Understanding the factors behind the (under-)reporting of sexual violence

Evidence from France

Author: Maëlle Stricot

Supervisors: Thomas Breda and Léa Toulemon

Master thesis

Public Policy and Development (PPD)

PARIS SCHOOL OF ECONOMICS

September 2021

Understanding the factors behind the (under-)reporting of sexual violence: Evidence from France

Maëlle Stricot[1]

September 2021

Abstract

Using data from the French victimisation survey, this thesis explores the determinants of reporting sexual violence to the police. My contribution to the economics of crime and economics of social norms literature is threefold. First, using binary regression models, I show that assault characteristics have higher explanatory power than victim's socio-demographic characteristics on this high-stakes individual decision. This descriptive analysis also highlights the significant influence of victims' social stigmatisation on the decision to file a complaint. Second, I employ a difference-in-difference strategy comparing the reporting rate for sexual and physical violence to analyse the effect of two high-profile sexual assault cases that pictured different representations of victims and the reporting process. On average, I find a 158% drop in sexual violence reporting following the allegations against Dominique Strauss-Kahn in May 2011, compared to a 111% increase after the ones against Harvey Weinstein in October 2017. Third, I provide evidence that changes in opinions and representations about sexual violence seem to be the primary mechanism explaining the increase in sexual violence reporting after the Weinstein case and #MeToo. These results suggest that social norms and representations surrounding sexual violence firmly influence the victim's reporting decision.

Keywords – Crime, Gender, Sexual assault, Media, Social norms, Decision making
JEL Codes – D81, D91, J16, K42, L82, Z13

[1]I am particularly grateful to Thomas Breda and Léa Toulemon for agreeing to supervise this master's thesis and providing valuable guidance and suggestions throughout this project. These discussions have helped guide my thinking and the development of my work. I would also like to thank Andrew Clark and Fabrice Etilé for their helpful comments during the Behavioural Economics Research Seminar, as well as Marion Monnet for her time and valuable suggestions. Finally, I am deeply thankful for my classmates' support and company during the two years of the master's program. A special recognition to Davi for his unconditional support and remarkable patience throughout this year, and without whom this work could not have been completed.

Contents

List of Figures

List of Tables

1 Introduction

Sexual violence is a major public health problem and a violation of human rights, affecting both developing and developed countries. In France, almost 15% of women and 4% of men aged 20 to 69 report having experienced rape, attempted rape or another form of sexual assault at least once in their lifetime (Hamel et al., 2016). Each year, this represents around 233,000 victims of sexual violence, 80% of which are women (Bernardi et al., 2019)[2].

The costs of sexual violence are sizeable. In addition to long-lasting and substantial harm on victims' health (Jina and Thomas, 2013), education (Rees and Sabia, 2013) and employment outcomes (Sabia et al., 2013), it entails high costs in terms of public service provision and reduced productivity in the workplace (UN Women, 2016). As an illustration, in the United States, it has been estimated that each rape costs between $87,000 and $240,776, with total estimated costs of $127 billion per year[3].

Reporting the assault to the authorities is a necessary condition for any police response, and thus the primary way for victims to get legal justice and obtain compensation for the harm suffered. Yet, sexual violence is arguably one of the most under-reported crimes (Bachman, 1998). For instance, estimates from victimisation surveys suggest that only one in ten rape victims files a complaint in France (Vanier and Langlade, 2018).

This under-reporting phenomenon has several detrimental consequences. First, it causes official crime statistics to underestimate the true prevalence of sexual crime, which criminologists refer to as the "dark figure" of crime (MacDonald, 2002). This underestimation leads to ineffective crime policy and a sub-optimal allocation of resources to prevent sexual violence. Second, crime under-reporting weakens law enforcement and criminal deterrence by reducing the crime detection rates (Garoupa, 2003), which is a significant issue as deterrence is one of the main ways for the criminal justice system to reduce crime (Robinson and Darley, 2002; Chalfin and McCrary, 2017). Indeed, an unreported crime eliminates the possibility that an offender will be arrested or convicted, lowering the perceived likelihood that such crime will be punished. This is all the more serious with sexual violence, as research suggests that recidivism is high among rapists and that would-be rapists are effectively deterred by the threat of formal sanctions such as arrest (Bachman et al., 1992). Hence, by decreasing the probability of apprehending perpetrators, under-reporting may also add to the costs of sexual violence and decrease social welfare. A better understanding of reporting a sexual crime is thus of prime importance from a policy perspective.

If the harm suffered by victims is so large, one might wonder why so few attempts are made to claim legal redress. Given these gaps in knowledge and the resulting policy implications, the overall purpose of this thesis is to investigate various factors that may affect the reporting behaviour of sexual assault victims, in the context of France. In particular, I focus on the role played by victims' characteristics, assault's circumstances and two high-profile sexual assault

[2] In French law, sexual violence refers to any sexual act committed with violence, coercion, threat or surprise. It can take different forms, such as sexual assault, rape, exhibitionism or sexual harassment – all criminally penalised.

[3] "Rape and sexual assault: A renewed call to action." The White House Council on Women and Girls (2014).

cases and seek to understand the effect of these factors on the decision to report to the police. Although these are certainly not the only factors affecting the reporting of sexual violence, they are the main ones I can study with the data at hand. Besides, a better understanding of these determinants is needed to assess the unequal situation of victims in terms of reporting propensities, which is the first step to design policy tools aimed at encouraging reporting.

A large body of work in psychology, criminology or sociology has already identified various factors that can influence the decision to report sexual assault (reviewed in Vanier and Langlade (2018)). The earliest studies have highlighted psychological factors, such as social stigma, shame or embarrassment (Amir, 1971; Macdonald, 1971). Using survey data from Anglo-Saxon countries, other studies focused on factors inherent to the assault. For instance, it has been shown that victims were more likely to report a rape that fitted with the "classic rape" stereotype, involving a violent attack by a stranger (Williams, 1984; Du Mont et al., 2003). In contrast, knowing the offender has been found to decrease the reporting propensity as victims may want to protect him or keep the case private (Gartner and Macmillan, 1995; Fisher et al., 2003). Victims' socio-demographic characteristics also received some attention in that literature. For instance, some studies showed that the propensity to report rape decreases with the victim's standard of living or education (Lizotte, 1985), while age has been found to affect reporting positively (Goudriaan et al., 2006). These studies give valuable insights but do not allow a comprehensive understanding of the victim's reporting behaviour as they focus on specific populations and are subject to endogeneity issues.

My thesis overcomes these limitations by using novel data and methods to study sexual violence reporting. First, I resort to under-exploited survey data that also enables me to investigate the effect of new factors. The sample is representative of the population, and I try to recover causal effects. Doing so, I complement previous findings and sometimes come to different conclusions. Besides, little quantitative research on the reporting of sexual violence has been conducted in France, with only one recent sociology paper of Vanier and Langlade (2018) on this issue.

Moreover, the decision to report sexual crime has received little attention in economics. An exception is Allen (2007), who investigates the decision to report rape using data from the US National Crime Victimisation Survey and finds that the reporting propensity increases with "social support availability" and "ancillary evidence" of the crime. Yet, as was shown in the famous "crime and punishment" models of Becker (1968) and Ehrlich (1973), the economic approach can be helpful to improve the understanding of factors determining certain high-stakes individual decisions (Levitt and Miles, 2006).

To shed light on the determinants of reporting sexual violence, I use data from the French victimisation survey *Cadre de Vie et Sécurité*, annually conducted from 2007 to 2019 among 16,000 households, which identifies victims of sexual violence in the two years preceding the survey. My first empirical strategy relies on a descriptive approach using binary regression models with OLS estimation. I show that assault characteristics influence the reporting behaviour more than individual characteristics. In particular, having suffered assaults that resulted in bodily injuries is associated on average with a statistically significant 14 percentage points increase in

the reporting propensity, all else being equal. This result illustrates the importance of evidence that would fit with the "classic rape" stereotype and support the victim's credibility.

However, some socio-demographic characteristics also turn out to be important in explaining reporting when restricting the analysis to specific categories of victims, such as women or victims of sexual violence outside their current household. The victim's socioeconomic category and the size of the urban unit where the household is located appear as relevant factors, with a higher propensity to report among lower classes or in urban areas. This notably reflects the key role of social stigma in the reporting decision. I also highlight the unequal situation of victims of sexual violence in terms of reporting by studying the reasons for not reporting. The results show that upper-class victims seem more likely to be deterred due to social stigmatisation, and lower-educated victims are more likely to foresee further harmful consequences.

Then, in order to investigate the effect of changing norms and representations regarding sexual violence, I exploit the plausibly exogenous timing of two high-profile sexual assault cases that received intense media coverage in France. The first is a criminal case following the allegations of sexual assault and attempted rape against Dominique Strauss-Kahn (DSK) in May 2011. The second is a series of public disclosures of sexual harassment and assault by Harvey Weinstein in October 2017, which gave rise to #MeToo. I employ a difference-in-difference strategy comparing the reporting rate for sexual and physical violence before and after each case. I show that by triggering different representations of victims and the reporting process, these two cases did not target the same victims and did not affect reporting in the same way.

On average, the results suggest a 158% drop in sexual violence reporting following the DSK case, compared to a 111% increase after Weinstein and the #MeToo movement. This represents a sizeable impact, especially since this latter effect is conservative due to a single survey year to measure reporting in 2018. When restricting the analysis only to women, the effect is even more important for the Weinstin case (135%), but the one found for DSK is no longer statistically significant. The heterogeneity analysis nevertheless points to a large decrease in sexual violence reporting for immigrant women in 2011, though the estimation is not entirely robust. In contrast, the positive impact of the Weinstein case and #MeToo is significantly more important among young French women, with a medium level of education or from the middle class, which does not fully support the hypothesis of a role model effect.

Despite the limitations of this difference-in-difference estimation, I attempt to reinforce confidence in the validity of this strategy with a set of robustness checks. I support the validity of the common trend assumption, and verify that a rewording of the survey questions does not drive my results. I also implement this estimation with administrative data on which I study the monthly number of complaints filed instead of the reporting rate, and find a similar result for the Weinstein case and #MeToo on sexual violence reporting. Besides, I show that changes in opinions and representations about sexual violence seem to be the main mechanisms explaining the increase in sexual violence reporting after 2017, rather than changes in the perceptions of police and justice effectiveness or in the processing of victims by police officers.

These results are related to two main streams of literature. First, I contribute to the large

economics of crime literature by complementing previous empirical studies investigating the determinants of the crime reporting decision (Allen, 2007; Soares, 2004; Bowles et al., 2009; Gingerich and Oliveros, 2018). More specifically, I mostly complement the literature on reporting gender-based violence by studying various factors affecting the reporting of sexual crimes, which has been under-studied in economics. The few works in that literature have mainly focused on the effect of exogenous shocks on the reporting of violence against women. For instance, previous studies showed that increased female representation in local government or among police officers led to greater reporting of crimes against women (Iyer et al., 2012; Miller and Segal, 2019), and that education-entertainment strategies such as TV shows or social media campaigns improved willingness to report domestic violence (Banerjee et al., 2019; Green et al., 2020).

More closely related to my paper is McDougal et al. (2018) and Bhatnagar et al. (2019). They show that a high-profile rape and murder case increased reporting of sexual crimes in India, mainly through changes in police behaviour. In addition, in a recent study using incident-level police records, Levy and Mattsson (2020) find that the #MeToo movement increased sexual crimes reporting in the United States by 8% in the six months after it started. However, police data may hide an increase in the effective number of assaults rather than in the reporting rate itself, which would lead to very different policy implications. I can overcome this limitation by studying the proportion of victims who lodged a complaint rather than the number of complaints filed. In addition, some high-profile cases could also discourage victims to report crimes or affect victims in different ways. Thanks to very detailed information collected in the French victimisation survey, I am able to check for heterogeneous effects on different subpopulations of victims.

Related to the previous point, my thesis also contributes to the economic literature on the of effect social norms in general and information provided by the media in particular on individual attitudes and behaviours. First, I supplement a growing literature on the economic analysis of social norms, and especially gender norms. Recent work has highlighted the importance of cultural norms in perpetuating gender inequality and violence against women and girls (Alesina et al., 2013; Bertrand et al., 2015; Jensen and Oster, 2009). Papers at the frontier of psychology and economics also showed that influencing individuals' perceptions of norms or attitudes is a way to create social change (reviewed in Tankard and Paluck (2016)), and that the media can be used for that purpose (Paluck, 2009). I contribute to this literature by showing how changes in social norms can affect a high-stakes behaviour.

Moreover, a wide literature in political economy has shown that news provision can influence many policy-relevant behaviours (reviewed in DellaVigna and La Ferrara (2015)), including offending (Dahl and DellaVigna, 2009; Jensen and Oster, 2009; Card and Dahl, 2011) and criminal justice (Lim et al., 2015; Lim, 2015; Philippe and Ouss, 2018). However, surprisingly little is known about the impact of media coverage of crime on crime reporting, and my paper aims to fill this gap. Furthermore, I bridge these literatures by identifying the effect of high-profile sexual assault cases that challenge gender norms and stereotypes about sexual violence.

The rest of the paper is organised as follows. Section II provides a theoretical framework that discusses the hypothetical effect of different factors on the decision to report sexual violence.

Section III describes the main data used for this analysis. Section IV presents the empirical strategy and results regarding the influence of victim's characteristics and assault's circumstances on the reporting behaviour. The effect of norms and representations of sexual violence is investigated in Section V, through the impact of two high-profile sexual assault cases. This section exposes the setting and empirical methodology, before describing the results and an associated set of robustness checks. Section VI discusses some possible mechanisms through which the effects may operate. Finally, Section VII concludes.

2 Theoretical framework

2.1 Cost-benefit analysis of the crime reporting decision

The crime reporting decision can be seen as a cost-benefit calculation under uncertainty, in which individuals derive a rational reporting strategy (Bowles et al., 2009). Victims compare the probabilistic expected payoff associated with reporting against the more certain payoff from not reporting when deciding whether to report their assault to the police; or, said in more economic terms by Allen (2007), when deciding to allocate an "information resource" about the crime they suffered by revealing it or not.

Regarding utility gains, reporting may provide victims with several psychological or material benefits. For instance, it can allow them to be recognised as victims by society, which is often an essential first step in the recovery process after an assault. It can also give victims the feeling of having fulfilled a moral or civic duty by preventing other attacks. Finally, reporting increases the likelihood that the offender will be apprehended and that the victim gets legal and financial reparation for the harm suffered. Sometimes, it is also the primary way to access public and social support systems, such as medical care or victim assistance services (Sable et al., 2006).

Victims will weigh those potential benefits against the pecuniary and non-pecuniary costs of reporting. Direct psycho-social costs like trauma, shame or embarrassment have long been recognised as significant for victims of sexual violence (Amir, 1971). These emotions could emanate from the victim herself as well as from others due to the strong stigmatisation of sexual assault victims. This decision also implies monetary costs such as time lost in the process, travel costs or associated expenses, which can be considered as the opportunity costs of reporting. These costs may be relatively high for some individuals, such as illegal immigrants who may be afraid to go to the police to report or domestically abused women who are financially dependent on their offender. Future potential costs should also be considered, such as fear of losing one's social status due to social recrimination, or the risk of retaliation by the offender. Besides, behavioural economics research on information avoidance suggests that people may not report as this information could hurt someone else (Golman et al., 2017). For instance, with domestic abuse or incest, the assault's disclosure may also have negative psychological consequences for other family members.

All these costs are reinforced by the high uncertainty of the final outcome, especially if the

victim believes that the probabilities of the offender's pursuit and apprehension are very low. Indeed, sexual crimes often remain unpunished, with only one complaint of sexual assault in ten that leads to the conviction of the aggressor in France on average (Bousquet, 2016). Under this premise, an increase in sexual violence reporting can thus be driven either by an increase in the expected benefits or a decrease in the expected costs.

Several factors exogenous to a victim's information disclosure might affect this cost-benefit analysis. The first two main factors are the victim's economic and social capital, which may amplify or mitigate the reporting costs. On the one hand, victims with a higher standard of living or education may be more likely to report to the police as they may be more aware of sexual violence and have a closer social connection to the legal system. The reporting costs may also decrease for victims less financially dependent on their partner in the case of domestic abuse. On the other hand, well-off or more educated victims may know that the process will be highly costly for low chances of success, and may also suffer from a stronger stigmatising effect, reducing their likelihood to report (Lizotte, 1985).

A third important mechanism is the victim's social support, which may decrease the costs associated with reporting (Allen, 2007). For instance, male victims may be less likely to report than female victims due to stronger stigma and less support (Pino and Meier, 1999). In contrast, a higher reporting propensity among older victims may be due to better social support than among younger ones (Goudriaan et al., 2006). The victim's relationship to the offender may also affect social support availability, as a victim who knows her attacker might be less inclined to report him for fear of reprisals or of not being believed by her entourage, or to keep the case private (Gartner and Macmillan, 1995; Fisher et al., 2003). As for the victim's marital status, a person in a relationship that could benefit from the moral support of her spouse could see these costs decrease, while it is most probably the opposite when the offender is the victim's partner. Moreover, the stigma associated with sexual violence may differ across territories, with potentially higher psycho-social costs in more rural or remote areas, where there are also likely to be fewer services or resources available to help victims.

Finally, this cost-benefit analysis may be strongly influenced by the availability of evidence that would support the victim's credibility and increase her probability of success (Allen, 2007). For instance, the reporting costs may be lower for assaults that involved weapon use or bodily injury, since this ancillary evidence could support the victim's declaration and increase the probability of pursuit and apprehension. This may also fit with the "classic rape" stereotype and decrease the reputation costs victims see in reporting (Williams, 1984; Du Mont et al., 2003). Conversely, being assaulted in a place familiar to the victim or by a person known before the attack is likely to decrease the reporting propensity as it will not match the classic stereotypical rape situation. Similarly, one could expect reporting to increase with the severity of the act suffered, being higher for rape than for other types of sexual assault. However, the significant psychological trauma associated with rape may conversely make it more costly to report than other sexual assaults (Lizotte, 1985).

2.2 Social change and high-profile sexual assault cases

Among the costs associated with reporting sexual violence to the police, the role of stigmatisation or social recrimination is probably one of the most important. This barrier reflects the social norms about sexual violence that prevail in a certain place and time, which can be defined as shared expectations of attitudes and behaviours considered acceptable in a given society or group. The influence of social stigma is thus likely to change as societal developments challenge gender norms, stereotypes and the degree of acceptability of sexual violence (Tankard and Paluck, 2016).

