IFS

Dan Anderberg
Helmut Rainer
Fabian Siudae

Working paper

20/29

Quantifying domestic violence in times of crisis

Quantifying Domestic Violence in Times of Crisis[*]

DAN ANDERBERG[a,b,c],   HELMUT RAINER[c,d,e]   and FABIAN SIUDA[e,f]

a) Royal Holloway University of London; b) Institute for Fiscal Studies;

c) CESifo; d) University of Munich; e) ifo Institute;

f) Vienna University of Economics and Business

August 27, 2020

Abstract

Recent contributions using police recorded calls-for-service and/or crime data to estimate impacts of COVID-19 lockdowns on the incidence of domestic violence (DV) have reported relatively modest effects. This may reflect a low reporting-propensity, exacerbated by the lockdown measures. Combining five years of daily Google Trends data for a set of DV-related search terms with daily data on DV crimes recorded by the London Metropolitan Police Service (MPS), we propose a method for generating a search-based DV-index, exploiting that both sets of data reflect the same inter-temporal variation in the (unobserved) DV incidence. Estimating the same model for the impact of lockdown on police-reported DV crimes and our search-based DV-index, we find a similar timing, but a substantially larger impact on the latter.

Keywords: Coronavirus; COVID-19; domestic violence; police recorded crime; Google Trends; signal-to-noise

JEL Classification: J12, I18

---

1 Introduction

While first and foremost a public health crisis, the COVID-19 pandemic is also an economic and social crisis that, *inter alia*, increased job insecurity and disrupted patterns of social interactions. One frequently voiced concern is that the lockdown policies and the general uncertainty created may have lead to a surge in the incidence of domestic abuse (see, e.g., Peterman et al. 2020; Van Gelder et al. 2020). Reports from victim's support charities, domestic abuse helplines operators, and frontline workers from across the world has fueled this concern, suggesting increases in domestic violence incidents following the implementation of lockdowns of anywhere between 25 percent and 80 percent (see, e.g., Allen-Ebrahimian 2020; Human Rights Watch 2020; UN Women 2020; Wagers 2020). Yet, in stark contrast to these alarming numbers, recent empirical studies exploiting daily counts of police-recorded domestic violence incidents have found either relatively modest or no increases. For instance, using data on 911 call from 14 large U.S. cities, McCrary & Sanga 2020 report finding a 12 percent increase in domestic violence call which peaked in early April and returned to pre-lockdown levels by end of April. Similarly, Leslie & Wilson 2020 exploit police calls data from 15 metropolitan areas in the U.S. and report finding a 10 percent increase in calls and a return back to 2019 levels by the end of April. For the U.K., Ivandic & Kirchmaier 2020 use data from the London Metropolitan Police Service, observing that the overall level of domestic abuse crimes remained stable when compared with the long-term.[1]

Given this major discrepancy between the increases in demands for domestic violence support reported by charities and practitioners and the recently presented research findings based on police-recorded call or crime data, this paper has two objectives: (1) to propose an algorithmic methodology for measuring temporal variation in domestic violence incidence based on internet search data, and (2) to highlight the potential limitations and biases in using police data for quantifying the scale of the domestic violence problem during this period of crisis.

From a policy perspective, there is an urgent need to quantify the impact of the unfolding COVID-19 pandemic on domestic violence (henceforth, DV): optimal policy responses to support victims of DV can only be implemented if the scale of the problem is known. However, the quantification of the prevalence of DV is difficult at the best of times due to data limitations, and the pandemic has exacerbated this difficulty. Victimization surveys have, under normal circumstances, become an accepted way of estimating incidence and prevalence rates for DV.

---

[1]See also Campedelli et al. 2020; Mohler et al. 2020; Payne & Morgan 2020; Piquero et al. 2020.

Unfortunately, these surveys are neither available in real-time nor do they provide granular enough information to adequately analyze the consequences of the COVID-19 crisis. By contrast, police records of DV incidents are often available at daily frequencies or even in real-time, and in many cases contain fine-level information on location. However, the vast majority of victims of DV do not report these crimes to the police (see, e.g., Podaná et al. 2010; UN Women 2020) and, importantly, reporting behavior itself may have been affected (see, e.g., Campbell 2020).

As a first step, we present evidence, based on daily counts of DV-related crimes recorded by the London Metropolitan Police Service (MPS), that the lockdown brought about a increase in recorded DV crimes of around 5-7 percent (at peak) compared to levels before the pandemic. This effect is however far below the increase in the average daily calls and contacts reported by Refuge, the operator of the National Domestic Abuse Helpline.

In an attempt to complement available data sources, we propose using DV-related internet search activity. Using such data comes with a set of methodological challenges in relation to selecting among alternative search terms and isolating the component of search activity that stems from victims of DV from other noisy components. We tackle these challenges by developing a simple statistical framework that uses pre-crisis data – in our case over five years – to relate daily DV-related internet search activity to daily police-recorded DV incidents (both observed). The intuition for the approach is that both reflect the same underlying (unobserved) temporal variation in DV incidence, leading to a positive correlation that is stronger for the most relevant/least noisy internet search terms. This allows us to use estimated signal-to-noise ratios to create a composite measure of DV-related search activity. Our algorithmic design further accounts for differential trends, seasonality etc. and the fact that victims may not search for help on the exact day of the incident. In our application of this procedure, the search terms that consistently get the largest weight in our composite DV index are key terms such as "abuse helpline", "abusive relationship", "domestic violence", and "psychological abuse".

We then present three results: First, empirical research investigating the relationship between weather and crime shows that temperature is positively correlated with aggressive behavior, including domestic violence (see, e.g., Butke & Sheridan 2010). Reassuringly, we find that higher temperatures are not only significant predictors of police-recorded DV-crimes but are also strongly correlated with our search-based DV index. Second, we observe a closely aligned timing of the increases in DV-crimes recorded by the London MPS and the increases in our search-based DV index: while the lockdown had no immediate impact, a significant effect emerged somewhere

between 3-5 weeks into the lockdown. Third, in level terms however, we find a 40 percent increase (at peak) in our search-based DV index, 5-8 times larger than the increase in police recorded crimes.

The remainder of the paper proceeds as follows. Section 2 provides some background on the COVID-19 lockdown and mobility in London. Section 3 estimates the impact of the lockdown on DV crimes recorded by the London MPS. In Section 4, we describe the construction of our search-based DV index, re-estimate the impact of the lockdown using the index, and present our robustness checks. Section 5 concludes.

