View metadata, citation and similar papers at core.ac.uk

brought to you by CORE

provided by ScholarlyCommons@Penn

University of Pennsylvania
ScholarlyCommons

Population Center Working Papers (PSC/PARC) Population Studies Center

1-12-2021

Harnessing the Potential of Online Searches for Understanding the Impact of COVID-19 on Intimate Partner Violence in Italy

Selin Köksal

Luca Maria Pesando

Valentina Rotondi

Ebru Şanlıtürk

Follow this and additional works at: https://repository.upenn.edu/psc_publications

Part of the Criminology Commons, Demography, Population, and Ecology Commons, Domestic and Intimate Partner Violence Commons, Family, Life Course, and Society Commons, Gender and Sexuality Commons, and the Social Control, Law, Crime, and Deviance Commons

This paper is posted at ScholarlyCommons. https://repository.upenn.edu/psc_publications/61
For more information, please contact repository@pobox.upenn.edu.

Harnessing the Potential of Online Searches for Understanding the Impact of COVID-19 on Intimate Partner Violence in Italy

Abstract

Despite the volume of studies leveraging big data to explore socio-demographic phenomena, we still know little about the intersection of digital information and the social problem of intimate partner violence (IPV). This is an important knowledge gap, as IPV remains a pressing public-health concern worldwide, with 35% of women having experienced it over their lifetime and cases rising dramatically in the wake of global crises such as the current COVID-19 pandemic. This study addresses the question of whether online data from Google Trends might help to reach “hard-to-reach” populations such as victims of IPV using Italy as a case-study. We ask the following questions: Can digital traces help predict instances of IPV — both potential threat and actual violent cases — in Italy? Is their predictive power weaker or stronger in the aftermath of crises such as COVID-19? Our results combined suggest that online Google searches using selected keywords measuring different aspects of IPV are a powerful tool to track potential threats of IPV before and after global-level crises such as the current COVID-19 pandemic — with stronger predictive power post-crisis — while online searches help to predict actual violence only in post-crises scenarios.

Keywords

digital data, Google Trends, intimate partner violence, Italy, COVID-19

Disciplines

Criminology | Demography, Population, and Ecology | Domestic and Intimate Partner Violence | Family, Life Course, and Society | Gender and Sexuality | Social and Behavioral Sciences | Social Control, Law, Crime, and Deviance | Sociology

Harnessing the Potential of Online Searches for Understanding the Impact of COVID-19 on Intimate Partner Violence in Italy

Selin Köksal[1], Luca Maria Pesando[2], Valentina Rotondi[3], and Ebru Şanlıtürk[1]

[1] *Bocconi University*
[2] *McGill University*
[3] *University of Applied Sciences and Arts of Southern Switzerland & University of Oxford*

Abstract

Despite the volume of studies leveraging big data to explore socio-demographic phenomena, we still know little about the intersection of digital information and the social problem of intimate partner violence (IPV). This is an important knowledge gap, as IPV remains a pressing public-health concern worldwide, with 35% of women having experienced it over their lifetime and cases rising dramatically in the wake of global crises such as the current COVID-19 pandemic. This study addresses the question of whether online data from Google Trends might help to reach "hard-to-reach" populations such as victims of IPV using Italy as a case-study. We ask the following questions: Can digital traces help predict instances of IPV — both potential threat and actual violent cases — in Italy? Is their predictive power weaker or stronger in the aftermath of crises such as COVID-19? Our results combined suggest that online Google searches using selected keywords measuring different aspects of IPV are a powerful tool to track potential threats of IPV before and after global-level crises such as the current COVID-19 pandemic — with stronger predictive power post-crisis — while online searches help to predict actual violence *only* in post-crises scenarios.

Keywords: Digital data, Google Trends, Intimate Partner Violence, Italy, COVID-19.

Introduction

Social media and online platforms including Facebook, Twitter, Instagram, and Google offer new and unprecedented means of communicating, networking, looking for information, and building communities for individuals all over the world. Online platforms are mechanisms by which millions of people search, spread, share, and exchange valuable information in domains ranging from work and occupation to personal relationships, health, illness, nutrition, sports, politics, etc. As such, bits of information obtained through online platforms – also called "digital traces" – have increasingly become valuable data sources to address some of the most pressing local and global social phenomena that we are confronted with every day (Lazer et al., 2020). For instance, digital-trace data have been used to study socio-demographic phenomena such as fertility (Billari et al., 2016; Billari and Zagheni, 2017; Rampazzo et al., 2018), migration (Alexander et al., 2020; Zagheni and Weber, 2012; Zagheni et al., 2017), health and mortality (Delpierre and Kelly-Irving, 2018; Öhman and Watson, 2019), gender dynamics (Fatehkia et al., 2018; Kashyap et al., 2020), family instability (Compton, 2019), and trust towards science and experts during the recent COVID-19 pandemic (Battiston et al., 2020), providing great opportunities yet also raising statistical, computational, and ethical challenges (Cesare et al., 2018).

Despite this proliferation of studies, we still know very little about the intersection of digital information and the social problem of intimate partner violence (IPV), defined as violence perpetrated against women by their partners within and outside of marriage. This is an important knowledge gap because IPV remains a pressing public-health concern worldwide, with about 35% of women having experienced IPV over their lifetime (WHO, 2013) and cases rising dramatically in the wake of global crises such as the Great Recession (Schneider et al., 2016) or the current COVID-19 pandemic (Lindberg et al., 2020; Peterman et al., 2020; Abel and McQueen, 2020; WHO, 2020; Arenas-Arroyo et al., 2021). As a matter of fact, during a situation of lockdown women's ability to escape abusive situations within their houses and their ability to reach their support networks are reduced. At the same time, confinement measures might increase alcohol consumption and other substances, while the increased economic uncertainty due to the global health crisis might also increase emotional stress, all elements which are associated with the perpetration of IPV (Storey, 2020; Card and Dahl, 2011; Aizer, 2010, 2011).

As tracking instances of IPV is challenging – and particularly so in times of crises during which reporting tends to be even lower – this study addresses the question of whether big data might help to reach this "traditionally difficult-to-reach" population (Xue, Macropol, Jia, Zhu and Gelles, 2019) using Italy as a case study. One of the key advantages of using digital traces – which are generated from the use of digital technologies, rather than based on reporting – is precisely that they can help address issues where reporting or social desirability biases that preclude systematic reporting are prevalent – such as IPV or abortion (Reis and Brownstein, 2010). If online information turns out to be a strong predictor of actual instances of violence, this would suggest that big data might be a key — and, to date, underappreciated — resource for tracking or even anticipating IPV and getting a "fair" picture of the brunt of domestic violence that women bear daily, allowing for better temporal resolution and spatial granularity than, for instance, household surveys.

To summarize, the current study addresses the following two research questions and tests the related hypotheses: [RQ1] Can digital traces from online sources such as Google Trends help track/predict instances of IPV in Italy? Evidence on the role of big data in other domains of social life suggests that they might [*HP1*]; [RQ2] Provided digital data can be of help, is their predictive power weaker, stronger, or unaltered in the wake of key macro-level discontinuities such as the current COVID-19 pandemic? As instances of IPV increase in times of crises, we hypothesize a stronger predictive power of digital data in the aftermath of crises. While actual reporting is lower, online connectivity during crises tends to be high, either due to rising unemployment, forced lockdowns at home, or both [*HP2*].

Findings from this paper: (i) inform whether digital information can be used to track/predict IPV for hard-to-reach populations in Italy; (ii) shed light on where, geographically, quick policy interventions might be needed the most and in what specific domain; (iii) reveal whether big data might provide "real time" bits that help target more immediate policy interventions in situations in which IPV cases cannot be reported easily/quickly, such as the current lockdowns; (iv) inform, ultimately, whether big data may allow to devise a sort of tracking system that serves as a precursor or signal for anticipating increases in IPV. Very importantly, as Internet penetration expands and digital divides by gender narrow, findings from this study will also have broad applicability to low- and middle-income countries (LMICs) in the years to come.

Background

At the onset of the pandemic, several media outlets pointed out the upsurge of IPV cases both in Italy[1] and in the rest of the world.[2] Moreover, the World Health Organization (WHO) and UN Women underlined the risk of IPV being intensified during lockdown periods as security, health, and economic concerns became more pronounced (Arenas-Arroyo et al., 2021). Findings from recent studies regarding the impact of lockdown measures on the incidence of IPV are aligned with these concerns, which are pervasive across countries (Every-Palmer et al., 2020). Research from Italy exploring the effectiveness of the media campaign "*Libera puoi*"[3] provides evidence of the immediate increase in number of calls during the first weeks of the lockdown, which remained at high levels until May 2020 (Colagrossi et al., 2020). The study also documents the abrupt rise in Google search volumes of the keyword *1522* — the main anti-violence helpline in Italy — which occurred right after the launch of the campaign.