One factor that could drive such societal changes is the national coverage of sexual allegations and cases. Indeed, there is extensive evidence that media exposure to particular information or behaviour can change the audience behaviour through a persuasion effect (Kearney and Levine, 2015), which would lead individuals to update their beliefs in a Bayesian manner. This effect could be even more important if the view promoted by the media is substantially different from the public's view initially in place (DellaVigna and La Ferrara, 2015)[4]. Therefore, by raising public awareness of sexual violence, high-profile sexual assault cases may influence individual attitudes and behaviours towards sexual violence, including the victims' reporting behaviour.

In this context, coverage of prominent cases may affect reporting in different ways. On the one hand, it can raise social awareness of the extent of this phenomenon and solve an information asymmetry problem among victims (Naess, 2020). This, in turn, can decrease the perceived costs of reporting by reducing the social stigma surrounding being a victim of sexual violence (Levy and Mattsson, 2020). It can also make victims more self-confident thanks to a role model effect that would encourage other victims to come forward (Iyer et al., 2012). Victims may also simply realise that what they suffered is a crime they can report to the police. Besides, coverage of sexual assaults stories can improve the judicial response and further encourage victims to report, as was found with a high-profile rape and murder case in India (Bhatnagar et al., 2019).

On the other hand, such coverage may conversely trigger a backlash effect[5], especially if it highlights relative social tolerance of sexual violence and inadequate criminal justice response, which would instead increase the perceived reporting costs. Victims may also feel that reporting their assault simultaneously as a high-profile case could decrease the credibility of their allegation (Lee and Suen, 2020), reducing their likelihood of filing a complaint. Therefore, the final sign of the effect is not clear *a priori* and may depend on the case in question as well as on the victims' environment and characteristics. Indeed, social norms may affect certain groups differently, who may interpret the same behaviour according to different moral and material standards (Bénabou et al., 2019).

[4]For instance, Jensen and Oster (2009) find that the introduction of cable television has led to a reduction in the acceptability of violence against women in India, mainly thanks to an imitation effect.

[5]Such a depressing effect was found with coverage of suicide cases, which actually led to a copycat effect when the aim was to raise awareness of the problem and decrease suicide rates (Stack, 2002).

3 Data overview

For this study, I use survey data from the French victimisation survey *Cadre de Vie et Sécurité* (CVS henceforth), a repeated cross-section annually conducted by Insee in the first quarter of each year since 2007. This survey covers crimes and offenses that households and their members may have experienced over the past two years. It consists of three main questionnaires: (i) a household questionnaire on violence suffered at the household level (e.g. burglary, property thefts, vandalism); (ii) an individual questionnaire on violence suffered by a randomly selected household member (e.g. robbery, theft, assault); (iii) an individual questionnaire on sensitive violence suffered by the same household member (sexual violence and domestic violence), which is the main module I exploit.

To mitigate the under-reporting bias that can be very strong with such sensitive issues, the survey adopts a specific methodology to record this violence. Whereas other forms of violence are covered in face-to-face questionnaires, the modules on sensitive violence are administered through a computer-assisted self-interview technology. The respondent fills in her answers directly on the computer without interacting with the enumerator[6].

Sexual violence is surveyed in two stages, depending on whether it occurred within the current household or not. The questions that identify victims ask respondents if they have ever been subjected to unwanted sexual touching or intercourse by someone through the use of violence, threats, coercion, or surprise over the past two years. The survey includes questions on the frequency of sexual violence, and then, for the most recent event, the location of the assault, the nature of the attack, the use of physical violence or threats and the offender's identity. It also contains several questions on the respondent's behaviour at the time of the attack, the consequences of the assault and the actions taken by the victims. In particular, it asks the person whether she went to the police to report the assault and whether she eventually filed a complaint, as well as the reasons why she did not. The data also includes detailed information on individual characteristics, such as age, gender, marital status, country of birth, education, employment status and socioeconomic category, as well as household characteristics, such as income, number of household members and living area.

Each CVS wave contains approximately 16,000 observations (one per household). After weighting, these data are representative of mainland France households. I use the thirteen waves that have been conducted on date with a similar sampling method, from 2007 to 2019. The modules on sensitive violence are only administered to people aged between 18 and 75 years old, totalling around 12,000 respondents each year. Pooling data from all survey years, the initial sample has 157,972 observations, comprised of all individuals who answered the self-administered questionnaire on sensitive violence. Table A1.3 describes the victims of sexual violence in the population by showing the weighted shares of victims among various socio-demographic

[6] In order for foreign respondents to be able to answer, a lighter version of the self-administered questionnaire is also translated into Arabic, Turkish, Portuguese and German, which are the four languages most spoken by non-French speakers in France (less than 1% of respondents in the data).

characteristics. Sexual violence affects 1.3% of the general population, being 2% among women and 0.6% among men. The proportion of victims is higher among the youngest individuals, single, born abroad, unemployed, disadvantaged or living in a priority urban district.

As I want to study the victim reporting behaviour, I restrict my primary sample to self-reported victims of sexual violence. I am left with 1,854 observations, 1,476 of which are individuals who were assaulted outside their current household at least once in the two years preceding the survey, 471 inside their current household, and 93 both. For simplicity, I use the terms "intra-household" and "extra-household" sexual violence to refer to the first two subsamples, and "total" when I do not make the distinction. Descriptive statistics of victims' socio-demographics and assault characteristics in the total sample are reported in Table 3.1 (see Tables A1.2 and A1.1 for the two subsamples), and Tables A1.1 and A1.2 describe the creation of each variable. For example, we see that more than 80% of self-reported victims of sexual violence are women, and almost half of them have been assaulted more than once in the last two years, while men are the perpetrators in nearly 90% of cases.

These data have several limitations. First, even though I get a relatively large sample of victims of sexual violence compared to other surveys run in the general population, the number of observations is still quite low. Hence, my analyses are likely to suffer from power issues. However, it can be argued that the absence of better data and their surprising under-exploitation compensate for this first limitation.

Second, the CVS survey has evolved over time, with deletions, modifications and additions of questions. For example, changes in response modalities occurred for questions regarding the type of assault experienced, the offender, or whether the assault resulted in physical injury. Several modalities had to be grouped together to harmonise the different waves in order to append the datasets and pool the analysis[7]. Also, several relevant questions did not appear until the 2009 survey, such as questions on the year of the assault and whether the aggressor used physical force or was under the influence of drugs or alcohol. A more important concern relates to a slight rewording of the questions on sexual violence in 2017 to harmonise them between the two modules of extra- and intra-household sexual violence. I perform a robustness check in Section 5.4.5 to verify that my results are not driven only by this rewording.

Third, I do not have access to the information on where the surveyed households are located for confidentiality reasons. The smallest geographical level available in the public version of the data is the *Zone d'Etudes et d'Aménagement du Territoire* (ZEAT), which divides metropolitan France into eight large zones[8]. This forces me to use relatively large regional fixed effects in my analyses, making the zero conditional mean assumption less plausible.

Finally, another possible limitation is that some respondents may still be reluctant to describe

[7]Likewise, the variable capturing the respondent's level of education underwent four modifications in the different categories used, so conversions to harmonise this variable across the different waves had to be made.

[8]More precise information on the exact geographical location of the respondents is available through a Secure Remote Center of Access to the Data (*Centre d'Accès Sécurisé Distant*), which I could not obtain for this dissertation.

their experience of sexual violence despite the anonymous nature of the survey, resulting in under-reporting bias in the data and potential problems of sample representativeness. However, the survey methodology aims precisely at avoiding this underestimation as much as possible, which reduces concerns on this side.

Table 3.1: Sample characteristics (total): socio-demographics and assault characteristics

	N	Mean	Sd	Min	Max
Socio-demographic variables					
Gender = Female	1854	0.821	0.383	0	1
Age	1854	40.391	14.746	18	75
Born in France	1854	0.817	0.387	0	1
In a couple	1854	0.420	0.494	0	1
Education					
High education	1852	0.260	0.439	0	1
Medium education	1852	0.441	0.497	0	1
Low eduction	1852	0.299	0.458	0	1
Employment status					
Employed	1854	0.486	0.500	0	1
Unemployed	1854	0.155	0.362	0	1
Inactive	1854	0.359	0.480	0	1
Socioeconomic category					
Upper class	1852	0.094	0.292	0	1
Middle class	1852	0.181	0.385	0	1
Lower class	1852	0.499	0.500	0	1
Other category	1852	0.225	0.418	0	1
Household income					
1st-5th deciles	1854	0.680	0.467	0	1
6th-10th deciles	1854	0.320	0.467	0	1
Number of household members	1854	2.339	1.343	0	9
Living in a priority district	1854	0.236	0.425	0	1
Size of the urban unit					
Country side	1852	0.134	0.341	0	1
$<$ 20,000 inhabitants	1852	0.120	0.325	0	1
20,000-100,000 inhab.	1852	0.164	0.371	0	1
$>$ 100,000 inhabitants	1852	0.424	0.494	0	1
Paris agglomeration	1852	0.157	0.364	0	1
Assault variables					
Rape	1709	0.340	0.474	0	1
Attempted rape	1706	0.261	0.440	0	1
Other sexual assault	1705	0.496	0.500	0	1
Occurred twice or more	1739	0.478	0.500	0	1
Several offenders	1673	0.096	0.294	0	1
Male offender	1469	0.882	0.322	0	1
Victim's partner or ex-partner	1658	0.292	0.455	0	1
Offender drugged or drunk	1607	0.402	0.490	0	1
Physical violence or weapon threat	1662	0.436	0.496	0	1
Physical injury	1751	0.392	0.488	0	1

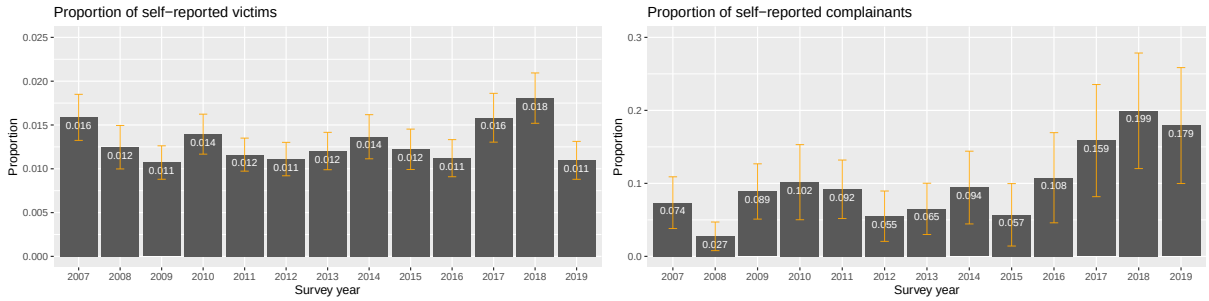
Reading: 82% of victims are women and the age of the average self-reported victim in this sample is 40 years and 4 months old.

4 Impact of victim's characteristics and assault's circumstances

4.1 Descriptive evidence

Figure 4.1 shows the evolution of the proportion of self-reported victims of at least one sexual assault in the two years preceding the survey[9]. This figure has remained relatively stable over time, ranging between 1-2% of the population. In contrast, the proportion of victims who filed a complaint has increased significantly, from 7.4% in 2007 to almost 18% in 2019. This complaint filing variable is built as a dummy that equals 1 if the victim went to the police and eventually filed a complaint, and 0 if the victim did not go to the police to report, or went to the police but did not file a complaint[10]. The vast majority of victims of sexual violence do not report to the police, with only 10% who lodged a formal complaint on average.

Figure 4.1: Evolution of the proportion of self-reported victims and self-reported complainants of sexual violence



Population: People aged between 18-75 and living in metropolitan France (left-hand side) and victims of sexual violence in the two years preceding the survey in that population (right-hand side). Weighted proportions, 10% confidence interval in orange. *Source*: CVS surveys, 2007-2019.

Yet, this reporting rate shows several disparities, as we can see in Table 4.1. One of the most striking concerns the level of education, which seems to be negatively correlated with the reporting propensity. In total, only 6.6% of victims with a diploma level higher than the baccalaureate report having filed a complaint, while this probability is twice as high for those who have not gone further than secondary school. This difference also holds for variables capturing the individual's standard of living, with a lower propensity to report among employed individuals, whose occupation is in the highest socioeconomic category or whose household income is in the top deciles. This pattern seems to support the hypothesis of stronger social stigma for these categories of victims. In addition, the likelihood to report extra-household sexual violence is higher among older victims, while younger victims are more likely to file a complaint when it

[9] In this study, sexual violence is defined as rape, attempted rape or any other sexual assault, but does not include exhibitionism or sexual harassment (as the questions in the CVS survey only concern the former categories).

[10] Table A1.4 shows that in the total sample of victims of sexual violence, around 16% of victims went to the police station to report. Once there, 63.5% lodged a formal complaint, 31.5% did a "main courante", and 9% gave up making a statement.

comes to intra-household sexual violence.

As for the living area, victims coming from rural areas are half as likely to report sexual violence than those living in urban areas. Again, social stigma may be at play, in addition to differential access to justice. Interestingly, inhabitants of priority neighbourhoods report sexual violence to the police significantly more often than those living outside priority areas, with respective reporting propensities of 0.20 and 0.09 on average.

Reporting also appears to differ depending on the circumstances of the assault, as shown in Table 4.2. The most striking differences are related to the use of violence or threat, the occurrence of physical injuries and the fact that the offender was drugged or drunk. The reporting rate ranges from 15% to 20% for assaults with one of these characteristics, compared to 4% to 8% for assaults with no such features, possibly because sexual violence is easier to prove in the former case. The propensity to report sexual violence is also slightly lower when the offender is known to the victim or is her partner than when he is not. A parallel can be drawn with a propensity to report intra-household sexual violence of 15.4% for single victims versus 8.5% for couple victims, given that the spouse is the perpetrator in more than half of cases in this situation (Table A1.2). Finally, victims of extra-household violence are more likely to report rape or attempted rape than other sexual assaults, while the opposite is true for victims of intra-household violence (although the differences are less pronounced).

The study of the reasons why victims did not report also provides interesting insight into the reporting behaviour. Figure 4.2 plots the proportion of victims who mentioned each of the five possible reasons for not reporting. Each modality is constructed as a combination of the non-exclusive responses to the two survey questions "Why didn't you go to the police station?" and "Why didn't you make a report?" for victims who did not take either of these actions. More than two-thirds of victims who did not seek legal redress thought this would have been pointless or preferred to find another solution. The prospect of facing additional hardship is also one of the main reasons for not reporting for 60% of them. Interestingly, Figure A2.1 shows that while this reason comes third for female victims, this is not the case among men, who rather mention that they thought it was not a serious assault.

As these alternatives are not mutually exclusive (i.e. victims can cite several reasons), I use principal component analysis (PCA) to transform the correlated response modalities into new less correlated variables with the least possible loss of information[11]. Table A2.1 and Figure A2.2 show the correlations between the different modalities. Overall, three main groups of reasons emerge following this analysis: (i) was not considered as a serious assault; (ii) did not think it was the best solution ("preferred to find another solution" or "would have been useless"); (iii) to avoid further harmful consequences ("to keep this private", "for fear of reprisals" or "to avoid further hardship").

[11]PCA is a technique generally used in exploratory data analysis for reducing the dimensionality of the data by projecting each data point onto only the first few principal components to obtain lower-dimensional data while preserving as much variability as possible (Jolliffe and Cadima, 2016).

Table 4.1: Proportion of self-reported complainants by individual and household characteristics

		Extra-household			**Intra-household**			**Total**		
		N	Prop.	(SE)	N	Prop.	(SE)	N	Prop.	(SE)
TOTAL		1476	0.102	(0.010)	471	0.099	(0.020)	1854	0.102	(0.010)
Gender	Women	1201	0.108	(0.011)	408	0.095	(0.022)	1523	0.105	(0.010)
	Men	275	0.075	(0.019)	63	0.087	(0.033)	331	0.074	(0.018)
Age	18-29	491	0.082	(0.014)	80	0.157	(0.030)	545	0.084	(0.013)
	30-54	705	0.101	(0.014)	285	0.084	(0.020)	936	0.096	(0.012)
	55-76	280	0.170	(0.028)	106	0.114	(0.036)	373	0.152	(0.025)
Marital status	Single	960	0.104	(0.013)	160	0.154	(0.034)	1076	0.111	(0.012)
	In a couple	516	0.106	(0.016)	311	0.085	(0.024)	778	0.096	(0.014)
Place of birth	In France	1212	0.098	(0.010)	375	0.088	(0.018)	1515	0.094	(0.009)
	Abroad	264	0.131	(0.027)	96	0.105	(0.034)	339	0.131	(0.025)
Educational level	High	392	0.074	(0.016)	106	0.053	(0.019)	481	0.066	(0.013)
	Medium	652	0.095	(0.015)	205	0.070	(0.021)	817	0.097	(0.013)
	Low	430	0.148	(0.021)	160	0.159	(0.042)	554	0.142	(0.019)
Employment status	Employed	699	0.078	(0.012)	243	0.068	(0.019)	901	0.075	(0.010)
	Unemployed	245	0.143	(0.027)	61	0.124	(0.022)	288	0.143	(0.025)
	Inactive	532	0.119	(0.018)	167	0.145	(0.034)	665	0.122	(0.017)
Socioeconomic category	Upper class	125	0.055	(0.025)	50	0.093	(0.014)	174	0.061	(0.023)
	Middle class	264	0.053	(0.014)	85	0.145	(0.042)	336	0.069	(0.015)
	Lower class	734	0.134	(0.017)	248	0.083	(0.025)	925	0.123	(0.015)
	Other	352	0.104	(0.018)	87	0.130	(0.042)	417	0.109	(0.017)
Household income	1st - 5th decile	455	0.119	(0.013)	155	0.106	(0.024)	594	0.116	(0.012)
	6th - 10th decile	1021	0.070	(0.013)	316	0.078	(0.027)	1260	0.070	(0.012)
City size	Countryside	169	0.053	(0.021)	88	0.055	(0.026)	249	0.057	(0.018)
	$<$ 20,000 inhab.	175	0.115	(0.025)	60	0.042	(0.028)	222	0.111	(0.024)
	20,000 - 100,000 inhab.	245	0.088	(0.021)	77	0.119	(0.037)	304	0.096	(0.019)
	$\geq$ 100,000 inhab.	646	0.120	(0.017)	181	0.141	(0.026)	786	0.122	(0.016)
	Parisian agglomeration	240	0.111	(0.024)	64	0.065	(0.025)	291	0.099	(0.021)
Priority district	Living in a QPV/ZUS	359	0.159	(0.026)	111	0.268	(0.044)	437	0.195	(0.025)
	Not living in a QPV/ZUS	1117	0.093	(0.011)	360	0.077	(0.019)	1417	0.089	(0.010)

Note: Victims of at least one sexual assault committed during the two years preceding the survey, living in metropolitan France and aged between 18 and 75. N: unweighted number of victims in each category. Prop.: weighted proportion of victims who filed a complaint within each category and standard errors in parentheses.
Source: CVS surveys, 2007-2019 (pooled data).

