terremoto"—Italian of earthquake reconstruction —by Google in the period between June 1 and June 22, 2019).

The above discussion suggests that internet users showed a growing interest in earthquake insurance over time. Now, our question is whether that interest actually turns into the purchase of coverage. In actual fact, another advantage of the Big Data from daily billions searches from Google Products' users is that GT supplies additional information valuable to correlate the searches. Actually, the "associated queries" option of GT suggests that when users searched for "assicurazione terremoto" they also searched for "assicurazione terremoto casa costo" (home earthquake cost insurance) and "Generali assicurazione terremoto" (Generali earthquake insurance), where Generali is the leading insurance group by direct premiums in Italy in 2017 and 2018 (Statista, 2019).

Available data on the number of insured houses in Italy, even if discontinuous and limited only to 4 years (Fig. 5), seem to strengthen the considerations that in addition to search for insurance information, part of the users actually look for purchasing the coverage. In Fig. 5, we note that the number of insured houses for earthquakes and floods was very low in 2009, involving only ~0.1% (35,000 units) of the houses in Italy. These data are coherent with an online interest almost zero in the first period analysed in this research. On September 2016, official data refer a quite significant increase of insured houses thus well-fitting the increasing web searches recorded in the second, third and early fourth period under investigation here. Lastly, the number of houses covered by earthquake insurance are about 608,000 in 2018 and 781,000 in 2019, with a clear increase in subscriptions. Again, these data seem well correlated with those recorded in the fourth period (the period "D" in Fig. 1) of investigation in this study, in which the highest search activity was performed (see also Table 1). This correlation between current insurance data and online search is the direction of previous studies that investigated the effect of the communication of earthquake and earthquake risk, particularly by media, showing the relationship between the information diffusion and the decision to purchase an insurance (Marti et al., 2018; Lindell and Perry, 2000).

Conclusions
The analysis discussed above reveals different aspects of the online interest in earthquake insurance, giving information about the potential demand and this becomes crucial in countries with low insurance penetration rate, such as Italy. However, although our research focus on Italy, the encouraging results show that the methodology can be considered in other national contexts to examine the specific causes driving the users' interest so as to fight the under-insurance phenomenon.

Our wDGT data about insurance show that information seeking via cyberspace is related to direct experience of strong earthquakes, foreshocks/aftershocks, vicarious experiences of Italian and overseas earthquakes, and policy discussions/decisions. We found that high-moderate search levels lasted up to 8–10 days, low levels of searching lasted from 3 to 4 months up to 3 years.

The GM data also show that information seeking spread across most of the Italian regions and the geographical spreading after an earthquake starts from the same affected regions (or those bordering them) and then the online activity involves the other regions. Direct and/or vicarious experiences are the main drivers of information seeking.

We also find that regions with low seismic hazard, not involved or only marginally involved in the strong earthquakes, show an interest in earthquake insurance. This finding is in line with the literature that focuses on the effects also of "no-loss" experience that nonetheless increases the salience of earthquake risk (Lin, 2019a). This means that individuals in low hazard zones also look for an insurance coverage, which is a complementary information to the mainstream in the literature about the presence of adverse selection for which only high-risk individuals want to purchase policies so as to limit the diffusion of coverages (Lin, 2019b). Conversely, GT data show that interest in some regions of Southern Italy characterized by a medium-high seismic hazard is absent. Our observations are indirectly validated by insurance data referring that the percentage of houses insured against "natural" disasters is very low in these regions (<2%) and that the percentage of subscriptions in municipalities with highest hazard (zone 1) is the lowest (ANIA, 2017).

A final finding of this study is that information seeking could be linked to insurance purchases. However, a further study with carefully designed surveys is required to test our data-driven hypothesis as the connection between searching for information and real preparedness actions can involve a distinct process (Paton et al., 2005). For example, individual actions appear to be conditioned by some aspects, such as the nature of people's experiences which is difficult to investigate through the aggregate-level of GT data. Combining with other social monitoring data, such as Twitter or Facebook, further studies could be devoted to explore the linkage between information seeking and effective insurance purchase.

Benefits for insurance industry, policymakers and other stakeholders. Literature suggests how earthquakes can supply useful windows of opportunities to be used to involve people in preparedness (e.g., Mattingly, 1987; Jain, 2016; McClure et al., 2016; Tan and Maharjan, 2018).