Evidence from the city of Dallas in Texas (US) suggests a short-term spike in domestic violence incident reports within the first two weeks of the lockdown period (Piquero et al., 2020). Further to that, another study conducted in 14 large metropolitan areas in the US found a 7.5% increase in IPV calls between March and May 2020, with the effect being more concentrated in the first five weeks of social-distancing measures (Leslie and Wilson, 2020). Interestingly, the rise in IPV cases was mainly driven by households without any previous violence history. A similar case applies to the city of Chicago, where increased time spent at home in challenging situations led to an overall decrease in total calls for police service, yet an increase in domestic violence-related police calls, especially among couples with no previous violence history and among married couples with children (Bullinger et al., 2020). Despite the increased number of violence-related calls for police service, a striking finding from the same study suggests that reported domestic-related crimes and arrests by police officers fell by 8.7% and 26.3% over the same period, respectively. This result — which led the authors to conclude that between March and April 2020 nearly 1,000 cases of domestic violence crimes went underreported — potentially stems from underfilling of official incident reports by police officers, an issue which further exacerbates the well-known plague of IPV underreporting. Overall, using police dispatch and crime data from 36 police and sheriff's departments

[1] https://www.corriere.it/cronache/20_aprile_13/coronavirus-donne-vittime-violenza-ora-chiedono-aiuto-via-mail-30b8152a-7c97-11ea-9e96-ac81f1df708a.shtml, https://www.istat.it/it/violenza-sulle-donne/speciale-covid-19

[2] https://www.nytimes.com/2020/04/06/world/coronavirus-domestic-violence.html,https://edition.cnn.com/2020/03/27/health/domestic-violence-coronavirus-wellness-trnd/index.html

[3] http://www.governo.it/it/media/campagna-di-comunicazione-libera-puoi/14459

and mobile tracking data, Hsu and Henke (2020) estimate that staying at home due to COVID-19 increased domestic violence in the US by over 5% from March 13 to May 24, 2020. Research from Peru and Argentina also provides evidence of the aggravated risk of domestic violence during lockdown periods and the rise in calls to helplines throughout Latin America (Aguero, 2021; Perez-Vincent et al., 2020). To the best of our knowledge, the impact of confinement measures on the risk of IPV remains relatively understudied in the European context. Italy, our country of interest, provides a suitable case study to address this research question by being among the earliest and worst-hit countries by the first wave of the COVID-19 pandemic.

Online data, most specifically social media data, provide a relatively recent — and, to date, underappreciated — source to analyze and interpret human behavior, as well as to now-cast and forecast individual and societal-level outcomes that have great relevance for policy design and intervention (Alexander et al., 2020). Epidemiology was one of the first disciplines to promote the use of big data for research purposes, by analyzing online search data to now-cast and forecast outbreaks such as influenza (Ginsberg et al., 2009), chicken pox (Pelat et al., 2009), and salmonella (Brownstein et al., 2009). The use of big data in epidemiological research came to be referred as *infodemiology* or *infoveillance* (Eysenbach, 2009) and, as IPV/gender-based violence is as a public health issue, IPV received the attention of infodemiology studies. In the case of Brazil, online search data on *feminicide* are found to be positively associated with female homicide rates but not with the introduction of feminicide-related laws, suggesting that in this context information-seeking behavior relies more on the experience from actual cases than policy campaigns (Martins-Filho et al., 2018). Social media data emerge as an especially useful source as forums, groups, and social networks allow users to share their experiences and establish emotional support among victims of IPV. For instance, Twitter has increasingly been used as a medium in IPV research based on big data, employing various computational methods, using tweets including IPV-related keywords or hashtags as units of analysis. Studies show evidence that there is an active Twitter community on violence against women, which tends to engage in conversations (Xue, Macropol, Jia, Zhu and Gelles, 2019); this community also highlights oft-neglected forms of violence such as reproductive coercion (McCauley et al., 2018) and provides important information on awareness campaigns, as well as a support platform (Purohit et al., 2015; Xue, Chen and Gelles, 2019). IPV studies using data from Pinterest (Carlyle et al., 2018) and Instagram (Carlyle et al., 2019), with predominantly female and young-adult users, respectively, corroborate the idea that social media platforms involve an experience-based narrative on different forms of violence and thus provide a valuable tool for policy makers and advocacy groups.

To the best of our knowledge, only few studies (e.g., Martins-Filho et al., 2018) to date have leveraged data on online searches from Google Trends to track dynamics of IPV — a contribution we intend to provide with the current study. Digital traces from Google Trends provide a useful source of information as, being the most commonly used search engine, Google is more widely used than Twitter, Pinterest or Instagram, thus providing an arguably less biased picture of socio-demographic phenomena under investigation. As a matter of example, as of December 2020 Google was the most popular search engine in Italy, with a 95.7% share of the search engine market compared to the 2.9% of Bing and the 0.81% of Yahoo!.[4] As of June 2020, social media penetration in Italy stood at 58%, with the most popular social network remaining Facebook, with 36.9 million users, followed by Instagram (27.7 million users), Linkedin (18.6 million users), and Pinterest (16.7 million users). As for the same period, Twitter counted only 10 million users, TikTok 6.6 million users and Reddit 2.8 million users.[5] Also, Google Trends provide a flexible tool to select and investigate a wide array of potential keywords measuring different facets of IPV.

The First Wave of the COVID-19 Outbreak in Italy

On Friday the 21st of February 2020, the first case of COVID-19 was diagnosed in a man living in the town of Codogno in the province of Lodi, a city located in the Northern region of Lombardy in Italy. The virus then spread across neighbouring regions in Northern Italy – including Veneto, Emilia Romagna, and Piedmont, all of which began to report rapid increases in cases. Two days later, on February 23, the government issued a decree which prohibited the movement of people outside 10 municipalities located in Lombardy and a municipality in Veneto. From March 8, restrictions to avoid any movement were extended to the whole of Lombardy and to other fourteen provinces in Northern Italy. On March 10 — our cutoff point for the definition of lockdown in the current study — a new decree issued by Prime Minister Giuseppe Conte extended these lockdown measures across the entire country of Italy until May 4, when the country started a mild reopening which was completed by mid-June 2020. The spread of the virus across the country – at least for the first wave of the COVID-19 pandemic – was uneven with the majority of cases (and deaths) being concentrated in Lombardy which, as of January 2021, counts more than 30% excess deaths (hence our focus on Lombardy in the latter part of the analysis). At the beginning of January 2021, Italy is the eighth country in the world and the fourth in Europe for total number of cases, and the fifth country in the world and the first in Europe for total number of deaths.

[4] https://gs.statcounter.com/search-engine-market-share/all/italy

[5] https://www.agcom.it/documents/10179/20440899/Documento+generico+16-10-2020/dcdbcb3a-720c-4878-9f10-dcd8912ba984?version=1.0

Data and Methods

This study combines data from various sources and tests the predictive power of online data in the Italian context by relating digital traces with information on actual IPV calls to official helplines. In a simplified framework, digital data provide information on our *predictors* of interest, while data on actual IPV cases provide information on our *outcomes* of study.

Starting from *predictors*, online data are obtained from Google Trends. Google Trends is an increasingly popular source in infodemiology studies and provides normalized data on the frequency of Google searches for a given time period, query, and location. Data obtained from Google Trends do not reflect the actual number or volume, but rather the frequency of online searches on Google on a scale from 0 to 100. Thus, zero does not necessarily mean a complete lack of Google searches; rather, it means that the frequency of the searches for a given parameter does not meet the minimum threshold set by Google. Google Trends provides search frequency data on a *daily* basis if the requested time range is 90 days or less, on a *weekly* basis if the time range is between 90 days and 5 years, and on a *monthly* basis if the requested time period is longer than 5 years.

Google Trends data are obtained through the R package *gtrendsR* (Massicotte and Eddelbuettel, 2020).[6] We selected nine main search queries which are: *1522* (the domestic violence helpline number in Italy), *abuse* (abuso), *home & abuse* (casa & abuso), *home & rape* (casa & stupro),[7] *feminicide* (femminicidio), *rape* (stupro), *domestic violence* (violenza domestica), *gender-based violence* (violenza di genere), and *sexual violence* (violenza sessuale). Note that we will use the English label in all figures and tables that follow. We obtained three different data sets from Google Trends, for these queries, for different time periods and locations, such that the unit of time matches the one pertaining to the official records. The first data set consists of daily data for Italy as a whole for the period between March 1 - June 30 for five years, from 2016 to 2020. Second, we created a data set composed of monthly data for the period March 1, 2013 to June 30, 2020, for all regions of Italy. We calculated four-months averages of Google search inquiries for each keyword and for each region. This step was performed in order to make the data compatible with the yearly-aggregated number of helpline calls at the regional level. Lastly, the third data set consists of daily Google Trends data, for the period between January 1, 2018 up to May 31, 2020 only for the region of Lombardy, Italy. As Google Trends only allows for daily data to be obtained for a period of three months (90 days), we repeatedly obtained daily Google Trends data for each three-month period to achieve

[6]Our code will be made publicly available.

