



# An Investigation of Models for Under-Reporting in the Analysis of Violence Against Women in Italy

Silvia Polettini<sup>1</sup> · Serena Arima<sup>2</sup> · Sara Martino<sup>3</sup>

Accepted: 16 September 2023  
© The Author(s) 2023

## Abstract

Violence against women is still one of the most widespread and persistent violations of human rights. Despite this, a significant gap of comprehensive, reliable and up-to-date figures on such a largely uncovered phenomenon remains. To develop efficient and effective policy and legal responses to gender-based violence, accurate data are necessary. Surveys specifically designed to quantify the number of victims of gender violence return prevalence estimates at a given time, and assess the under-detection of violence and its drivers. However, the last Italian Women's Safety Survey was conducted by ISTAT in 2014. Given the substantial under-reporting affecting official counts of violence reports to the police, and the lack of recent survey data, up-to-date prevalence estimates cannot be produced. Designing ad hoc techniques suitable to pool data arising from different sources, first of all official police reports, and accounting for the under-reporting, is crucial to understand and measure violence against women to return a realistic picture of this greatly underrated phenomenon and assess its scope. We use publicly available registry data on violence reports in 2020 as a primary source to provide improved estimates of gender violence in the Italian regions, by introducing a Bayesian model that supplements the observed counts with a pool of auxiliary information, including socio-demographic indicators, data on calls from 1522 helpline number and prevalence estimates from previous surveys, while explicitly modelling the reporting process using covariates and external information. We propose using statistical models for the analysis of misreported data to improve the understanding of the problem from a methodological point of view and to get insights into the complex dynamics of the phenomenon in Italy.

**Keywords** Bayesian model · Compound Poisson distribution · MCMC · Under-reporting · Violence against women

## 1 Introduction

Violence against women, with its various forms of physical, sexual, and psychological abuses, is still one of the most widespread and persistent violations of human rights. It has been argued (Dobash et al. 2014) that such kind of violence is an extension of patriarchal society to express power and control of men over women, their partners in particular.

---

Extended author information available on the last page of the article

Despite the debate in the literature (Johnson, 2006; Melton & Sillito, 2012), with some researchers (e.g. Straus, 1999) having argued that in intimate partner relationships women are as likely as men to perpetrate violence against their partners, the feminist literature stresses that intimate partner violence emerges primarily as an asymmetrical problem of men's violence to women, and findings illustrate that "women's violence does not equate to men's in terms of frequency, severity, consequences and the victim's sense of safety and well-being", see Dobash & Dobash (2004) and Dugan et al. (1999).

Gender-based violence involves the violation of many of the fundamental rights declared in the Charter of Fundamental Rights of the European Union, and commitments to its eradication have been undertaken worldwide, first of all the United Nations Convention on the Elimination of all forms of Discrimination Against Women.<sup>1</sup> In 2011, the Council of Europe has drawn a legal framework at pan-European level to protect women against all forms of violence, signing the Istanbul Convention,<sup>2</sup> whose pillars are to prevent, prosecute and eradicate violence against women and domestic violence. The EU Gender Equality Strategy for 2020–2025 emphasizes gender-based violence as one of our societies' biggest challenges, deeply rooted in gender inequality. Finally, eliminating all forms of violence against all women and girls in the public and private spheres is one of the targets of the 2030 Agenda for Sustainable Development, adopted by all United Nations Member States in 2015.

Violence against women is not only an effect of gender inequality, but also serves to reinforce the gender-power unbalance: indeed the perpetration of violence may be seen as a strategy to achieve subordination of women. Despite the growing coverage from the media in the most striking aspects of the phenomenon, especially femicide and rape, violence against women is still not part of the social awareness, and it remains a largely uncovered phenomenon mainly because it is a social construct with multifaceted characteristics, that only emerges when the victims seek help and report their experiences. There is a variety of reasons (taboos, gender prejudices, stigma, self-blame, social desirability, economic instability, insecurity, fear, privacy concerns,...) why the propensity to report an abuse is low, especially in relation to intimate partner violence (Leon et al., 2022). An interesting analysis is in Orchowski et al. (2022), where Twitter hashtag #WhyIDidntReport were examined to study barriers to reporting for victims of rape. Moreover, such propensity, correlates with socio-demographic characteristics.

The under-reporting hinders the planning of appropriate awareness-rising policies and prevention interventions, such as the funding of local premises that can practically support women. By consequence, it contributes to protect and perpetrate the violent environments.

Results from the 2014 EU Violence against women survey show that 33% of women in the EU have experienced physical or sexual violence and that 67% did not report the most serious incident of partner violence to the police or other organizations.<sup>3</sup> WHO, based on a sample of 852 million women in 2000–2008, estimated that nearly 1 in 3 women have experienced a form of violence at least once in their lifetime.<sup>4</sup> Through a comprehensive

<sup>1</sup> Convention on the Elimination of All Forms of Discrimination against Women (CEDAW). United Nations General Assembly, New York, 18 December 1979.

<sup>2</sup> Council of Europe Convention on Preventing and Combating Violence Against Women and Domestic Violence: Istanbul, 11.V.2011, Council of Europe Treaty Series - No. 210, 2011.

<sup>3</sup> <https://fra.europa.eu/en/content/media-memo-violence-against-women-eu-wide-survey>

<sup>4</sup> <https://www.who.int/news/item/09-03-2021-devastatingly-pervasive-1-in-3-women-globally-experience-violence>

review of studies conducted between 2000 and 2018, Sardinha et al. (2022) have obtained a collection of data covering 90% of the global female population of 15 years or older, from which they estimate that among ever-partnered women aged 15–49 years, 27% have experienced physical or sexual, or both, intimate partner violence in their lifetime (Sardinha et al., 2022). These figures confirm that physical and sexual violence remain pervasive in the lives of women and adolescent girls worldwide.

There is agreement in the literature on the topic (Copp et al., 2019, Capaldi et al., 2012, Garcia-Moreno et al., 2006) that gender violence is significantly associated with social and cultural circumstances and varies across different life stages. Education and labour participation, increasing women access to opportunities and promoting their economic independence and social integration, should reduce the risk of victimization. Indeed Ventura et al. (2022) analyze homicide mortality by sex and its association with socio-demographic characteristics through a longitudinal retrospective study covering years 2012–2018 and find for both sexes a negative association (notably, stronger for males than for females) between education and homicide mortality. At the same time, both a higher reporting rate and an opposite backlash effect can be envisaged (see e.g. Whaley, 2001, and references therein), and mixed conclusions arise in the literature. Risk factors for the perpetration of violence include contextual characteristics of partners, developmental characteristics and behaviors of the partners, relationship influences and interactions. Exposure to violent relationship (see e.g. Dugan et al., 1999) clearly affects the event rates. Association between violence-supportive attitudes and the perpetration of violence against women is described in several studies (see e.g. Flood and Pease, 2009, Ambrosetti et al., 2013, and references therein). Not only men's sexist, misogynist or patriarchal attitudes, but also adherence to conservative gender and sexuality norms are associated to both tolerance and use of violence against women; moreover, gender roles norms and women's beliefs towards gender roles and sexuality mould the perceptions of their experiences, their response to violence and thus affect the propensity to disclose a violence. Broadly speaking, all the above relations are the effect of the cultural context. Factors such as labour market participation and socioeconomic status seem to impact on attitudes toward violence against women; associations between economic and social disadvantage and higher violence rates, that can be attributed to both violence-prone attitudes and higher exposure to violence, are reported in Flood & Pease (2009). Also, educational attainment and age are considered as factors determining individuals' perceptions of violence against women and therefore to affect the propensity to report abuses.

Specialized surveys are the primary source of information on the phenomenon, however the last victimization survey for Italy dates back to 2014. Lack of data make an up-to-date assessment of the phenomenon difficult in terms of both its drivers and the under-reporting. The most recent data source on violence against women currently is the official register of crime statistics administered by the Italian Ministry of Interior and released yearly. However, despite their valuable informative content, such registry data are prone to quality issues including in particular under-reporting. A discussion of these issues, comparing evidences from police registers and victimization surveys in the Netherlands, is reported in Wittebrood and Junger (2002). As a consequence these data are not sufficient to return the picture of such a complex phenomenon, but can be fruitfully exploited if supplemented by other data sources.