2   Timing of Lockdown

Starting mid-March the UK government's implemented a string of measures to limit the spread of the coronavirus. On March 16, the Prime Minister announced that everyone should begin social distancing. Later the same week, schools, theatres, nightclubs, cinemas, gyms and leisure centres were ordered to close. Finally, on the evening of March 23, a stay-at-home order effective immediately was announced. All non-essential shops and services were ordered to close. People were instructed to stay home, except for exercise once a day, shopping for essential items, any medical need, providing care to a vulnerable person, or travelling to work that could not be done from home. The police were granted powers to issue fines and send people home.

The impact of the policy-measures on people's movement was strong. A sharp drop in mobility followed after social distancing was announced, and after the announcement of the full stay-at-home order, mobility was down to 10-20 percent of the pre-lockdown level.[2] The easing of the lockdown was gradual from mid-May. Nevertheless, mobility remained below 50 percent of pre-lockdown levels through to the end of June.

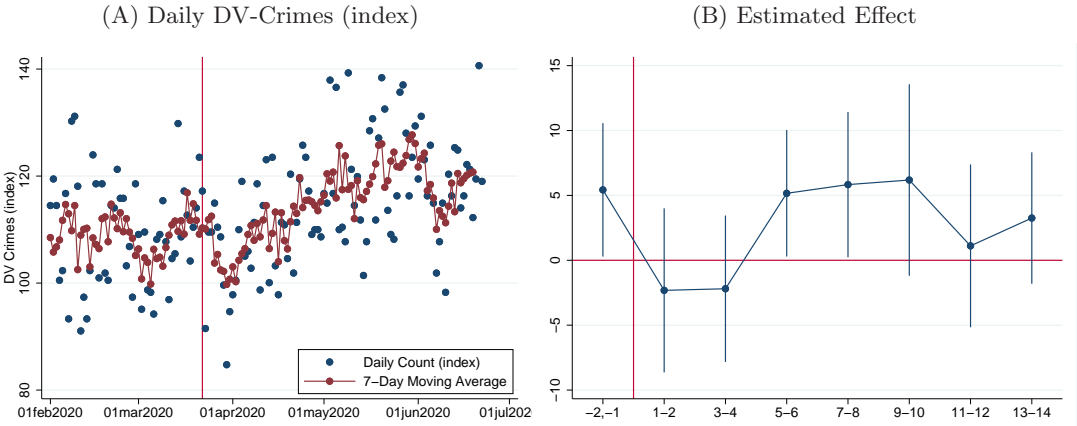## 3   Domestic Violence Crimes Recorded by the London MPS

Our data contains the daily counts of DV-related crimes recorded by the MPS, running from 1st April 2015 through to 22nd June, 2020. Hence the unit of time, $t = 1, ..., T$, in our data is a day, and we use t_0 to denote the time of the lockdown, March 23rd. We present further details of the count data in Appendix A where we illustrate that the recorded crime counts exhibit an

---

[2]See https://data.london.gov.uk/dataset/coronavirus-covid-19-mobility-report for mobility measures from Citimapper, Google, Apple, and Transport for London.

overall growing trend, a seasonality pattern where DV-incidence is increasing during the late spring, and with incidence being higher at weekends (see Figure A.1). Overall, these strong patterns imply that a robust analysis of daily incidence needs to account for long-term growth, seasonality and day-of-week patterns.

We will use the DV-crime data in index form rather than in absolute counts. The average daily count over the sample period up to 8th March 2020 was 222, and we will use this pre-crisis period as base (index 100 = 222 daily DV-crimes). This allows us to interpret estimated effects etc. in proportional terms. Focusing then on the time since February of this year, Panel A of Figure 1 highlights the daily DV-crimes index and the timing of the lockdown. The figure suggests a fairly steady increase in incidence starting after April 1st and right through to the end of May. Due to general seasonal patterns, the observed rise however cannot be immediately interpreted.

Figure 1: Daily DV-crimes since February 2020 and estimated bi-weekly effects of lockdown

(A) Daily DV-Crimes (index)　　　　(B) Estimated Effect

Notes: Panel (A): The sample consists of daily counts of domestic violence-related crimes recorded by the London MPS between 1 February 2020 and 22 June 2020 used in index form (index 100 = 222 DV-crime counts, and corresponds to the average daily count between 1st April 2015 and 8th March 2020). Panel (B): The figure plots the coefficients from a regression estimating the effect of the lockdown by two-week interval on the index of DV-crimes recorded by the London MPS. The underlying regression controls for year-, month-, and day-of-the-week effects, as well as for controls for temperature and rainfall.

One factor that may have played a role was the weather. Hot weather is a well-documented factor that increases the DV-incidence (Sanz-Barbero et al. 2018). April of this year saw particularly warm and dry weather. To consider the impact of weather as a factor, we use data on

daily average temperature (in $^\circ C$) and rainfall (in mm) in London over the sample period (see Appendix B for details).

3.1 Empirical Model and Findings

To assess the impact of the lockdown on police-reported DV-crimes, we estimate a regression that accounts for an overall trend, seasonality, and day-of-week effects. Denoting the index of daily DV-crimes, P_t, we model P_t as

$$P_t = \alpha + \beta_y + \gamma_m + \delta_d + \zeta x_t + f\left(t - t_0\right) I_{t \geq t_0} + \varepsilon_t, \quad t = 1, ...T, \tag{1}$$

where β_y, γ_m and δ_d are year-, month-, and day-of-the-week fixed-effects, controlling for trend, seasonality, and weekly cycles respectively. x_t includes controls for temperature and rainfall. $I_{t \geq t_0}$ is a dummy for the date t being within the lockdown period, and $f\left(t - t_0\right)$ is a flexible, but continuous, function of lockdown duration. Note that $f\left(0\right)$ is not restricted to be zero. Hence it allows for an immediate lockdown effect. Our baseline specification for $f\left(\cdot\right)$ is a quadratic function, possibly with a distinct effect for weekends,

$$f\left(\tau\right) = \phi_0 + \phi_1 \tau + \phi_2 \tau^2 + \phi_3 I_{wkend}, \tag{2}$$

where I_{wkend} is a weekend (Saturday/Sunday) indicator.

Table 1 presents estimates of the impact of the lockdown on the daily incidence of DV-related crimes – in index form – recorded by the MPS. Specification (i) provides a first basic version, ignoring weather and separate weekend effects. The estimates suggest, if anything, a negative immediate effect. However, a positive effect emerged over the following weeks and peaked after about 50 days of lockdown $(= -\phi_1 / \left(2\phi_2\right))$. This would be around mid-May, aligning well with the visual impression from Panel A of Figure 1. The finding that the impact of the lockdown grew with duration naturally accords with the notion that household tensions built up gradually.