For the insurance industry, the time windows of opportunities that we found should be used on the right time:

i. to set a system of people's behaviour "surveillance" by using free and easy-to-use GT and GM data, thus evaluating as changes in earthquake(s) occurrence(s), information, communication and advertising strategies influence the users' interest at national as well as regional scale;

ii. to stimulate subscriptions with proper local advertising campaigns in the regions/areas even where damage not occurred, but people are aware of the need to mitigate the risk, following the GT and GM real-time data observations;

iii. to perform advertising campaigns in low–moderate hazards areas to increase portfolio and reduce premiums in areas with high seismic hazard level, thus encouraging subscriptions and cutting down insurance insolvency risk. In these perspectives, the advertising and communication plans could take further advantage by the availability of data concerning the insurance subscriptions in Italy on both regional and monthly basis, at least. These data will allow correlating the interest in earthquake insurance as inferred by Big Data with the effective number of policy subscriptions;

iv. to improve efficacy, people's involvement should be accomplished by mutual agreement between insurance companies and local or national agencies. This mutual agreement would optimize and finalize risk education actions (e.g., "Io non rischio"—I don't take risks national communication campaign—iononrischio.protezionecivile.it/), so also improving the citizens' risk perception, the latter being one of the main factors responsible for under-insurance in Italy as well as in most of countries where the insurance rate is low.

Policymakers should be attentive, support and share the actions indicated above to undertake a policy of mid-term insurance awareness of citizenship, only at the end of which eventually propose legislative initiatives that look at the introduction of a national insurance system. In fact, the numerous attempts to decree a national insurance system have failed in the last 20 years probably because they aimed to oblige people without having first developed adequate information and education of the population on the subject.

Limitations (seen as such and as spur of new research activities). Some limitations rise from the GT data. As mentioned before, GT misses the detailed documentation of how their data are generated, and it allows users to access only to the post-processed data of search activity volumes. Indeed, GT eliminates repeated queries from the same users over a short time period to reduce the count of continues searching (Schootman et al., 2015). Furthermore, limits of GT may prevent the complete use of data for network analysis or cause–effect investigations because we cannot know about individuals' behaviour (users flow, engagement, etc.). However, we can overcome these intrinsic limits of GT with other surveys and social media data. The findings of this study encourage further interdisciplinary research alliance in not only the earthquake research community, but also other disaster research communities to take full advantage of Big Data.

Data availability

All data analysed during the current study are freely down-loadable at the Google Trend web site http://google.com/trends. Furthermore, the data are also available in the POSTECH Hydroclimatology Research Group Data Repository at the following web site https://hydroclimatology.postech.ac.kr/home/data/earthquake_italy_data.

References

Amendola A, Ermoliev Y, Ermolieva T (2000) Earthquake risk management: a case study for an Italian Region. In: Linnerooth-Bayer J, Amendola A (eds) Proceedings of the second European conference on global change and catastrophe risk management, earthquake risks in Europe. IIASA, Laxenburg, pp. 267–295

ANIA (2017) (Associazione Nazionale fra le Imprese Assicuratrici) ANIA trends—Assicurazione incendio e catastrofi naturali delle abitazioni civili. Polizze attive al mese di settembre 2016 (in Italian). https://www.ania.it/pubblicazioni/-/categories/52460?p_r_p_resetCur=true&p_r_p_categoryId=52460. Accessed 2 Jun 2020

ANIA (2019) (Associazione Nazionale fra le Imprese Assicuratrici) ANIA trends—Assicurazione incendio e catastrofi naturali delle abitazioni civili. Polizze

attive al mese di marzo 2019 (in Italian). https://www.ania.it/pubblicazioni/-/categories/52460?p_r_p_resetCur=true&p_r_p_categoryId=52460. Accessed 2 Jun 2020

Bao G, Liao Z, Hine D (2019) Managerial cognition, emergency preparedness and firm's emergency response performance. J Risk Res 22(12):1490–1502. https://doi.org/10.1080/13669877.2018.1485171

Becker JS, Paton D, Johnston DM, Ronan KR (2013) Salient beliefs about earthquake hazards and household preparedness. Risk Anal 33(9):1710–1727. https://doi.org/10.1111/risa.12014

Becker JS, Paton D, Johnston DM, Ronan KR, McClure J (2017) The role of prior experience in informing and motivating earthquake preparedness. Int J Disast Risk Red 22:179–193. https://doi.org/10.1016/j.ijdrr.2017.03.006

Breakwell GM (2007) The psychology of risk. Cambridge University Press, Cambridge

Bronfman NC, Cisternas PC, Repetto PB, Castañeda JV (2019) Natural disaster preparedness in a multi-hazard environment: characterizing the socio-demographic profile of those better (worse) prepared. PLoS ONE 14(4): e0214249. https://doi.org/10.1371/journal.pone.0214249