A draft law has just been issued<sup>5</sup> dealing with national statistics on gender based violence. As a consequence, a whole system of information on the subject will be available under the umbrella of the Italian National Strategic Plan on Male Violence Against Women, with more detailed, accurate and timely information both from registry data originating from justice databases, police activities, hospitals and emergency rooms, and survey data. Such a shift in information will make it possible in the future to investigate the spatio-temporal trends of the phenomenon using data from two specialized surveys (Survey on users of Centers Against Domestic Violence and Women's Safety Survey), data from helpline calls, reports received by the Police forces and court data, and data from hospitals and emergency rooms. Pooling all the available information is crucial to understand and measure such a hidden phenomenon better. Indeed surveys may allow to estimate the prevalence of violence against women and to define a profile of the women at risk, but they do not allow a continuous monitoring of the phenomenon, nor deriving estimates at a fine geographical detail. Moreover, given the sensitivity of the subject, survey results depend on "research methods, definitions of violence, sampling techniques, interviewer training and skills, and cultural differences that affect respondents' willingness to reveal intimate experiences" (Venis & Horton, 2002).

In this paper we describe a first attempt to exploit the Italian official register of crime statistics, administered by the Italian Ministry of Interior, in conjunction with auxiliary data both to assess the intensity of the phenomenon and the extent of the under-reporting at the regional level in Italy. We focus on data for years 2019 and 2020, considering reports on battering, stalking, and sexual violence.

Our interest is investigating the drivers of both violence and under-reporting, trying to capture the extent of under-reporting and its variation across the Italian regions. To our knowledge, no previous studies have addressed the issue of under-reporting in this context.

The paper is organized as follows: in Sect. 2 we discuss the available information for Italy, in Sect. 3 we motivate our approach, while in Sect. 4 we review the main statistical models for under-reported counts presented in the literature. Section 5 describes the model adopted in relation to data obtained from the official register of crime statistics. In Sect. 6, we describe the application of the model to the Italian data and discuss the results obtained. We conclude with a short conclusion in Sect. 7.

## 2 The Italian Figures

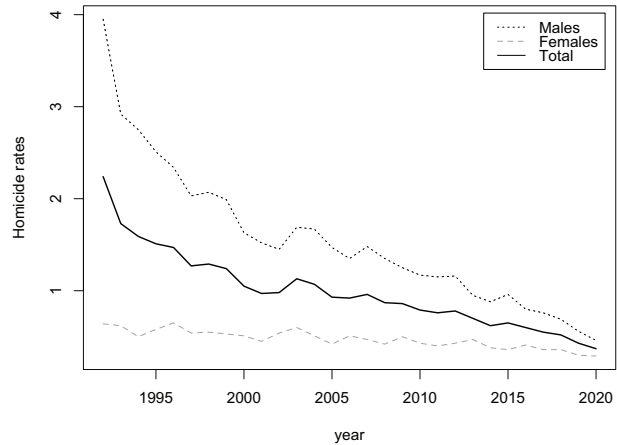
According to data from the last women victimization survey, conducted by Istat in 2014, a 31.5% of women aged 16–70 had suffered a form of physical or sexual violence across their lifetime, with large improvements with respect to the previous survey (carried out in 2006); however, the more serious types of violence (rapes and attempted rapes) are reportedly unchanged, whereas the seriousness of the violence suffered is increasing.<sup>6</sup>

In the absence of updated prevalence figures, official registry time series data on homicides, being not affected by under-reporting, may help in describing the phenomenon of violence against women over time. Although homicide rate in Italy is currently one of the

<sup>5</sup> Law 5/5/2022, n. 53 "Disposizioni in materia di statistiche in tema di violenza di genere".

<sup>6</sup> [https://www.istat.it/it/files//2019/11/Violence-against-women-\\_2014.pdf](https://www.istat.it/it/files//2019/11/Violence-against-women-_2014.pdf)

**Fig. 1** Victims of voluntary homicides by gender, years 2002–2020 (rates per 100,000 inhabitants)



lowest within the European context (0.48 in 2020, with a EU average rate of 0.89<sup>7</sup>), a large gender difference occurs. As shown in Fig. 1, in Italy the homicide rates have substantially decreased over time, but the reduction in male rates has been much higher than for women, corresponding to a decrease in deaths due to organised and common crime. On the other hand, women homicides are mostly family or intimate partner related: according to Istat<sup>8</sup> a 61.3% of woman homicides in 2019 and a 57.8% in 2020 were perpetrated by partner or ex-partner, with a striking sex imbalance: in 2019, 83.8% of women homicides and just 27.9% of men homicides have occurred in this context.<sup>9</sup>

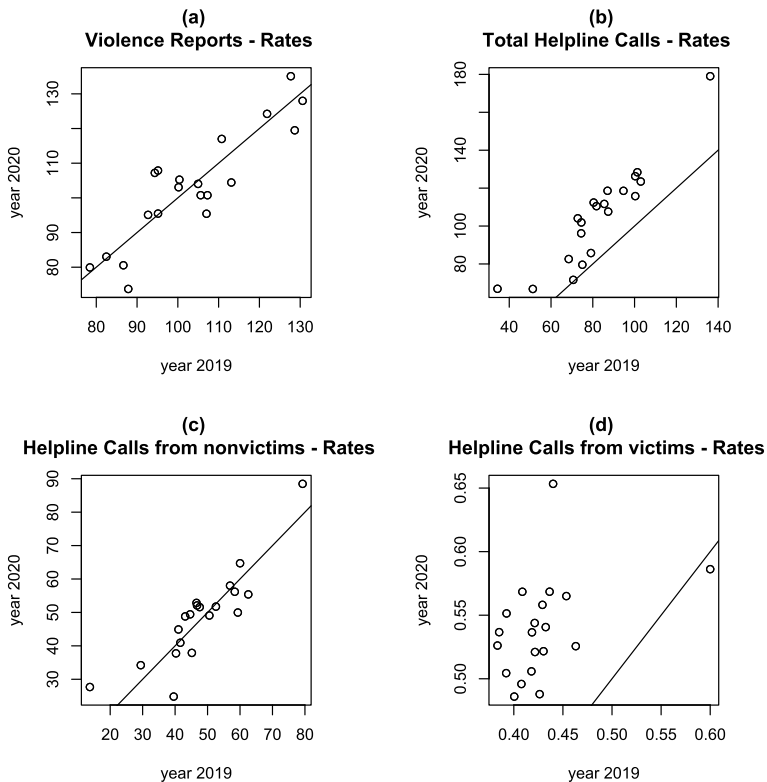
In 2020, lockdown and self-isolation policies, together with the strong psychological and socio-economic effects of the COVID-19 outbreak have exposed women and children worldwide to higher risks of abusive domestic relationships, producing what can be called an “invisible pandemic” (see Usta et al., 2021; Merenda et al., 2021; Viero et al., 2021; Seidenbecher et al., 2023, among the others). In Italy, whereas the decrease in female homicide rates was small in 2020 (about 3% compared to 2019), it was much higher (nearly 18%) for males over the same period. This can be interpreted an effect of the lockdown policies, that protected males while leaving women exposed to the major risk factors. According to Istat,<sup>10</sup> the male excess ratio of victims of homicides in 2020 (170 men over 116 women killed, namely about three men for every two women) is the lowest ever recorded; just three years prior, in 2017, the same ratio was 2 men every woman. Analysing the official register of crime statistics, Istat also reports a substantial gender difference in homicides, in that women are most often killed by partners/ex partners (75% in 2020), whereas the majority of homicides involving men are perpetrated outside the family context. Inspecting the relation between the victim and the perpetrator, the rate of female homicide committed by partner/ex partner is unchanged during 2019–2020 whereas for men the rate of homicides in which the perpetrator is unknown to the victim has fallen by 11% over the same period (see also the data available at <https://www.istat.it/it/archivio/263847>).