[7]These refer to search queries that include both words, in no particular order.

the time frame required. Once the consecutive periods of three months were pieced together, we addressed the issue of each part of data being normalized within itself. Thus, we downloaded weekly data for the given time frames (at once to ensure that the normalization is within the necessary time frame) and matched the weekly and daily data, where a value is present for the same date in both data sets. Based on the matching dates, we calculated a weekly adjustment factor (Risteski and Davcev, 2014). By multiplying the daily data with the relevant weekly adjustment factor, we ensured the consistency of the data set of daily data and then normalized it for a range of 0-100 data.

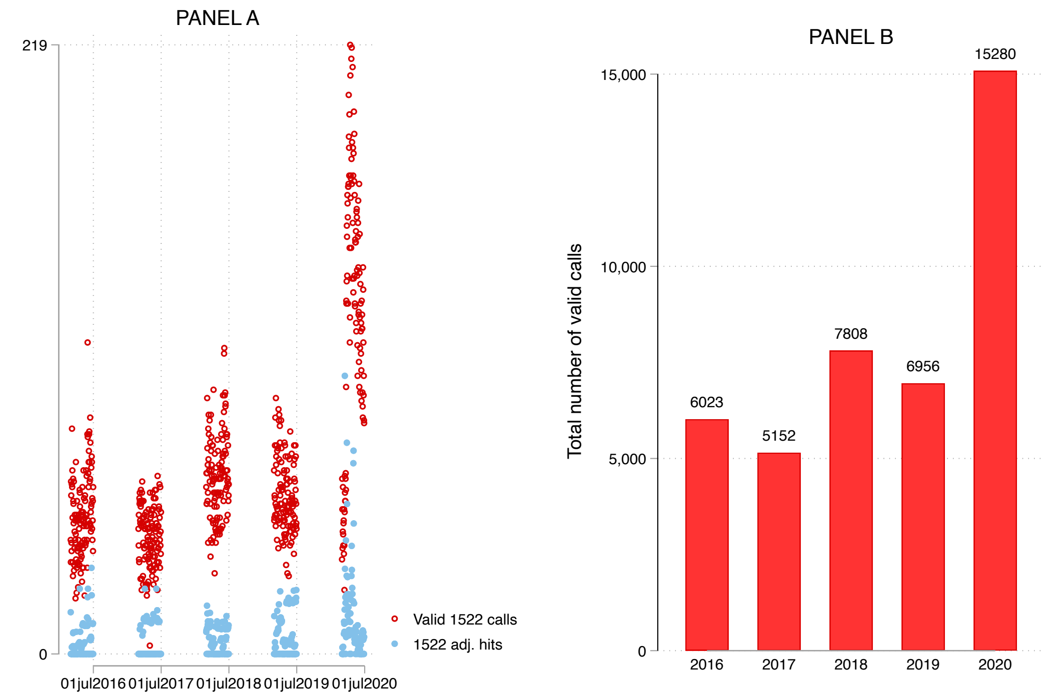
Moving to *outcomes*, we rely on three sources of data. First, we obtained the daily number of calls (valid calls) to the 1522 anti-violence helpline (*1522 - Numero anti violenza e stalking*) from the Equal Opportunity Department (Presidency of the Italian Council of Ministers) for Italy as a whole daily from March 1, to June 30, from 2016 to 2020. Panel A of Figure 1 plots the daily number of valid 1522 calls and daily number of 1522 Google hits between March and July from 2016 to 2020. As shown in the Figure, during the period between March and July 2020, the number of daily valid calls to the Italian helpline number increased considerably compared to the same period of the previous years. Moreover, and as shown also in Colagrossi et al. (2020), Google search volumes for the keyword 1522— the main anti-violence helpline in Italy — rose considerably right after the Italian national lockdown and the launch of the campaign "*Libera puoi*". As a matter of comparison, during the first wave of the COVID-19 outbreak in Italy (between the 1st of March and the 30th of June 2020) there were 15,280 valid calls to 1522, +119.7% over the same period in 2019. Of these calls, 32% came from victims of violence or stalking seeking for help, 24% came from people seeking information about the helpline 1522, 6% came from people reporting violence. The remaining were related to general information seeking (37%)[8] and emergency (1%).

The second source of data is the number of monthly anti-violence 1522 calls, collected for the period March and June between years 2013-2020 and aggregated at the regional level. These data are publicly available at The Italian National Institute of Statistics (ISTAT, henceforth) website.[9] Figure 2 visualizes the population-adjusted number of 1522 calls for the period of March-June and for years 2013 (earliest) and 2020 (latest). While the highest call rate is between 0.20 and 0.25 per 1,000 people in 2013, it climbs up to 0.30-0.35 per 1,000 in 2020 during the lockdown period. In particular, we observe a notable increase in helpline call rates in those regions that were severely impacted by the first wave of the pandemic such as Lombardy, Piedmont and Emilia-Romagna. Lazio is the region exhibiting the highest rate of

[8]Such as search for legal information, out-of-target calls related to requests for other useful phone numbers, information about national shelters for victims of violence, and other reasons.

[9]Available at: https://www.istat.it/en/archivio/246618.

Figure 1: Daily number of valid 1522 calls and daily number of 1522 Google hits (Panel A) and total number of 1522 calls (Panel B) by year (over the period 1st of March-30th of June)

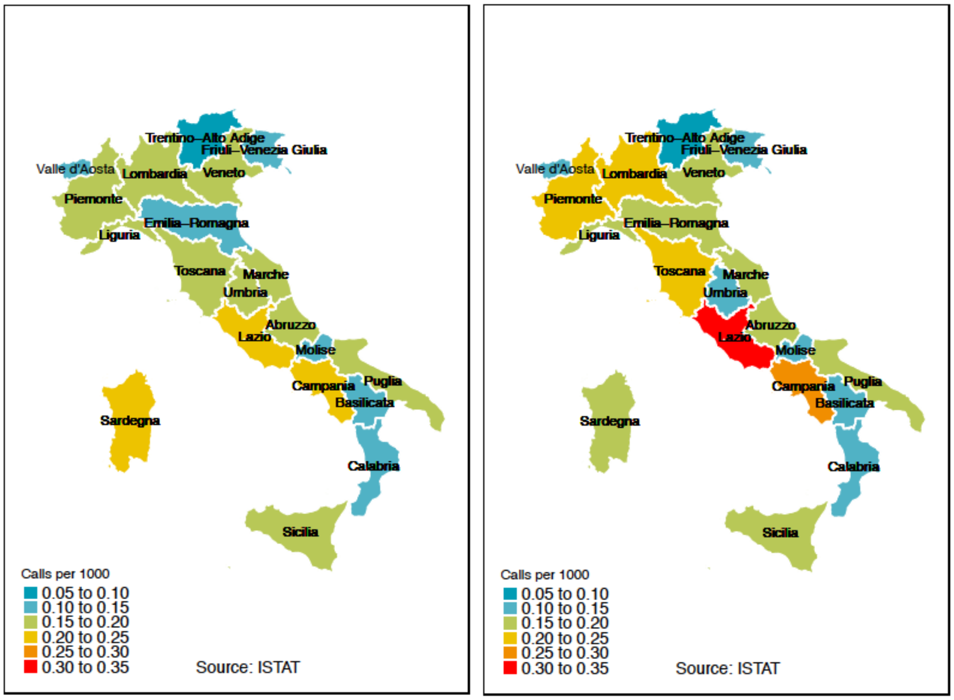


increase in helpline calls between 2013 and 2020.

The third one are data from AREU (Azienda Regionale Emergenza Urgenza – Regional Agency for Emergency Urgency), which provides data on daily calls to the AREU emergency number in Lombardy (112) between January 1, 2018 and May 30, 2020, alongside the reason behind the emergency call. This additional variable (reason) helps us identify calls that were received by AREU for reasons that can be traced back to accident or violence-related purposes.[10] Note that, as the first source of data, data from AREU are daily, yet they pertain to Lombardy only — the first source provides daily information for Italy as a whole with no regional identifier, hence no possibility to conduct region-specific analyses — and they record a different outcome (all emergency calls versus 1522 calls). The second source of data tracks the same outcome — 1522 calls — yet data for Lombardy are yearly, rather than daily. Overall, data from AREU add value to the analysis for two main reasons. First, AREU data pertain specifically to Lombardy, the region hardest-hit by the COVID-

[10]Albeit not perfect, we imposed the following restrictions to identify violence-related calls: (i) we kept only "home" as the location from which the call was made; (ii) we kept only women as the sex of the caller; (iii) we restricted the age range to 10-85; (iv) we only kept violence-related motives (e.g., we excluded respiratory motives, among others).