Carneiro HA, Mylonakis E (2009) Google trends: a web-based tool for real-time surveillance of disease outbreaks. Clin Infect Dis 49(10):1557–1564. https://doi.org/10.1086/630200

Choi H, Varian H (2012) Predicting the present with Google Trends. Econ Rec 88:2–9. https://doi.org/10.1111/j.1475-4932.2012.00809.x

CRED (2015). Database. http://www.emdat.be/database. Accessed 26 May 2020

Dowrick DJ (2003) Earthquake risk reduction actions for New Zealand. Bull NZ Soc Earthqu Eng 36(4):249–259. https://doi.org/10.5459/bnzsee.36.4.249-259

Gonzales P, Ajami N (2017) Social and structural patterns of drought-related water conservation and rebound. Water Resour Res 53(12):10619–10634. https://doi.org/10.1002/2017WR021852

ISTAT (Istituto Nazionale di Statistica) (2018) Mobilità interna e migrazioni internazionali della popolazione residente (in Italian). www.istat.it. Accessed 2 May 2020

Jain SK (2016) Earthquake safety in India: achievements, challenges and opportunities. Bull Earthq Eng 14(5):1337–1436. https://doi.org/10.1007/s10518-016-9870-2

Kagan Y (2006) Why does theoretical physics fail to explain and predict earthquake occurrence? In: Bhattacharyya P, Chakrabarti BK (eds) Modelling critical and catastrophic phenomena in geoscience. Lecture notes in physics, vol 705. Springer, Berlin, Heidelberg, pp. 303–359

Kam J, Stowers K, Kim S (2019) Monitoring of drought awareness from Google trends: a case study of the 2011–17 California drought. Weather Clim Soc 11 (2):419–429. https://doi.org/10.1175/WCAS-D-18-0085.1

Kim S, Shao W, Kam J (2019) Spatiotemporal patterns of US drought awareness. Palgrave Commun 5:107. https://doi.org/10.1057/s41599-019-0317-7

Kunreuther H (2000) Linking insurance and mitigation to manage natural disaster risk. In: Dionne G (eds) Handbook of insurance. Huebner International series on risk, insurance, and economic security, vol 22. Springer, Dordrecht, pp. 593–618

Kunreuther HC, Ginsberg R, Miller L et al. (1978) Disaster Insurance Protection: Public Policy Lessons. John Wiley, New York, NY

Kunreuther H, Roth R, Davis J et al. (eds) (1988) Paying the price: the status and role of insurance against natural disasters in the United States. Joseph Henry Press, Washington

Lee O (1999) Science knowledge, world views, and information sources in social and cultural contexts: making sense after a natural disaster. Am Educ Res J 36 (2):187–219. https://doi.org/10.3102/00028312036002187

Lin X (2019a) Feeling is believing? Evidence from earthquake shaking experience and insurance demand. J Risk Insur 87(2):351–380. https://doi.org/10.1111/jori.12291

Lin X (2019b) Risk awareness and adverse selection in catastrophe insurance: evidence from California's residential earthquake insurance market. SSRN: https://doi.org/10.2139/ssrn.3126862

Lindell MK, Perry RW (2000) Household adjustment to earthquake hazard: a review of research. Env Behav 32(4):461–501. https://doi.org/10.1177/00139160021972621

Marti M, Stauffacher M, Matthes J, Wiemer S (2018) Communicating earthquake preparedness: the influence of induced mood, perceived risk, and gain or loss frames on homeowners' attitudes toward general precautionary measures for earthquakes. Risk anal 38(4):710–723. https://doi.org/10.1111/risa.12875

Mattingly S (1987) Picture windows of opportunity. Earthq Volcanoes (USGS) 19 (5):174–177

McClure J, Henrich L, Johnston D, Doyle EE (2016) Are two earthquakes better than one? How earthquakes in two different regions affect risk judgments and preparation in three locations. Int J Disaster Risk Reduct 16:192–199. https://doi.org/10.1016/j.ijdrr.2016.03.003

Moon J-W, Hwang H, Chung J-B (2019) Factors affecting awareness of preparedness after moderate earthquakes. An analysis of the Pohang earthquake

in Korea. Disaster Prev Manag 29(3):405–420. https://doi.org/10.1108/DPM-07-2019-0209

Mulas MG, Perotti F, Coronelli D, Martinelli L, Paolucci R (2013) The partial collapse of "Casa dello Studente" during L'Aquila 2009 earthquake. Eng Fail Anal 34:566–584. https://doi.org/10.1016/j.engfailanal.2013.02.031