<sup>7</sup> [https://www.istat.it/it/files//2022/11/REPORT-VITTIME-DI-OMICIDIO\\_2021.pdf](https://www.istat.it/it/files//2022/11/REPORT-VITTIME-DI-OMICIDIO_2021.pdf)

<sup>8</sup> [https://www.istat.it/it/files//2022/11/REPORT-VITTIME-DI-OMICIDIO\\_2021.pdf](https://www.istat.it/it/files//2022/11/REPORT-VITTIME-DI-OMICIDIO_2021.pdf)

<sup>9</sup> <https://www.istat.it/it/files//2021/03/perpetrators-victims-homicide-2018-19.pdf>

<sup>10</sup> [https://www.istat.it/it/files//2021/11/EFFETTI\\_PANDEMIA\\_-VIOLENZA\\_D\\_GENERE.pdf](https://www.istat.it/it/files//2021/11/EFFETTI_PANDEMIA_-VIOLENZA_D_GENERE.pdf)



**Fig. 2** Inspecting the effect of pandemic on available data on violence: **a** Official reports of violence (rates for 100,000 inhabitants); **b** Total calls to the helpline number 1522 (rates for 100,000 inhabitants); **c** calls to the helpline number 1522 for reasons other than violence (rates for 100,000 inhabitants); **d** proportion of victims among calls to the helpline number 1522

The surge in the number of calls to the Italian National Helpline Service observed in 2020 (+79.5% with respect to the previous year<sup>11</sup>) testifies the increased vulnerability of women to domestic violence and abuse; at the same time, the systematic rise in calls during the strict lockdown periods can be ascribed to fewer opportunities for women to benefit from social and protective networks and to report an abuse to the police. Indeed, the above mentioned analysis carried by Istat illustrates a systematic and peculiar decrease in the number of reported cases of violence to the police authorities for year 2020, concentrated during the lockdown periods. Analysing the 1522 data released by Istat<sup>12</sup> and pertaining to users classified as "victims", we notice that the absolute number of such calls has increased compared to the previous years. At the same time, among victims referring to the helpline number, those for which the reporting could not be further investigated is much larger than previous years and is particularly high during the first semester, possibly due to insecurity associated with the presence of the perpetrator at the time of the call. Among calls for

<sup>11</sup> <https://www.istat.it/it/archivio/257704>

<sup>12</sup> <https://www.istat.it/it/archivio/278050>

which the operator could investigate for the user having reported to the police, the percentage of users that do report violence to the police is lower for year 2020 and 2021 in comparison to the previous years, and overall is about 17% considering both reports and withdrawals. Note that the same figures for 2018 and 2019 are about 19 and 20%, respectively. These figures confirm that the propensity to report is quite low, especially in 2020, and that often referring to the police is not the primary option. Figure 2 compares the available data for 2019 and 2020: whereas yearly rates of reported violence for the crimes of battering, stalking, and sexual violence are comparable for the two years as shown in panel (a), the calls to 1522 have a rise in 2020 (panel (b)) which is due to a substantial and generalised increase in calls from victims of violence (panel (d)), a rise that is not apparent in panel (c) where only calls from users other than victims are considered.

### 3 Hierarchical Models for Misreported Data Borrowing Information From Multiple Sources

Due to the substantial under-reporting affecting violence counts, and to the lack of recent survey data (the last available surveys date back to 2014 and 2006, respectively, and cannot provide up-to-date figures on the phenomenon), estimates of the prevalence of the phenomenon in Italy have not been produced recently. Police reports are currently the only available official figures to assess violence against women in Italy, yet they are well known to drastically downplay the scale of violent crime against women, as already commented.

Technically, not accounting for under-reporting is shown to bias inferences (see e.g. Chen et al., 2022) and conclusions (Bettio et al., 2020).

The incorporation of the misreporting mechanism into the probabilistic modeling of the outcome of interest, as a further source of uncertainty to be accounted for, represents a way to pool the available information on crime reports to produce prevalence estimates in the absence of specific survey data, possibly at finer geographical detail than it would be permitted by the same surveys.

From the statistical point of view, it is important to exploit all the available information (Gelles, 2000; Wiśniowski et al., 2020) to give a picture of a hard to measure phenomenon, aware of the limitations and issues typical of each data source. Several approaches have been exploited in the literature to cope with the under-reporting problem: an established demographic literature (e.g. Brass, 1996) deals with indirect methods for correcting vital statistics; in this paper we refer instead to hierarchical models that allow us to model the reporting mechanism explicitly.

We describe a first investigation of the approach by using publicly available registry data on violence reports as a primary source to provide estimates of the phenomenon in the Italian regions. Then we introduce a Bayesian model that supplements the observed counts with a pool of auxiliary information, including socio-demographic indicators and prevalence estimates from previous surveys. Using these data from years 2019 and 2020, we address the under-reporting issue by considering the approach pursued e.g. in Stoner et al. (2019) and Chen et al. (2022) using covariates and external information from Women's Safety Survey<sup>13</sup> to inform the reporting process. This allows us to assess the extent of violence against women and the level of under-reporting and their heterogeneity across regions. We investigate the

<sup>13</sup> Istat, Indagine Multiscopo sulle Famiglie: Sicurezza delle Donne, 2014.

effectiveness of the auxiliary data in informing the data generating process and especially the misreporting process.

#### 4 Statistical Models for Under-Reported Counts

Let us denote by  $T_i$  the number of events of violence against women occurred over a set of  $m$  regions. We assume

$$T_i \sim \text{Poisson}(E_i\theta_i), \quad i = 1, \dots, m \quad (1)$$

where  $\theta_i$  is the event rate and  $E_i$  is the number of women in the  $i$ -th area. Let us further assume that we can relate the rates  $\theta_i$  to a set of covariates  $X_1, \dots, X_p$  through a regression model such as:

$$\log(\theta_i) = \beta_0 + \beta_1 X_{1i} + \dots + \beta_p X_{pi}. \quad (2)$$

As discussed, it is widely recognised that the events of violence are only partially reported, so that we observe reported counts  $Y_i, Y_i \leq T_i, i = 1, \dots, m$ . The aim of the analysis is usually twofold, namely estimating the event intensities  $\theta_i, i = 1, \dots, m$  and the effects of the covariates, and predicting the true counts  $T_i, i = 1, \dots, m$  given the available information.

Several extensions of the Poisson model have been proposed for counts subject to under-reporting; all of them rely on expert information and/or auxiliary data on the reporting process. Part of the literature is based on censored likelihoods, that involve a binary variable indicating the presence of under-reporting in each area: for instance, Bailey et al. (2005) extended the censored Poisson regression model in Caudill and Mixon (1995) assuming the counts of suspected areas are the lower bound for the true non-observed counts. This approach allows estimation of the incidence rate and the probability of under-reporting, but requires precise information about which areas are misreported and indeed Bailey et al. (2005) use ad-hoc procedures to partition the areas into observed or censored. As stressed in Stoner et al. (2019), the approach based on censored likelihoods cannot allow for the severity of the under-reporting, which limits the quality of the prediction of the true counts. Moreover, it requires that at least some of the geographical units are completely observed to estimate the underlying counts.

A different approach relies on the specification of a hierarchical model that assumes that all areas are potentially under-reported. The true counts are described in terms of Poisson stopped-sum distributions (see Johnson et al., 2005, Section 4.11). A vast literature (see e.g. Whittemore and Gong, 1991, Winkelmann, 1996, Stoner et al., 2019, Dvorzak and Wagner, 2016, Li et al., 2003 to mention but a few) follows this approach, under which the severity, rather than the mere chances of under-reporting, can be estimated along with the true count prevalence.

The compound Poisson model (CPM) that we describe next is a particular case within this class: let us assume that each individual  $j$  in region  $i$  has probability  $\epsilon_i$  to report the event, independently,  $j = 1, \dots, T_i, i = 1, \dots, m$ , and independent on  $T_i$ , whereby the observed counts can be defined as  $Y_i = \sum_{j=1}^{T_i} Z_{ij}$  and  $Y_i | T_i, \epsilon_i \sim \text{Bin}(T_i, \epsilon_i)$ .