Figure 2: 1522 helpline calls by regions (per 1,000 people, 2013-2020)



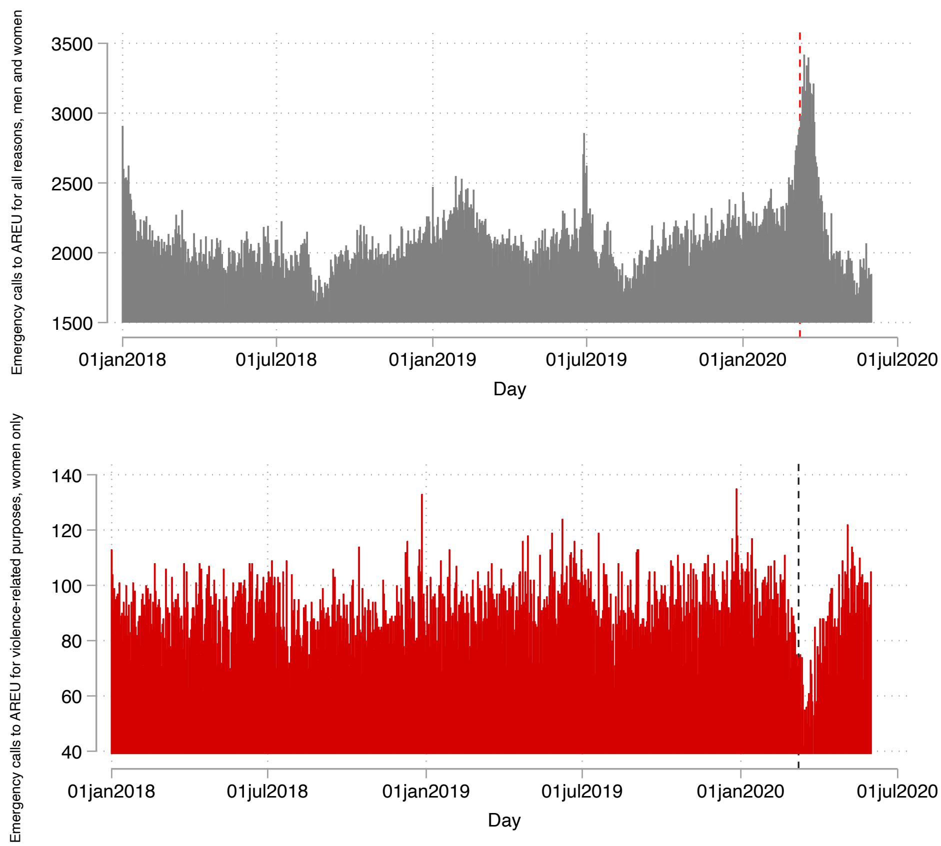
(a) 1522 helpline calls by regions in March-June 2013 (per 1,000)

(b) 1522 helpline calls by regions in March-June 2013 (per 1,000)

19 pandemic. Second, from a theoretical standpoint, emergency calls — i.e., calls to request an ambulance, mostly — measure actual or "realized violence," while 1522 calls also measure "potential threat" or potential risk of experiencing IPV, as shown above. This is confirmed by Figure 3 (bottom panel), which shows a marked drop in calls to AREU made by women for accident or violence-related purposes in the immediate post-lockdown period — a piece of evidence which stands in contrast with trends for 1522 calls shown in Figure 1 and discussed in Colagrossi et al. (2020), and with trends in calls to AREU made by both men and women for all reasons combined (top panel). One hypothesis is that in the wake of strict confinement measures the threat of violence might increase importantly, hence women resort to the main IPV helpline (1522) to seek help and information, rather than requesting an ambulance (AREU), which would rather occur in the presence of actual violence. Alternatively, it may be the case that women calling do not provide the exact reason underlying the emergency (e.g., due to stigma), thus leading to mis-recording or mis-reporting of information on reasons behind the calls.

We first estimate the model reported in eq. (1) using Ordinary Least Squares (OLS) regression with standard errors robust to heteroskedasticity:

Figure 3: Daily number of calls to AREU from men and women for all reasons combined (top panel) and from women for accident or violence-related purposes (bottom panel)



$$Y_t = GoogleSearch_t + post + \lambda_y + \varepsilon_t \tag{1}$$

where Y_t indicates the number of valid calls received by 1522 in day t, $GoogleSearch_t$ represents the frequency of Google inquiries for the aforementioned keywords in day t, post is a dummy variable assuming value 1 after March the 10th, when Italy enforced the lockdown. Year fixed effects (λ_y) are included to allow for heterogeneity across different years. Google searches are either contemporaneous to the outcome or lagged (one week) in order to account for the potential time lapse between Google search and call to helplines. As a robustness check we also compute the average Google hits for each selected keyword for the week and use it instead of daily searches.

As far as the regional-level analysis is concerned, to mitigate the potential endogeneity due to several confounding factors which are correlated with both Google searches and number of 1522 calls, we draw on a set of regional controls such as educational attainment, unemployment rate, and Gross Domestic Product (GDP) per capita. These data are also obtained from ISTAT. As of December 2020, educational attainment and annual unemployment data at the regional level are available until the year 2019. Therefore, we replaced data on these two controls for 2020 with data from 2019 as closest proxy. Moreover, we replaced regional GDP per capita of the years 2019 and 2020 with the one of 2018 since the data are present only up to 2018.[11] We therefore estimate the following model using Ordinary Least Squares (OLS) regression with clustered standard errors at the regional level:

$$Y_{rt} = GoogleSearch_{rt} + X_{rt} + \varepsilon_{rt} \tag{2}$$

where Y_{rt} indicates violence outcome (1522 or emergency number calls) for region r and for year t, $GoogleSearch_{rt}$ represents the frequency of Google inquiries for the aforementioned keywords in region r and averaged for year t [12] , X_{rt} are the set of regional control variables and ε_{rt} is the error term. Number of calls and control variables are adjusted to the regional population. We also include weekly lags in order to account for the potential time lapse between Google search and call to helplines, as above. Lastly, analyses with AREU data are conducted at the level of Lombardy — rather than Italy — following eq. (1), i.e., including a weekly lag and accounting for year fixed-effects and a dummy for post-lockdown period.

Table 1 summarizes our data sources, alongside the spatio-temporal coverage and the empirical strategy.

Table 1: Data sources, coverage, and empirical specifications

Data source	Geographical Unit	Time Unit	Controls	Model
Italian Equal Opportunity Department, valid 1522 calls	Italy	Daily	Post lockdown (dummy), year fixed-effects	OLS. SE robust to heteroskedasticity
ISTAT, number of calls to 1522	Regions within Italy	Yearly	Educational attainment, Unemployment rate, GDP	OLS. SE clustered at the regional level
AREU, number of calls	Only Lombardy	Daily	Post lockdown (dummy), year fixed-effects	OLS. SE robust to heteroskedasticity

Results

Italy

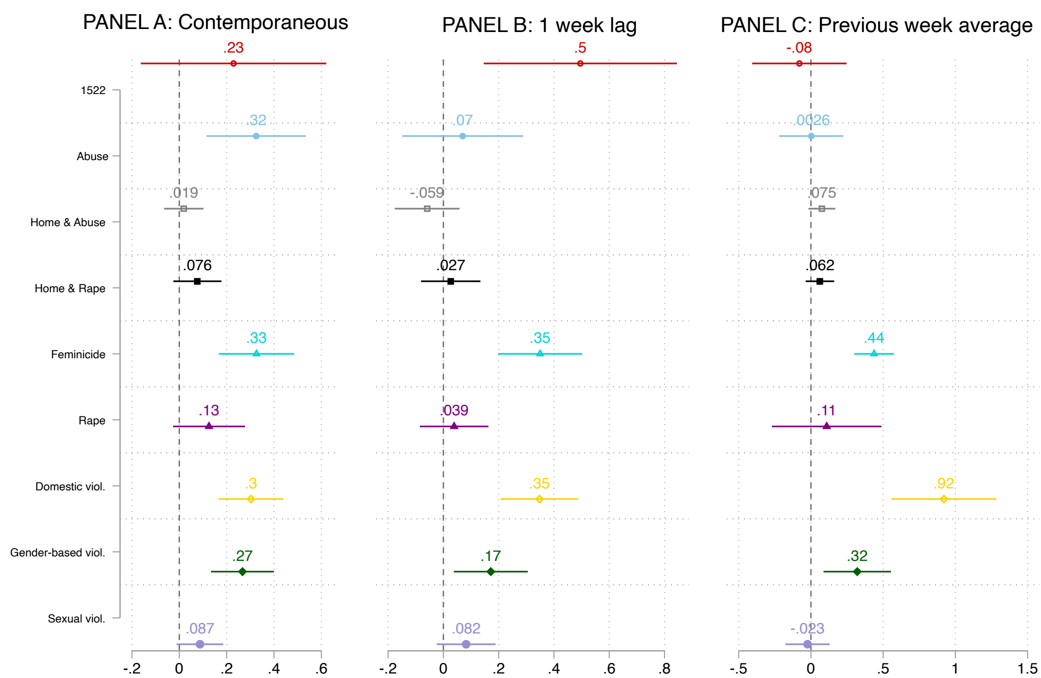
Figure 4 reports the coefficient plot from regressions of daily 1522 valid calls, for Italy as a whole, on our selected keywords (one for each row). In Panel A (and in Table A1, Appendix) Google hits are contemporaneous to our outcome of interest, i.e.,

[11] We intend to update our regional-level controls as the data become available at ISTAT.