Naoi M, Seko M, Sumita K (2010) Community rating, cross subsidies and underinsurance: why so many households in Japan do not purchase earthquake insurance. J Real Estate Finance Econ 40(4):544–561. https://doi.org/10.1007/s11146-009-9224-9

Nigg J (2000) Predicting earthquakes: science, pseudoscience, and public policy paradox. Prediction: science, decision making, and the future of nature. Island Press, Covelo, pp. 135–156

Nuti SV, Wayda B, Ranasinghe I et al. (2014) The use of Google trends in health care research: a systematic review. PLoS ONE 9(10):e109583. https://doi.org/10.1371/journal.pone.0109583.g001

OECD (2015) Disaster risk financing: a global survey of practices and challenges. OECD Publishing, Paris, 10.1787/9789264234246-en. Accessed 2 May 2020

OECD (2018) Financial management of earthquake risk. www.oecd.org/finance/Financial-Management-of-Earthquake-Risk.htm. Accessed 2 May 2020

Palm RI (2020) Earthquake insurance in California: environmental policy and individual decision-making. Routledge

Palm R, Hodgson M (1992) Earthquake insurance: mandated disclosure and homeowner response in California. Ann Am Assoc Geogr 82(2):207–222. https://doi.org/10.1111/j.1467-8306.1992.tb01905.x

Panda A, Lambert P, Surminski S (2020) Insurance and financial services across developing countries: an empirical study of coverage and demand. Centre for Climate Change Economics and Policy Working Paper 367/Grantham Research Institute on Climate Change and the Environment Working Paper 336. London School of Economics and Political Science, London

Paton D, Smith L, Johnston D (2005) When good intentions turn bad: promoting natural hazard preparedness. Aust J Emerg Manag 20(1):25–30

Pothon A, Gueguen P, Buisine S, Bard PY (2019) California earthquake insurance unpopularity: the issue is the price, not the risk perception. Nat Hazard Earth Syst 19(8):1909–1924. https://doi.org/10.5194/nhess-19-1909-2019

Rogers GO, Sorensen JH (1991) Adoption of emergency planning practices for chemical hazards in the United States. J Hazard Mater 27(1):3–26

Rovida A, Locati M, Camassi R, Lolli B, Gasperini P (2019) Catalogo Parametrico dei Terremoti Italiani (CPTI15), versione 2.0. Istituto Nazionale di Geofisica e Vulcanologia (INGV). https://doi.org/10.13127/CPTI/CPTI15.2. Accessed 20 May 2020

Schootman M, Toor A, Cavazos-Rehg P et al. (2015) The utility of Google Trends data to examine interest in cancer screening. BMJ Open 5(6):e006678. https://doi.org/10.1136/bmjopen-2014-006678

Seko M (2019) Perceived preparedness and attitude of Japanese households toward risk mitigation activities following the great East Japan earthquake: earthquake insurance purchase and seismic retrofitting. In: Housing markets and household behavior in Japan. Advances in Japanese business and economics, vol 19. Springer

Slovic P (1987) Perception of risk. Science 236:280–285

Spittal MJ, McClure J, Siegert RJ, Walkey FH (2008) Predictors of two types of earthquake preparation: survival activities and mitigation activities. Environ Behav 40(6):798–817. https://doi.org/10.1177/0013916507309864

Statista (2019) Non-life insurance market in Italy. https://www.statista.com/

Tan YJ, Maharjan R (2018) What googling trends tell us about public interest in earthquakes. Seismol Res Lett 89(2A):653–657. https://doi.org/10.1785/0220170116

Tosi P, De Rubeis V, Sbarra P, Sorrentino D (2007) Hai Sentito Il Terremoto (HSIT). https://doi.org/10.13127/HSIT. Accessed 28 May 2020

Wang M, Liao C, Yang S et al. (2012) Are people willing to buy natural disaster insurance in China? Risk awareness, insurance acceptance, and willingness to pay. Risk Anal 32(10):1717–1740. https://doi.org/10.1111/j.1539-6924.2012.01797.x

Zhu D, Xie X, Xie J (2012) When do people feel more risk? The effect of ambiguity tolerance and message source on purchasing intention of earthquake insurance. J Risk Res 15(8):951–965. https://doi.org/10.1080/13669877.2012.686051

Additional information

Supplementary information is available for this paper at https://doi.org/10.1057/s41599-020-0532-2.

Correspondence and requests for materials should be addressed to F.T.G.

Reprints and permission information is available at http://www.nature.com/reprints

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.