The real scale of domestic violence during Covid-19

Victims of domestic abuse may struggle to contact the police. But they *are* likely to seek help on the internet. By using internet search data to measure domestic violence during the Covid-19 pandemic, **Dan Anderberg**, **Helmut Rainer** and **Fabian Siuda** found an increase several times larger than that suggested in official police records

Many types of crisis – such as economic recessions, natural disasters or disease outbreaks – raise stress levels and thus carry the risk of increasing domestic violence.[1–4] In the case of Covid-19, the pandemic

are available in real time nor provide information that is temporally granular enough to quantify the immediate impacts of policies such as those implemented during the pandemic.

An alternative data source that might have great potential as a proxy for the scale of domestic violence during crises like Covid-19 is information from domestic violence helplines or women's support charities. However, thus far this information has rarely been systematically collected or made available for research.

By contrast, police records of domestic violence incidents are often available at daily frequencies and in real time. Several recent studies have used such data to assess the impact of lockdown restrictions on domestic violence during the pandemic. Surprisingly to many observers, these studies found relatively modest or no increases in family violence after lockdown measures were implemented.[6–10]

There is, however, a fundamental problem with this evidence. While domestic violence frequently goes underreported to the police in normal times, there is every reason to believe that lockdown restrictions imposed during the pandemic may have further exacerbated underreporting. Indeed, the pandemic and associated lockdowns conceivably left victims of domestic violence trapped at home with their perpetrators, limiting their opportunity to safely report incidents to the police. As a consequence, any analysis of domestic violence incidents recorded by the police runs the risk of underestimating the domestic violence problem during crises like Covid-19.

To tackle this measurement problem, we propose the use of an alternative approach to measure daily variation in domestic violence in real time: an algorithm that draws upon internet search activity for terms related to domestic violence help-seeking. Compared to making a police report, internet search is less likely to have been affected by self-isolation and quarantine measures. It allows for more anonymity, carries fewer consequences for both victim and perpetrator, and is much easier to hide from an abusive partner than an emergency call to the police.

Our algorithm, laid out in a recent paper,[11] uses five years of pre-2020 data to regress daily internet search activity in England

Daxiao Productions/Bigstock.com

arguably led to a global "shadow pandemic" of violence against women.[5] Effective policy responses require up-to-date reliable data on the scale of the problem. However, quantifying the prevalence of domestic violence is difficult at the best of times due to

data limitations, and the Covid-19 pandemic exacerbated the problem in various ways.

Victimisation surveys have, under normal circumstances, become an accepted way of estimating prevalence rates for domestic violence. However, these surveys neither

Dan Anderberg is professor of economics at Royal Holloway University of London. He got his PhD from the University of Lund and focuses his research on family economics and public policy.

for 35 terms related to domestic violence help-seeking on daily incidents of domestic violence crimes recorded by the London Metropolitan Police (both observed). Both data sources reflect the same underlying (unobserved) daily variation in domestic violence incidence, driven, for instance, by factors such as national holidays, variation in weather conditions or high-profile events. Thus, those two data sources have a positive correlation that is stronger for the most relevant/least noisy internet search terms. We make use of this positive correlation to construct our domestic violence index.

First, we regress the time series on year, month, and day-of-the week effects and obtain the residuals from these regressions – that is, the component not explained by these factors. This cleans the data from time trends and patterns in online search and domestic violence that regularly occur over time. The unexplained components from these regressions hold all information about the variable's temporal variation that is not driven by these long-run trends, and seasonal and weekly effects. We do this for the police-reported cases of domestic violence as well as all internet search terms individually. In the next step, we use the time series of these unexplained components for further analysis.

Second, we estimate the positive correlation between the unexplained components in the internet search activity and the unexplained components in reports of domestic violence. Since the domestic violence and internet search activity do not necessarily have to occur on exactly the same day but may potentially occur on contiguous days, we estimate correlations between the internet search intensities on a given day and the measured crimes on up to 3 days before and 3 days after.

We also calculate the signal-to-noise ratio for each of those components, which – in simple terms – is a measure of how much explanatory power each internet search term has in predicting domestic violence, relative to the variation it is unable to explain. The signal-to-noise ratio is larger for search terms that experience a high level of correlation with the prevalence of domestic violence, and is smaller for search terms that leave larger shares of the variation in domestic violence unexplained. Thus, the signal-to-noise ratio is a measure of how useful each term is in explaining domestic violence in the data. Conversely, a high signal-to-noise ratio indicates that the particular search term is more frequently used by victims of domestic violence.