Table 4.2: Proportion of self-reported complainants by assault characteristics

		Extra-household			Intra-household			Total		
		N	Prop.	(SE)	*N*	Prop.	(SE)	*N*	Prop.	(SE)
TOTAL		1476	0.102	(0.010)	471	0.099	(0.020)	1854	0.102	(0.010)
Type of assault*	Rape	414	0.188	(0.023)	202	0.099	(0.022)	581	0.164	(0.019)
	Attempted rape	255	0.169	(0.032)	200	0.102	(0.026)	446	0.143	(0.021)
	Other sexual assault	705	0.048	(0.009)	154	0.148	(0.030)	846	0.073	(0.011)
Occurrence	Once	810	0.105	(0.013)	148	0.106	(0.029)	907	0.098	(0.012)
	Twice or more	584	0.098	(0.015)	286	0.089	(0.022)	832	0.103	(0.013)
Number of offenders	Only one	1226	0.107	(0.011)	387	0.098	(0.020)	1513	0.099	(0.009)
	More	112	0.101	(0.028)	31	0.201	(0.015)	160	0.135	(0.029)
Offender's gender**	Male/only men	1175	0.111	(0.011)	168	0.144	(0.032)	1296	0.117	(0.011)
	At least one woman	157	0.087	(0.025)	19	0	(0.000)	173	0.082	(0.023)
Link victim-offender**	Known before	981	0.098	(0.012)	-	-	-	-	-	-
	Unknown before	364	0.128	(0.025)	-	-	-	-	-	-
Location of the assault	In a public space	167	0.116	(0.026)	-	-	-	-	-	-
	Elsewhere	1183	0.104	(0.011)	-	-	-	-	-	-
Victim's partner**	Yes	-	-	-	210	0.093	(0.026)	-	-	-
	No	-	-	-	184	0.123	(0.027)	-	-	-
Victim's partner or ex-partner**	Yes	-	-	-	-	-	-	484	0.117	(0.020)
	No	-	-	-	-	-	-	1174	0.102	(0.011)
Physical violence/weapon use***	Yes	536	0.215	(0.023)	198	0.183	(0.037)	724	0.207	(0.021)
	No	848	0.051	(0.009)	151	0.008	(0.008)	938	0.043	(0.008)
Physical injuries	Yes	501	0.209	(0.023)	197	0.184	(0.033)	687	0.196	(0.019)
	No	901	0.049	(0.009)	235	0.029	(0.015)	1064	0.044	(0.007)
Offender drugged or drunk***	Yes	525	0.148	(0.019)	125	0.211	(0.049)	646	0.154	(0.018)
	No	796	0.078	(0.011)	214	0.059	(0.018)	961	0.074	(0.010)

Note: Victims of at least one sexual assault committed during the two years preceding the survey, living in metropolitan France and aged between 18 and 75 years. *N*: unweighted number of victims in each category. Prop.: weighted proportion of victims who filed a complaint within each category and standard errors in parentheses.

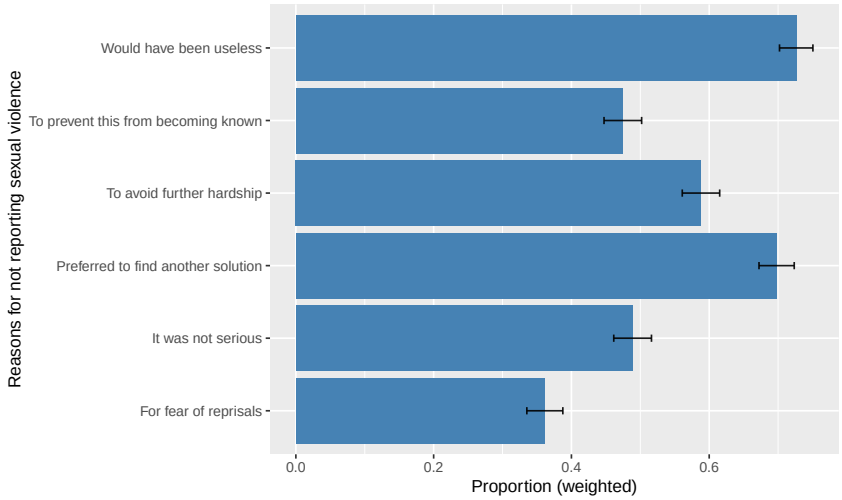
* In the "intra-household sexual violence" and "total" samples, victims may have experienced several different sexual offences (i.e. non-exclusive categories).

** Main perpetrator if more than one aggressor.

*** Only for the years 2009-2019 for victims of sexual violence inside their current household (questions not asked in 2007 and 2008).

Source: CVS surveys, 2007-2019 (pooled data).

Figure 4.2: Reasons for not reporting sexual violence



Population: People aged between 18-75 and living in metropolitan France, who suffered sexual violence in the two years preceding the survey and did not report the assault to the police. Weighted proportions, 10% confidence interval in black. *Source*: CVS surveys, 2007-2019 (pooled data).

4.2 Empirical methodology

To investigate in more detail the influence of socio-demographic and assault characteristics on the victim reporting behaviour, I run a series of regressions to assess the *ceteris paribus* effect of each variable and find which ones drive the reporting decision. I estimate the following equation on the main sample of self-reported victims of sexual violence:

$$Y_{itr} = \alpha + \beta X_i + \gamma Z_i + \mu_t + \phi_r + \epsilon_i \tag{4.1}$$

- Y_{itr} is a dummy variable that equals 1 if the victim i living in region r at survey time t filed a complaint about the assault, and 0 otherwise.
- X_i is a vector of socio-demographic characteristics that includes the following set of variables[12]: Age and its squared; Gender (Female = 1); An indicator for whether the victim is in a relationship; An indicator for whether the victim is born in France; Level of education (high, medium or low); Employment status (employed, unemployed or inactive); Socioeconomic category (upper class, middle class, lower class or other); Number of household members; An indicator for whether the household income is in the 1st-5th deciles; Size of the urban unit (rural area, less than 20,000 inhabitants, between 20,000-100,000 inhabitants, more than 100,000 inhabitants, Paris urban unit); An indicator for whether the victim lives in a priority district.
- Z_i is a vector of assault characteristics that controls for: Type of assault (rape, attempted rape or other sexual assault)[13]; Number of occurrences (once or more than once); Number of offenders (one or more); An indicator for whether the offender was the victim's partner

[12] See Appendix Tables A1.1 and A1.2 for the exact construction of each variable.

[13] As victims of intra-household sexual violence may have experienced several different sexual offences, these categories are not mutually exclusive and there is no reference for the "intra-household" and "total" samples.

or ex-partner; An indicator for whether the assault led to bodily injuries (in addition to rape and sexual assault injuries)[14].

- μ_t are year fixed-effects to eliminate time-varying shocks that affect all individuals simultaneously[15], such as changes in laws, economic situation, or the general evolution of mentalities regarding sexual violence over time.
- ϕ_r are region fixed effects that absorb any time-invariant variation in the outcome variable caused by factors that vary across regions[16], such as different norms or different access to and organisation of police and justice services.

The main specification is a linear probability model (LPM) using OLS estimates, with robust standard errors. I choose to use LPM since coefficients are easier to interpret than with Probit models. Yet, LPM may lack of internal consistency as it can predict probabilities outside of the [0,1] range[17]. Figures A4.1 and A4.2 indeed show that it is the case. Therefore, I have also run Probit model regressions, which lead to qualitatively and quantitatively similar results. Indeed, when comparing the LPM coefficients with the marginal effects derived from the Probit models displayed on Tables A4.1 and A4.2, we see that LPM remains a good approximation of Probit.

4.3 Main specification results

4.3.1 The decision to report

To start the analysis, I compare the influence of socio-demographic and assault characteristics on the reporting decision. To do so, I estimate Equation 4.1 on different sets of variables that I gradually add to the model. Results for either type of sexual violence are reported in Table 4.3. Column (1) only includes socio-demographic variables, Column (2) only assault variables and Column (3) include them all. Columns (4) and (5) respectively add region and year fixed-effects.

First, looking at the adjusted R-squared of the various regressions, we see that assault characteristics play a more important role than victim's characteristics: socio-demographics explain only 3.3% of the variation in the reporting propensity, while assault characteristics explain 6.7% of it. The nature of the dependent variable may explain why none of the R-squared is very large, as it is a binary variable with a small number of "ones".

In addition, much of the reporting behaviour is likely to be explained by unobserved factors also correlated with some explanatory variables, such as psychological reasons or personality

[14]Other variables presented in the descriptive statistics are not included in the model due to a relatively important number of missing values, such as the offender's gender, whether the assault involved physical violence or weapon threat or the offender was drugged or drunk. In fact, the offender was a man almost 90% of the time and the latter two characteristics are also correlated to the fact that the assault resulted in physical injuries.

[15]These year fixed-effects correspond to the year of the survey rather than the year of the assault, due to missing data for the assault year that would further reduce the sample size.

[16]These region fixed-effects use the *Zones d'études et d'aménagement du territoire* variable, which divides metropolitan France into eight large zones.

[17]LPM also contains heteroscedasticity (i.e. the variance of the error term depends on X), which is why I use robust standard errors. Besides, LPM implies constant marginal effect of each explanatory variable that appears in its original form, which may not be completely realistic.

traits[18], and we cannot interpret the results causally since the exogeneity assumption does not hold. Nevertheless, the examination of correlations also provides interesting insights into the phenomenon of sexual crime reporting. Furthermore, even if the survey methodology minimises this problem, results could be biased if sexual violence under-reporting in the data is non-random but correlated with various factors also included in the model. This issue is difficult to address as it is complicated to get the true number of victims, but this should be kept in mind when interpreting the results.

Table 4.3: Determinants of the reporting propensity for victims of sexual violence (total) - LPM

	Dependent variable: Filed a complaint (0/1)				
	(1)	(2)	(3)	(4)	(5)
Age	−0.0003		−0.002	−0.001	0.001
	(0.007)		(0.007)	(0.007)	(0.007)
Age-squared	0.00003		0.00004	0.00004	0.00001
	(0.0001)		(0.0001)	(0.0001)	(0.0001)
Female	0.040		0.007	0.008	0.007
	(0.025)		(0.030)	(0.030)	(0.029)
Born abroad	0.005		0.025	0.027	0.023
	(0.032)		(0.036)	(0.036)	(0.035)
In a couple	−0.011		0.013	0.014	0.013
	(0.022)		(0.025)	(0.025)	(0.025)
Education (Ref: High education)					
Medium education	0.019		−0.004	−0.005	0.001
	(0.026)		(0.027)	(0.027)	(0.028)
Low education	0.047		0.014	0.014	0.019
	(0.031)		(0.034)	(0.034)	(0.035)
Employment status (Ref: Employed)					
Unemployed	0.046		0.034	0.035	0.045
	(0.032)		(0.035)	(0.035)	(0.035)
Inactive	0.008		0.015	0.017	0.026
	(0.053)		(0.055)	(0.057)	(0.055)
Socioeconomic category (Ref: Upper class)					
Middle class	0.020		0.021	0.023	0.021
	(0.038)		(0.038)	(0.038)	(0.036)
Lower class	0.041		0.038	0.038	0.037
	(0.039)		(0.040)	(0.039)	(0.037)
Other	0.050		0.046	0.045	0.037
	(0.060)		(0.063)	(0.064)	(0.062)
Income decile (Ref: 6th-10th decile)					
1st-5th decile	0.005		−0.001	−0.003	−0.004
	(0.021)		(0.022)	(0.022)	(0.022)
Number of household members	−0.002		−0.004	−0.004	−0.003
	(0.009)		(0.010)	(0.010)	(0.010)
Size of the urban unit (Ref: Rural areas)					
Less than 20,000	0.065*		0.083**	0.085**	0.077**
	(0.035)		(0.036)	(0.037)	(0.036)

To be continued on next page

[18]For instance, victims may be deterred to report sexual violence to the police due to shame and embarrassment or because of the mental manipulation exerted by the aggressors on their victims, these phenomena being well-documented in psychology.

Table 4.3: Determinants of the reporting propensity for victims of sexual violence - LPM

20,000-100,000	0.021		0.025	0.029	0.018
	(0.026)		(0.028)	(0.029)	(0.029)
More than 100,000	0.070**		0.070**	0.074**	0.065**
	(0.028)		(0.029)	(0.030)	(0.029)
Paris urban unit	0.048		0.050	0.119**	0.106*
	(0.029)		(0.033)	(0.055)	(0.058)
Living in QP/ZUS	0.101***		0.099**	0.098**	0.083**
	(0.039)		(0.040)	(0.040)	(0.040)
Assault type (not exclusive cat.)					
Rape		0.024	0.021	0.021	0.029
		(0.032)	(0.032)	(0.032)	(0.033)
Attempted rape		0.017	0.015	0.016	0.013
		(0.027)	(0.027)	(0.027)	(0.027)
Other sexual assault		−0.032	−0.027	−0.028	−0.038
		(0.027)	(0.027)	(0.027)	(0.027)
Offender (dummies)					
Several offenders		−0.0003	−0.007	−0.005	0.009
		(0.039)	(0.040)	(0.041)	(0.039)
Victim's partner or ex-partner		−0.019	−0.019	−0.023	−0.015
		(0.027)	(0.026)	(0.026)	(0.027)
Occurred twice or more		−0.029	−0.033	−0.031	−0.028
		(0.022)	(0.021)	(0.021)	(0.021)
Physical injuries		0.156***	0.140***	0.138***	0.135***
		(0.027)	(0.027)	(0.027)	(0.026)
Intercept	−0.077	0.076***	−0.031	−0.101	−0.172
	(0.127)	(0.028)	(0.137)	(0.152)	(0.162)
Region fixed effects	No	No	No	Yes	Yes
Year fixed effects	No	No	No	No	Yes
Observations	1,721	1,584	1,578	1,578	1,578
R^2	0.043	0.071	0.103	0.106	0.132
Adjusted R^2	0.033	0.067	0.088	0.087	0.107
F Statistic	4.064***	17.154***	6.838***	5.540***	5.198***
	(df = 19; 1701)	(df = 7; 1576)	(df = 26; 1551)	(df = 33; 1544)	(df = 45; 1532)

Note: Robust standard errors are reported in parentheses. * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$.

Let us now examine the results in more detail, starting with the effect of socio-demographics. Column (1) of Table 4.3 shows that the patterns observed in the descriptive statistics are reflected in the signs of the coefficients on education, employment status and socioeconomic category. Indeed, the reporting propensity is higher for victims with a low rather than a high level of education, who are unemployed rather than employed, or coming from a lower rather than upper socioeconomic category. However, none of this coefficient is statistically significant at any conventional level, contrasting with the findings of Lizotte (1985).

To verify that this is not due to a heterogeneous effect that offsets for different categories of victims, I repeat this analysis on the extra- and intra-household victim subsamples. I also restrict it to female victims only, for whom the dynamics may be very different from those of men. The last columns of Tables A3.1 and A3.2 show that when it comes to extra-household (resp. intra-household) sexual violence, the probability of lodging a complaint is 8.5 percentage points (resp. 12.5 p.p.) higher for victims coming from a lower (resp. middle) rather than upper

socioeconomic class, significant at the 10% level.

Besides, Column (1) of Table A3.3 shows that female victims with a low rather than a high education level, or coming from a low rather than a high socioeconomic background, are 0.06 more likely to file a complaint on average (significant at the 10% level). This confirms the hypothesis of stronger social stigma for higher-educated or more advantaged victims. However, these effects disappear when controlling for assault characteristics. The changes in coefficients on education when gradually adding the assault-type variables also suggest that higher-educated victims are more likely to report having suffered assaults other than rape/attempted rape than lower-educated ones. A possible interpretation is that this could be due to a difference in the type of sexual crime these two groups of victims suffer.

Contrary to Allen (2007), the victim's age does not have a significant influence on the reporting decision when we consider all types of victims, but it does have a statistically significant and positive effect when we restrict the analysis only to women. The impact of the victim's marital status should be analysed depending on the type of sexual violence. While being in a couple is associated with a higher reporting propensity for extra-household violence (though not statistically different from zero), it decreases the probability of lodging a complaint by 11.6 percentage points for intra-household violence (significant at the 10% level), which seems quite logical as the offender is the victim's spouse more than half of the time (Tables A3.2, A3.1, A1.2).

Furthermore, as expected, the size of the urban unit has a significant impact on reporting. Table 4.3 displays a statistically significant and 0.06 to 0.11 higher probability of reporting for victims living in urban rather than rural areas, reinforcing the role of social stigma and potentially reflecting unequal access to legal services across territories. Also in line with the descriptive statistics, living in a priority district is associated with a statistically significant increase in the reporting propensity. Though insignificant, the negative sign for the coefficient on the number of household members for intra-household sexual violence (Table A3.2) is consistent with the information avoidance theory. Indeed, compared to extra-household sexual violence where this effect is almost zero, victims of intra-household sexual violence may want to avoid hurting other household members by disclosing the assault.

Regarding assault characteristics, the estimated coefficient on the dummy for whether the assault resulted in physical injury is positive and statistically significant for all specifications. Column (4) shows that having suffered an assault that led to physical injuries is associated on average with a 13.5 percentage points increase in the reporting propensity (significant at the 1% level). This underpins the role of "ancillary evidence" mentioned by Allen (2007), suggesting that this factor would support the victims' credibility and increase the offender's probability of apprehension, in addition to decreasing the reputational costs victims see in reporting. Besides, this evidence seems to have more influence than the severity of the crime itself, given that the average reporting rate for a rape not resulting in physical injury is 6.7%, compared to 13.8% for a sexual assault other than rape/attempted rape with additional physical injury.

None of the other assault characteristics has a statistically significant effect on the total sample of victims. However, for extra-household violence (Column (4) of Table A3.1), having

suffered a sexual assault more than once negatively correlates the likelihood of filing a complaint by almost 5 percentage points (significant at the 5% level), and having been assaulted by more than one assailant or by someone known to the victim before the assault reduces it by 8.4 and 7.1 percentage points respectively (significant at the 10% level). Unlike the descriptive statistics, having been attacked by one's partner for intra-household sexual violence does not reduce the reporting propensity, though the coefficients are not statistically different from zero (Table A3.2). Moreover, contrary to previous findings in the literature (Williams, 1984), having been assaulted in the street or on public transport does not influence the complaint behaviour here. Finally, Table 4.3 suggests that having suffered rape increases the reporting propensity, while having suffered other sexual assaults decreases it. Focusing on the subsample of extra-household sexual violence (for which those categories are mutually exclusive), Table A3.1 shows that this difference is statistically significant, with a 0.11 to 0.12 higher probability to report rape or attempted rape than other sexual assaults (significant at the 1% level).