[12] Due to lack of adequate number of observations, inquiries for *home & abuse* and *home & rape* are excluded from regional analysis.

daily 1522 valid calls, while in Panel B (Table A2, Appendix) Google hits are lagged by one week. In Panel C the explanatory variables are computed as the average of the hits for the previous week (Table A3, Appendix). Google hits, i.e., the frequency of queries, for keywords *feminicide*, *domestic violence*, and *gender-based violence* are consistently positively and significantly correlated with helpline calls all across the three different models. *Abuse* is positively and significantly associated with daily calls only in Panel A, while searches for *1522* are associated with actual 1522 calls only in Panel B. Table A4 in the Appendix shows that for *domestic violence* the association between searches and calls is significantly higher in the post-lockdown period compared to the pre-lockdown period, yet no differential associations are observed for the remaining keywords.

Figure 4: Coefficient plot from regressions of daily 1522 valid calls on Google searches, by selected keywords (whole Italy).



Notes: Data on the outcome from the Equal Opportunity Department, Presidency of Italian Council. In Panel A the explanatory variables are contemporaneous the the outcome; in Panel B they are lagged by one week; in Panel C they are computed as the average over the previous week.

Regions

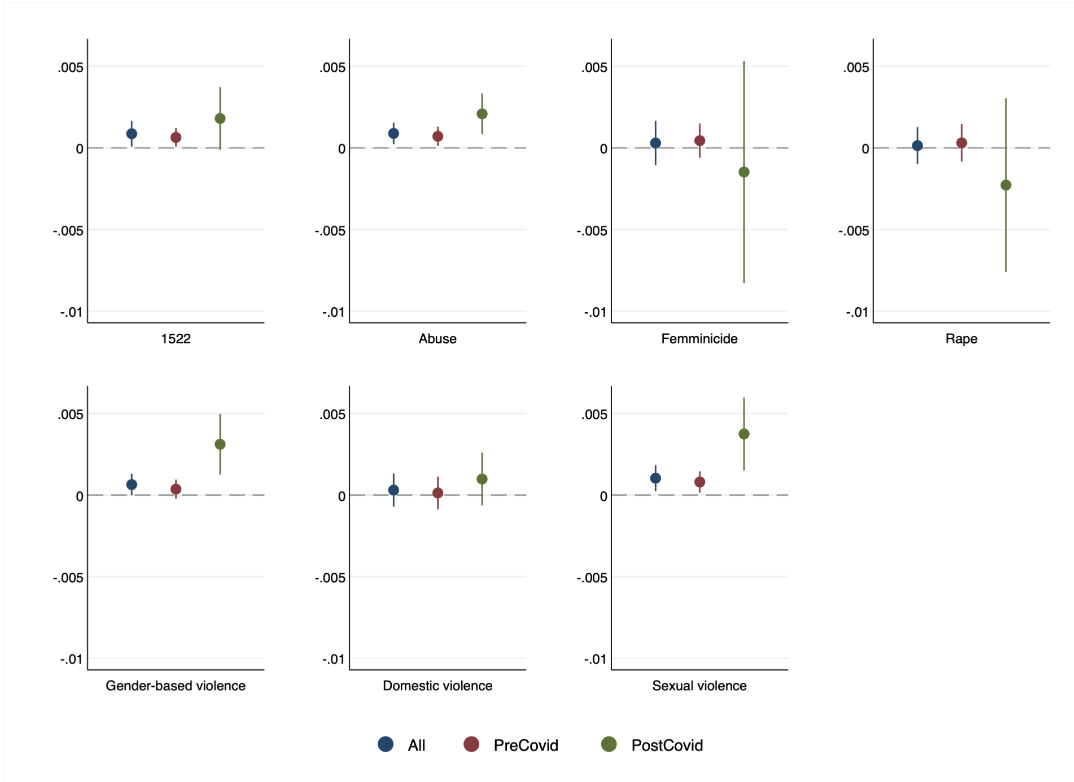
Figure 5 presents results obtained from regional-level data on yearly-aggregated number of calls.[13] Search frequencies for keywords *1522*, *abuse*, *gender-based violence* and *sexual violence* are positively and significantly associated with the number of anti-violence helpline calls, while results for *domestic violence* suggest positive yet weakly significant associations. On the other hand, searches for keywords *feminicide* and *rape* appear to be insignificant in predicting the number of 1522 calls for the March-June between 2013 and 2020 (full results in Appendix Table A5). Focusing specifically on post-COVID findings (green marker), the relationship between number of helpline calls and Google search frequencies on keywords *1522*, *abuse*, *gender-based violence* and *sexual violence* appears to be even stronger — in terms of both magnitude and statistical significance — during the lockdown period (full results in Appendix Table A6). Together, these findings signal the relevance of Google search engines for women seeking help during confinement periods in which traditional help mechanisms become harder to reach.

AREU calls in Lombardy

We last test the relationship between online searches and daily calls to the emergency number in Lombardy using daily data from AREU. Note, once again, that emergency calls to AREU differ from 1522 calls as the former are aimed at requesting an ambulance, thus measuring actual violent cases that require immediate help and assistance. Consistently with such discrepancy, our results with AREU data — reported in Figure 6 — are quite different from the above. While the sign of the estimated coefficients is for the most part positive, only searches for the keyword *feminicide* positively and significantly predict emergency calls to AREU for the whole period considered (blue marker) — full results reported in Appendix Table A7. However, the evidence changes drastically when restricting the focus on the post-lockdown period (green marker). As a matter of fact, the estimated coefficient on online searches gets two to six times bigger in magnitude for all keywords except for *rape*, and the coefficient becomes statistically significant for the keywords *1522, abuse, domestic violence* and *sexual violence*. Note that three of the four keywords — namely *1522*, *abuse*, and *sexual violence* — are the same ones that become more strongly significant when predicting 1522 calls in the post-lockdown period in Figure 5 — full results reported in Appendix Table A8. These results imply that the tendency to seek IPV-related help online and reported IPV emergencies are

[13] Analyses could only be conducted for seven out of the nine keywords, as too few queries could be produced for the keywords *home & abuse* and *home & rape* for the selected geographical unit and time frame.

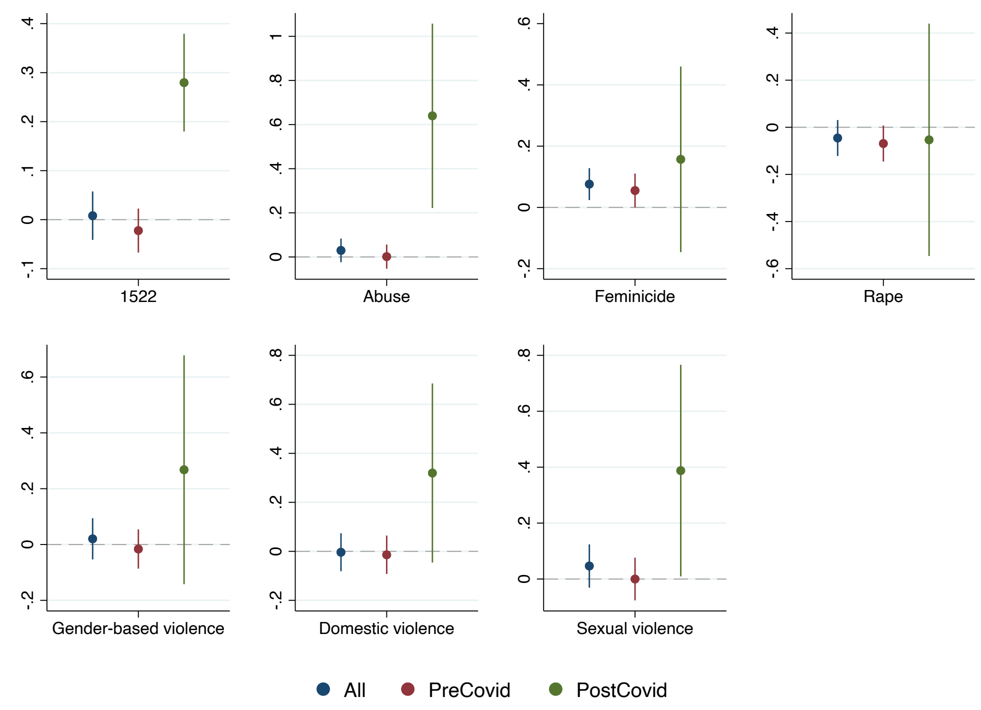
Figure 5: Coefficient plot from regressions of yearly 1522 calls on Google search inquiries, by selected keywords (regional-level).