Third, we create a composite domestic violence index based on internet search activity using the estimated signal-to-noise ratios as weights for the individual search terms, including their leads and lags. In Table 1 we show some examples of Google search terms that experience high daily variation in search volume. These search

Table 1: Selected Google search terms

Search terms	Relative weight in search index
Group 1: Seeking support	
Refuge	1.294
Abuse helpline	1.268
Shelter	0.715
Group 2: Searching on abuse	
Domestic violence	4.207
Domestic abuse	3.317
Abusive relationship	2.884
Psychological abuse	1.625
Emotional abuse	1.184
Group 3: Police/legal protection	
Domestic violence law	1.575
Domestic violence police	0.809
Abuse police	0.718

Notes: The table lists selected Google search terms used in the construction of the composite domestic violence search intensity index. The second column reports the relative weight placed on that term, averaged over the ±*K* days used in the construction of the composite index.

terms fall in the general categories "seeking support", "searching on abuse", and "police/legal protection". In column 2 of Table 1, we show the relative weights given to those search terms in the index, which are based on the estimated signal-to-noise ratios.
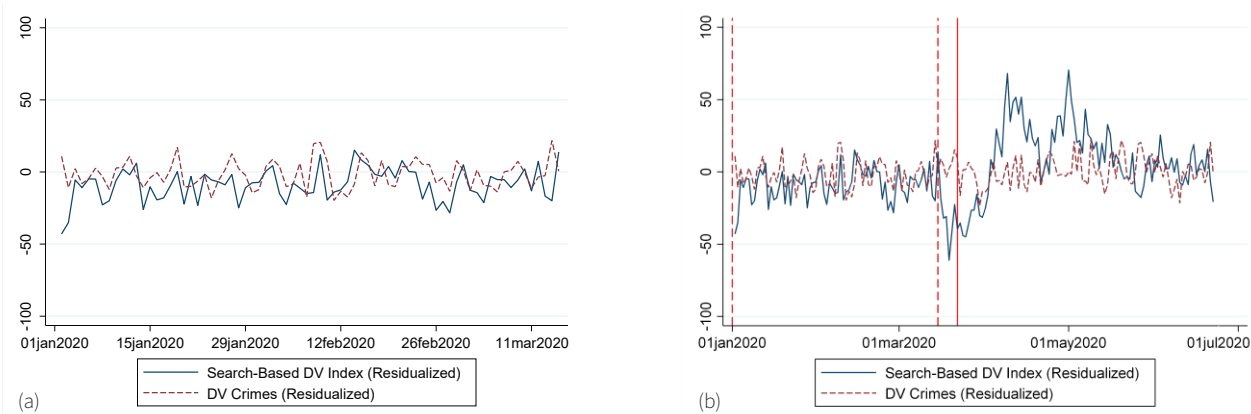

(a)

(b)

Figure 1 Time series for testing period and 2020 until end of first lockdown: (a) testing period, January–March 2020; (b) testing period followed by first lockdown. The figure shows the residuals of the normalised daily counts of domestic violence (DV) crimes recorded by the Metropolitan Police and of the search-based DV index after removing year, month and day-of-the-week fixed effects from each series. The initial normalisation rescaled both variables to have a mean of 100 over the algorithm training period from 1 April 2015 to 31 December 2019. The residualised series are shown in (a) for the testing period from 1 January to 15 March 2020. In (b) the series plotted in (a) is extended to include the lockdown period. The dashed vertical red lines indicate the start and end of the training period, while the solid vertical red line indicates the beginning of the lockdown.

Helmut Rainer is professor of economics at the University of Munich and director of the Center for Labor and Demographic Economics of the ifo Institute. He received his PhD from the University of Essex. His research interests are in applied microeconomics.

Fabian Siuda is assistant professor at the Vienna University of Economics and Business. He got his PhD from the University of Munich and focuses on labour and demographic economics.

After "training" our index with 5 years of pre-2020 data, we validate our index against police-reported domestic violence crimes for the pre-lockdown period in 2020, effectively the first 75 days of that year. Figure 1(a) shows the time series for the internet-search-based index and the official police records for the period from January to mid-March 2020 after both have been cleaned for long-run trends, and seasonal and weekly effects. Both time series behave very similarly in the time leading up to the first lockdown.

Finally, and most importantly, we analyse the impact of the London lockdown on the incidence of domestic abuse using our search-based index and we contrast our findings to a corresponding analysis using police-reported domestic violence incidents as outcome measure. Figure 1(b) shows the time series for the internet search-based index and the police-recorded crimes for the period from January 2020 to June 2020. After the implementation of the lockdown measures (solid red vertical line) we observe a strong increase in the search-based index while the series for police-reported domestic violence crimes remains fairly flat.

After controlling for trends and seasonal effects, we find only a small increase in police-recorded domestic violence crimes of around 5–7% (at peak) following the London lockdown in spring 2020. The effect of the lockdown remains small and positive until mid-June with an additional 10–15 extra daily domestic violence crimes relative to the average of 200.