Therefore, these results suggest that overall, the victim's living area and occurrence of physical injuries are the factors that most influence the victim's decision to file a complaint. The results for victims of extra-household violence confirm that assault characteristics play a more important role than victim's socio-demographics in this decision, both in terms of magnitude and significance of the coefficients. Besides, the fact that most coefficients are insignificant for intra-household violence is likely to be due to power issues given the small sample size.

4.3.2 Reasons for not reporting

After having better understood the factors influencing the decision to report sexual violence, I focus on the reasons for not reporting. To do so, I restrict the analysis to victims who did not report the assault to the police, and re-estimate Equation 4.1 on this subsample. The new outcomes of interest are dummy variables that equal 1 if the victim mentioned a given category of reasons and 0 otherwise, the three non-exclusive categories being: (i) was not considered as a serious assault; (ii) did not think it was the best solution and (iii) to avoid further harmful consequences[19]. Table 4.4 reports the estimates from the regressions of these three groups of reasons on all socio-demographic and assault characteristics, and including region and year fixed effects to capture any time or regional trends in the reasons for not reporting.

Many interesting patterns emerge from this table. First, confirming the descriptive statistics, the propensity to say that it was not a serious assault is 17 percentage points lower for female than male victims (significant at the 1% level). At the same time, women are significantly more likely than men to say that they did not think it was the best solution or wanted to avoid further harmful consequences. The victim's age also has a small positive effect on the likelihood of mentioning this last category of reason. However, this probability decreases with the number of household members, reflecting the role of family support in avoiding this feeling.

As for the educational level, Column (3) shows that compared to highly-educated victims,

[19] See Section 4.1 for more details on the construction of these categories.

those with a low or medium level of education are respectively 13.8 and 15.4 percentage points more likely to cite further harmful consequences as a reason for not reporting sexual violence (significant at the 1% level). It may be that higher-educated victims are closer to the legal system and have more resources to engage in this process, reducing relative concerns of further hardship. Besides, the victim's socioeconomic category seems to influence the probability of saying it was not serious, with a 0.12 lower chance of mentioning this reason for middle-class compared to upper-class individuals (significant at the 10% level). This result again reinforces the hypothesis of greater social stigmatisation of victims from the highest socioeconomic category when it comes to reporting sexual violence.

Regarding assault characteristics, the likelihood of saying that the assault was not serious or reporting was not the best solution is significantly higher for victims who suffered an assault other than rape/attempted rape. At the same time, the prospect of further harmful consequences is more important with attempted rape. Interestingly, victims are more likely to cite this last reason when they were assaulted more than once, suggesting that they may be afraid of potential reprisals from the offender. Individuals who experienced assaults involving multiple rather than a single offender are also significantly less likely to say that it was not a serious assault or that reporting was not the best solution. Finally, having suffered physical injury has a statistically significant impact on the likelihood of saying it was not serious, with a 0.13 lower chance of mentioning that reason. Almost mechanically, it has a significant positive effect on the probability of fearing further adverse consequences, increasing this probability by 19 percentage points.

Table 4.4: Determinants of the reasons for not reporting sexual violence - Full specification (LPM)

	Dependent variable: Reason (0/1)		
	Was not considered as a serious assault	Did not think it was the best solution	To avoid further harmful consequences
	(1)	(2)	(3)
Age	−0.001	−0.001	0.017**
	(0.008)	(0.006)	(0.008)
Age-squared	0.00002	−0.00002	−0.0002**
	(0.0001)	(0.0001)	(0.0001)
Female	−0.169***	0.070*	0.114**
	(0.047)	(0.038)	(0.049)
Born abroad	0.049	−0.020	0.010
	(0.055)	(0.040)	(0.054)
In a couple	0.017	−0.010	0.017
	(0.040)	(0.029)	(0.040)
Education (Ref: High education)			
Medium education	0.050	0.004	0.138***
	(0.044)	(0.026)	(0.043)
Low education	0.015	−0.042	0.154***
	(0.056)	(0.036)	(0.053)
Employment status (Ref: Employed)			

To be continued on next page

Table 4.4: Determinants of the reasons for not reporting sexual violence - Full specification (LPM)

Unemployed	−0.081	0.024	−0.043
	(0.058)	(0.033)	(0.048)
Inactive	−0.040	0.022	0.035
	(0.072)	(0.044)	(0.059)
Socioeconomic category (Ref: Upper class)			
Middle class	−0.118*	0.032	0.020
	(0.067)	(0.042)	(0.065)
Lower class	−0.046	0.009	0.051
	(0.066)	(0.040)	(0.061)
Other	−0.008	−0.010	−0.017
	(0.091)	(0.054)	(0.079)
Income decile (Ref: 6th-10 decile)			
1st-5th decile	−0.009	−0.011	−0.031
	(0.041)	(0.024)	(0.038)
Number of household members	0.017	−0.010	−0.031*
	(0.017)	(0.015)	(0.018)
Size of the urban unit (Ref: Rural areas)			
Less than 20,000	0.040	−0.029	0.055
	(0.064)	(0.047)	(0.061)
20,000-100,000	0.151**	0.053	−0.020
	(0.063)	(0.043)	(0.061)
More than 100,000	0.087	0.074*	0.032
	(0.055)	(0.038)	(0.051)
Paris urban unit	0.195	−0.005	0.014
	(0.173)	(0.130)	(0.142)
Living in QP/ZUS	0.028	0.039	0.066
	(0.053)	(0.036)	(0.047)
Assault type (not exclusive cat.)			
Rape	−0.058	0.011	0.066
	(0.050)	(0.043)	(0.048)
Attempted rape	−0.043	0.001	0.084*
	(0.048)	(0.042)	(0.046)
Other sexual assault	0.080*	0.074*	0.016
	(0.047)	(0.043)	(0.046)
Offender (dummies)			
Several offenders	−0.137**	−0.162***	−0.062
	(0.066)	(0.054)	(0.061)
Victim's partner or ex-partner	0.012	0.054	0.007
	(0.046)	(0.033)	(0.042)
Occurred twice or more	−0.030	0.030	0.058*
	(0.038)	(0.023)	(0.035)
Physical injuries	−0.134***	0.020	0.187***
	(0.039)	(0.027)	(0.032)
Intercept	0.529**	0.817***	0.013
	(0.261)	(0.193)	(0.235)
Region fixed effects	Yes	Yes	Yes
Year fixed effects	Yes	Yes	Yes
Observations	1,335	1,352	1,352
R^2	0.112	0.121	0.153
Adjusted R^2	0.081	0.090	0.124
F Statistic	3.623***	3.985***	5.260***
	(df = 45; 1289)	(df = 45; 1306)	(df = 45; 1306)

Note: Robust standard errors are reported in parentheses. $^{*}p < 0.1$; $^{**}p < 0.05$; $^{***}p < 0.01$.

5 Impact of two high-profile sexual assault cases

5.1 Setting and first evidence

As mentioned earlier, the social stigma surrounding being a victim of sexual misconduct is likely to be a significant barrier to reporting. Yet, norms perception is a dynamic process and social norms are amenable to change (Tankard and Paluck, 2016). In this context, events that challenge misrepresentations and norms regarding sexual violence are likely to affect sexual assault reporting. To investigate this mechanism, I exploit the longitudinal dimension of the CVS survey by studying the evolution of reporting over time. In particular, I focus on the impact of two prominent sexual assault cases that received intense media coverage in France and can be considered as an exogenous shock to gender norms and stereotypes about sexual violence.

The first is the criminal case involving the economist and politician Dominique Strauss-Kahn (DSK) in May 2011, following the allegations of sexual assault and attempted rape by Nafissatou Diallo, a hotel maid at the Sofitel in New York. At the time, DSK led voting intentions for the French 2012 presidential election, explaining why this affair has received a lot of media attention in France. The second is the series of public disclosures of sexual harassment and assault in work-related contexts by the film producer Harvey Weinstein in October 2017. Although the case originated in the United States, it also had a significant media impact in France, notably with the creation of the #BalanceTonPorc movement just before the global #MeToo movement.

By picturing a different image of sexual violence, victims and the reporting process, these high-profile cases may have influenced the victim's decision to report sexual violence to the authorities in different ways, in line with the various theoretical predictions made in Section 2.2. The DSK case portrayed a rather bad image of the reporting process and the victim's credibility, exposing all the difficulties and costs associated with the decision to report. This may have led to a backlash effect and a decrease in sexual assault reporting. On the contrary, the Weinstein affair is likely to have encouraged reporting as the succession of assaults disclosures made it possible to reduce the social stigma endured by victims. This case also involved well-known victims from a more privileged social background and whose word is probably more credible, which may have affected the overall credibility of victims' allegations. However, victims could be impacted differently. Assuming a role model effect, immigrant women or women from a disadvantaged background are likely to have been most affected by the DSK affair. In contrast, the Weinstsein case may have mainly reached white women of high socioeconomic status.

The first way to describe a potential impact of changing social norms on the reporting behaviour is to compare the time evolution of the share of self-reported victims of sexual violence to the share of victims who have filed a complaint. For instance, with no change in the victimisation rate, an increase in the reporting rate among different socioeconomic categories could indicate a drop in social stigma associated with reporting sexual violence.

Table 5.1 shows that while the victimisation rate has remained fairly stable over time (ranging from 1.2% to 1.4% of the population), the reporting rate exhibits notable fluctuations. Starting

from 7.7% in 2007-2011, it slightly decreased to 6.8% in 2012-2015 and more than doubled at the end of the period, with a reporting rate of more than 16% in 2016-2019. We can also see that the share of self-reported victims among various socioeconomic categories has not changed significantly over time. Conversely, the share of self-reported complainants has witnessed significant changes.

The increase in the reporting rate is the most important for victims with low and medium education, from lower socioeconomic category, with lower income and living in urban areas. For instance, it rose from 9% to 25.5% between the second and third periods for the least educated victims. The other categories also experienced a large increase in the reporting rate, which has doubled over that time for victims with the highest standard of living or from rural areas. These developments seem to indicate certain democratisation of the reporting behaviour, potentially due to a #MeToo effect following the Weinstein affair, a hypothesis that I will test below.

Table 5.1: Evolution of the share of self-reported victims and share of self-reported complainants by socio-demographic categories

		Survey period		
		2007-2011	2012-2015	2016-2019
Share of self-reported victims of sexual violence				
		$N = 65006$	$N = 48801$	$N = 44165$
Total	Prop. of self-reported victims	0.013 (0.001)	0.012 (0.001)	0.014 (0.001)
Level of education	High	0.010 (0.001)	0.011 (0.001)	0.012 (0.001)
	Medium	0.012 (0.001)	0.011 (0.001)	0.014 (0.001)
	Low	0.016 (0.001)	0.016 (0.002)	0.017 (0.002)
Socioeconomic category	Upper class	0.006 (0.001)	0.006 (0.001)	0.006 (0.001)
	Middle class	0.009 (0.001)	0.009 (0.001)	0.013 (0.001)
	Lower class	0.014 (0.001)	0.013 (0.001)	0.014 (0.001)
	Other	0.023 (0.002)	0.025 (0.003)	0.033 (0.004)
Household income	1st - 5th decile	0.019 (0.001)	0.016 (0.001)	0.020 (0.001)
	6th - 10th decile	0.008 (0.001)	0.008 (0.001)	0.009 (0.001)
Territory	Rural area	0.010 (0.001)	0.008 (0.001)	0.006 (0.001)
	Urban area	0.014 (0.001)	0.014 (0.001)	0.016 (0.001)
Share of self-reported victims who filed a complaint				
		$N = 789$	$N = 536$	$N = 529$
Total	Prop. who filed a complaint	0.077 (0.010)	0.068 (0.012)	0.161 (0.023)
Level of education	High	0.087 (0.027)	0.056 (0.019)	0.051 (0.020)
	Medium	0.062 (0.012)	0.062 (0.020)	0.175 (0.035)
	Low	0.094 (0.020)	0.090 (0.026)	0.255 (0.051)
Socioeconomic category	Upper class	0.011 (0.011)	0.020 (0.017)	0.164 (0.072)
	Middle class	0.061 (0.022)	0.072 (0.028)	0.076 (0.032)
	Lower class	0.066 (0.013)	0.089 (0.021)	0.228 (0.039)
	Other	0.133 (0.028)	0.055 (0.022)	0.132 (0.036)
Household income	1st - 5th decile	0.087 (0.013)	0.083 (0.018)	0.186 (0.030)
	6th - 10th decile	0.055 (0.017)	0.038 (0.013)	0.119 (0.030)
Territory	Rural area	0.043 (0.018)	0.039 (0.023)	0.092 (0.051)
	Urban area	0.087 (0.012)	0.071 (0.014)	0.169 (0.024)

Note: Victims of at least one sexual assault committed during the two years preceding the survey, living in metropolitan France and aged between 18-75. Weighted proportions and standard errors in parentheses.
Source: CVS surveys, 2007-2019.

5.2 Empirical methodology

I use a difference-in-difference strategy to estimate the impact of these two high-profile cases on the reporting behaviour of sexual assault victims. More precisely, I compare the evolution of the reporting rate for sexual violence and non-household physical violence before and after coverage of each affair. Victims of non-household physical violence (referred to as "physical violence" hereafter) are defined as victims who have been physically assaulted by someone not currently living with them over the past two years. For example, this may refer to situations of physical violence between road users, of a discriminatory nature or following a request from a stranger. However, this does not include physical violence for theft or attempted theft, or physical violence by someone currently living with the respondent, measured in other categories of the CVS survey. The intuition is that while the DSK and Weinstein affairs may have affected reporting of sexual violence, they are unlikely to have affected reporting of other violence.

Although the choice of victims of physical violence as a control group may be questionable, especially given the changes in representations of physical violence over time, it is probably a relevant counterfactual. First, the reporting rate for physical violence varies between 20% and 30% in the data[20]. This figure is larger than the reporting rate for sexual violence (10% on average), but is among the lowest reporting rates for non-sexual violence[21]. Moreover, as physical violence is also considered the most serious type of individual violence, along with sexual violence, it seems logical to use this control group rather than other types of violence such as thefts, insults or threats. Besides, 79.5% of victims of sexual violence in the sample are women, and 20.5% are men. While one may be worried that an opposite difference is observed for victims of physical violence, the gap is not as large: 47.2% of victims of physical violence are women and 52.8% are men, notably because this category also includes gender-based physical violence. Finally, some factors of non-reporting are likely to be similar between these two types of violence, such as the availability of evidence or knowing the perpetrator.

For this analysis, I merge data from the face-to-face questionnaire on individual violence with data from the self-administered questionnaire on sensitive violence. I restrict the sample to individuals who experienced either sexual or physical violence, pooling the data from all survey years, which amounts to 6,167 observations. I estimate the following equation:

$$Y_{itr} = \beta_0 + \beta_1 SV_i + \beta_2 Post_t + \beta_3 SV_i \times Post_t + \gamma X_{it} + \phi_r + \epsilon_i \tag{5.1}$$

Where Y_{itr} is the legal remedy implemented by victim i living in region r and assaulted in year t.

[20] As a comparison, the reporting rate for theft or attempted theft with violence or threat ranges from 30% to 50%, and the reporting rate for theft or attempted theft without violence nor threat ranges from 30% to 40%.

[21] The second-lowest reporting rate concerns victims of intra-household physical violence, which varies from 7% to 15%. However, since almost half of the victims of intra-household sexual violence also suffered intra-household physical violence (mainly by the same offender), and since in this case, the survey does not make it possible to precisely identify for which type of violence the victim has made a complaint, it is more complicated to use it as a control group. Conversely, the questions on non-household physical and sexual violence are decoupled as they are asked in two different questionnaires, making it possible to identify the complaint behaviour associated with each of these types of violence.

As before, it is a binary variable that equals 1 if the victim filed a complaint and 0 otherwise, but this time using the *assault* year instead of the *survey* year. Since I only have information on the year of the assault and not the year or date of the complaint, I am making the assumption that victims who reported having filed a complaint did so in the same year as their assault[22].

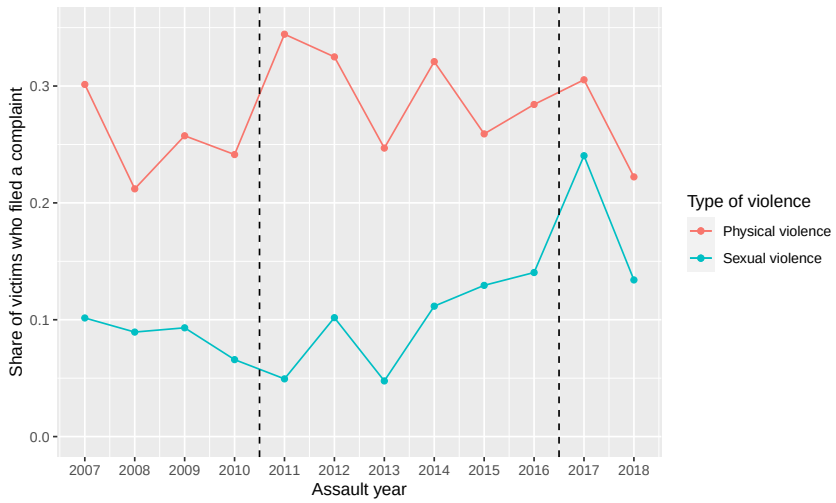
SV_i is a binary variable indicating that the individual is in the treatment group, i.e. it equals 1 if the victim i suffered a sexual assault and 0 if she was physically assaulted[23].

$Post_t$ is an indicator of treated time periods, i.e. it equals 1 if the assault took place in the year of the case or in subsequent years. To distinguish the effect of the two affairs, I restrict the time frame to 2007-2011 for the analysis of the DSK case and to 2012-2018 for the Weinstein case. Using data from Google Trends, Figure A5.2 shows that the DSK case was not mentioned much after 2011; in contrast, the Weinstein case and #MeToo seem to have lasted longer in the media. Thus, $Post_t = 1$ for assault year 2011 in the former case and $Post_t = 1$ for assault years 2017 and 2018 in the latter case. I further discuss this time frame choice and its underlying assumptions in Section 5.4.4, where I estimate the effect of both cases simultaneously.