Combining the above with the Poisson assumption (1) and marginalising over  $T_i$  this is the so called compound Poisson model (CPM), under which

$$Y_i | \theta_i, \epsilon_i \sim \text{Poisson}(E_i\theta_i\epsilon_i). \quad (3)$$



Moreover, for the number of non-reported events it holds that

$$T_i - Y_i \mid \theta_i, \epsilon_i \sim \text{Poisson}(E_i(1 - \theta_i)\epsilon_i). \tag{4}$$

Within the CPM, Stoner et al. (2019) and Chen et al. (2022) propose a Bayesian hierarchical model for count data including covariates to inform not only the true count-generating process but also the under-reporting mechanism, while also allowing for complex spatio-temporal structures. They model the reporting probability hierarchically through a logistic regression using suitable covariates:

$$\text{logit}(\epsilon_i) = Z_i' \gamma. \tag{5}$$

An issue of identifiability arises for model (3): indeed, whereas the product  $\theta_i \epsilon_i$  is identified from the observations,  $\theta_i$  and  $\epsilon_i$  are not, since the same likelihood arises for different combinations of  $\theta_i$  and  $\epsilon_i$  that yield the same value of  $\theta_i \epsilon_i$ . In non-identified models, the Bayesian approach is particularly appropriate in that, as long as the prior is proper, inferences are guaranteed because the posterior distribution is well defined and therefore MCMC algorithms can be designed to simulate from the posterior (see Florens and Simoni, 2021), where the role played by identification in the Bayesian analysis of statistical models and the differences with the frequentist approach are discussed). In the absence of any completely reported observations, the non-identifiability issue can be addressed by introducing external information on one or both of the models for  $\theta_i$  and  $\epsilon_i$ , as discussed in the literature: models can be fitted by employing validation data or by specifying priors on the parameters based on expert knowledge to inform the reporting process. See the discussion in Stoner et al. (2019), de Oliveira et al. (2022) and references therein for details. de Oliveira et al. (2017) propose a more flexible approach in which areas are clustered according to the probability of under-reporting based on expert opinion and auxiliary data. Identifiability is shown to be guaranteed by specifying an informative prior on the under-reporting probability for the best data quality cluster.

### 5 A Model for Under-Reported Violence Against Women in Italy

From the official register of crime statistics administered by the Italian Ministry of Interior, we consider the reported number of female victims of battering, stalking, and sexual violence at the regional level, that we denote by  $Y_i$ ,  $i = 1, \dots, 20$ . Due to the exceptionality of year 2020, we also consider data for year 2019 to investigate the possible changes occurred.

Pursuing the approach outlined in the previous section, for each year we consider the following model:

$$Y_i \mid \theta_i, \epsilon_i \sim \text{Poisson}(E_i \theta_i \epsilon_i) \tag{6}$$

$$\text{log}(\theta_i) = \beta_0 + \beta_1 X_{1i} + \dots + \beta_p X_{pi} \tag{7}$$

$$\text{logit}(\epsilon_i) = \gamma_0 + \gamma_1 Z_{1i} + \dots + \gamma_r Z_{ri} \tag{8}$$

in which the observed counts are modelled by a CPM, and the event intensities as well as the reporting probability are regressed on a set of auxiliary covariates.

A model explicitly incorporating the reporting probabilities  $\epsilon_i$ , would allow us to assess the extent of under-reporting and its variation across regions. Through the relation (4), the

model can be exploited to return predictions of the partially reported counts  $T_i$ . Finally, the analysis may illustrate which covariates are the most important drivers of the intensity of violence and of the under-reporting, respectively.

To model the event intensities we selected prevalence estimates from the last Women's Safety Survey and the proportion, among of calls to 1522, of those reporting episodes of violence, the social deprivation index, and a measure of woman employment. A retrospective study in Yakubovich et al. (2020) based on individual data documents that exposure to neighborhood deprivation over the first 18 years of life is associated with women's increased risk of experiencing intimate partner violence in early adulthood. We considered 5-year averages of the index to account for the large sampling error of the available regional estimates of the deprivation index for some geographical units, and to reflect the short-term effect of the variable across time.

Labour participation guarantees women's economic independence and is associated with higher socio-economic status; it also reflects gender role attitudes. Low employment rate can be considered a measure of women's vulnerability. Also, labour participation gap between women with preschool-age children and women without children highlights inequalities in women's economic independence and socio-economic status and reflects gender role attitudes. In our model we tested either of these variables in turn.

The literature on the subject (see e.g. Usta et al., 2021) emphasizes that during the Covid-19 outbreak, in analogy with other crises, with growing household tension, women are more affected by abuses not only because they are stuck with the perpetrator and shelters and other support institutions are more difficult to access, but also because they tend to lose their financial empowerment. Moreover, in Italy women are expected to bear the household responsibility so they might have sacrificed their jobs in response to the pandemic, especially in households with school and preschool-age children. A partial evidence of this can be found in the decrease in the women employment rate in 2020 compared to 2019, and in the considerable reduction in almost all regions of the ratio of the employment rate for women aged 25–49 with preschool-age children to women without children.

The underlying reporting mechanism is modelled by a logistic regression using the public trust in the judicial system obtained by the Istat Survey on Aspects of daily life, the rate of calls arriving at the 1522 helpline service by non-victims, and a "South" binary covariate. The latter was introduced as a proxy of the social and cultural differences and stereotypes that may affect the social construction of gender-based violence. The trust in the judicial system is expected to be strongly related to a person's propensity to report, while calls to the 1522 helpline service by non-victims may reflect awareness of the problem and thus be related to the propensity to report an abuse, or reversely may be a first attempt to seek for help in contexts characterized by conservative gender roles norms and beliefs. We also considered the female to male ratio of higher education rates as an indicator of gender equality (see Di Noia, 2002, and references therein), but its effect was not significant in the fitted model.

All the data are either available on the subject-specific Istat website<sup>14</sup> or can be obtained from Istat databases, in particular the system of "BES" (Equitable and Sustainable Well-being) indicators.<sup>15</sup>

Except for the intercept in (8), for the regression coefficients we specified vague normal distributions setting  $\beta_j \sim N(0, 10^2)$ ,  $j = 0, \dots, p$  and  $\gamma_j \sim N(0, 10^2)$ ,  $j = 1, \dots, r$ .

<sup>14</sup> <https://www.istat.it/it/violenza-sulle-donne>

<sup>15</sup> <https://www.istat.it/en/well-being-and-sustainability/the-measurement-of-well-being/indicators>

**Table 1** Posterior means of regression coefficients in the models for the intensity and the reporting process, and their 0.95 credible intervals

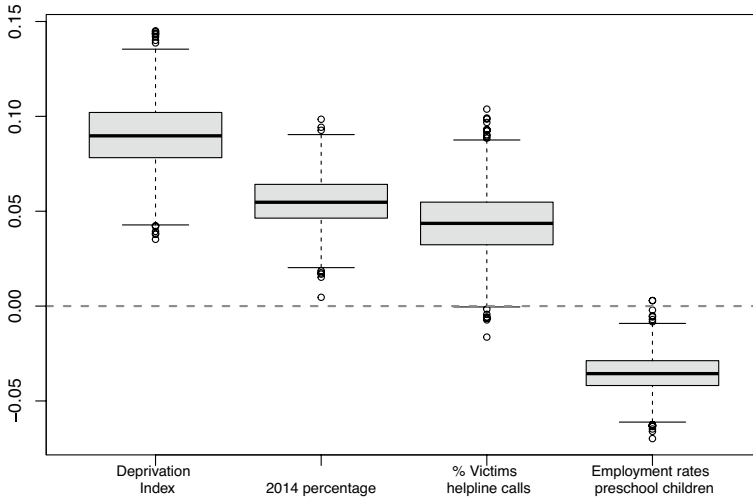
	Year 2019		Year 2020	
	Post. mean	0.95 CI	Post. mean	0.95 CI
<i>Intensity model</i>				
Deprivation index	0.090	(0.057, 0.125)	0.164	(0.127, 0.201)
2014 percentage	0.055	(0.030, 0.081)	0.066	(0.036, 0.096)
% Victims helpline calls	0.044	(0.011, 0.076)	0.059	(0.029, 0.090)
Employment rate ratio	- 0.035	(- 0.054, - 0.016)		
Women Employment rate			- 0.093	(- 0.130, - 0.055)
<i>Reporting model</i>				
South	- 0.309	(-0.393, -0.223)	- 0.485	(-0.600, -0.364)
Trust in the judicial system	0.070	(0.042, 0.096)	0.100	(0.077, 0.123)
1522 calls - nonvictims	- 0.073	(-0.101, -0.045)	- 0.053	(-0.081, -0.026)