Notes: Data on the outcome from ISTAT.

more aligned during the lockdown period as Google, and Internet in general, have become major sources for information-seeking during confinement periods. Overall, all our results combined seem to suggest that online searches are a powerful tool to track potential threats of IPV before and after global-level crises such as the current COVID-19 pandemic — with stronger predictive power post-crisis — while online searches help to predict actual violence *only* in post-crises scenarios.

Figure 6: Coefficient plot from regressions of daily calls to AREU on Google search inquiries, by selected keywords (Lombardy), one-week lag.



Notes: Data on the outcome from AREU. Data from January 1, 2018 to May 30, 2020.

Conclusions

This study addressed the question of whether big data might help to reach "hard-to-reach" populations such as victims of intimate partner violence. Our focus on digital traces relies on the premise that data that are generated from the use of digital technology have the potential to address issues where reporting or social desirability biases are prevalent, allowing for better temporal resolution and spatial granularity than, for instance, household surveys. These are critical aspects when studying time-sensitive social phenomena that require immediate interventions, such as IPV, and especially so in the wake of global-level crises that increase levels of uncertainty, economic instability, and stress within families and couples. Leveraging the temporal discontinuity brought about by the current COVID-19 crisis, we have focused on Italy, one of the countries hardest-hit by the first wave of the pandemic. Also, we have focused specifically on digital traces from Google Trends to explore a source of data which is to date underutilized yet holds strong potential in terms of representativeness and population coverage, being the most commonly used search engine — and far more widely used than other platforms such as Twitter, Pinterest, Instagram, or Yahoo.

We explored two related research questions. First, we investigated whether online searches help predict instances of IPV in Italy over the last half decade. Second, we assessed whether the predictive power of online searches is higher, lower, or unaltered in the wake of global-level crises such as the current COVID-19 pandemic. To do so, we relied on search frequencies for multiple keywords measuring different facets of IPV, and we combined three different data sources varying in terms of temporal coverage (yearly versus daily data) and level of analysis, thus providing analyses at the country-level, regional-level, and for Lombardy only — the hardest-hit region by the pandemic. By combining different sources of data, we were also able to characterize instances of IPV into potential violence (or "threat" of violence) and actual violence as measured by calls to request an ambulance.

Starting from the first research question, our findings at the country-level suggest that online searches using selected keywords such as *feminicide*, *domestic violence* and *gender-based violence* well predict daily calls to domestic-violence helplines irrespective of specification (i.e., regardless of whether online searches are measured contemporaneously or during the previous week). The same overall finding is confirmed by regional-level analyses predicting yearly — rather than daily — calls. On top of *gender-based violence*, regional analyses show that also keywords such as *1522*, *abuse*, and *sexual violence* well predict helpline calls, even after controlling for regional-level controls such as GDP per capita, unemployment rate, and educational attainment. Conversely, analyses on daily calls to the emergency number in Lombardy provide little evidence of predictive power of online searches, except for the keyword *feminicide*. These findings combined suggest that digital traces are a powerful tool to track the risk of potential violence in "normal" or non-crises times, while they are less effective at tracking actual violent cases reported.

Moving to the second research question, while country-level analyses show little evidence of differential associations in the post-lockdown period, regional- and Lombardy-level analyses do suggest far stronger associations — both in terms of magnitude and statistical significance — in the post-lockdown period, with a high degree of concordance in terms of relevant keywords, namely *1522*, *abuse*, and *sexual violence*. Two are the implications. First, these findings underscore the key relevance of search engines — and of online connectivity in general — for women seeking help during confinement periods in which traditional help mechanisms become harder to reach. Second, findings from Lombardy which showed little to no associations in non-crises times in fact reveal that online searches can be a powerful tool to track actual violence in situations of global crises (i.e., post-lockdown) such as the current COVID-19 pandemic.

Overall, results from this study suggest that Google searches using selected keywords measuring different aspects of IPV serve as a powerful tool for tracking potential threats of IPV before and after global-level crises such as the COVID-19 pandemic — with stronger predictive power post-crisis — while online searches help predict actual violence in post-crises scenarios only, likely pointing towards a more active use of the Internet and the online resources that the Internet offers. As a matter of fact, while actual reporting of IPV tends to be lower in times of crises, online connectivity tends to be high, either due to rising unemployment, forced lockdowns at home, or both. We thus conclude that big data might be a very important — and, to date, widely underappreciated — resource for tracking or even anticipating IPV and getting a real-time picture of the brunt of domestic violence that women bear every day, but especially so in the wake of global-level crises.

Appendices

1 Results using daily data from the Equal Opportunity Department (Presidency of the Italian Council of Ministers), Italy

Table A1: Google hits and number of valid calls: Daily data, Italy

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
					Valid 1522 calls				
1522	0.229								
	(0.199)								
Abuse		0.325***							
		(0.107)							
Home & Abuse			0.019						
			(0.042)						
Home & Rape				0.076					
				(0.052)					
Feminicide					0.326***				
					(0.081)				
Rape						0.126			
						(0.077)			
Domestic viol.							0.303***		
							(0.070)		
Gender-based viol.								0.267***	
								(0.068)	
Sexual viol.									0.087*
									(0.050)
Post lockdown	83.355***	86.015***	86.370***	85.443***	85.612***	85.914***	80.551***	85.156***	85.930***
	(5.085)	(4.925)	(4.876)	(5.009)	(4.982)	(4.942)	(5.203)	(5.000)	(4.932)
Year FE	✓	✓	✓	✓	✓	✓	✓	✓	✓
Observations	608	608	608	608	608	608	608	608	608

Note: OLS. Robust standard error reported in brackets. * p<0.10, ** p<0.05, *** p<0.01

Table A2: Google hits and number of valid calls: Daily data, Italy (1 week lag)

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
	Valid 1522 calls								
1522 (l)	0.495***								
	(0.178)								
Abuse (l)		0.070							
		(0.111)							
Home & Abuse (l)			-0.059						
			(0.060)						
Home & Rape (l)				0.027					
				(0.055)					
Feminicide (l)					0.350***				
					(0.078)				
Rape (l)						0.039			
						(0.063)			
Sexual viol. (l)							0.082		
							(0.054)		
Gender-based viol. (l)								0.172**	
								(0.068)	
Domestic viol. (l)									0.348***
									(0.072)
Post lockdown	83.114***	86.114***	86.375***	86.214***	85.059***	86.190***	86.542***	85.073***	82.914***
	(5.103)	(4.938)	(4.896)	(4.901)	(4.961)	(4.918)	(4.955)	(5.152)	(4.532)
Year FE	✓	✓	✓	✓	✓	✓	✓	✓	✓
Observations	601	601	601	601	601	601	601	601	601

Note: OLS. Robust standard error reported in brackets. * p<0.10, ** p<0.05, *** p<0.01

Table A3: Google hits and number of valid calls: Weekly data, Italy

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
	Valid 1522 calls								
1522 (1 week avg.)	-0.080								
	(0.166)								
Abuse (1 week avg.)		0.003							
		(0.113)							
Home & Abuse (1 week avg.)			0.075						
			(0.047)						
Home & Rape (1 week avg.)				0.062					
				(0.050)					
Feminicide (1 week avg.)					0.437***				
					(0.070)				
Rape (1 week avg.)						0.109			
						(0.193)			
Domestic viol. (1 week avg.)							0.921***		
							(0.185)		
Gender-based viol. (1 week avg.)								0.321***	
								(0.119)	
Sexual viol. (1 week avg.)									-0.023
									(0.078)
Post lockdown	87.204***	86.058***	83.530***	85.378***	86.294***	85.887***	70.571***	85.056***	86.241***
	(4.730)	(5.715)	(5.815)	(5.016)	(4.894)	(5.051)	(6.019)	(5.074)	(5.001)
Year FE	✓	✓	✓	✓	✓	✓	✓	✓	✓
Observations	608	608	608	608	608	608	608	608	608