X_{it} is a vector of socio-demographic characteristics that were already included in Equation 4.1 and also matter in the reporting decision[24]. Finally, ϕ_r are region fixed-effects that capture the effect of factors fixed over time and specific to regions.

The coefficient of interest is β_3, which represents the differential impact of the high-profile case (using the assault year as a proxy for the beginning of each case) on the reporting of sexual violence relative to the reporting of physical violence.

Figure 5.1: Pre-trends for filing a complaint - raw data



Population: Victims of sexual violence or physical violence in the two years preceding the survey. Weighted proportions. *Source*: CVS surveys, 2007-2019.

[22] As a reminder, each CVS survey carried out in year N records the assaults suffered in years $N-1$ and $N-2$. This assumption is thus necessarily verified for people assaulted in $N-1$, but not necessarily for those assaulted in $N-2$. Still, I need to make this assumption to be able to use these observations.

[23] When the individual suffered both types of violence, I include them in the treatment group, given fewer observations for sexual violence than physical violence.

[24] As I apply this DiD strategy on repeated cross-sections and not panel data, the individuals change each year. The coefficients of covariates X_{it} can therefore be estimated since they vary over time in the two groups.

The identifying assumption is that without these two affairs and their coverage, the reporting rate for sexual and physical crimes would have evolved in the same way from the pre-period to the post-period. A first argument favouring the parallel trend assumption is that the representations associated with these two types of violence are likely to have evolved in the same way, with greater recognition of such violences over time. Moreover, a parallel evolution is also plausible due to the improvement in the quality of police services in general and an increased incentive to report these two types of violence, for which the complaint rate is relatively low.

The pre-trends shown in Figure 5.1 provide some support to this assumption, though it is not a necessary nor sufficient condition for identification[25]. This graph plots the evolution of the proportion of victims who filed a complaint about sexual and physical assault respectively over time. Although the levels differ a lot, the two reporting rates followed a somewhat parallel evolution up to 2010. We see an apparent change between 2010 and 2011, the year during which the DSK case occurred, with a large increase in physical violence reporting that contrasted with a small decrease in sexual violence reporting. From 2012, the trends become parallel again (except between 2014 and 2015). Finally, there is a new shift in 2017, the year of the Weinstein affair, with a much larger increase in sexual violence reporting[26]. I also reinforce confidence in the validity of this strategy by performing a test of parallel pre-trends in Section 5.4.1.

For an omitted variable to explain the results, it would need to have a non-linear change after the affair that affects the reporting of sexual crimes more than physical crimes. Given that there were no changes to laws concerning sexual violence over that period in France, and as these two affairs are certainly the most important events that may have affected sexual violence representations at that time, I would argue that such possibility is unlikely.

However, this methodology suffers from several limitations to measure the effect of the Weinstein case and #MeToo. First, I use the post-DSK years as a pre-trend to estimate the post-Weinstein effect, and the 2012-2016 period could therefore also be affected by the DSK case. Still, I show that it is not likely to be the case in Section 5.4.4.

Moreover, as this affair started in October, it can only have affected the last three months of 2017. But as I will show later with administrative data, there appears to have been a relatively large spike in complaints from October onwards, which supports the rationale for not excluding this year since I cannot start the treatment in October 2017 with survey data. Nevertheless, including the whole of 2017 in the treated period may underestimate the actual effect. I also perform a robustness check by restricting the treated period to 2018 only in Section 5.4.4.

Besides, as the last CVS survey was in 2019, I can only use one survey year to measure the complaint rate in 2018 (the 2019 survey), whereas I use two survey years to compute all other

[25]For instance, past trends can be identical but the control group may be affected by a group-specific shock at the period of the treatment. Alternatively, past trends may be different, but the common trend assumption at the period of the treatment be verified.

[26]In line with Section 4, it is interesting to note that the complaint rate for physical violence crimes alone is 27.6% on average, compared to 20% for sexual violence crimes involving physical injury. Hence, among crimes involving physical violence, the additional existence of sexual violence seems to decrease the reporting likelihood.

annual complaint rates (e.g. the 2013 and 2012 surveys to measure the complaint rate in 2011). Therefore, the effect measured for 2018 is likely to be conservative as it could be underestimated by construction, which may explain the drop in sexual violence reporting in 2018 compared to 2017 in Figure 5.1.

5.3 Results

5.3.1 Main results

Table 5.2 displays the results from the estimation of Equation 5.1 using linear probability models and OLS estimates, with robust standard errors. Columns (1) and (3) present the estimated coefficients on the total sample of victims, while Columns (2) and (4) restrict the analysis to female victims only.

Table 5.2: Difference-in-difference estimation for the effect of the two high-profile cases

	Dependent variable: Filed a complaint (0/1)			
	All victims	Women only	All victims	Women only
	(1)	(2)	(3)	(4)
Sexual violence	-0.193^{***}	-0.185^{***}	-0.204^{***}	-0.205^{***}
	(0.024)	(0.030)	(0.023)	(0.028)
Assault post 2011	0.092^{**}	0.044		
	(0.037)	(0.054)		
Sexual violence*Assault post 2011	-0.118^{***}	-0.092		
	(0.045)	(0.061)		
Assault post 2017			-0.032	-0.083^{**}
			(0.029)	(0.041)
Sexual violence*Assault post 2017			0.138^{***}	0.163^{***}
			(0.053)	(0.061)
Intercept	0.087	0.279	0.001	0.174
	(0.145)	(0.214)	(0.086)	(0.120)
Socio-demographic controls	Yes	Yes	Yes	Yes
Region fixed effects	Yes	Yes	Yes	Yes
Period	2007-2011	2007-2011	2012-2018	2012-2018
Observations	2,158	1,318	2,692	1,589
Adjusted R^2	0.070	0.091	0.077	0.090

Note: OLS estimates from LPM. Robust standard errors in parentheses *p<0.1; **p<0.05; ***p<0.01

Column (1) shows that the probability of filing a complaint about sexual assault decreases by 11.8 percentage points on average for assaults that occurred in 2011 (statistically significant at the 1% level). With an average baseline reporting rate for sexual violence of 0.075, this amounts to a consequent 157% drop in sexual crimes reporting. To the extent that we can assume there were no further changes during this period that affected the treatment group and could act as a confounder, this result can be interpreted as an important adverse effect of the DSK case on the reporting behaviour of victims of sexual violence, in line with the ex-ante hypothesis.

However, it may also be driven by a shock specific to the control group at the time of treatment, since we notice a significant increase in the probability to file a complaint about

physical assault in 2011. As no major changes to laws or in society relating to physical violence occurred during this period, I would argue that this is unlikely. The estimated coefficients on the subsample of women nevertheless suggest that male victims of physical violence partly drive these results. Indeed, Column (2) shows that the coefficient on $SV_i \times Post_{2011}$ is no longer statistically different from zero for women, but neither is the increase in reporting for physical violence in 2011. Still, Figure A5.1 suggests that the parallel trend assumption is not more likely to hold when we restrict the two groups to female victims only. Overall, identification seems better achieved without restricting the sample to women for the DSK case, and I support the validity of using this treatment and control group in the robustness checks.

Regarding the Weinstein case, Column (3) shows that, on average, the likelihood of filing a complaint about sexual violence increases by 13.8 percentage points for assaults that occurred in 2017 or 2018 (statistically significant at the 1% level). Compared to the baseline reporting rate (0.124), this corresponds to a sizeable 111% increase in sexual violence reporting. This time, this result is mainly driven by a greater increase in the rate of complaints about sexual rather than physical violence. When focusing on female victims only, the impact is even more important, with a 16.3 percentage points increase in sexual assault reporting in 2017-2018, significant at the 1% level. With an average baseline reporting rate of 0.12 in this subsample, this is an effect of 135%. This time, Figure A5.1 shows that the common trend assumption seems more plausible when restricting the analysis to women. Therefore, this finding supports the hypothesis of a strong positive effect of the Weinstein case and subsequent #MeToo movement on the reporting of sexual violence. As discussed above, this effect is likely to be a lower bound due to a possible underestimation of the reporting rate on the 2017-2018 treated period.

5.3.2 Heterogeneity analysis

To better understand how these two high-profile cases affected reporting, I study whether and to what extent the effects differ according to various socio-demographic characteristics, in line with Section 4. Previous results suggested that female victims of physical violence are a more appropriate control group to analyse the impact of the Weinstein case, but not necessarily for the DSK case. Yet, as we have seen with the determinants to filing a complaint, the dynamics can be quite different depending on gender and women are the main victims of sexual violence. Therefore, I perform this heterogeneity analysis restricting the sample only to women.

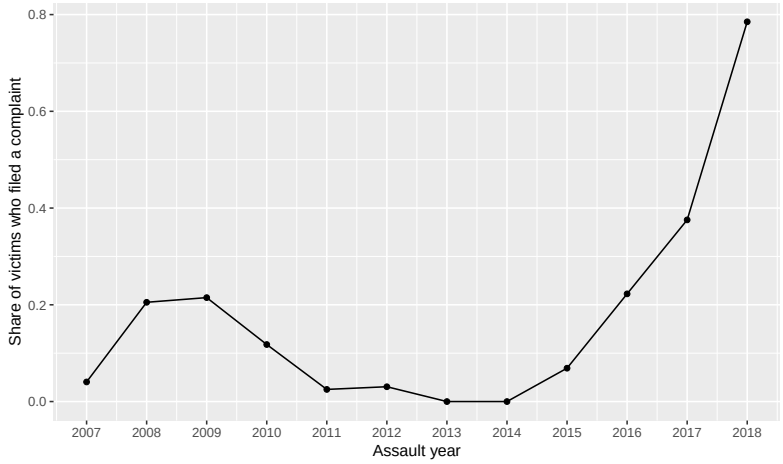
Table 5.3 presents the results by place of birth. Compared to previous results, there was a considerable and statistically significant reduction of 40 percentage points of sexual violence reporting in 2011 among victims born abroad. Even though this finding seems in line with the hypothesis of a stronger effect of the DSK case among immigrant victims, it is complicated to interpret because of the small number of female victims born abroad in the sample and weak parallel pre-trends for this subgroup of victims. In addition, Figure 5.2 shows that the reporting rate for sexual violence was already following a decreasing trend before 2011 for foreign-born women. Conversely, the rise in sexual violence reporting from 2017 was mainly driven by victims

Table 5.3: DiD estimation – Effect by place of birth (female victims)

	Dependent variable: Filed a complaint (0/1)			
	Born in France		Born abroad	
	(1)	(2)	(3)	(4)
Sexual violence	−0.181***	−0.201***	−0.202**	−0.227***
	(0.030)	(0.029)	(0.096)	(0.065)
Assault post 2011	0.025		0.201	
	(0.054)		(0.154)	
Sexual violence*Assault post 2011	−0.043		−0.403**	
	(0.062)		(0.168)	
Assault post 2017		−0.112***		0.155
		(0.039)		(0.150)
Sexual violence*Assault post 2017		0.131**		0.155
		(0.056)		(0.200)
Intercept	0.301	0.233*	−0.172	−0.149
	(0.185)	(0.120)	(0.385)	(0.216)
Socio-demographic controls	Yes	Yes	Yes	Yes
Region fixed effects	Yes	Yes	Yes	Yes
Year restrictions	2007-2011	2012-2018	2007-2011	2012-2018
Observations	1,140	1,384	178	205
Adjusted R^2	0.077	0.097	0.111	0.241

Note: OLS estimates from LPM. Robust standard errors in parentheses *p<0.1; **p<0.05; ***p<0.01

Figure 5.2: Evolution of the reporting rate for sexual violence among foreign-born women



Population: Foreign-born women victims of sexual violence in the two years preceding the survey. Weighted proportions. *Source*: CVS surveys, 2009-2019.

born in France according to Table 5.3, which supports the role model effect hypothesis for the Weinstein case. Interestingly, Table 5.4 shows that this increase was most important for victims in the first fifth income deciles, with a 0.21 higher chance of reporting (significant at the 5% level) – which is not entirely consistent with the ex-ante hypothesis.

Further results of the heterogeneity analysis can be found in the Appendix. I focus on the Weinstein case and subsequent #MeToo movement, for which the estimated effects are more robust. Table A6.1 suggests that this case mainly affected younger women, and Tables A6.2 and A6.3 that the effect was significantly larger for medium-educated and middle-class

Table 5.4: DiD estimation – Effect by household income (female victims)

	Dependent variable: Filed a complaint (0/1)			
	1st - 5th decile		6th - 10th decile	
	(1)	(2)	(3)	(4)
Sexual violence	−0.213***	−0.220***	−0.157***	−0.142***
	(0.038)	(0.037)	(0.055)	(0.043)
Assault post 2011	0.063		−0.004	
	(0.068)		(0.111)	
Sexual violence*Assault post 2011	−0.075		−0.058	
	(0.079)		(0.126)	
Assault post 2017		−0.083		−0.051
		(0.059)		(0.060)
Sexual violence*Assault post 2017		0.206**		0.057
		(0.084)		(0.084)
Intercept	0.277	0.300	0.419	−0.006
	(0.256)	(0.198)	(0.412)	(0.141)
Socio-demographic controls	Yes	Yes	Yes	Yes
Region fixed effects	Yes	Yes	Yes	Yes
Period	2007-2011	2012-2018	2007-2011	2012-2018
Observations	865	1,054	453	535
Adjusted R^2	0.107	0.113	0.055	0.059

Note: OLS estimates from LPM. Robust standard errors in parentheses *p<0.1; **p<0.05; ***p<0.01

victims. Moreover, Table A6.4 shows that it principally reached inhabitants of large French cities, potentially due to different degrees of exposure to media coverage and stigma experienced by victims across territories.

These findings are not fully consistent with the hypothesis of a role model effect since victims from a more privileged social background were not significantly more affected by the Weinstein case. This notably goes against the criticism of the #MeToo movement as having primarily reached women of high socioeconomic status, and could be explained by the fact that the Weinstein affair and #MeToo have had more resonance within the "general public". Alternatively, another possible interpretation is that the stigma experienced by upper-class or well-off victims may be too strong for this high-profile case to have affected their reporting behaviour.

5.4 Robustness checks

This section presents the results from several robustness checks to support the validity of my empirical strategy, and notably the assumption that no time-varying differences exist between the treatment and control groups. I also use alternative specifications and perform a similar estimation with administrative data to compare the results.

5.4.1 Testing for a common pre-trend

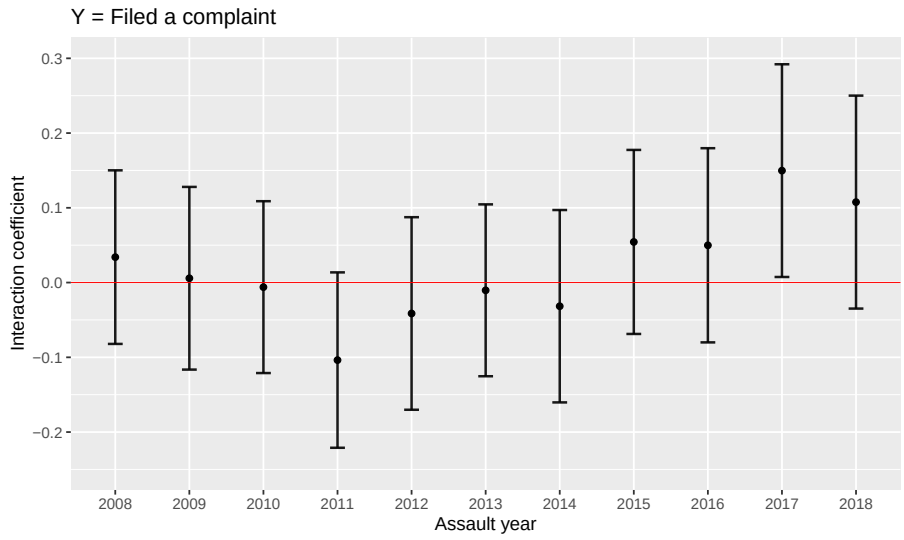
A first way to assess the validity of the common trend assumption is to compare changes in the outcome for victims of sexual violence and victims of physical violence repeatedly before the high-profile cases occurred. If the outcome trend moves in parallel before the case in question, it

likely would have continued moving in tandem in the absence of the high-profile case. To test whether the pre-trends shown Figure 5.1 are parallel, I run the following specification on the total sample of victims, where the variables are similar to the ones of the difference-in-difference specification described in Equation 5.1:

$$Y_{itr} = \beta_0 + \beta_1 SV_i + \beta_2 Year_t + \beta_3 SV_i \times Year_t + \gamma X_{it} + \phi_r + \epsilon_i \tag{5.2}$$

Figure 5.3 plots the estimated coefficients $\hat{\beta}_3$ for each possible assault year. None of them is statistically different from zero before 2011 and between 2012 and 2016, which is actually what we want to find. This statistically insignificant pre-trend test plays a crucial role in validating the parallel trends assumption underlying this analysis, though it does not establish the validity of the parallel trends counterfactual (Kahn-Lang and Lang, 2020). Besides, Figure A7.1 shows that pre-trends are also parallel when restricting the analysis to female victims only, but that the estimated effects for 2011 and 2017-2018 are less pronounced. This check also acts as a placebo test as it shows that the effect I measure is not a random event occurring in other years.

Figure 5.3: Placebo tests for DiD analysis



Note: Estimated coefficients β_3 from the equation $Y_{itr} = \beta_0 + \beta_1 SV_i + \beta_2 Post_t + \beta_3 SV_i \times Year_t + \gamma X_i + \phi_r + \epsilon_i$ with 10% confidence intervals. Sample: victims of sexual violence or of physical violence in the two years preceding the survey.

5.4.2 Using a different control group

In addition, I perform this difference-in-difference estimation with alternative comparison groups. More precisely, I estimate Equation 5.1 using victims of theft or attempted theft (in general and only with violence or threats) as a control group. Contrary to physical violence, there are slightly more female (51.1%) than male victims (48.9%) of theft or attempted theft in the sample. Hence, I do not restrict the sample to female victims only.

The purpose of this analysis is twofold. First, I want to show that victims of physical

violence are still the most relevant comparison group concerning the validity of the parallel trends counterfactual, even without restricting the analysis only to women. Second, finding similar estimates of the impact of the two high-profile cases on sexual violence reporting with these alternative control groups would help support previous results and the common trend assumption.

Table A7.1 presents the results of the new difference-in-difference estimations on the total sample of victims. Although the overall magnitudes differ, the signs of the estimated coefficients on $SV_i \times Post_t$ are the same as with the main estimation. However, these results are complicated to interpret as the parallel trend assumption is less likely to hold with these alternative comparison groups. Indeed, Figure A7.2 shows that this assumption is way more difficult to believe here based on the pre-trends data.