The findings from the 2014 Women’s Safety Survey and more recent data from the 1522 helpline service were exploited to inform the reporting process. Indeed, as discussed previously, the model parameters  $\theta_i, \epsilon_i$  are not identified by the observed data, but using proper priors informed by external information, MCMC algorithms converge appropriately to the posterior. To tackle the issue, as suggested in Stoner et al. (2019), it suffices to specify an informative prior on the intercept of the logit model for the reporting probability based on expert opinion or induced from other studies. This is convenient, as one can rely on previous data or expert opinions about the mean reporting rate. A proposal in Chen et al. (2022) is to elicit a prior on the reporting probability  $\pi_0$  that would be observed at the “average” value of all the covariates, by defining a beta prior with prespecified mode and  $q$ -quantile. In turn, this induces a prior for the intercept  $\gamma_0$  of the logit model. We have worked on the logit scale and specified a normal prior for  $\gamma_0 \sim N(-1.516, 0.05^2)$  for 2019 and  $\gamma_0 \sim N(-1.658, 0.05^2)$  for 2020 corresponding to a prior mean on the reporting probability at the “average” value of the covariates,  $\pi_0$  of 0.18 for year 2019 and 0.16 for year 2020. This implies that with prior probability 0.99, the reporting rate is within 0.162 and 0.200 for year 2019 and within 0.143 and 0.178 for year 2020.

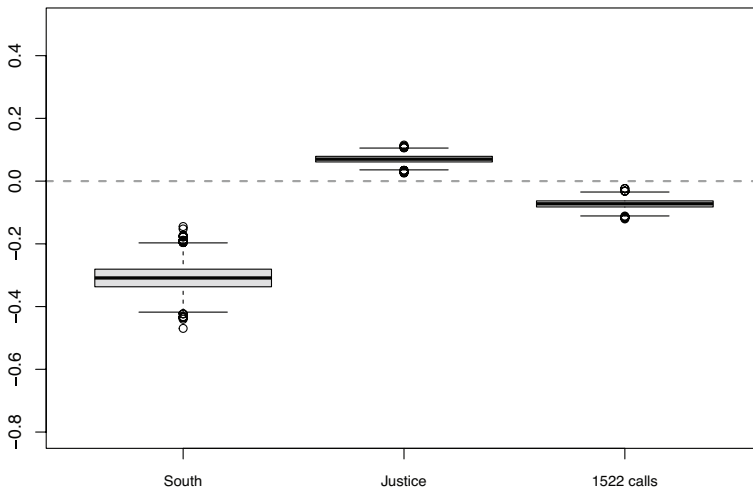
## 6 Results

As the posterior distributions are not available in closed form, Markov Chain Monte Carlo (MCMC) techniques are used to draw samples from the posterior distributions of model parameters. The model has been estimated using NIMBLE (de Valpine et al., 2017). We allow for 2 chains, each with 400,000 iterations (200000 burn-in and thinned over 100). Trace plots of the model parameters were inspected to check convergence of the MCMC chains.

All the covariates have been standardised. For the sake of clarity we report the covariates used in the components (7) and (8) of the model:



**Fig. 3** Year 2019: posterior distributions of the Poisson regression parameters for the intensity of violence against women



**Fig. 4** Year 2019: Posterior distributions of the logistic regression parameters for the probability of reporting a violence against women

- $X_1$ : Social deprivation Index
- $X_2$ : Estimated prevalence of violence in the last 12 months from the 2014 Women’s Safety Survey
- $X_3$ : Proportion of calls from victims among valid calls to the national helpline number 1522

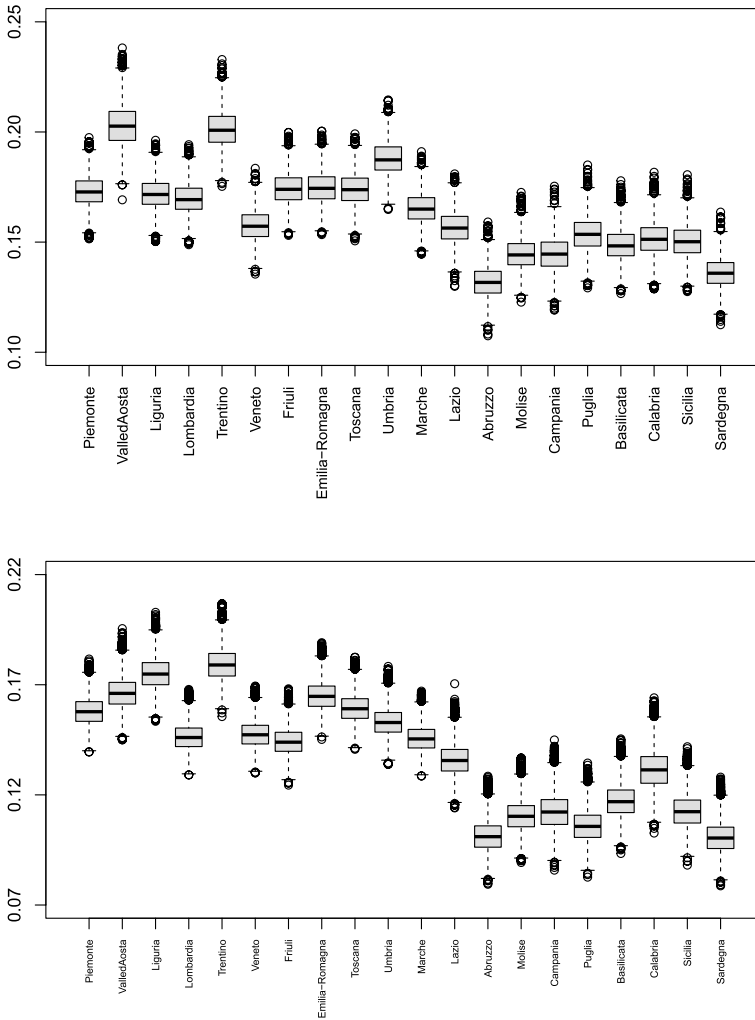
- $X_4$ : Employment variable: ratio of the employment rate for women aged 25-49 with pre-school-age children to women without children for year 2019; woman employment rate for year 2020
- $Z_1$ : Indicator variable of Southern region
- $Z_2$ : trust in the judicial system
- $Z_3$ : Rate of calls to the helpline number 1522 for purposes other than reporting a violence (calls received from non victims over the total population).

Posterior summaries for the regression coefficients of the two model components (7) and (8) are reported in Table 1 for both years.

For year 2019, the application reveals that the Social Deprivation Index has a strong positive impact on the risk of violence (95% credibility interval, CI: (0.057, 0.125)). The estimates from the 2014 Women's Safety Survey (CI: (0.030, 0.081)) explain the risk of violence, as well as the proportion of call to 1522 reporting a violence episode (CI: (0.011, 0.076)). A negative effect emerges for the employment rate ratio for women with preschool age children to woman without children (CI: (-0.054, -0.016)). On the other hand, an analogous model including the women employment rate had no significant effect for this variable (95% CI (-0.034, 0.040)). The posterior distributions of the regression coefficients are reported in Fig. 3.