In specification (ii) we add controls for weather, confirming a strong effect of temperature: a one degree Celsius increase in the daily (average) temperature is associated with a 0.8 percent increase in DV-crimes per day. Rainfall is estimated to have a negative, but less precisely estimated, impact. Hence the prolonged period of above-average temperature and dry weather observed for April and May does account for some part of the rise in reported DV-crimes over this period. Finally, specification (iii) allows for the lockdown to have a differential effect

Table 1: The effect of lockdown on DV-reported crimes recorded by the London MPS

	(i)	(ii)	(iii)
Lockdown (ϕ_0)	-7.752**	-6.389*	-4.694
	(3.791)	(3.683)	(3.610)
Days of Lockdown (ϕ_1)	0.433**	0.370**	0.379**
	(0.169)	(0.162)	(0.158)
Days Sq. (ϕ_2)	-0.00403**	-0.00328**	-0.00332**
	(0.00164)	(0.00157)	(0.00157)
Temperature ($^\circ C$)		0.840***	0.842***
		(0.0683)	(0.0678)
Precipitation (mm)		-3.053**	-2.979**
		(1.519)	(1.517)
Weekend $\times$ Lockdown (ϕ_3)			-7.114***
			(2.285)
Observations	1,910	1,910	1,910

Standard errors in parentheses

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

Notes: The dependent variable is the daily count of DV-related crimes recorded by the London MPS between 1 April 2015 and 22 June 2020 in index form (100 = average daily count over the period 1 April 2015 to 8 March 2020). All regressions include year, month, and day-of-week fixed effects.

on weekends. The strong negative effect here indicates that recorded DV-crimes during the lockdown had a much lower weekend-weekday difference than pre-crisis.[3]

In order to avoid influence of the parametric form, we can replace the function $f(t - t_0)$ with a set of dummies for two-week periods relative to the time of lockdown, starting with the two weeks leading up to the formal lockdown.[4] The results are presented in Panel B of Figure 1. The estimates suggest that there was no significant effect on the number of DV-related crimes recorded by the MPS early on in the lockdown. There was however a noticeable increase in recorded DV-crimes from the end of April through into early June (weeks 5-10). Nevertheless, the estimated effects are generally quite small – an increase of $5 - 7$ percent, at peak, compared to the pre-lockdown average.

To summarize: (i) The lockdown led to an increase in the DV-related crimes recorded by the MPS; the impact was not immediate, but emerged gradually and remained positive until

---

[3]The estimated coefficient translates into about 15 DV-crimes per day, implying that the weekend-weekday difference during the lockdown was only about half of pre-crisis difference. See Figure A.1.

[4]Note that this means that the left-out "reference period" (1st April 2015 - 8th March 2020) corresponds to the period used as base for the indexing of the DV-crime variable (index = 100).

mid-June when daily counts returned to normal levels; (ii) The lockdown significantly reduced the normal weekend-weekday gap in reported DV-crimes; (iii) The unusually warm and dry weather that occurred through April and May may have played a role, but does not account for the observed increase; (iv) The impact on the recorded DV-crimes was nevertheless modest – about 10-15 extra DV-crimes per day relative to a normal average of over 200 crimes per day.

4 Using Google Search Data

The vast majority of victims of domestic abuse however do not contact the police even under normal circumstances. Hence it is not clear whether the above fairly modest estimated effect on DV-reported crimes is representative of the wider impact or, potentially, a substantial underestimate. For instance, the charity Refuge that operates the National Domestic Abuse Helpline reported a steady increase in the average daily call and contacts after lockdown, reaching 80 above pre-lockdown levels by early May.[5] Data on helpline calls and contacts however tends to be patchy and selectively reported, and inherently difficult to independently verify. Hence there is clear motivation for exploring alternative sources of data that can capture help-seeking behavior by victims. Research has highlighted the potential usefulness of internet search data, with a particular readily available source being Google Trends. Recent work by Koutaniemi & Einiö 2019 explored this in the context of Finland, who explored the monthly seasonal patterns of police-reported abuse and DV-related Google searches. We build on this work here in order to study the potential impact of the current crisis, but at a finer temporal granularity.

4.1 A Framework

As above, let $t \in \{1, ..., T\}$ denote time, where a unit of time is a day. Lockdown occurs at some time t_0 and continues to the end of the sample period. Hence the overall sample period is split into two regimes, $R \in \{0, 1\}$, with $R_t = 0$ (no lockdown) if $t < t_0$ and $R_t = 1$ (lockdown) if $t \geq t_0$. Let n_t denote the number of DV-incidents/victims at time t. Whilst not directly observed, n_t will have some distribution, and the concern is that this distribution will have changed from the pre- to within-lockdown regime. Hence let $f_R(n)$ be the probability mass function for n in regime R.

A given victim of abuse i at time t, may seek help through alternative routes. Let $p_{it} \in$

$\{0,1\}$ indicate whether she contacts the police, leading to a recorded crime. Similarly, let $y_{it} \in \{0,1\}$ denote whether she seeks support via an internet search. Note that the two help-seeking responses are not mutually exclusive, so for any given victim, either none, or either, or potentially even both may occur. For expositional convenience we will assume that p_{it} and y_{it} are statistically independent, but nothing in the below hinges on this assumption. In the data we observe the count of incidents recorded by the police at time t. This is, we observe $P_t \equiv \sum_{i=1}^{n_t} p_{it}$. Similarly, we observe search intensity $Y_t \equiv \sum_{i=1}^{n_t} y_{it}$.[6] One of the issues below will be the construction of the search intensity measure Y_t.