5.4.3 Using a different treatment group

Then, I perform a placebo test using a fake treatment group, which should be a group that was not affected by the two high-profile cases. To do so, I implement the difference-in-difference estimation using victims of non-household physical violence as a control group, and victims of theft or attempted theft with or without violence as a fake treatment group. The pre-trends displayed in Figure A7.3 follow a relatively parallel evolution. Besides, Table A7.2 shows that the estimated coefficients on $Theft_i \times Post_t$ are not statistically significant, meaning that there was no significant increase or decrease in the reporting rate for theft or attempted theft in 2011 or in 2017-2018. This is reassuring, as a placebo test that reveals zero impact gives support to the common trend assumption for the main difference-in-difference strategy.

5.4.4 Using other specifications

I also test the sensitivity of my estimates to the specification. First, I re-estimate Equation 5.1 with Probit models, whose results are displayed in Table A7.3. When comparing these marginal effects with the estimated coefficients of Table 5.2, we see that Probit leads to qualitatively similar results but of larger magnitudes. The main difference is that the decrease in sexual violence reporting in 2011 becomes statistically significant (at the 10% level) for the subsample of women. Besides, Figure A7.4 shows that the predicted probabilities mainly fall within the [0,1] range with LPM, which further validates the use of this specification.

Moreover, in my main specification, I split the sample into two different periods to study the effect of the two affairs as I assume a temporary effect for the DSK case. Indeed, the changes in sexual violence representations triggered by this affair are likely to have been short-lived as this was an isolated case that probably did not last long enough for a drastic change of norms to occur in society. Figure 5.1 also shows that the reporting rate returned to its initial level in 2012, illustrating this temporary effect only for the year 2011 for DSK. On the other hand, the Weinstein case and #MeToo movement had a greater impact on norms and stereotypes about sexual violence. As changes in norms take time to develop, they are likely to have led to changes in trends rather than levels. However, I do not have enough post-MeToo years to study whether

the change in level caused by the Weinstein case is permanent or transitory.

As an alternative, I also implement a specification that includes all the periods and simultaneously estimate the two high-profile cases' effects. This would be the correct estimation if I assumed that the DSK case permanently affected sexual violence reporting, as this should be considered when analysing the Weinstein case's impact. With this specification, Table 5.5 shows that the two estimated coefficients on $SV_i \times Post_{2011}$ are negative but not statistically different from zero when using both the total sample and the subsample of women. In contrast, the estimated coefficients on $SV_i \times Post_{2017}$ are positive and statistically significant, as well as of the same magnitude as on Table 5.2. This confirms the finding of an increase in sexual violence reporting following the Weinstein case, but seems to show that the DSK case did not permanently affect reporting. In addition, this suggests that the effect I estimate for the Weinstein case and #MeToo in my main specification is not so much affected by the impact of the DSK case on the pre-Weinstein period.

Table 5.5: DiD estimation - all time periods

	Dependent variable: Filed a complaint (0/1)	
	All victims	Women only
	(1)	(2)
Sexual violence	-0.196^{***}	-0.184^{***}
	(0.023)	(0.029)
Assault post 2011	0.047^{**}	0.045
	(0.020)	(0.029)
Sexual violence*Assault post 2011	−0.021	−0.028
	(0.030)	(0.039)
Assault post 2017	−0.039	-0.079^{**}
	(0.029)	(0.040)
Sexual violence*Assault post 2017	0.149^{***}	0.160^{***}
	(0.052)	(0.059)
Intercept	0.005	0.178^{*}
	(0.076)	(0.108)
Socio-demographic controls	Yes	Yes
Region fixed effects	Yes	Yes
Period	2007-2018	2007-2018
Observations	4,850	2,907
Adjusted R^2	0.068	0.081

Note: OLS estimates from LPM. Robust standard errors in parentheses *p<0.1; **p<0.05; ***p<0.01

Finally, I perform this difference-in-difference estimation with an alternative treated time period for the effect of the Weinstein case. In my main specification, $Post_t = 1$ for assault years 2017 and 2018. However, only the last three months of 2017 can have been affected by this affair, which started in October. Therefore, I exclude 2017 from the treated period to compare the results, which appear in Table 5.6. The estimated coefficients on $SV_i \times Post_{2018}$ are still positive but of lower magnitude. The effect is statistically significant only in the subsample of women, with a 14.6 percentage points increase in the likelihood of filing a complaint about sexual violence. Yet, this lower effect is expected as this specification excludes three 2017 months that may also

have been affected by the Weinstein case from the treated period. The fact that we still find a large and statistically significant effect for women is thus evidence of an extensive impact of this affair on sexual violence reporting, as this estimated effect is necessarily underestimated.

Table 5.6: DiD estimation - excluding 2017 from treated period for the Weinstein case

	Dependent variable: Filed a complaint (0/1)	
	All victims	Women only
	(1)	(2)
Sexual violence	-0.175^{***}	-0.180^{***}
	(0.023)	(0.027)
Assault post 2018	-0.056	-0.122^{**}
	(0.042)	(0.053)
Sexual violence*Assault post 2018	0.076	0.146^{*}
	(0.067)	(0.081)
Intercept	-0.002	0.161
	(0.087)	(0.120)
Socio-demographic controls	Yes	Yes
Region fixed effects	Yes	Yes
Period	2012-2018	2012-2018
Observations	2,692	1,589
Adjusted R^2	0.074	0.088

Note: OLS estimates from LPM. Robust standard errors in parentheses *p<0.1; **p<0.05; ***p<0.01

5.4.5 Investigating the rewording of the question on sexual victimisation

As mentioned in Section 3, an important threat with the CVS data relates to a rewording of the questions on sexual violence in 2017[27]. Since this change occurred during the same year as the Weinstein case, this could bias the results I find for this affair's impact. To address this point, I exploit the fact that the change was not exactly the same in the two modules of intra- and extra-household sexual violence. Indeed, the question on identifying victims of intra-household sexual violence remained virtually the same compared to the question on extra-household sexual violence. Therefore, I re-estimate Equation 5.1 by splitting the treatment group of victims of sexual violence depending on whether it occurred within the current household or not. Results could be biased if the only statistically significant effect I find is an increase in sexual assault reporting from 2017 in the treatment group of victims of extra-household sexual violence.

Table A7.4 indeed shows that the coefficient on $SV_i \times Post_{2017}$ is statistically significant when using this sub-treatment group, while it is not when using victims of intra-household sexual violence. However, the decrease in reporting in 2011 is not statistically significant either with this

[27]For the module on extra-household sexual violence, the question shifted from "*Apart from the people currently living with you, has anyone ever forced or attempted to force you to have sexual touching or sex against your will?*" before 2017 to "*Apart from the people currently living with you, has anyone ever forced or attempted to force sexual touching or unwanted sex on you using violence, threats, coercion or surprise?*" from 2017 onwards. For the module on intra-household sexual violence, the wording was changed from "*Apart from these episodes of violence, has a person currently living with you ever touched you or had unwanted sex with you, using violence, threats, coercion or surprise?*" before 2017 to "*Apart from these episodes of violence, has a person currently living with you ever touched you in a sexual way or engaged in unwanted sexual intercourse, or attempted to do so, using violence, threats, coercion or surprise?*" from 2017 onwards.

second sub-treatment group, while it is statistically different from zero when using victims of extra-household sexual violence as a treatment group. As this latter effect cannot be attributed to a question rewording, this gives more confidence that the impact on sexual violence reporting found after 2017 is not due to rewording either. In addition, as with the analysis of the reporting determinants, the absence of significance with intra-household sexual violence may be due to the low number of victims in this category (382 intra-household victims vs. 1229 extra-household victims in this sample). Since this difference in numbers questions whether the sample can really be separated for this analysis, I also address this issue of rewording in the next subsection.

5.4.6 Using administrative data

As a final check, I perform a similar difference-in-difference estimation with administrative data. I use data from the police's procedural registration database, publicly available from 1996 to 2021 on *data.gouv.fr*[28]. This database provides aggregated data on the number of crimes and offences recorded per month and per department for different crime categories (with 103 categories, referred to as the "état 4001"). However, there is no information on the victims nor details on the assault. These crimes may have been recorded following a complaint, testimony, flagrant offence, denunciation, etc., but also on the initiative of the security forces. Compared to survey data, administrative data allows monitoring specific or rare offences by recording the entire universe of the number of complaints filed.

For this analysis, I use the data on the number of crimes and offences recorded per month in metropolitan France. The purpose is to compare the difference-in-difference estimation results obtained with survey and administrative data, studying a different outcome than the reporting rate but capturing a similar phenomenon. I restrict the analysis only to certain crime categories comparable to those I used with CVS data. My treatment group consists of the three following categories: rape, sexual harassment and other sexual assaults against adults (i.e. sexual violence). My control group consists of attempted homicides for reasons other than robbery and criminal or correctional intentional injuries that did not result in death (i.e. physical violence)[29]. My dataset is made of one observation per type of crime × month, and I estimate the following equation:

$$Y_{ct} = \beta_0 + \beta_1 SV_c + \beta_2 Post_t + \beta_3 SV_c \times Post_t + \epsilon_c \tag{5.3}$$

Where Y_{ct} is the natural logarithm of the number of reported crimes of type c in month t. A major limit is that I cannot distinguish whether an increase in the number of reported crimes is due to an increase in the reporting propensity or in the incidence of crime. However, without information on the year when the crime was committed (in addition to when it was reported), I cannot correct the number of reported crimes in administrative data using the annual number of committed crimes of the same category in the CVS data, which would have allowed me to recover

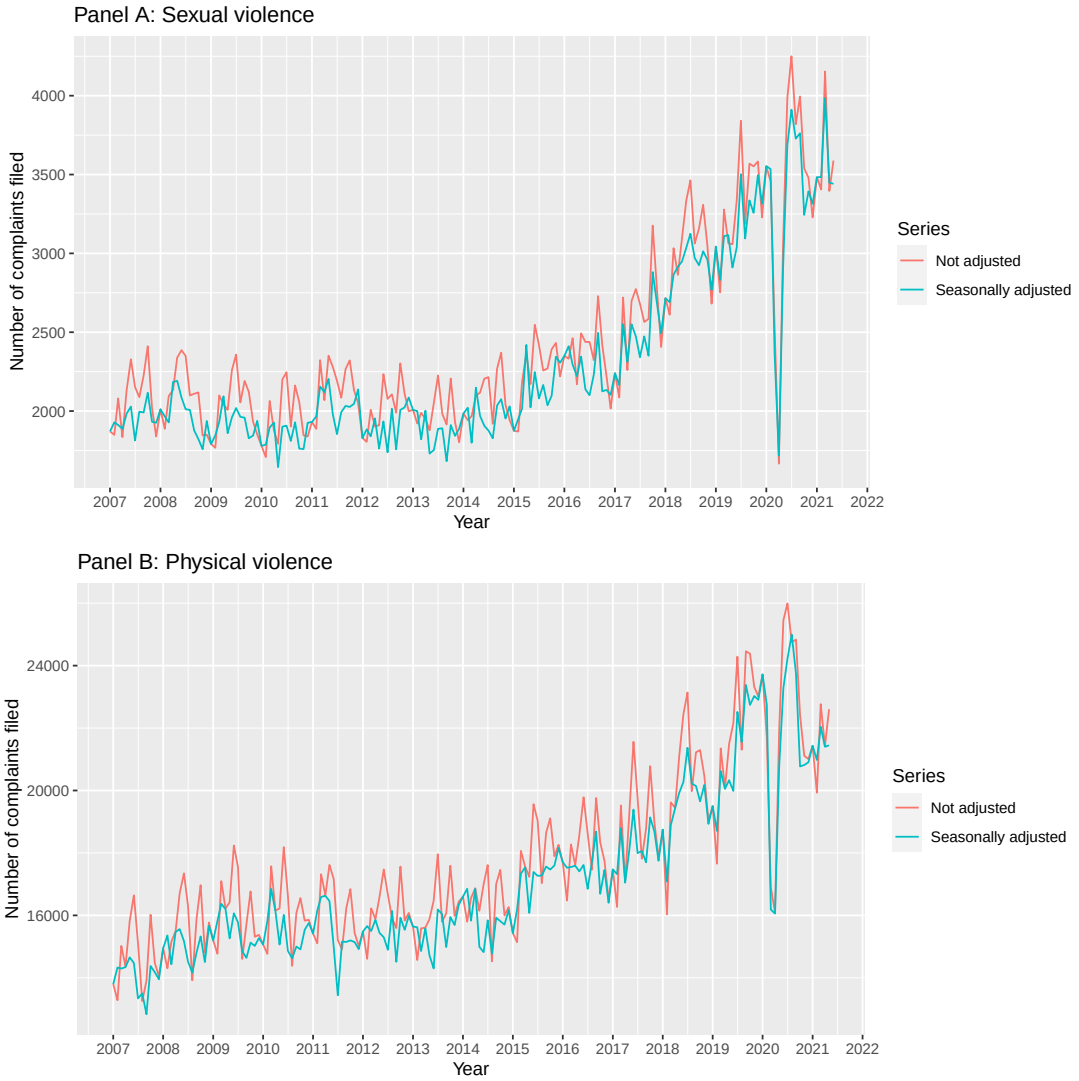
[28] *Chiffres départementaux mensuels relatifs aux crimes et délits enregistrés par les services de police et de gendarmerie depuis janvier 1996.*

[29] These offences are also reported in the same unit of account, i.e. in terms of the number of victims.

a reporting rate. Therefore, since it is not possible to make this correction, the analysis assumes that the effective number of crimes has remained constant.

SV_c is an indicator for the treatment group, i.e. it equals 1 for sexual offences and 0 for physical offences. $Post_t$ is an indicator of treated time periods, i.e. it equals 1 if the crime or offence was recorded in the month of the case or subsequent months. To stay as close as possible to the previous estimation, I also restrict the time frame to 2007-2011 for the DSK case and to 2012-2018 for Weinstein and #MeToo. But this time, I take advantage of the fact that administrative data gives monthly complaint figures to define the two treated periods more precisely, starting from the month of each affair. Thus, $Post_t = 1$ from May 2011 to December 2011 for the DSK case and $Post_t = 1$ from October 2017 to December 2018 for the Weinstein case. The main coefficient of interest is β_3. As before, the identifying assumption is that the outcome in the treatment and control groups would have followed the same time trend in the absence of the two high-profile cases.

Figure 5.4: Evolution of the number of reports in administrative data



Source: Base des crimes et délits enregistrés par la police et la gendarmerie, SSMSI, 2007-2021

Even if I do not observe a complaint rate, this estimation can provide interesting insights into the influence of the DSK and Weinstein cases on the victims' reporting behaviour. Indeed, I assume there is a stock of perpetrated sexual violence for which victims did not file a complaint. As previous results suggested that the Weinstein case and #MeToo increased the likelihood of reporting sexual violence, we can therefore expect a jump in the number of complaints filed in administrative data from October 2017. This spike could be due to both increased complaints about past sexual crimes or victims speaking out more freely about sexual violence. However, given that the DSK case seems to have reduced this probability, it is likely to be more difficult to detect an effect on administrative data for this affair without observing complaint rates.

Figure 5.4 presents the evolution of the number of reports made in the two groups since 2007. The red line plots the raw data, while the one in blue is seasonally adjusted (i.e. it removes changes due to seasonal and calendar influences to explore the trends). There is a lot of fluctuation, even in the deseasonalised series. On Panel A, we see that the number of complaints filed for sexual violence remained relatively stable until 2015, but with a jump in May 2011 on the seasonally adjusted series – start of the DSK case. The slight increase from 2015 (also observed for physical violence) is complicated to interpret as it corresponds to the introduction of a new system for recording offences in the national police in April 2015, the series breaks of which have not been corrected (Interstats, 2015).

The trend then remained more or less constant until 2017. We notice a jump in sexual violence reporting in October 2017, corresponding to the Weinstein case's beginning. From this date onwards, there is a clear break in the previously stable trend, with an increase in complaints filed over time that confirms the hypothesis discussed in Section 5.4.4 of permanent changes in sexual violence reporting following this affair. Although less pronounced, we also observe a gradual increase in reports for physical violence throughout the time on Panel B[30].

Results of the difference-in-difference estimation appear in Table 5.7. The estimated coefficient on $SV_c \times Post_{2011}$ is positive but not statistically different from zero. In contrast, the coefficient on $SV_c \times Post_{2017}$ is statistically significant at the 1% level, with an average increase in the number of sexual assault reports of 15.5% from October 2017 to December 2018 compared to before[31]. This estimated effect for the Weinstein case is consistent with the results previously found with survey data, even though the outcome of interest is not exactly the same. This confirms the finding of a sizeable impact of the Weinstein affair and #MeToo on sexual violence reporting. As a comparison, using a similar methodology, Levy and Mattsson (2020) find an 8% increase in sexual crimes reporting during the first six months of #MeToo in the United States. However, the assumption that the number of effective crimes stayed constant is a strong one. In reality, I cannot exclude that this effect could also be driven by an increase in the incidence of sexual crimes rather than the reporting propensity, with a backlash effect of #MeToo.

[30] These graphs also clearly depict the effect of the March 2020 containment in France during the first wave of COVID-19, with a massive drop in the number of reports.

[31] This finding is also consistent with reports from the French Ministry of the Interior indicating that the number of registered complaints about sexual violence in police data increased by +19% in 2018 and +12% in 2019 (Baux et al., 2020).

Despite this limitation, this result importantly suggests that the increase in the reporting rate I found with survey data following the Weinstein case was not due to the rephrasing of the questions about sexual violence in the CVS survey, but rather to an actual increase in reporting.

Table 5.7: DiD estimation using administrative data

	Dependent variable: Filed a complaint (0/1)	
	(1)	(2)
Sexual violence	-2.038^{***}	-2.060^{***}
	(0.016)	(0.017)
Assault post 2011	0.026	
	(0.031)	
Sexual violence*Assault post 2011	0.056	
	(0.044)	
Assault post 2017		0.154^{***}
		(0.028)
Sexual violence*Assault post 2017		0.155^{***}
		(0.040)
Intercept	9.656^{***}	9.746^{***}
	(0.011)	(0.012)
Period	2007-2011	2012-2018
Observations	120	168
Adjusted R^2	0.994	0.991
F Statistic	$6{,}149.773^{***}$ (df = 3; 116)	$5{,}982.261^{***}$ (df = 3; 164)

Note: Standard errors in parentheses *p<0.1; **p<0.05; ***p<0.01

6 Discussion and potential mechanisms

Several limits arise with this difference-in-difference strategy. The previous section reinforced confidence in the validity of the common trend assumption and the relevance of using this treatment and control group. However, a crucial limit concerns the rewording of the questions on sexual violence from 2017, which could be a confounder for the effect of the Weinstein case and #MeToo. Still, the results found with administrative data strengthen the credibility of the results obtained with survey data.