In sharp contrast, although exhibiting a similar lockdown timing structure, we find a 40% increase (at peak) in our search-based domestic violence index. The effect is seven to eight times larger than the increase in police-recorded crimes and much closer to the increase in helpline calls reported by the UK's National Domestic Abuse Helpline at the time. We show that our results are seemingly robust to concerns that the measured internet search activity during the lockdown might have been driven by searches conducted,

for instance, by concerned neighbours or triggered by increased media attention. We do so by showing that our results are unaffected by removing the most "obvious" search terms (e.g. "domestic violence") from the construction of our search-based index. In a different robustness check, we reduce the time window of the internet search around the incidence date of domestic violence. Both robustness checks provide results very similar to our baseline results described before.

If we assume that the increase in the search-based index accurately captures the effect of the London lockdown on domestic violence incidence, whereas the lower increase in police-recorded domestic violence crimes reflects a reduced reporting rate by victims, we are able to estimate the number of "missing" police-recorded crimes over the lockdown period. The estimate we obtain indicates that the Metropolitan Police would have recorded an additional 4,700 domestic violence crimes over the lockdown period had the rate of reporting to the police itself not been reduced by the lockdown.

The broader lesson from our analysis is that it cautions against relying solely on police-recorded crimes or calls for service to assess the scale of the domestic violence problem during crises like Covid-19. The use of complementary data sources is important, as it allows researchers to move towards demarcating the lower and upper bounds of likely impacts on domestic violence. Our algorithm for measuring temporal variation in domestic violence incidence using internet search activity provides another viable strategy to complement assessments based on police records. Although our analysis by no means provides a definite answer to how to construct a real-time indicator of domestic violence, it will hopefully serve as a starting point that can be extended and further validated. ∎

Authors' note
This article is based on prior research by the authors.[11]

The estimate we obtain indicates that the Metropolitan Police would have recorded an additional 4,700 domestic violence crimes over the lockdown period had the rate of reporting to the police itself not been reduced by the lockdown.

References
1. Anastario, M., Shehab, N. and Lawry, L. (2009) Increased gender-based violence among women internally displaced in Mississippi 2 years post-Hurricane Katrina. *Disaster Medicine and Public Health Preparedness*, **3**(1), 18–26.
2. Anderberg, D., Rainer, H., Wadsworth, J. and Wilson, T. (2016) Unemployment and domestic violence: Theory and evidence. *Economic Journal*, **126**(597), 1947–1979.
3. Bermudez, L. G., Stark, L., Bennouna, C., Jensen, C., Potts, A., Kaloga, I. F. *et al.* (2019) Converging drivers of interpersonal violence: Findings from a qualitative study in post-hurricane Haiti. *Child Abuse & Neglect*, **89**, 178–191.
4. Onyango, M. A., Resnick, K., Davis, A. and Shah, R. R. (2019) Gender-based violence among adolescent girls and young women: A neglected consequence of the West African Ebola outbreak. In D. A. Schwartz, J. N. Anoko and S. A. Abramowitz (eds), *Pregnant in the Time of Ebola* (pp. 121–132). Berlin: Springer.
5. UN Women (2021) *Measuring the Shadow Pandemic: Violence against Women during COVID-19*. United Nations. bit.ly/3tJzfEz
6. Campedelli, G. M., Aziani, A. and Favarin, S. (2021) Exploring the immediate effects of COVID-19 containment policies on crime: An empirical analysis of the short-term aftermath in Los Angeles. *American Journal of Criminal Justice*, **46**, 704–727.
7. Ivandic, R., Kirchmaier, T. and Linton, B. (2021) The role of exposure in domestic abuse victimization: Evidence from the COVID-19 lockdown. ssrn.com/abstract=3686873
8. Leslie, E. and Wilson, R. (2020) Sheltering in place and domestic violence: Evidence from calls for service during COVID-19. *Journal of Public Economics*, **189**, 104241.
9. Mohler, G., Bertozzi, A. L., Carter, J., Short, M. B., Sledge, D., Tita, G. E. *et al.* (2020) Impact of social distancing during COVID-19 pandemic on crime in Los Angeles and Indianapolis. *Journal of Criminal Justice*, **68**, 101692.
10. Piquero, A. R., Jennings, W. G., Jemison, E., Kaukinen, C. and Knaul, F. M. (2021) Domestic violence during the COVID-19 pandemic: Evidence from a systematic review and metaanalysis. *Journal of Criminal Justice*, **74**, 101806.
11. Anderberg, D., Rainer, H. and Siuda, F. (2022) Quantifying domestic violence in times of crisis: An internet search activity-based measure for the COVID-19 pandemic. *Journal of the Royal Statistical Society*, *Series A*, **185**, 498–518.