Note: OLS. Robust standard error reported in brackets. * p<0.10, ** p<0.05, *** p<0.01

Table A4: Google hits and number of valid calls: Daily data, Italy, with post-lockdown interaction

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
					Valid 1522 calls				
1522	0.252**								
	(0.108)								
Post lockdown=1 × 1522	-0.035								
	(0.316)								
Abuse		0.265***							
		(0.078)							
Post lockdown=1 × Abuse		2.182							
		(1.705)							
Home + Abuse			-0.010						
			(0.032)						
Post lockdown=1 × Home & Abuse			0.208						
			(0.229)						
Home & Rape				0.022					
				(0.038)					
Post lockdown=1 × Home & Rape				0.213					
				(0.183)					
Feminicide					0.316***				
					(0.079)				
Post lockdown=1 × Feminicide					0.113				
					(0.414)				
Rape						0.084			
						(0.074)			
Post lockdown=1 × Rape						0.811			
						(0.726)			
Domestic viol.							0.034		
							(0.043)		
Post lockdown=1 × Domestic viol.							0.854***		
							(0.188)		
Gender-based viol.								0.229***	
								(0.047)	
Post lockdown=1 × Gender-based viol.								0.220	
								(0.321)	
Sexual viol.									0.067
									(0.042)
Post lockdown=1 × Sexual viol.									0.223
									(0.341)
Post lockdown=1	83.526***	75.660***	84.124***	84.090***	85.093***	79.357***	65.819***	81.480***	82.294***
	(5.655)	(10.450)	(5.790)	(5.365)	(5.812)	(8.592)	(6.935)	(7.909)	(8.034)
Year FE	✓	✓	✓	✓	✓	✓	✓	✓	✓
Observations	608	608	608	608	608	608	608	608	608

Note: OLS. Robust standard error reported in brackets. * p<0.10, ** p<0.05, *** p<0.01

2 Results using yearly data from ISTAT, regional-level

Table A5: Google hits and 4 months-aggregated calls at the regional level

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	Valid 1522 calls						
	b/se	b/se	b/se	b/se	b/se	b/se	b/se
1522	0.001***						
	(0.000)						
Abuse		0.001***					
		(0.000)					
Feminicide			0.000				
			(0.000)				
Rape				0.000			
				(0.000)			
Domestic violence					0.000		
					(0.000)		
Gender-based violence						0.000***	
						(0.000)	
Sexual violence							0.001***
							(0.000)
Regional controls	✓	✓	✓	✓	✓	✓	✓
Year FE	✓	✓	✓	✓	✓	✓	✓
Observations	144	152	152	158	104	136	136

Note: OLS, as indicated. Clustered standard errors at regional level reported in brackets. * p<0.10, ** p<0.05, *** p<0.01

Table A6: Google hits and 4 months-aggregated calls at the regional level, with post-lockdown interaction

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
				Valid 1522 calls			
	b/se	b/se	b/se	b/se	b/se	b/se	b/se
1522	0.001***						
	(0.000)						
Post lockdown=1 × 1522	0.001						
	(0.001)						
Abuse		0.001***					
		(0.000)					
Post lockdown=1 × Abuse		0.001**					
		(0.001)					
Feminicide			0.000				
			(0.000)				
Post lockdown=1 × Feminicide			-0.002				
			(0.003)				
Rape				0.000			
				(0.000)			
Post lockdown=1 × Rape				-0.003			
				(0.003)			
Domestic violence					0.000		
					(0.000)		
Post lockdown=1 × Domestic violence					0.001		
					(0.001)		
Gender-based violence						0.000*	
						(0.000)	
Post lockdown=1 × Gender-based violence						0.003***	
						(0.001)	
Sexual violence							0.001***
							(0.000)
Post lockdown=1 × Sexual violence							0.003***
							(0.001)
Post lockdown=1	0.020	0.028	0.112***	0.120***	0.062**	0.031*	0.023
	(0.037)	(0.026)	(0.041)	(0.042)	(0.028)	(0.018)	(0.028)
Regional controls	✓	✓	✓	✓	✓	✓	✓
Year FE	✓	✓	✓	✓	✓	✓	✓
Observations	144	152	152	158	104	136	136

Note: OLS, as indicated. Clustered standard errors at regional level reported in brackets. * p<0.10, ** p<0.05, *** p<0.01

3 Results using daily data from AREU, Lombardy

Table A7: Google hits and calls to AREU, Lombardy

	AREU calls, Lombardy						
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
1522	0.008 (0.025)						
Abuse		0.030 (0.027)					
Feminicide			0.076*** (0.027)				
Rape				-0.045 (0.039)			
Gender-based violence					0.020 (0.038)		
Domestic violence						-0.004 (0.040)	
Sexual violence							0.047 (0.039)
Constant	88.975*** (0.538)	88.718*** (0.616)	88.600*** (0.555)	89.346*** (0.610)	89.004*** (0.538)	89.005*** (0.538)	88.996*** (0.538)
Year FE	✓	✓	✓	✓	✓	✓	✓
Observations	875	875	875	875	875	875	875

Note: OLS, as indicated. Robust standard errors in brackets. * p<0.10, ** p<0.05, *** p<0.01.

Table A8: Google hits and calls to AREU, Lombardy, with post-lockdown interaction

	AREU calls, Lombardy						
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
1522	-0.028						
	(0.023)						
Post lockdown=1 x 1522	0.306***						
	(0.055)						
Abuse		-0.001					
		(0.027)					
Post lockdown=1 x Abuse		0.659***					
		(0.211)					
Feminicide			0.059**				
			(0.026)				
Post lockdown=1 x Feminicide			0.095				
			(0.154)				
Rape				-0.050			
				(0.037)			
Post lockdown=1 x Rape				-0.014			
				(0.247)			
Gender-based violence					-0.008		
					(0.035)		
Post lockdown=1 x Gender-based violence					0.270		
					(0.208)		
Domestic violence						-0.037	
						(0.039)	
Post lockdown=1 x Domestic violence						0.368**	
						(0.187)	
Sexual violence							0.007
							(0.038)
Post lockdown=1 x Sexual violence							0.396**
							(0.193)
Post lockdown=1	-11.790***	-13.358***	-9.467***	-9.402***	-9.360***	-9.420***	-9.474***
	(2.662)	(3.254)	(2.700)	(2.878)	(2.549)	(2.533)	(2.527)
Constant	89.110***	89.011***	88.693***	89.382***	89.006***	89.004***	89.004***
	(0.535)	(0.617)	(0.556)	(0.605)	(0.539)	(0.539)	(0.538)
Year FE	✓	✓	✓	✓	✓	✓	✓
Observations	875	875	875	875	875	875	875

Note: OLS, as indicated. Robust standard errors in brackets. * p<0.10, ** p<0.05, *** p<0.01.

References

Abel, T. and McQueen, D. (2020), 'The COVID-19 pandemic calls for spatial distancing and social closeness: not for social distancing!', *International Journal of Public Health* **7**, 1366.

Aguero, J. M. (2021), 'Covid-19 and the rise of intimate partner violence', *World Development* **137**(105217), 1–7.

Aizer, A. (2010), 'The gender wage gap and domestic violence', *American Economic Review* **100**(4), 1847–59.

Aizer, A. (2011), 'Poverty, violence, and health the impact of domestic violence during pregnancy on newborn health', *Journal of Human Resources* **46**(3), 518–538.

Alexander, M., Polimis, K. and Zagheni, E. (2020), 'Combining social media and survey data to nowcast migrant stocks in the United States', *Population Research and Policy Review* (0123456789).

Arenas-Arroyo, E., Fernandez-Kranz, D. and Nollenberger, N. (2021), 'Intimate partner violence under forced cohabitation and economic stress: Evidence from the covid-19 pandemic', *Journal of Public Economics* **194**, 104350.

Battiston, P., Kashyap, R. and Rotondi, V. (2020), 'Reliance on scientists and experts during an epidemic: Evidence from the Covid-19 outbreak in Italy', *SSM - Population Health* p. 100721.
URL: *http://www.sciencedirect.com/science/article/pii/S235282732030358X*

Billari, F. C., D'Amuri, F. and Marcucci, J. (2016), 'Forecasting Births Using Google', *First International Conference on Advanced Research Methods and Analytics* .

Billari, F. C. and Zagheni, E. (2017), Big Data and Population Processes: A Revolution?, *in* A. Petrucci and R. Verde, eds, 'Statistics and Data Science: New Challenges, New Generations. 28-30 June 2017. Proceedings of the Conference of the Italian Statistical Society.', Firenze, pp. 167–178.