Help-Seeking Behavior Across Regimes

Each help-seeking behavior is guided by the net benefit to victim i from taking that action, which may be regime-specific. Hence let V_k^R denote the systematic (or "common") net systematic benefit to a victim from taking action $k \in \{p,y\}$ in regime $R \in \{0,1\}$. In addition, a given victim i perceives an individual-specific utility component to either taking or not taking each action, and if these are i.i.d. extreme value distributed, the probability of any given victim i in regime R taking action k will take the standard logit form,[7]

$$\pi_k^R = \Pr\left(k_{it} = 1 | R\right) = \frac{\exp\left(V_k^R\right)}{1 + \exp\left(V_k^R\right)}, \text{ for } k \in \{p,y\} \text{ and } R \in \{0,1\}. \tag{3}$$

This highlights the key issue of potentially changing reporting propensity: only if $V_p^1 = V_p^0$, and hence $\pi_p^1 = \pi_p^0$, will the proportional change – from pre- to within-lockdown – in P_t accurately reflect the proportional change in the abuse incidence level. However, there can be many reasons why the police-reporting propensity of victims may have changed during lockdown. The same can of course also be argued for victims seeking help via the internet. Hence we cannot *a priori* assume that $V_k^1 = V_k^0$ for either action. Hence define $\Delta V_k \equiv V_k^1 - V_k^0$. Under the weaker assumption that $\Delta V_k \leq 0$ for $k = \{p,y\}$ (no increased propensity for either help-seeking behaviour), the proportional change in either action from before- to within-lockdown serves as a lower bound for the corresponding proportional change in the abuse incidence level. As we will show, we observe a substantially larger proportional increase in help-seeking via the internet

---

[6]In the empirical application, both P_t and Y_t will be in index form. As this merely re-scales each by a multiplicative constant, the statistical properties are preserved.

[7]Specifically, a given victim i obtains an additive random utility ε_{i1}^k from *taking action k* and an additive random utility ε_{i0}^k of *not taking action k* which are assumed to be i.i.d. extreme value distributed across individuals and actions.

than via the police, indicating that, at the very least, $\Delta V_p < \Delta V_y$. A potential threat to the assumption $\Delta V_k \leq 0$ for both actions would be "substitutability": if the lockdown decreased the perceived benefit to contacting the police, this could potentially have shifted help-seeking onto alternative routes. In contrast, "complementary" between the actions would imply that lockdown would tend to change both net values in the same direction.

Relating Internet Searches to Police Reports

Under the above framework, the daily internet search activity and police report counts will be correlated as both reflect the underlying temporal variation in DV-incidence. To see this, consider the covariance between P_t and Y_t within either given regime $R_t \in \{0, 1\}$. Using the law of iterated expectations, it is easily shown that,

$$Cov\left(P_t, Y_t | R_t\right) = \pi_p^{R_t} \pi_y^{R_t} Var\left(n_t | R_t\right) > 0. \tag{4}$$

Intuitively, P_t and Y_t are positively correlated as both tend to be large on days when n_t is large.

In practice, we observe a set J of DV-related Google search terms, and for each term $j \in J$ we observe a daily search intensity Y_{jt} (in index form). Hence in order to create a single composite index Y_t we need to apportion relative weight across the various terms. To do so, we will use pre-lockdown data and draw on (4). This equation can be taken to apply for each search term $j \in J$, whereby the *relative* covariances of the various search terms Y_{jt} with P_t will indicate the *relative* frequency with which abuse victims use the J terms: using π_{jy}^0 to denote the pre-lockdown propensity for a victim to do a search on term $j \in J$ it follows from (4) that for two alternative terms, j and j', $Cov\left(P_t, Y_{j't} | R_t = 0\right) / Cov\left(P_t, Y_{jt} | R_t = 0\right) = \pi_{j'y}^0 / \pi_{jy}^0$.[8]

However, measured search intensities can be expected to contain a fair amount of noise, e.g. due to random searches by non-victims. Hence consider the regression specification,

$$Y_{jt} = \alpha_j + \lambda_j P_t + v_{jt}, \text{ for } j \in J, \tag{5}$$

where v_{jt} represents noise. As usual, the ordinary least squares estimator of λ_j is $\widehat{\lambda}_j = \widehat{Cov}\left(P_t, Y_{jt}\right) / \widehat{Var}\left(P_t\right)$. Hence, applying this on pre-lockdown data will allow us to identify

---

[8]Note that if data were pooled across regimes and $\pi_{jy}^R / \pi_{j'y}^R$ remained constant, the relative covariance $Cov\left(P_t, Y_{jt}\right) / Cov\left(P_t, Y_{j't}\right)$ would only correspond to the relative search frequency if the component frequencies π_{jy}^R and $\pi_{j'y}^R$ remained constant also in level terms. This can be seen by noting that, by the law of total covariance, $Cov\left(P_t, Y_{jt}\right) = E\left[\pi_p^R \pi_{jy}^R Var\left(n|R\right)\right] + Cov\left(\pi_p^R E\left[n|R\right], \pi_{jy}^R E\left[n|R\right]\right)$.

search terms that are more commonly used by victims – as indicated by their relative values of $\widehat{\lambda}_j$ – and that contain relatively less noise. We will use this approach to construct our composite search index Y_t. In particular, we will estimate (a version of) equation (5) for each search term $j \in J$, and terms with an estimated positive correlation, $\widehat{\lambda}_j > 0$, will be given a weight in the composite index that corresponds to its implied signal-to-noise ratio.

Data and Algorithm

The exact algorithm used in constructing the composite index Y_t accounts for two further complications. First, to account for the possibility that police reports and internet searches have different growth over time, seasonality etc., rather than directly relating Y_{jt} to P_t, we relate the *unexpected component* of Y_{jt} to the corresponding *unexpected component* of P_t after removing year-, month-, and day-of-the-week effects. Second, while victims can be expected to contact the police at the time of a DV-incident, on-line help-seeking may be distributed around the time of the event, either in the days following the event or, if tensions are building in advance, in the days before. To account for this, we relate the unexpected component of P_t to the unexpected components of $Y_{j\tau}$ for a set of days τ *around* t.

It should be noted that while we use DV-crime data for the London MPS, the Google Trends data is for England. There are two reasons why our methods can be expected to be robust to this geographical discrepancy. First, the MPS is by far the largest territorial police force in England, covering over 8 million people, or about 15 percent of the entire population of England.[9] Second, whilst we refer to some dates as having "unexpectedly" high levels of DV-crimes, this is only in relation to the year, month and day-of-the-week that are controlled for. In fact, many of the days involved are highly predictable and include, for instance, all New Year's Days, many bank holiday weekends etc. which are, of course, common across the whole of England. Hence, one way to view the algorithm is that it uses the crime data to statistically identify high-risk days and then identifies search terms that spike on nearby days.

To implement our algorithm, we selected a set J of 35 potentially DV-relevant search terms. For each search term $j \in J$, we used Google Trends to generate a daily index of search intensity Y_{jt}, spanning our full sample period. We eliminated all terms that showed low daily variation, having zero entries for a majority of days. This left us with a reduced set $J_0 \subset J$ containing 23

---

[9]Non-London-based DV-related searches will in this respect be absorbed into the noise term, v_{jt}.

search terms (see below) from which we generate our composite measure Y_t.[10]

Using these remaining terms, we apply the following algorithm.