With respect to the reporting probability (see Fig. 4), the "South" covariate has the strongest impact (0.95 CI: (-0.393, -0.223)), suggesting that the under-reporting is, other things being equal, higher in the southern regions. Moreover, the trust in the judicial system has a strongly significant effect on the violence reporting probability: regions in which people's trust is higher show higher propensities to report (CI: (0.042, 0.096)). The effect of calls to the 1522 helpline (CI: (-0.101, -0.045)) is negative: an increase in the number of calls seems to be associated with a lower reporting probability, which could be the case if calls to 1522 convey barriers in reporting abuses to the police.

Similar comments can be done for 2020, with some distinctions. The Social Deprivation Index is confirmed to have a strong impact on the risk of violence (95% credibility interval, CI: (0.127, 0.201)), with positive coefficient. Notably, the effect of this variable is stronger than for 2019. The estimates from the 2014 Women's Safety Survey (CI: (0.036, 0.096)) as well as the proportion of calls to 1522 reporting a violence episode (CI: (0.029, 0.090)) have positive significant effects. A negative effect emerges for the women employment rate (CI: (-0.130, -0.055)), meaning that the risk of violence is lower in regions where the women employment rate is larger. On the other hand, a model including the employment rate ratio for women with pre-school age children to woman without children in place of the employment rate had no significant effect for this variable (95% CI (-0.005, 0.030)). This result is opposite to the conclusions drawn for 2019, where the latter variable has a strong effect on the risk whereas the coefficient for the women employment rate resulted not significant. During the pandemic, a large proportion of women, not only those with pre-school children, have sacrificed their jobs in response to the crisis and therefore the variable comparing the occupation rates of women with pre-school children to women without children does not probably convey a measure of women empowerment. Concerning the reporting probability, the "South" covariate has the strongest impact (0.95 CI: (-0.389, -0.162)), with larger absolute size if compared to 2019. Similarly, trust in the judicial system has a larger effect on reporting, compared to 2019, and is significant with 0.95 CI (0.0768, 0.1241). The effect of calls to the 1522 helpline (CI: (-0.069 - 0.019)) is again negative but smaller in absolute size if compared to 2019.

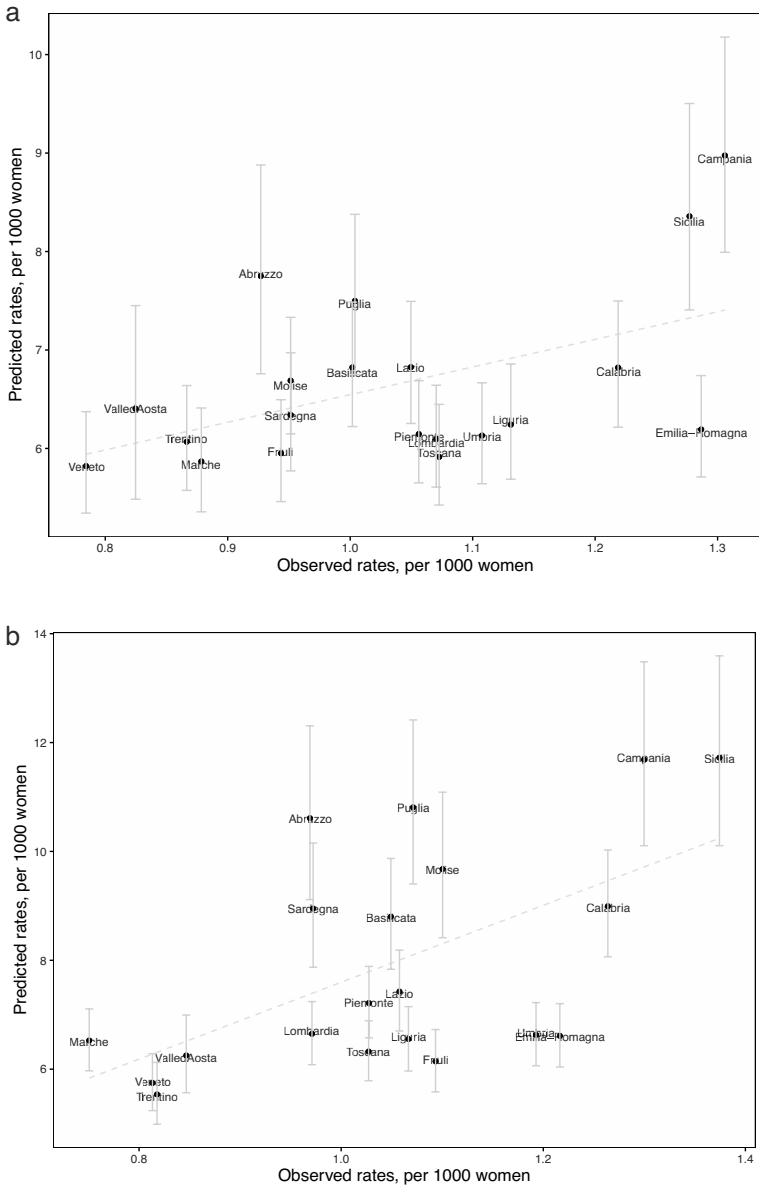


**Fig. 5** Posterior distributions for the reporting probabilities  $\epsilon_i$  of each region: years 2019 (top) and 2020 (bottom)

The posterior distributions of the reporting probabilities  $\epsilon_i$  are shown in Fig. 5. In both years, the geographical variation is evident, and reflects previous findings on gender stereotypes in Italy based on survey data.<sup>16</sup> Especially for 2020, the North/South divide is striking, with two main levels of the reporting probabilities. The variability of the reporting probability across regions appears to be more pronounced in 2020.

Estimates of true unreported counts  $T_i$  are obtained by drawing from their predictive distribution. Figure 6 compares the observed and the predicted counts for the two years. The superimposed regression line has a coefficient of 2.81 for year 2019 and 7.06 for 2020. We

<sup>16</sup> <https://www.istat.it/it/files/2019/11/Report-stereotipi-di-genere.pdf>



**Fig. 6** Observed vs predicted counts of violence: year 2019 (top) and 2020 (bottom). The superimposed dashed regression line has slope 2.81 for 2019 and 7.06 for 2020

notice a smaller inflation in some areas (Emilia Romagna, Toscana, Umbria) and a larger inflation in some regions (Campania, Puglia, Sicilia, Abruzzo). Such inflation is larger in 2020, especially for the Southern regions. This is the effect of the reporting model component, which for 2020 has stronger effects and has a lower mean value.

## 7 Conclusions

We considered a Bayesian model that explicitly describes the reporting process and illustrated some preliminary findings from the analysis of the official reports on gender based violence in the framework of partially observed responses, using information available at the regional level. To our knowledge, no previous studies have addressed formally the issue of under-reporting in the analysis of such data. The findings are largely consistent for the two years. The importance of the Social Deprivation Index in explaining the event rate confirms the role of the socioeconomic context in both the mentioned aspects of perpetuating violence-prone attitudes and exposing to more violent environments, and we found a larger effect for 2020. This is consistent with findings in the literature: see e.g. Hammett et al. (2022), where a general correlation between socioeconomic deprivation and intimate partner violence is found and moreover the association between violence and Covid-related stress is found to be stronger for individuals living in more deprived areas. Larger labour market participation seems to correlate with lower violence rates, again possibly due to a combination of economic status, culture, attitudes. Of interest is the differential impact of the two alternative covariates that we selected for occupation over the two years. Indeed including the employment rate ratio for women with pre-school age children to woman without children in place of the employment rate gave a non significant effect for year 2020, and vice-versa for year 2019. As mentioned, this might be the result the extreme peculiarities of 2020 as far as the access of woman to the labour market during the pandemic and lockdown period is concerned.

Regarding the probability to report an abuse, the results obtained are comparable among the two years although stronger for 2020 as commented. The findings support the hypothesis that such a probability depends on both the trust in the judicial system as well as on the cultural geographical divide. The negative estimated association of the calls to 1522 from non-victims with the propensity to report might be ascribed to the fact that not explicitly reporting an abuse or asking for generic information in a call to the helpline might signal, rather than greater awareness, a combination of safety concerns, insecurity, conservative gender norms and women's attitudes in general that prevent users of 1522 to disclose abuses to the operator even when the former have actually occurred. In fact, 1522 data indicate that a large proportion of calls from victims are not reported to the police, so that the helpline might be considered as the preferred if not the only means to receive support and escape violence. This effect is however lower in 2020, where indeed the proportion of victims among users of the helpline is higher.