Brownstein, J. S., Freifeld, C. C. and Madoff, L. C. (2009), 'Digital disease detection—harnessing the Web for public health surveillance', *The New England journal of medicine* **360**(21), 2153.

Bullinger, L. R., Carr, J. B. and Packham, A. (2020), 'COVID-19 and Crime: Effects of Stay-at-Home Orders on Domestic Violence', *Forthcoming at: American Journal of Health Economics* .

Card, D. and Dahl, G. B. (2011), ‘Family violence and football: The effect of unexpected emotional cues on violent behavior’, *The Quarterly Journal of Economics* **126**(1), 103–143.

Carlyle, K. E., Guidry, J. P. and Burton, C. (2018), ‘Recipes for prevention: An analysis of intimate partner violence messages on pinterest’, *Journal of interpersonal violence* p. 0886260518812073.

Carlyle, K. E., Guidry, J. P., Dougherty, S. A. and Burton, C. W. (2019), ‘Intimate partner violence on instagram: Visualizing a public health approach to prevention’, *Health Education & Behavior* **46**(2_suppl), 90S–96S.

Cesare, N., Lee, H., McCormick, T., Spiro, E. and Zagheni, E. (2018), ‘Promises and Pitfalls of Using Digital Traces for Demographic Research’, *Demography* **55**(5), 1979–1999.

Colagrossi, M., Deiana, C., Geraci, A. and Giua, L. (2020), ‘Hang up on stereotypes: Domestic violence and anti-abuse helpline campaign’, *Available at SSRN 3643922* .

Compton, R. (2019), ‘A Data-Driven Approach to the Fragile Families Challenge: Prediction through Principal-Components Analysis and Random Forests’, *Socius: Sociological Research for a Dynamic World* **5**, 237802311881872.

Delpierre, C. and Kelly-Irving, M. (2018), ‘Big Data and the Study of Social Inequalities in Health: Expectations and Issues’, *Frontiers in Public Health* **6**(October), 1–5.

Every-Palmer, S., Jenkins, M., Gendall, P., Hoek, J., Beaglehole, B., Bell, C., Williman, J., Rapsey, C. and Stanley, J. (2020), ‘Psychological distress, anxiety, family violence, suicidality, and wellbeing in New Zealand during the covid-19 lockdown: A cross-sectional study’, *PLOS One* **15**(11), e0241658.

Eysenbach, G. (2009), ‘Infodemiology and infoveillance: Framework for an emerging set of public health informatics methods to analyze search, communication and publication behavior on the Internet’, *Journal of medical Internet research* **11**(1), e11.

Fatehkia, M., Kashyap, R. and Weber, I. (2018), ‘Using Facebook ad data to track the global digital gender gap’, *World Development* **107**, 189–209.

Ginsberg, J., Mohebbi, M. H., Patel, R. S., Brammer, L., Smolinski, M. S. and Brilliant, L. (2009), ‘Detecting influenza epidemics using search engine query data’, *Nature* **457**(7232), 1012–1014.

Hsu, L.-C. and Henke, A. (2020), 'Covid-19, staying at home, and domestic violence', *Review of Economics of the Household* (https://doi.org/10.1007/s11150-020-09526-7), 1–11.

Kashyap, R., Fatehkia, M., Al Tamime, R. and Weber, I. (2020), 'Monitoring global digital gender inequality using the online populations of Facebook and Google', *Demographic Research* **43**, 779–816.

Lazer, D. M. J., Pentland, A., Watts, D. J., Aral, S., Athey, S., Contractor, N., Freelon, D., Gonzalez-Bailon, S., King, G., Margetts, H., Nelson, A., Salganik, M. J., Strohmaier, M., Vespignani, A. and Wagner, C. (2020), 'Computational social science: Obstacles and opportunities', *Science* **369**(6507), 1060–1062.

Leslie, E. and Wilson, R. (2020), 'Sheltering in place and domestic violence: Evidence from calls for service during Covid-19', *Journal of Public Economics* **104241**, 1–7.

Lindberg, L., VandeVusse, D., Mueller, A., Kirstein, J. and Mariell (2020), Early Impacts of the COVID-19 Pandemic: Findings from the 2020 Guttmacher Survey of Reproductive Health Experiences, Technical Report June, Guttmacher Institute.

Martins-Filho, P. R. S., Mendes, M. L. T., Reinheimer, D. M., do Nascimento-Júnior, E. M., Vaez, A. C., Santos, V. S. and Santos, H. P. (2018), 'Femicide trends in Brazil: Relationship between public interest and mortality rates', *Archives of Women's Mental Health* **21**(5), 579–582.

Massicotte, P. and Eddelbuettel, D. (2020), 'gtrendsr: Perform and display Google Trends queries'. R package version 1.4.6.9000.

McCauley, H. L., Bonomi, A. E., Maas, M. K., Bogen, K. W. and O'Malley, T. L. (2018), '# MaybeHeDoesntHitYou: Social media underscore the realities of intimate partner violence', *Journal of Women's Health* **27**(7), 885–891.

Öhman, C. J. and Watson, D. (2019), 'Are the dead taking over Facebook? A Big Data approach to the future of death online', *Big Data and Society* **6**(1), 1–13.

Pelat, C., Turbelin, C., Bar-Hen, A., Flahault, A. and Valleron, A.-J. (2009), 'More diseases tracked by using Google Trends', *Emerging infectious diseases* **15**(8), 1327.

Perez-Vincent, S. M., Carreras, E., Gibbons, M. A., Murphy, T. E. and Rossi, M. A. (2020), 'Covid-19 lockdowns and domestic violence'.

Peterman, A., Potts, A., O'Donnell, M., Thompson, K., Shah, N., Oertelt-Prigione, S. and van Gelder, N. (2020), 'Pandemics and violence against women and children', *Center for Global Development Working Paper* **528**.

Piquero, A. R., Riddell, J. R., Bishopp, S. A., Narvey, C., Reid, J. A. and Piquero, N. L. (2020), 'Staying home, staying safe? A short-term analysis of Covid-19 on Dallas domestic violence', *American Journal of Criminal Justice* pp. 1–35.

Purohit, H., Banerjee, T., Hampton, A., Shalin, V. L., Bhandutia, N. and Sheth, A. P. (2015), 'Gender-based violence in 140 characters or fewer: A# bigdata case study of Twitter', *arXiv preprint arXiv:1503.02086* .

Rampazzo, F., Zagheni, E., Weber, I., Testa, M. R. and Billari, F. (2018), Mater certa est, pater numquam: what can facebook advertising data tell us about male fertility rates?, *in* 'Proceedings of the International AAAI Conference on Web and Social Media', Vol. 12.

Reis, B. Y. and Brownstein, J. S. (2010), 'Measuring the impact of health policies using internet search patterns: The case of abortion', *BMC Public Health* **10**.

Risteski, D. and Davcev, D. (2014), Can we use daily internet search query data to improve predicting power of egarch models for financial time series volatility, *in* 'Proceedings of the International Conference on Computer Science and Information Systems (ICSIS'2014), October 17–18, 2014, Dubai (United Arab Emirates)'.

Schneider, D., Harknett, K. and McLanahan, S. (2016), 'Intimate partner violence in the Great Recession', *Demography* **53**(2), 471–505.

Storey, J. E. (2020), 'Risk factors for elder abuse and neglect: A review of the literature', *Aggression and violent behavior* **50**, 101339.

WHO (2013), Global and regional estimates of violence against women: prevalence and health effects of intimate partner violence and non-partner sexual violence, Technical report, WHO, Geneva.

WHO (2020), 'COVID-19 and violence against women What the health sector/system can do', (March).
URL: *https://www.who.int/reproductivehealth/publications/emergencies/COVID-19-VAW-full-text.pdf?ua=1*

Xue, J., Chen, J. and Gelles, R. (2019), 'Using data mining techniques to examine domestic violence topics on Twitter', *Violence and gender* **6**(2), 105–114.

Xue, J., Macropol, K., Jia, Y., Zhu, T. and Gelles, R. J. (2019), 'Harnessing big data for social justice: An exploration of violence against women-related conversations on Twitter', *Human Behavior and Emerging Technologies* **1**(3), 269–279.

Zagheni, E. and Weber, I. (2012), You are where you e-mail: Using email Data to Estimate International Migration Rates, *in* 'Proceedings of the 4th Annual ACM Web Science Conference', New York, pp. 348–351.

Zagheni, E., Weber, I. and Gummadi, K. (2017), 'Leveraging Facebook's Advertising Platform to Monitor Stocks of Migrants', *Population and Development Review* **43**(4), 721–734.