1. We regress P_t, on year-, month-, and day-of-the-week dummies using pre-lockdown data, $t \leq t_0$ and obtain the residual, denoted $\widehat{\varepsilon}_t$. These represent the *unexpected daily variation* in DV-crimes.

2. We correspondingly regress each search term intensity Y_{jt}, $j \in J_0$, on year-, month-, and day-of-the-week dummies again using $t \leq t_0$ and obtain the residuals, denoted $\widehat{\epsilon}_{jt}$. These represent the *unexpected daily variation* in the search intensity for term j.

3. Still using $t \leq t_0$, we relate $\widehat{\varepsilon}_t$ to $\widehat{\epsilon}_{j,\tau}$ for a set of $\pm K$ days around t by estimating $\widehat{\epsilon}_{j,t+s} = \alpha_j^s + \lambda_j^s \widehat{\varepsilon}_t + \omega_{j,t+s}$ for each $j \in J$ and $s \in \{-K, ..., +K\}$, and we compute (j,s)-specific signal-to-noise ratios, denoted $\sigma_{js} = \left(\widehat{\lambda}_j^s\right)^2 \widehat{Var}\left(\widehat{\varepsilon}_t\right) / \left[\left(\widehat{\lambda}_j^s\right)^2 \widehat{Var}\left(\widehat{\varepsilon}_t\right) + \widehat{Var}\left(\omega_{j,t+s}\right)\right]$.

4. Using the estimated signal-to-noise ratios as weights we construct a composite index , $Y_t = \sum_{j \in J_0} \sum_s \sigma_{js} Y_{j,t+s}$, from the individual search terms for the full sample period.

Hence the final daily composite index Y_t is a weighted average of the original J_0 search indices, along with their leads and lags. In our leading case we use a window of ± 3 days.[11] We rescale Y_t to have a mean of 100 over the period 1 April 2015 to 8 March 2020 (as for P_t above).

The 23 terms in J_0 included in the final composite index were (where the number in parenthesis is their relative weight – averaged over the ± 3 days, with overall mean of unity – in the index): "London Abuse" (0.063), "Abuse Help" (0.099), "Abuse Police" (0.611), "Abuse Support" (0.377), "Domestic Violence" (4.484), "Domestic Violence Law" (1.749), "Domestic Violence Support" (0.296), "Refuge" (1.550), "Reporting Abuse" (0.199), "Victim Support" (0.015), "Womens Refuge" (0.261), "Psychological Abuse" (1.883), "Emotional Abuse" (1.367), "Controlling Relationship" (0.882), "Abusive Partner" (0.065), "Abusive Relationship" (3.037), "Coercive control" (0.188), "Womens Aid" (0.457), "Shelter" (0.876), "Abuse Helpline" (0.933), "Abuse Protection" (0.407), "Domestic Abuse" (2.367), and "Domestic Violence Police" (0.835).

---

[10]The 12 terms eliminated as having low daily variation were "domestic violence protection", "threat of violence", " refuge centre", "domestic violence help", "report domestic abuse", "London refuge", "domestic violence charges", "partner violence", "refuge helpline", "violence refuge", "national domestic violence helpline", "domestic abuse charity".

[11]In our leading case, we thus estimate $23 \times 7 = 161$ signal-to-noise ratios and just over two-thirds (110) of these was positive and hence used in construction of the composite index.

4.2 Findings

Here we repeat the analysis from Section 3, but with the DV-related search index Y_t as outcome variable rather than the index of MPS-recorded DV-crimes. Starting in Panel A of Figure 2, we plot the index from February to the end of our sample period. This shows a gradual and sustained increase over first four weeks of the lockdown.

Figure 2: Daily values of the composite index of DV-related search intensity since February 2020 and estimated bi-weekly effects of lockdown
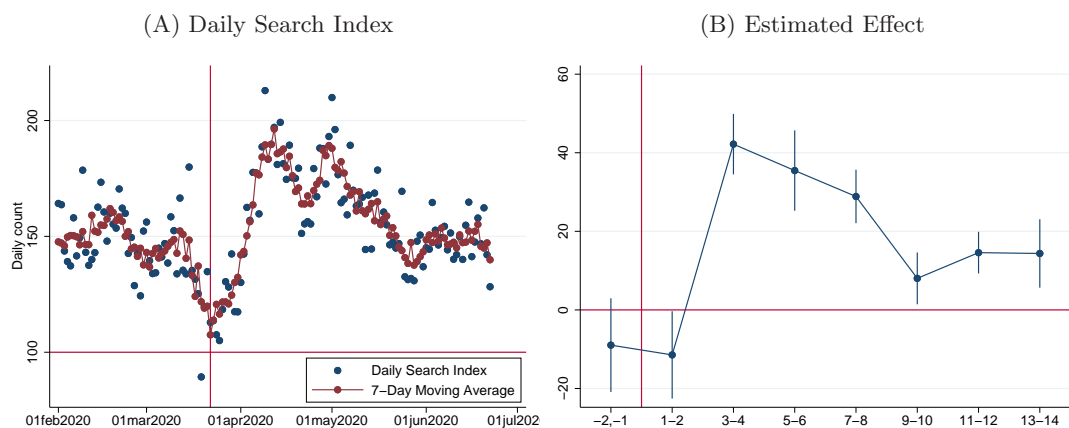
Notes: Panel (A): The sample consists of daily composite index of DV-related searches 1 February 2020 and 22 June 2020 (indexed such that 100 is the average search daily search intensity between 1st April 2015 and 8th March 2020). Panel (B): The figure plots the coefficients from a regression estimating the effect of the lockdown by two-week interval on the index of DV-related searches. The underlying regression controls for year-, month-, and day-of-the-week effects, as well as for controls for temperature and rainfall.

However, as in the case of DV-crimes, we will want to control for seasonality etc. in order to determine the effect of the lockdown of DV-related searches. Hence we repeat the estimation of (1), but with Y_t as outcome variable in place of P_t. Note that both outcome variables were used in the same index form, making them directly comparable.

Table 2 shows the results using the same specifications as in Table 1. While the timing is more or less identical to that of the DV-crimes – peaking about 50 days into the lockdown – the magnitude is about 7-8 times larger with, after controlling for trend, seasonality etc, a predicted increase at the peak of about 35-40 percent.