In the absence of completely reported data, models like the one introduced allow us to produce regional estimates of the under-reporting and, in turn, of the extent of violence against women, and their geographical heterogeneity.

We stress that precise prior information on the intercept  $\gamma_0$  for the reporting model (8) is essential to produce the estimates for the  $\epsilon_{i,s}$  and in turn the predictive distribution of the true counts  $T_i$ ,  $i = 1, \dots, m$ . Information on this has been derived from the last Victimization Survey, that however dates back to 2014; for this survey, compared to the previous one, a wide change in the reporting attitudes had been found, so that a certain increase in the reporting rates might be expected. The available 1522 helpline data on the reporting rates of victims have also been exploited to elicit the prior. This data source might represent a subpopulation of the victims. Another limitation of the study is related to the analysis of possibly different types and definitions of violence. As noticed e.g. in Müller and Schrötte (2004), depending on the scope of the definition, the range of prevalence varies widely. Concerning our data



source, we notice that the official reports are clearly based on the categories of violence that agree with the criminal law and most likely are associated to the use of physical force, bodily injuries and threats.

We conducted a sensitivity analysis to check the prior assumption on the intercept in the model for the reporting probability (8). The prior choice has an impact in determining the overall reporting level, which is expected due to the parameter unidentifiability, and must be informative as discussed. We observed that if we select a different prior mean for  $\gamma_0$  while holding the prior variance for  $\gamma_0$  fixed, the overall reporting level is affected, but only slight changes in the posterior variability of the reporting probabilities are observed. However, the regression coefficients in both the intensity and the reporting models are not affected by moderate changes in the prior mean for  $\gamma_0$ . On the other hand, the prior variance for  $\gamma_0$  impacts mainly on the variability of the estimates. In particular, estimates of the regression parameters ( $\beta_j$ ,  $j = 1, \dots, 4$ ) as well as those of the variables affecting the reporting probabilities ( $\gamma_j$ ,  $j = 1, \dots, 3$ ) do not significantly vary for moderate changes of the prior variance. On the other hand, the variances of posterior distribution of the reporting probability  $\epsilon_i$  ( $i = 1, \dots, 20$ ) increase with the prior variance. This result is expected since, as stressed in the paper, identifiability of the model is achieved when strong prior information is specified for  $\gamma_0$ .

The proposed model can be extended in several directions. In particular, random effects could be added to the model to account for lack of fit and spatial variation. However, we did not pursue this approach here due to the small number of areas considered. To overcome the low granularity of the data, we plan to analyze the counts at the municipality level and consider a more informative panel of auxiliary information.

Given that sexist and sexually hostile attitudes are associated with violence against women (Flood & Pease, 2009), a vast research has recently focused on gender based violence online. Online expressions of misogyny and abuse quantify the diffusion of these attitudes; moreover, they may even reinforce the violent environment. In this context, research is being conducted on the frequency of some terms in online searches and social media communities as proxies of gender based violence. For instance, Zaleski et al. (2016) analyze the digital discourse of rape culture by studying comments threads posted in response to sexual assault news articles and find strong evidence of victim blaming and perpetrator support. Köksal et al. (2022) shows that Google Trends data on online searches may help predicting episodes of violence against women. This kind of information may also be exploited as useful covariates in the class of models presented in this paper.

**Acknowledgements** We are grateful to Alessandra Capobianchi and Franco Turetta from Istat for helpful discussion and to the anonymous reviewers for their valuable suggestions.

**Funding** Open access funding provided by Università degli Studi di Roma La Sapienza within the CRUI-CARE Agreement. The authors acknowledge the financial support provided by the MiuR-PRIN Grant No 2022Z85NCT (Violence against women: modelling misreported information in social data, PI: Serena Arima) and Sapienza University grant No RP12218169970CD7. Open access funding was provided by Università degli Studi di Roma La Sapienza within the CRUI-CARE Agreement.

## Declarations

**conflict of interest** The authors have no competing interests to declare that are relevant to the content of this article.

**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article

are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

## References

- Ambrosetti, E., Amara, N. A., & Condon, S. (2013). Gender-based violence in Egypt: Analyzing impacts of political reforms, social, and demographic change. *Violence Against Women*, 19(3), 400–421. <https://doi.org/10.1177/1077801213486329>. PMID: 23676450.
- Bailey, T. C., Carvalho, M. S., Lapa, T. M., Souza, W. V., & Brewer, M. J. (2005). Modeling of under-detection of cases in disease surveillance. *Annals of Epidemiology*, 15(5), 335–343.
- Betto, F., Ticci, E., & Betti, G. (2020). A fuzzy index and severity scale to measure violence against women. *Social Indicators Research*, 148(1), 225–249.
- Brass, W. (1996). Demographic data analysis in less developed countries: 1946–1996. *Population Studies*, 50(3), 451–467.
- Capaldi, D. M., Knoble, N. B., Shortt, J. W., & Kim, H. K. (2012). A systematic review of risk factors for intimate partner violence. *Partner Abuse*, 3(2), 231–280.
- Caudill, S. B., & Mixon, F. G. (1995). Modeling household fertility decisions: Estimation and testing of censored regression models for count data. *Empirical Economics*, 20(2), 183–196.
- Chen, J., Song, J. J., & Stamey, J. D. (2022). A Bayesian hierarchical spatial model to correct for misreporting in count data: Application to state-level COVID-19 data in the united states. *International Journal of Environmental Research and Public Health*. <https://doi.org/10.3390/ijerph19063327>
- Copp, J. E., Giordano, P. C., Longmore, M. A., & Manning, W. D. (2019). The development of attitudes toward intimate partner violence: An examination of key correlates among a sample of young adults. *Journal of Interpersonal Violence*, 34(7), 1357–1387.
- de Oliveira, G. L., Argiento, R., Loschi, R. H., Martins Assuno, R., & Ruggeri, F. (2022). Bias correction in clustered underreported data. *Bayesian Analysis*, 17(1), 95–126.
- de Oliveira, G. L., Loschi, R. H., & Assunção, R. M. (2017). A random-censoring Poisson model for underreported data. *Statistics in Medicine*, 36(30), 4873–4892.
- de Valpine, P., Turek, D., Paciorek, C., Anderson-Bergman, C., Lang, D. T., & Bodik, R. (2017). Programming with models: Writing statistical algorithms for general model structures with NIMBLE. *Journal of Computational and Graphical Statistics*, 26, 403–413. <https://doi.org/10.1080/10618600.2016.1172487>
- Di Noia, J. (2002). Indicators of gender equality for American states and regions: An update. *Social Indicators Research*, 59(1), 35–77. <https://doi.org/10.1023/A:1016094427841>
- Dobash, R. P., & Dobash, R. E. (2004). Women's violence to men in intimate relationships: Working on a puzzle. *The British Journal of Criminology*, 44(3), 324–349. <https://doi.org/10.1093/bjc/azh026>
- Dobash, R. P., Dobash, R. E., Wilson, M., & Daly, M. (2014). The myth of sexual symmetry in marital violence\*. *Social Problems*, 39(1), 71–91. <https://doi.org/10.2307/3096914>
- Dugan, L., Nagin, D. S., & Rosenfeld, R. (1999). Explaining the decline in intimate partner homicide: The effects of changing domesticity, women's status, and domestic violence resources. *Homicide Studies*, 3(3), 187–214. <https://doi.org/10.1177/1088767999003003001>
- Dvorzak, M., & Wagner, H. (2016). Sparse Bayesian modelling of underreported count data. *Statistical Modelling*, 16, 24.
- Flood, M., & Pease, B. (2009). Factors influencing attitudes to violence against women. *Trauma, Violence, & Abuse*, 10(2), 125–142. <https://doi.org/10.1177/1524838009334131>. PMID: 19383630.
- Florens, J.-P., & Simoni, A. (2021). Revisiting identification concepts in Bayesian analysis. *Annals of Economics and Statistics*, 144, 1–38.
- Garcia-Moreno, C., Jansen, H. A., Ellsberg, M., Heise, L., Watts, C. H., et al. (2006). Prevalence of intimate partner violence: Findings from the who multi-country study on women's health and domestic violence. *The Lancet*, 368(9543), 1260–1269.
- Gelles, R. J. (2000). Estimating the incidence and prevalence of violence against women: National data systems and sources. *Violence Against Women*, 6(7), 784–804. <https://doi.org/10.1177/107780120006007005>
- Hammett, J. F., Halmos, M. B., Parrott, D. J., & Stappenbeck, C. A. (2022). Covid stress, socioeconomic deprivation, and intimate partner aggression during the covid-19 pandemic. *BMC Public Health*, 22(1), 1–11.
- Johnson, M. P. (2006). Conflict and control-Gender symmetry and asymmetry in domestic violence. *Violence Against Women*, 12(11), 1003–1018. <https://doi.org/10.1177/1077801206293328>