In addition, the DSK and Weinstein cases may have directly affected the incidence of sexual crimes, as well as their severity and nature. This may also have influenced the likelihood of both reporting these crimes in the CVS survey and reporting it to the police. Therefore, for this analysis to be valid, I must assume that these affairs did not affect the likelihood of reporting sexual violence in CVS. This point relates to the limitation already discussed in Section 4.3 that under-reporting sexual violence in the survey could also be correlated with factors that influence complaint filing. However, as mentioned earlier, the CVS methodology precisely aims at minimising this bias.

A third important threat is that causal inference may be problematic if the media coverage of these two high-profile cases is in fact the result of changing social norms regarding sexual violence and its representations. However, as these are prominent cases involving well-known

personalities, such bias seems less likely than with coverage of miscellaneous sexual assaults. Indeed, contrary to more "ordinary" violence, these two affairs would have been covered in all cases, making the exogeneity assumption more plausible for this type of event. Nevertheless, the way the media cover this type of affair may have changed over time due to the evolving norms and representations about sexual violence. Thus, there was probably an evolution in how sexual assaults are dealt with between the DSK and Weinstein cases.

Furthermore, my initial hypothesis was that by challenging social norms and representations about sexual violence, these two high-profile cases must have influenced the victims' reporting behaviour. Overall, my findings suggest that it is the case, but mainly for the Weinstein affair. Besides, the estimated effect of this high-profile case is a conservative effect due to the inability to start the treatment period in October 2017 in survey data and a single survey year to measure reporting in 2018, probably resulting in an underestimation. The heterogeneity analysis also showed that these affairs did not affect the same individuals, potentially due to the different representations of victims and the reporting process that they depicted. Some of these results confirmed the role of the "social stigma" channel, though it is certainly not the only mechanism at play. Indeed, as mentioned in Section 2.2, coverage of prominent cases may affect reporting in different ways and through different mechanisms.

In this last section, I try to understand better how these two high-profile cases (and especially the Weinstein case) affected the reporting of sexual violence by focusing on three potential mechanisms: (i) the opinions and representations about sexual violence; (ii) the perceptions of the police and justice effectiveness and (iii) the processing of victims by police officers. I show that this effect does not seem to come from the two latter mechanisms but mainly from the former, supporting my initial hypothesis.

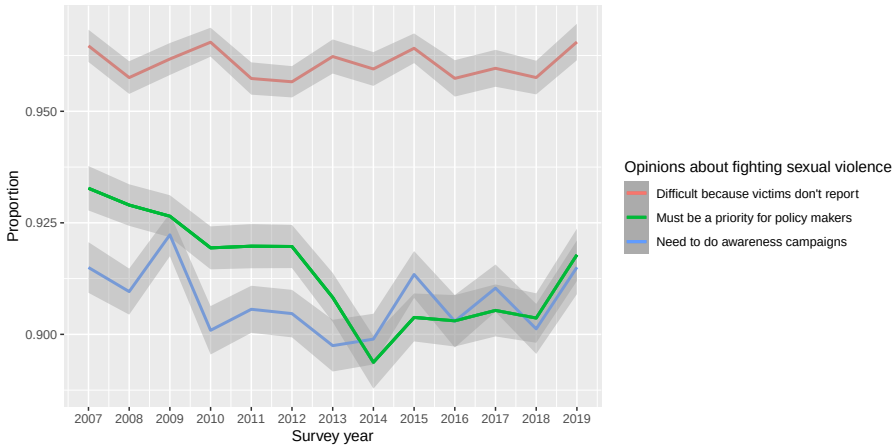
6.1 Changes in representations regarding sexual violence

Firstly, I expect that these two-high profile cases affected the general representations of sexual violence in society, and as a result, the stigma associated with being a victim of sexual misconduct. Although the CVS survey does not directly measure beliefs and representations about sexual violence, some questions can serve as a proxy for this. Indeed, the first three questions of the module on sensitive violence ask respondents their opinions on the fight against sexual crimes, and in particular: (i) whether it should be a priority for the public authorities, (ii) whether it is particularly difficult because victims are reluctant to report it, and (iii) whether awareness campaigns should be conducted to raise awareness of these problems.

Figure 6.1 plots the evolution of the proportion of individuals who agree with these different statements over the years. Almost all respondents (96% on average) agree that fighting sexual violence is difficult as victims are reluctant to report it, with no major change over the period. Society is therefore well aware of this under-reporting problem. The evolution of the share of respondents who think that fighting sexual violence must be a priority for policymakers provides interesting insights to capture a potential effect on sexual violence representations. There was a

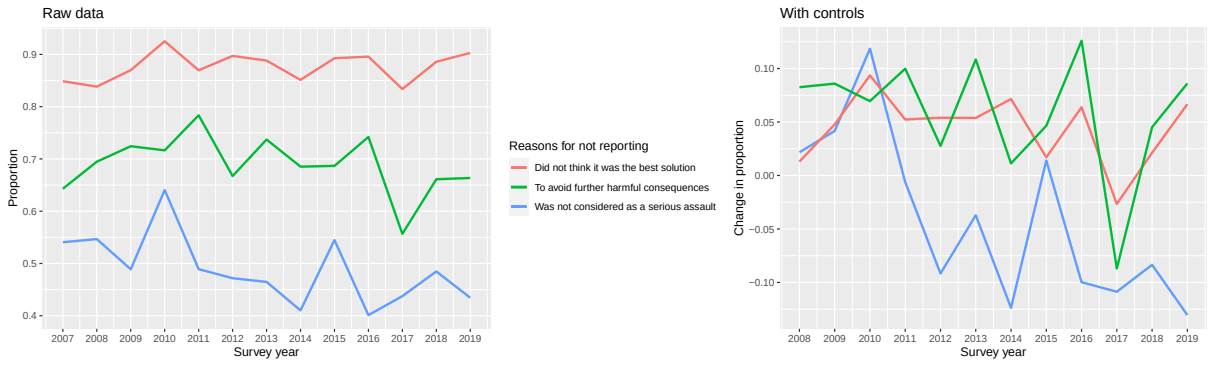
large drop in this proportion for respondents surveyed after 2012, which then rose significantly again after 2018. Although this development may be due to many different reasons, it is interesting to note that it corresponds exactly to the timing of the DSK and Weinstein cases. It also goes in the expected direction, consistent with the negative image of sexual assault victims conveyed by the DSK case and the more positive one displayed by the Weinstein case.

Figure 6.1: Evolution of the general opinions regarding the fight against sexual violence



Population: People aged between 18-75 and living in metropolitan France. Weighted proportions. 10% confidence interval in dark. *Source*: CVS surveys, 2007-2019.

Figure 6.2: Evolution of the three main categories of reasons for not reporting sexual violence



Population: People aged between 18-75 and living in metropolitan France, who suffered sexual violence in the two years preceding the survey and did not report the assault to the police. Weighted proportions.
Source: CVS surveys, 2007-2019.

To investigate whether these changes in general opinions and representations about sexual violence also resulted in less social stigma for the victims, I come back to the analysis of why victims did not report. Figure 6.2 shows the evolution of the proportion of victims who did not report sexual violence for the three main categories of reasons. Panel A plots the raw data, while Panel B plots the fitted values from regressions of Equation 4.1. No clear pattern emerges from these graphs that would depict a sudden drop in social stigma over time, probably due to too few observations resulting in large fluctuations. Nevertheless, we can see a more or less decreasing

trend for the share of victims who did not consider the assault as serious, which is in line with a potential change of representations and mentalities about sexual violence over time, and especially in the last years following #MeToo.

6.2 Changes in perceptions of police and justice effectiveness

Furthermore, the victim's decision to report is likely to be affected by the criminal treatment of sexual violence, which may constitute a significant obstacle when one knows that only one complaint of sexual assault in ten leads to the conviction of the aggressor in France (Bousquet, 2016). Hence, by potentially influencing the judicial response, the effect of these two high-profile cases on sexual violence reporting could also be driven by a change in the perceptions of the police and justice effectiveness (Bhatnagar et al., 2019).

To investigate this mechanism, I use the questions on the satisfaction with the police and justice action from the individual face-to-face questionnaire of the CVS survey. I perform a similar difference-in-difference estimation as before with victims of sexual violence and victims of physical violence, but this time using the share of victims satisfied with the legal system as the outcome of interest. More precisely, I use the questions on (i) "satisfaction with the general performance of the police in French society today", (ii) "effectiveness of the police in combating crime in your neighbourhood" and (iii) "satisfaction with the overall performance of the justice system and the courts in dealing with crime". I re-estimate Equation 5.1 with these new outcomes to study a potential effect of the Weinstein case on the perceptions of the police and justice effectiveness[32].

Table 6.1: DiD estimation for the satisfaction with the police and justice

	Dependent variable:		
	Satisfied with police in general (0/1)	Satisfied with police in local neighbourhood (0/1)	Satisfied with justice in general (0/1)
	(1)	(2)	(3)
Sexual violence	0.023	0.085***	0.069***
	(0.031)	(0.032)	(0.027)
Assault post 2017	−0.027	0.061*	−0.011
	(0.035)	(0.036)	(0.027)
Sexual violence*Assault post 2017	0.086	−0.044	−0.014
	(0.063)	(0.062)	(0.055)
Intercept	0.595***	0.514***	0.105
	(0.117)	(0.113)	(0.079)
Socio-demographic controls	Yes	Yes	Yes
Region fixed effects	Yes	Yes	Yes
Period	2012-2018	2012-2018	2012-2018
Observations	2,716	2,714	2,716
Adjusted R^2	0.024	0.024	0.032

Note: OLS estimates from LPM. Robust standard errors in parentheses *p<0.1; **p<0.05; ***p<0.01

[32] As the questions regarding the general actions of the police and justice in society were only introduced in 2010 and 2013 respectively, I can only perform this analysis for the Weinstein case.

Table 6.1 shows that the estimated coefficients on $SV_i \times Post_{2017}$ are negative for satisfaction with the justice system and the local police, and positive for satisfaction with police action in general, but not statistically different from zero for all specifications. This insignificance suggests that the increase in sexual violence reporting from 2017 is not driven by a change in the police and justice effectiveness perceptions. However, there are many possible explanations for this result. For example, perhaps the police are better trained to deal with sexual assault cases or allegations after #MeToo, but victims are also better able to identify abnormal treatment by the police. In addition, this absence of effect could also be due to the lower relevance of comparing these two groups for this difference-in-difference strategy, as shown with the pre-trends regarding satisfaction with justice in Figure A8.1.

6.3 Changes in the processing of victims by police officers

Related to the previous point, high-profile sexual assault cases could also affect reporting by improving how victims are received and listened to by police officers. As a reminder, only two-thirds of victims who went to the police station then decided to lodge a formal complaint (Table A1.4). Therefore, the behaviour of the police is also an important channel to consider. To study this mechanism, I use the three questions from the CVS module on sensitive violence that asked respondents who went to the police station whether they were satisfied by: (i) the time and attention they were given, (ii) the advice they were given and (iii) the conditions of confidentiality of their exchange. Figure 6.3 shows that no clear trend emerges over time regarding the victims' satisfaction with the police exchange.

Figure 6.3: Evolution of satisfaction with the exchange with the police

Population: People aged between 18-75 and living in metropolitan France, who suffered sexual violence in the two years preceding the survey and went to the police to report. Weighted proportions.
Source: CVS surveys, 2007-2019.

Table 6.2 presents the evolution of the share of victims satisfied with their experience to the police by dividing the sample into three main periods and by whether the victims filed a complaint. Interestingly, the figures on the total sample of victims who went to the police suggest

Table 6.2: Evolution of the share of victims satisfied with their experience to the police by reporting behaviour

	Total			**Filed a complaint**			**Did not file a complaint**		
Satisfied with	2007-2011	2012-2015	2016-2019	2007-2011	2012-2015	2016-2019	2007-2011	2012-2015	2016-2019
	$N = 135$	$N = 83$	$N = 88$	$N = 83$	$N = 47$	$N = 67$	$N = 41$	$N = 31$	$N = 19$
The time and attention	0.769	0.734	0.645	0.742	0.793	0.604	0.790	0.654	0.764
given	(0.047)	(0.057)	(0.061)	(0.062)	(0.058)	(0.069)	(0.068)	(0.088)	(0.096)
The advice given	0.706	0.600	0.591	0.696	0.647	0.561	0.691	0.540	0.721
	(0.048)	(0.066)	(0.068)	(0.062)	(0.085)	(0.077)	(0.082)	(0.089)	(0.096)
The exchange	0.747	0.764	0.635	0.770	0.817	0.584	0.723	0.697	0.764
confidentiality	(0.045)	(0.058)	(0.066)	(0.048)	(0.057)	(0.070)	(0.073)	(0.099)	(0.096)

Note: Victims of at least one sexual assault committed during the two years preceding the survey, living in metropolitan France and aged between 18-75, and who went to the police station to report their assault. Weighted proportions and standard errors in parentheses. *Source:* CVS surveys, 2007-2019.

that there was a decrease in satisfaction with the police exchange over time, though these changes are on average not statistically different from zero. However, a statistically significant decrease is found for victims who filed a complaint, with a reduction in satisfaction with the time and attention given from 80% to 60% between the second and third periods, and in satisfaction with the exchange confidentiality from 82% to 58%. Surprisingly, satisfaction increased for victims who did not eventually file a complaint, even though this development must be contrasted with the small sample size. Therefore, this implies that the increase in sexual violence reporting in 2017 and 2018 is not due to a better exchange with the police either.

7 Conclusion

This paper demonstrates that several factors can affect the decision to report sexual violence to the police. Using data from the French victimisation survey and binary regression models, I first find that assault circumstances play a more important role than victim's socio-demographics in this decision. The results illustrate the deterring effect of social stigmatisation associated with reporting sexual violence. Conversely, they show the incentive effect of assault characteristics that would fit with the "classic rape" stereotype and support the victim's credibility in the decision to file a complaint.

Employing a difference-in-difference strategy, I also study the effect of two high-profile sexual assaults cases that pictured different representations of victims and the reporting process on the complaint behaviour. Overall, the results suggest that the Weinstein case led to a significant increase in the likelihood of reporting sexual violence. In contrast, the DSK case seems to have decreased this propensity on average. The heterogeneity analysis provides additional evidence that the Weinstein case and subsequent #MeToo movement did not affect all victims in the same way. In particular, the results do not fully support the hypothesis of a role model effect for this high-profile case.

Finally, in an attempt to better understand the mechanisms underlying these impacts, I show that the increase in sexual violence reporting after the Weinstein affair and #MeToo does not seem to stem from an improvement in the perceptions of the police and justice systems

effectiveness, or a shift in the treatment of victims by the police. Instead, changes in norms and representations about sexual violence appear to be the main explanation.

Nonetheless, I also point out a few limitations regarding the empirical designs and results presented in this paper, which may provide an opportunity for future improvement. The first is related to the data and numerous statistical power issues that make it complicated to precisely estimate the effects due to the small sample sizes. Moreover, although the survey adopts a specific methodology to avoid under-reporting sexual victimisation, we can still question the sample representativeness. Results could be biased if misreporting in the data is also affected by the main factors studied regarding complaint filing. Furthermore, I reinforce confidence in the validity of my second empirical strategy with a set of robustness checks to support the validity of the common trend assumption and show that my results are unlikely to be driven by a rewording of the questions on sexual violence. However, the results are less robust when restricting the sample to specific subcategories of victims for the effect of the DSK case. Besides, the estimated effect of the Weinstein case is a conservative one. Finally, this study is specific to the French context, and the effects I found might differ depending on country-specific characteristics.

These limitations should nevertheless be contrasted with the fact that, despite its importance, sexual violence remains an under-explored topic, especially in the economic literature. Further research should therefore continue exploring this topic in other contexts and as new data become available to develop better identification strategies.

Summing up, sexual violence and its under-reporting is a major public policy issue. Encouraging victims to report is a crucial first step to understand better the extent of the problem and design efficient policies to prevent and contain such violence. This thesis shows the key role of norms and representations surrounding sexual violence in the victim's reporting decision. In terms of policy implications, the findings suggest that playing on those norms and representations, for instance through awareness-raising campaigns and educational interventions, could be an effective way to create social change and improve the reporting of sexual violence. In addition, the results also suggest that, depending on the representations of sexual violence it depicts, policymakers should seek to increase or reduce coverage of sexual assault allegations or cases within the media industry in order to act on those harmful norms and stereotypes.

References

Alesina, A., Giuliano, P., and Nunn, N. (2013). "On the origins of gender roles: Women and the plough". *The quarterly journal of economics*, 128(2):469–530.

Allen, W. D. (2007). "The reporting and underreporting of rape". *Southern Economic Journal*, pages 623–641.

Amir, M. (1971). *Patterns in forcible rape*. Chicago: University of Chicago Press.

Bachman, R. (1998). "The factors related to rape reporting behavior and arrest: New evidence from the national crime victimization survey". *Criminal justice and behavior*, 25(1):8–29.

Bachman, R., Paternoster, R., and Ward, S. (1992). "The rationality of sexual offending: Testing a deterrence/rational choice conception of sexual assault". *Law and Society Review*, pages 343–372.

Banerjee, A., Ferrara, E. L., and Orozco, V. (2019). "Entertainment, education, and attitudes toward domestic violence". In *AEA Papers and Proceedings*, volume 109, pages 133–37.

Baux, D., Bernardi, V., Delobel, L., Estival, A., Filatriau, O., Hama, S., Pramil, J., Quélennec, Y., Razafindranovona, T., and Tir, D. (2020). Insécurité et délinquance en 2019: Bilan statistique. *SSMSI Ministère de l'Intérieur*.

Becker, G. S. (1968). "Crime and punishment: An economic approach". *Journal of Political Economy*, 76(2):169–217.

Bénabou, R., Falk, A., and Tirole, J. (2019). "Narratives, imperatives, and moral persuasion". *NBER working paper*.

Bernardi, V., Guedj, H., Moreau, A., Razafindranovona, T., and Zilloniz, S. (2019). Rapport d'enquête «Cadre de vie et sécurité» 2019: Victimation, délinquance et sentiment d'insécurité.