In specification (ii) we further control for weather. Here we find, reassuringly, that higher temperatures generate more DV-related searches according to our composite index: a one degree

Table 2: The effect of lockdown on DV-related search intensity

	(i)	(ii)	(iii)
Lockdown (ϕ_0)	-10.56	-9.945	-10.09
	(6.990)	(6.927)	(7.086)
Days of Lockdown (ϕ_1)	1.928***	1.907***	1.906***
	(0.298)	(0.297)	(0.297)
Days Sq. (ϕ_2)	-0.0205***	-0.0203***	-0.0202***
	(0.00294)	(0.00294)	(0.00294)
Temperature ($°C$)		0.275***	0.274***
		(0.104)	(0.104)
Precipitation (mm)		1.361	1.355
		(1.630)	(1.633)
Weekend $\times$ Lockdown (ϕ_3)			0.617
			(3.712)
Observations	1,905	1,905	1,905

Standard errors in parentheses

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

Notes: The outcome variable is a composite index of DV-related search intensity at daily frequency (100 = average daily intensity over the period 1 April 2015 to 8 March 2020). The sample period is 1 April 2015 to 22 June 2020. All regressions include year, month, and day-of-week fixed effects.

Celsius increase in the daily temperature is associated with close to a 0.3 percent increase in DV-related searches. The estimated effect of rainfall is, in contrast, highly imprecise.[12] In specification (iii) we find no indication of any differential impact of the lockdown on DV-related searches on weekends versus weekdays. This is expected given that our index Y_t is a composite of $\pm K$ days around t.

Panel B of Figure 2 (just as Figure 1) shows the effect of the lockdown by two-week intervals, starting in the two weeks prior to the lockdown. The two figures again exhibit similar timing, suggesting that the first few weeks of the lockdown remained relatively quiet. However, after that, our index shows a sharp increase approaching mid-April. At this stage, according to our index, DV-related searches were about 40 percent higher than their pre-lockdown average. Over the following two months, searches gradually fall back down towards pre-lockdown levels, but remains significantly about the pre-lockdown level.

To summarize, our analysis of the Google Trends data thus suggests a closely aligned timing with the DV-reported crimes, showing no immediate increase after lockdown, but instead a peak

---

[12]The lower estimates and precision is natural given that the weather measurements are local to London whereas the search data is for the whole of England.

in April and May. However, the estimated impact on DV-related searches is much larger than that for police-recorded DV-crimes, suggesting a marked decrease in police-reporting propensity by victims during the lockdown.

4.3 Robustness

In our main specification we used a 7-day (± 3) window when constructing our DV-related search intensity measure. In Table 3 we show that our results are not sensitive to the window used. Specification (i) reiterates the main specification from Table 2. Specifications (ii) and (iii) successively narrow down the window to ± 2 and ± 1, respectively, whilst specification (iv) uses on lagged searches (up to 3). In all cases, the overall estimated of the lockdown remain stable.

Table 3: The effect of lockdown on DV-related search intensity: Robustness to index construction

	(1)	(2)	(3)	(4)	(5)
	7 Day Window	5 Day Window	3 Day Window	3 Lags	No "Dom. Viol."
Lockdown (ϕ_0)	-10.09	-9.099	-18.98**	-16.08**	-14.55**
	(7.086)	(7.542)	(8.530)	(7.640)	(6.894)
Days of Lockdown (ϕ_1)	1.906***	1.620***	1.734***	2.178***	1.896***
	(0.297)	(0.320)	(0.378)	(0.354)	(0.292)
Days Sq. (ϕ_2)	-0.0202***	-0.0178***	-0.0188***	-0.0229***	-0.0196***
	(0.00294)	(0.00321)	(0.00404)	(0.00376)	(0.00292)
Weekend $\times$ Lockdown (ϕ_3)	0.617	2.970	3.366	2.871	0.576
	(3.712)	(4.032)	(5.272)	(4.859)	(3.864)
Temperature ($^\circ C$)	0.274***	0.178*	0.0383	0.422***	0.261**
	(0.104)	(0.108)	(0.148)	(0.139)	(0.114)
Precipitation (mm)	1.355	0.334	-1.232	-2.399	3.248*
	(1.633)	(1.901)	(2.909)	(2.226)	(1.887)
Observations	1,905	1,905	1,905	1,905	1,905

Standard errors in parentheses

* $p < .1$, ** $p < .05$, *** $p < .01$

Notes: The outcome variable is a composite index of DV-related search intensity at daily frequency (100 = average daily intensity over the period 1 April 2015 to 8 March 2020). The sample period is 1 April 2015 to 22 June 2020. All regressions include year, month, and day-of-week fixed effects.

After the lockdown, the potential impact on domestic violence was much debated in the media etc. A concern is hence that this might have fueled a general interest in the issue, and hence more Google searches. As the most likely search term in that case would have been "domestic violence", we might be worried that our composite index places a large weight on this particular term.[13] Hence, in specification (v) we set the weight on "domestic violence" to zero when constructing our composite index, verifying that this was not driving our results.

---

[13]Note however that the signal-to-noise ratios used as weights were determined entirely from the pre-lockdown data.

5 Conclusions

Many types of crises—be it disease outbreaks like the current one, severe economic downturns, or natural disasters—carry the risk of increasing domestic violence (Anastario et al. 2009; Anderberg et al. 2016; Bermudez et al. 2019; Onyango et al. 2019). To be effective during such crises, policy responses require the most current and reliable evidence on the scale of the problem. However, conventional data sources have severe limitations in this respect, a prime example being police-recorded DV-incidents. Although available in real-time in some countries, regions and cities, incidents recorded by the police only represent the "tip of the iceberg". In times of crisis, under-reporting is further likely exacerbated as DV victims may be trapped with and/or economically dependent on their perpetrator. The reports from charities and practitioner and the empirical evidence from our internet search-based DV index jointly suggest that police-based evidence may be seriously underestimating the consequences of the current pandemic.

The current paper by no means provides a definite answer to how to best construct a real-time indicator of DV, but can hopefully serve as a very basic starting point for further thought and analysis. Although we believe that the type of methods that we have proposed hold the promise of generating DV indicators that are contemporaneously available, have a fine temporal resolution, allow for international or regional comparisons, and exhibit a demonstrated validity, it would seem equally important to engage with governmental and non-governmental organizations supporting DV victims. If gathered systematically, information on helpline calls and online contacts could also become an important jigsaw piece in the complex task of quantifying DV in times of crisis and beyond.