- Johnson, N. L., Kotz, S., & Kemp, A. W. (2005). *Univariate Discrete Distributions*. Wiley Series in Probability and Statistics. Hoboken: Wiley.
- Köksal, S., Pesando, L. M., Rotondi, V., & Şanlıtürk, E. (2022). Harnessing the potential of google searches for understanding dynamics of intimate partner violence before and after the COVID-19 outbreak. *European Journal of Population*, 1–29, 517–545.
- Leon, C. M., Aizpurua, E., & Rollero, C. (2022). None of my business? An experiment analyzing willingness to formally report incidents of intimate partner violence against women. *Violence Against Women*, 28(9), 2163–2185.
- Li, T., Trivedi, P. K., & Guo, J. (2003). Modeling response bias in count: A structural approach with an application to the national crime victimization survey data. *Sociological Methods & Research*, 31(4), 514–544. <https://doi.org/10.1177/0049124103251951>
- Melton, H. C., & Sillito, C. L. (2012). The role of gender in officially reported intimate partner abuse. *Journal of Interpersonal Violence*, 27(6), 1090–1111. <https://doi.org/10.1177/0886260511424498>. PMID: 22203618.
- Merenda, A., Garro, M., & Schirizzi, M. (2021). The invisible pandemic: Domestic violence and health and welfare services in Italy and in the United Kingdom during covid-19. *International Journal of Humanities Social Sciences and Education*, 8, 11–20.
- Müller, U., & Schröttle, M. (2004). Health, well-being and personal safety of women in Germany (Violence against Women-Additional Survey). Bielefeld, Germany: Federal Ministry for Family Affairs, Senior Citizens, Women and Youth. <https://www.bmfsfj.de/resource/blob/93906/9c0076fc66b1be6d0eb28258fe0aa569/frauenstudie-englisch-gewalt-gegen-frauen-data.pdf>
- Orchowski, L. M., Grocott, L., Bogen, K. W., Ilegbusi, A., Amstadter, A. B., & Nugent, N. R. (2022). Barriers to reporting sexual violence: A qualitative analysis of #whyididntreport. *Violence Against Women*, 28(14), 3530–3553. <https://doi.org/10.1177/10778012221092479>. PMID: 35946129.
- Sardinha, L., Maheu-Giroux, M., Stöckl, H., Meyer, S. R., & García-Moreno, C. (2022). Global, regional, and national prevalence estimates of physical or sexual, or both, intimate partner violence against women in 2018. *The Lancet*, 399(10327), 803–813. [https://doi.org/10.1016/S0140-6736\(21\)02664-7](https://doi.org/10.1016/S0140-6736(21)02664-7)
- Seidenbecher, S., Dobrowolny, H., Wolter, S., Klemen, J., Meyer-Lotz, G., Gescher, D. M., Steiner, J., & Frodl, T. (2023). In P. C. Guest (Ed.), *Consequences of the Lockdown: Domestic Violence During the COVID-19 Pandemic* (pp. 53–72). Cham: Springer. [https://doi.org/10.1007/978-3-031-28012-2\\_3](https://doi.org/10.1007/978-3-031-28012-2_3)
- Stoner, O., Economou, T., & da Silva, G. D. M. (2019). A hierarchical framework for correcting under-reporting in count data. *Journal of the American Statistical Association*, 114(528), 1481–1492. <https://doi.org/10.1080/01621459.2019.1573732>
- Straus, M. (1999). The controversy over domestic violence by women - a methodological, theoretical, and sociology of science analysis. In: Arriaga, X., Oskamp, S. (eds.) *Violence in Intimate Relationships. Claremont Symposium on Applied Social Psychology*, Claremont CA, pp 17–44
- Usta, J., Murr, H., & El-Jarrah, R. (2021). Covid-19 lockdown and the increased violence against women: Understanding domestic violence during a pandemic. *Violence and gender*, 8(3), 133–139.
- Venis, S., & Horton, R. (2002). Violence against women: A global burden. *The Lancet*, 359(9313), 1172. [https://doi.org/10.1016/S0140-6736\(02\)08251-X](https://doi.org/10.1016/S0140-6736(02)08251-X)
- Ventura, M., Di Napoli, A., Petrelli, A., Pappagallo, M., Mirisola, C., & Frova, L. (2022). Male and female differences in homicide mortality: Results of an Italian longitudinal study, 2012–2018. *Frontiers in Public Health*. <https://doi.org/10.3389/fpubh.2022.919335>
- Viero, A., Barbara, G., Montisci, M., Kustermann, K., & Cattaneo, C. (2021). Violence against women in the covid-19 pandemic: A review of the literature and a call for shared strategies to tackle health and social emergencies. *Forensic science international*, 319, 110650.
- Whaley, R. B. (2001). The paradoxical relationship between gender inequality and rape: Toward a refined theory. *Gender and Society*, 15(4), 531–555.
- Whittemore, A. S., & Gong, G. (1991). Poisson regression with misclassified counts: Application to cervical cancer mortality rates. *Journal of the Royal Statistical Society. Series C (Applied Statistics)*, 40(1), 81–93.
- Winkelmann, R. (1996). Markov chain Monte Carlo analysis of underreported count data with an application to worker absenteeism. *Empirical Economics*, 21(4), 575–587. <https://doi.org/10.1007/BF01180702>
- Wittebrood, K., & Junger, M. (2002). Trends in violent crime: A comparison between police statistics and victimization surveys. *Social Indicators Research*, 59(2), 153–173. <https://doi.org/10.1023/A:1016207225351>
- Wiśniowski, A., Sakshaug, J. W., Perez Ruiz, D. A., & Blom, A. G. (2020). Integrating probability and non-probability samples for survey inference. *Journal of Survey Statistics and Methodology*, 8(1), 120–147.
- Yakubovich, A. R., Heron, J., Feder, G., Fraser, A., & Humphreys, D. K. (2020). Long-term exposure to neighborhood deprivation and intimate partner violence among women: A UK birth cohort study. *Epidemiology*, 31(2), 272.

Zaleski, K. L., Gundersen, K. K., Baes, J., Estupinian, E., & Vergara, A. (2016). Exploring rape culture in social media forums. *Computers in Human Behavior*, *63*, 922–927. <https://doi.org/10.1016/j.chb.2016.06.036>

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

## Authors and Affiliations

Silvia Poletti<sup>1</sup>  · Serena Arima<sup>2</sup> · Sara Martino<sup>3</sup>

✉ Silvia Poletti  
silvia.poletti@uniroma1.it

Serena Arima  
serena.arima@unisalento.it

Sara Martino  
sara.martino@ntnu.no

<sup>1</sup> Dipartimento di Scienze Sociali ed Economiche, Sapienza Università di Roma, P.le Aldo Moro, 5, 00185 Roma, Italy

<sup>2</sup> Dipartimento di Storia, Società e Studi sull'Uomo, Università del Salento, Studium 2000, Edificio 5, Via di Valesio, 73100 Lecce, Italy

<sup>3</sup> Department of Mathematical Science, Norwegian University of Science and Technology (NTNU), Alfred Getz' vei 1, 7034 Trondheim, Norway