References

Allen-Ebrahimian, B. (2020). *China's domestic violence epidemic.* Axios. Retrieved on August 5, 2020 from: https://www.axios.com/china-domestic-violencecoronavirus-quarantine-7b00c3ba-35bc-4d16-afdd-b76ecfb28882.html.

Anastario, M., Shehab, N., & 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. Retrieved from https://doi.org/10.1097/DMP.0b013e3181979c32

Anderberg, D., Rainer, H., Wadsworth, J., & Wilson, T. (2016). Unemployment and domestic violence: Theory and evidence. *The Economic Journal*, *126*(597), 1947–1979. Retrieved from https://doi.org/10.1111/ecoj.12246

Bermudez, L. G., Stark, L., Bennouna, C., Jensen, C., Potts, A., Kaloga, I. F., . . . others (2019). Converging drivers of interpersonal violence: Findings from a qualitative study in post-hurricane Haiti. *Child Abuse & Neglect*, *89*, 178–191. Retrieved from https://doi.org/10.1016/j.chiabu.2019.01.003

Butke, P., & Sheridan, S. C. (2010). An analysis of the relationship between weather and aggressive crime in Cleveland, Ohio. *Weather, Climate, and Society*, *2*(2), 127–139. Retrieved from https://doi.org/10.1175/2010WCAS1043.1

Campbell, A. M. (2020). An increasing risk of family violence during the Covid-19 pandemic: Strengthening community collaborations to save lives. *Forensic Science International: Reports*, 100089. Retrieved from 10.1016/j.fsir.2020.100089

Campedelli, G. M., Aziani, A., & Favarin, S. (2020). *Exploring the effect of 2019-nCoV containment policies on crime: The case of Los Angeles.* Università Cattolica del Sacro Cuore. Retrieved from https://arxiv.org/pdf/2003.11021.pdf

Human Rights Watch. (2020). *UK failing domestic abuse victims in pandemic.* Human Rights Watch. Retrieved on August 6, 2020 from: https://www.hrw.org/news/2020/06/08/uk-failing-domestic-abuse-victims-pandemic.

Ivandic, R., & Kirchmaier, T. (2020). *Domestic abuse in times of quarantine.* CEP CentrePiece Summer 2020. Retrieved from https://cep.lse.ac.uk/pubs/download/cp578.pdf

Koutaniemi, E. M., & Einiö, E. (2019). Seasonal variation in seeking help for domestic violence based on Google search data and Finnish police calls in 2017. *Scandinavian Journal of Public Health*, *1-6*. Retrieved from https://journals.sagepub.com/doi/abs/10.1177/1403494819834098

Leslie, E., & Wilson, R. (2020). Sheltering in place and domestic violence: Evidence from Calls for Service during COVID-19. *Journal of Public Economics*. Retrieved from https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3600646 (Forthcoming)

McCrary, J., & Sanga, S. (2020). *The impact of the coronavirus lockdown on domestic violence.* Northwestern University. Retrieved from https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3612491

Mohler, G., Bertozzi, A. L., Carter, J., Short, M. B., Sledge, D., Tita, G. E., . . . Brantingham, P. J. (2020). Impact of social distancing during COVID-19 pandemic on crime in Los Angeles and Indianapolis. *Journal of Criminal Justice*, *68*. Retrieved from https://www.sciencedirect.com/science/article/pii/S0047235220301860

Onyango, M. A., Resnick, K., Davis, A., & Shah, R. R. (2019). Gender-based violence among adolescent girls and young women: A neglected consequence of the West African Ebola outbreak. In *Pregnant in the time of ebola* (pp. 121–132). Springer. Retrieved from https://doi.org/10.1007/978-3-319-97637-2_8

Payne, J. L., & Morgan, A. (2020). *COVID-19 and violent crime: A comparison of recorded offence rates and dynamic forecasts (ARIMA) for March 2020 in Queensland, Australia.* Australian National University. Retrieved from https://ideas.repec.org/p/osf/socarx/g4kh7.html

Peterman, A., Potts, A., O'Donnell, M., Thompson, K., Shah, N., Oertelt-Prigione, S., & van Gelder, N. (2020). Pandemics and violence against women and children. *Center for Global Development working paper*, *528*. Retrieved from https://www.cgdev.org/publication/pandemics-and-violence-against-women-and-children

Piquero, A. R., Riddell, J. R., Bishopp, S. A., Narvey, C., Reid, J. A., & Piquero, N. L. (2020). Staying home, staying safe? A short-term analysis of COVID-19 on Dallas domestic violence. *American Journal of Criminal Justice*. Retrieved from https://doi.org/10.1007/s12103-020-09531-7

Podaná, Z., et al. (2010). Reporting to the police as a response to intimate partner violence. *Sociologickỳ časopis/Czech Sociological Review*, *46*(03), 453–474. Retrieved from https://www.jstor.org/stable/41132867

Sanz-Barbero, B., Linares, C., Vives-Cases, C., González, J. L., López-Ossoriod, J. J., & Díaz, J. (2018). Heat wave and the risk of intimate partner violence. *Science of The Total Environment*, *644*, 413-419. Retrieved from https://doi.org/10.1016/j.scitotenv.2018.06.368

UN Women. (2020). *Covid-19 and ending violence against women and girls.* Policy Brief. Retrieved on August 6, 2020 from: https://www.unwomen.org/-/media/headquarters/attachments/sections/library/publications/2020/issue-brief-covid-19-and-ending-violence-against-women-and-girls-en.pdf?la=en&vs=5006.

Van Gelder, N., Peterman, A., Potts, A., O'Donnell, M., Thompson, K., Shah, N., & Oertelt-Prigione, S. (2020). Covid-19: Reducing the risk of infection might increase the risk of intimate partner violence. *EClinicalMedicine*, *21*. Retrieved from 10.1016/j.eclinm.2020.100348

Wagers, S. M. (2020). *Domestic violence growing in wake of coronavirus outbreak.* The Conversation. Retrieved on August 5, 2020 from: https://theconversation.com/domestic-violence-growing-in-wake-of-coronavirus-outbreak-135598.

Appendix

In this appendix we present further descriptive details of the data used.

A   MPS Domestic Violence Crime Data

Data on the daily count of DV-related crimes recorded by the MPS was obtained by a Freedom of Information request. Our data covers the period April 1, 2015, to June 22, 2020. The data exhibit some general time patterns. Figure A.1 shows the average daily count of DV-related crimes by year, month and day of the week. Panel A shows that the daily average has increased from about 205 in 2015 to 245 in 2019, which corresponds to an average annual increase of 4.5 percent. The data for 2020 covers only the time up to June 22 and, of course, incorporates the lockdown period. The steady growth over time makes simple comparisons – for instance comparing a given week to the corresponding week a year before – somewhat problematic. Panel B shows a strong seasonal pattern, with reported DV-incidence being lower in the first and fourth quarter and higher between late spring and end of summer. Finally, panel C shows a strong day-of-the-week pattern, with incidence being about 10 percent higher on weekends than during weekdays.

B   Weather Data

We use data on daily average temperature (in °C) and rainfall (in mm) from the London Heathrow weather station covering the full sample period. As noted, April and May of this year were unusually warm and dry. Panel A of Figure A.2 shows the daily average temperature (in °C) with the horizonal red lines indicating the average temperature by month over the past five years. The second half of May was also unusually warm. Panel B shows rainfall per day, indicating that the key period from the beginning of the lockdown through to early June saw barely any rainfall at all.

Figure A.1: Trends, seasonality and weekly patterns of DV-reported crimes

(A) Trends

(B) Seasonality

(C) Day-of-Week
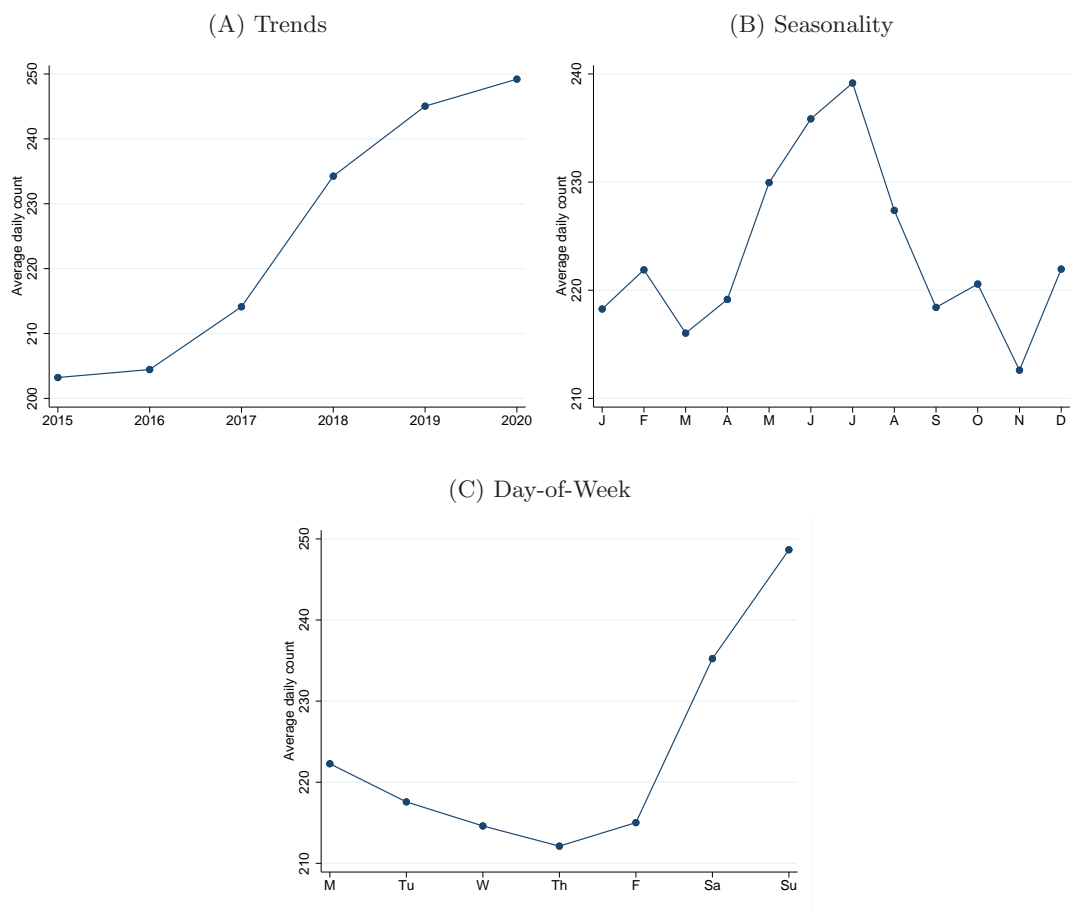
Notes: The sample consists of daily counts of domestic violence-related crimes recorded by the London MPS between 1 April 2015 and 22 June 2020.

Figure A.2: Daily average temperature and rainfall since February 2020

(A) Temperature
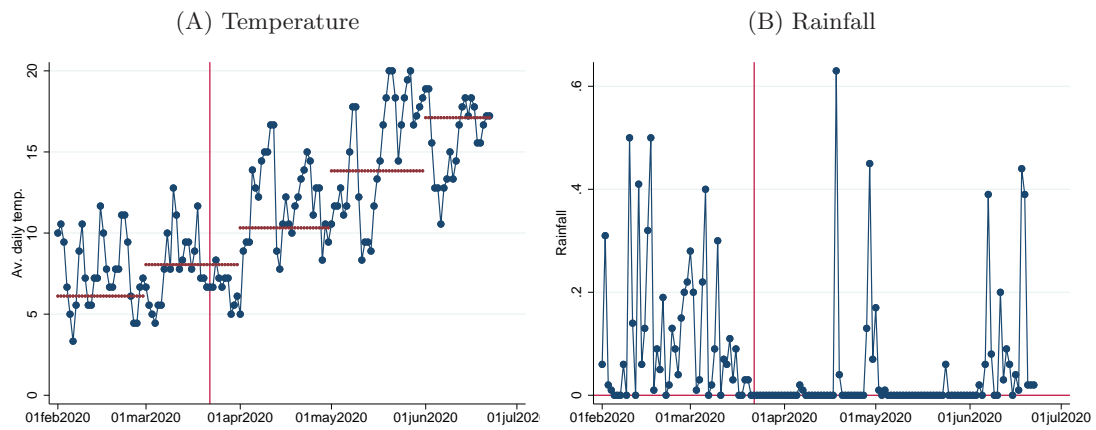
(B) Rainfall

Notes: The figure shows the daily average temperature in degrees Celsius and the daily rainfall in *mm*. The data is from National Climatic Data Centre and is for the London Heathrow weather station.