Bertrand, M., Kamenica, E., and Pan, J. (2015). "Gender identity and relative income within households". *The Quarterly Journal of Economics*, 130(2):571–614.

Bhatnagar, A., Mathur, A., Munasib, A., and Roy, D. (2019). "Sparking the# MeToo revolution in India: The "Nirbhaya" case in Delhi". Working paper. *AEI Paper & Studies*.

Bousquet, D. (2016). Rapport final d'évaluation du 4e plan interministériel de prévention et de lutte contre les violences faites aux femmes. *Haut conseil à l'égalité entre les femmes et les hommes*.

Bowles, R., Reyes, M. G., and Garoupa, N. (2009). "Crime reporting decisions and the costs of crime". *European journal on criminal policy and research*, 15(4):365.

Card, D. and Dahl, G. B. (2011). "Family violence and football: The effect of unexpected emotional cues on violent behavior". *The quarterly journal of economics*, 126(1):103–143.

Chalfin, A. and McCrary, J. (2017). "Criminal deterrence: A review of the literature". *Journal of Economic Literature*, 55(1):5–48.

Dahl, G. and DellaVigna, S. (2009). "Does movie violence increase violent crime?". *The Quarterly Journal of Economics*, 124(2):677–734.

DellaVigna, S. and La Ferrara, E. (2015). "Economic and social impacts of the media". In *Handbook of media economics*, volume 1, pages 723–768. Elsevier.

Du Mont, J., Miller, K.-L., and Myhr, T. L. (2003). "The role of "real rape" and "real victim" stereotypes in the police reporting practices of sexually assaulted women". *Violence Against Women*, 9(4):466–486.

Ehrlich, I. (1973). "Participation in illegitimate activities: A theoretical and empirical investigation". *Journal of political Economy*, 81(3):521–565.

Fisher, B. S., Daigle, L. E., Cullen, F. T., and Turner, M. G. (2003). "Reporting sexual victimization to the police and others: Results from a national-level study of college women". *Criminal justice and behavior*, 30(1):6–38.

Garoupa, N. (2003). "Optimal law enforcement when victims are rational players". In *Conflict and Governance*, pages 123–134. Springer.

Gartner, R. and Macmillan, R. (1995). "The effect of victim-offender relationship on reporting crimes of violence against women". *Canadian journal of criminology*, 37(3):393–429.

Gingerich, D. W. and Oliveros, V. (2018). "Police violence and the underreporting of crime". *Economics & Politics*, 30(1):78–105.

Golman, R., Hagmann, D., and Loewenstein, G. (2017). "Information avoidance". *Journal of Economic Literature*, 55(1):96–135.

Goudriaan, H., Wittebrood, K., and Nieuwbeerta, P. (2006). "Neighbourhood characteristics and reporting crime: Effects of social cohesion, confidence in police effectiveness and socio-economic disadvantage". *British journal of criminology*, 46(4):719–742.

Green, D. P., Wilke, A. M., and Cooper, J. (2020). "Countering violence against women by encouraging disclosure: A mass media experiment in rural Uganda". *Comparative Political Studies*, 53(14):2283–2320.

Hamel, C., Debauche, A., Brown, E., Lebugle, A., Lejbowicz, T., Mazuy, M., Charruault, A., Cromer, S., and Dupuis, J. (2016). « Viols et agressions sexuelles en France: premiers résultats de l'enquête Virage ». *Population & Sociétés*, (10):1–4.

Interstats (2015). Interstats méthode n°4 : Statistiques sur les crimes et délits enregistrés par la police et la gendarmerie : correction des ruptures techniques liées aux modes de collecte. Technical report, SSMSI Ministère de l'Intérieur.

Iyer, L., Mani, A., Mishra, P., and Topalova, P. (2012). "The power of political voice: Women's political representation and crime in India". *American Economic Journal: Applied Economics*, 4(4):165–93.

Jensen, R. and Oster, E. (2009). "The power of TV: Cable television and women's status in India". *The Quarterly Journal of Economics*, 124(3):1057–1094.

Jina, R. and Thomas, L. S. (2013). "Health consequences of sexual violence against women". *Best Practice & Research Clinical Obstetrics & Gynaecology*, 27(1):15–26.

Jolliffe, I. T. and Cadima, J. (2016). "Principal component analysis: a review and recent developments". *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, 374(2065):20150202.

Kahn-Lang, A. and Lang, K. (2020). "The promise and pitfalls of differences-in-differences: Reflections on 16 and pregnant and other applications". *Journal of Business & Economic Statistics*, 38(3):613–620.

Kearney, M. S. and Levine, P. B. (2015). "Media influences on social outcomes: The impact of MTV's 16 and pregnant on teen childbearing". *American Economic Review*, 105(12):3597–3632.

Lee, F. X. and Suen, W. (2020). "Credibility of crime allegations". *American Economic Journal: Microeconomics*, 12(1):220–59.

Levitt, S. D. and Miles, T. J. (2006). "Economic contributions to the understanding of crime". *Annu. Rev. Law Soc. Sci.*, 2:147–164.

Levy, R. and Mattsson, M. (2020). "The effects of social movements: Evidence from #MeToo". Working paper. *Available at SSRN 3496903*.

Lim, C. S. (2015). "Media influence on courts: Evidence from civil case adjudication". *American Law and Economics Review*, 17(1):87–126.

Lim, C. S., Snyder Jr, J. M., and Strömberg, D. (2015). "The judge, the politician, and the press: newspaper coverage and criminal sentencing across electoral systems". *American Economic Journal: Applied Economics*, 7(4):103–35.

Lizotte, A. J. (1985). "The uniqueness of rape: reporting assaultive violence to the police". *Crime & Delinquency*, 31(2):169–190.

Macdonald, J. M. (1971). *Rape offenders and their victims*. Charles C. Thomas Publisher.

MacDonald, Z. (2002). "Official crime statistics: their use and interpretation". *The Economic Journal*, 112(477):F85–F106.

McDougal, L., Krumholz, S., Bhan, N., Bharadwaj, P., and Raj, A. (2018). "Releasing the tide: how has a shock to the acceptability of gender-based sexual violence affected rape reporting to police in India?". *Journal of interpersonal violence*, pages 1–23.

Miller, A. R. and Segal, C. (2019). "Do female officers improve law enforcement quality? Effects on crime reporting and domestic violence". *The Review of Economic Studies*, 86(5):2220–2247.

Naess, O.-A. E. (2020). "Under-reporting of crime". Working paper.

Paluck, E. L. (2009). "Reducing intergroup prejudice and conflict using the media: a field experiment in Rwanda". *Journal of personality and social psychology*, 96(3):574.

Philippe, A. and Ouss, A. (2018). "“No hatred or malice, fear or affection”: Media and sentencing". *Journal of Political Economy*, 126(5):2134–2178.

Pino, N. W. and Meier, R. F. (1999). "Gender differences in rape reporting". *Sex roles*, 40(11):979–990.

Rees, D. I. and Sabia, J. J. (2013). "Forced intercourse, mental health, and human capital". *Southern Economic Journal*, 80(2):324–344.

Robinson, P. H. and Darley, J. M. (2002). "The role of deterrence in the formulation of criminal law rules: At its worst when doing its best". *Georgetown Law Journal*, 91:949–1002.

Sabia, J. J., Dills, A. K., and DeSimone, J. (2013). "Sexual violence against women and labor market outcomes". *American Economic Review*, 103(3):274–78.

Sable, M. R., Danis, F., Mauzy, D. L., and Gallagher, S. K. (2006). "Barriers to reporting sexual assault for women and men: Perspectives of college students". *Journal of American College Health*, 55(3):157–162.

Soares, R. R. (2004). "Crime reporting as a measure of institutional development". *Economic Development and cultural change*, 52(4):851–871.

Stack, S. (2002). "Media coverage as a risk factor in suicide". *Injury Prevention*, 8(suppl 4):iv30–iv32.

Tankard, M. E. and Paluck, E. L. (2016). "Norm perception as a vehicle for social change". *Social Issues and Policy Review*, 10(1):181–211.

UN Women (2016). The economic costs of violence against women. Technical report, UN Women.

Vanier, C. and Langlade, A. (2018). « Comprendre le dépôt de plainte des victimes de viol: facteurs individuels et circonstanciels ». *Déviance et Société*, 42(3):501–533.

Williams, L. S. (1984). "The classic rape: when do victims report?". *Social Problems*, 31(4):459–467.

Appendix

A1 Data description

Table A1.1: Sample characteristics (extra-household sexual violence): socio-demographic and assault characteristics

Extra-household	N	Mean	Sd	Min	Max
Socio-demographics					
Gender = Female	1476	0.814	0.389	0	1
Age	1476	39.436	14.934	18	75
Born in France	1476	0.821	0.383	0	1
In a couple	1476	0.350	0.477	0	1
Education					
High education	1474	0.266	0.442	0	1
Medium education	1474	0.442	0.497	0	1
Low education	1474	0.292	0.455	0	1
Employment status					
Employed	1476	0.474	0.499	0	1
Unemployed	1476	0.166	0.372	0	1
Inactive	1476	0.360	0.480	0	1
Socioeconomic category					
Upper class	1475	0.085	0.279	0	1
Middle class	1475	0.179	0.383	0	1
Lower class	1475	0.498	0.500	0	1
Other category	1475	0.239	0.426	0	1
Household income					
1st-5th deciles	1476	0.692	0.462	0	1
6th-10th deciles	1476	0.308	0.462	0	1
Number of household members	1476	2.148	1.310	0	9
Living in a priority district	1476	0.243	0.429	0	1
Size of the urban unit					
Country side	1475	0.115	0.319	0	1
$<$ 20,000 inhabitants	1475	0.119	0.323	0	1
20,000-100,000 inhabitants	1475	0.166	0.372	0	1
$>$ 100,000 inhabitants	1475	0.438	0.496	0	1
Paris agglomeration	1475	0.163	0.369	0	1
Assault variables					
Rape	1374	0.301	0.459	0	1
Attempted rape	1374	0.186	0.389	0	1
Other sexual assault	1374	0.513	0.500	0	1
Occurred twice or more	1394	0.419	0.494	0	1
Several offenders	1338	0.084	0.277	0	1
Male offender	1332	0.882	0.323	0	1
Occurred in a public space	1350	0.124	0.329	0	1
Known offender	1345	0.729	0.444	0	1
Victim's ex-partner	1336	0.189	0.391	0	1
Offender drugged or drunk	1321	0.397	0.490	0	1
Physical violence or weapon threat	1384	0.387	0.487	0	1
Physical injury	1402	0.357	0.479	0	1

Reading: 81% of extra-household sexual violence victims are women and the age of the average self-reported victim in this subsample is 39 years and 4 months old.

Table A1.2: Sample characteristics (intra-household sexual violence): socio-demographic and assault characteristics

Intra-household	N	Mean	Sd	Min	Max
Socio-demographics					
Women	471	0.866	0.341	0	1
Age	471	43.244	13.379	18	75
Born in France	471	0.796	0.403	0	1
In a couple	471	0.660	0.474	0	1
Education					
High education	471	0.225	0.418	0	1
Medium education	471	0.435	0.496	0	1
Low education	471	0.340	0.474	0	1
Employment status					
Employed	471	0.516	0.500	0	1
Unemployed	471	0.130	0.336	0	1
Inactive	471	0.355	0.479	0	1
Socioeconomic category					
Upper class	470	0.106	0.309	0	1
Middle class	470	0.181	0.385	0	1
Lower class	470	0.528	0.500	0	1
Other category	470	0.185	0.389	0	1
Household income					
1st-5th deciles	471	0.671	0.470	0	1
6th-10th deciles	471	0.329	0.470	0	1
Number of household members	471	3.068	1.196	1	9
Living in a priority district	471	0.236	0.425	0	1
Size of the urban unit					
Country side	470	0.187	0.391	0	1
$<$ 20,000 inhabitants	470	0.128	0.334	0	1
20,000-100,000 inhab.	470	0.164	0.371	0	1
$>$ 100,000 inhabitants	470	0.385	0.487	0	1
Paris agglomeration	470	0.136	0.343	0	1
Assault variables					
Rape	419	0.482	0.500	0	1
Attempted rape	416	0.481	0.500	0	1
Other sexual assault	415	0.371	0.484	0	1
Occurred twice or more	434	0.659	0.475	0	1
Several offenders	418	0.074	0.262	0	1
Male offender	187	0.898	0.303	0	1
Victim's partner	394	0.533	0.500	0	1
Offender drugged or drunk	355	0.366	0.482	0	1
Physical violence or weapon threat	349	0.567	0.496	0	1
Physical injury	432	0.456	0.499	0	1

Reading: 87% of intra-household sexual violence victims are women and the age of the average self-reported victim in this subsample is 43 years and 2 months old.

Figure A1.1: Construction of socio-demographic variables

Category	Value	Type	CVS variable(s) used	Definition
SOCIO-DEMOGRAPHIC CHARACTERISTICS				
Gender	Female	Dummy	sexe = Féminin	Sexe de l'individu
	Male	Dummy	sexe = Masculin	
Age	Age	Numeric	age	Age à la date de l'enquête
	Aged 18-29	Dummy	age>=18 & age<30	
	Aged 30-54	Dummy	age>=30 & age<55	
	Aged 55-76	Dummy	age>=55 & age<=76	
Marital status	In a couple	Dummy	couple = Oui, avec une personne qui vit dans le logement ; Oui, avec une personne qui ne vit pas dans le logement	Vie en couple et non état matrimonial légal
	Single	Dummy	couple = Non	
Place of birth	Born in France	Dummy	Inais = En France	Indicateur du lieu de naissance
	Born abroad	Dummy	Inais = A l'étranger	
Educational level	High education	Dummy	dip12 = Diplôme des professions sociales et de la santé bac+2 ; BTS, DUT ou équivalent ; Diplôme de 1er cycle universitaire ; Diplôme de 2e cycle universitaire ; Diplôme d'ingénieur / d'une grande école ; Diplôme de 3e cycle universitaire, Doctorat	Nouvelle variable dip12 créée en regroupant les modalités des variables du diplôme détaillé le plus élevé des enquêtes de 2010, 2011 2013 et 2014
	Medium education	Dummy	dip12 = CAP, BEP ou diplôme de même niveau ; Brevet supérieur/professionnel ou de technicien ; Capacité en droit, DAEU ; Baccalauréat professionnel / technologique / général	
	Low education	Dummy	dip12 = Aucun diplôme ; Certificat d'études primaires ; Brevet des collèges, BEPC, brevet élémentaire	
Employment status	Employed	Dummy	situa = Occupe un emploi ; Apprenti(e) sous contrat ou en stage rémunéré	Situation principale vis-à-vis du travail
	Unemployed	Dummy	situa = Chômeur (inscrit(e) ou non à l'ANPE)	
	Inactive	Dummy	situa = Étudiant(e), élève, en formation ou en stage non rémunéré ; Retraité(e) ou retiré(e) des affaires ou en préretraite ; Femme ou homme au foyer ; Inactif(ve) pour cause d'invalidité ; Autre situation d'inactivité	
Socioeconomic category	Upper class	Dummy	cs42 = Chefs d'entreprise de 10 salariés ou plus ; Professions libérales ; Cadres de la fonction publique ; Professeurs, professions scientifiques ; Professions de l'information, des arts et des spectacles ; Cadres administratifs et commerciaux d'entreprise ; Ingénieurs et cadres techniques d'entreprise ; Anciens cadres ; Anciennes professions intermédiaires ; Agriculteurs sur grande exploitation	Catégorie socioprofessionnelle détaillée
	Middle class	Dummy	cs42 = Artisans ; Commerçants et assimilés ; Professeurs des écoles, instituteurs et assimilés ; Professions intermédiaires de la santé et du travail social ; Clergé, religieux ; Professions intermédiaires administratives de la fonction publique ; Professions intermédiaires administratives et commerciales des entreprises ; Techniciens ; Contremaîtres, agents de maîtrise ; Agriculteurs sur moyenne exploitation ; Anciens agriculteurs exploitants ; Anciens artisans, commerçants, chefs d'entreprise	
	Lower class	Dummy	cs42 = Employés civils et agents de service de la fonction publique ; Policiers et militaires ; Employés administratifs d'entreprise ; Employés de commerce ; Personnels des services directs aux particuliers ; Ouvriers qualifiés et non qualifiés ; Chauffeurs ; Ouvriers agricoles ; Agriculteurs sur petite exploitation ; Anciens employés ; Anciens ouvriers	
	Other	Dummy	cs42 = Chômeurs n'ayant jamais travaillé ; Militaires du contingent ; Élèves, étudiants ; Personnes diverses sans activités professionnelles (sauf retraités)	
Household income	1st-5th decile	Dummy	nivie10_ocde >= 1 & <= 5	Décile de revenue simulé par UC échelle équivalente OCDE
	6th-10th decile	Dummy	nivie10_ocde >= 5 & <= 10	
Number of household members		Numeric	npers	Nombre de personnes dans le ménage
City size	Country side	Dummy	strate = Commune rurale	Strate d'unité urbaine du lieu de résidence
	< 20,000 inhabitants	Dummy	strate = Unité urbaine de moins de 20 000 habitants	
	20,000-100,000 inhab.	Dummy	strate = Unité urbaine de 20 000 à 99 999 habitants	
	> 100,000 inhabitants	Dummy	strate = Unité urbaine de plus de 100 000 habitants	
	Parisian agglomeration	Dummy	strate = Agglomération Parisienne	
Priority district	Living in a QPV/ZUS	Dummy	qp | Izus = Oui	Résidence en Quartier prioritaire ou Zone urbaine sensible
	Not living in a QPV/ZUS	Dummy	qp | Izus = Non	

Note: The variable names and question numbers listed here refer to those used in the last CVS survey of 2019 (but may have changed from year to year and have been harmonised with the creation of new variables in this case).

Figure A1.2: Construction of assault and outcome variables

Category	Value	Type	Definition	CVS questions used
ASSAULT CHARACTERISTICS				
Type of assault	Rape	Dummy	Viol, rapport sexuel forcé	S11B et Q0401, Q25bis, Q47
	Attempted rape	Dummy	Tentative de viol, de rapport sexuel forcé	
	Other sexual assault	Dummy	Attouchements sexuels, autre forme d'agression sexuelle	
Occurrence	Once	Dummy	Est arrivé une fois au cours des deux dernières années	S8 et Q02bis
	Twice or more	Dummy	Est arrivé deux fois ou plus au cours des deux dernières années	
Number of offenders	Only one	Dummy	Agression commise par une seule personne	S15 et Q03a, Q19, Q20, Q43a, Q44a
	More	Dummy	Agression commise par plusieurs personnes	
Offender's gender	Male / only men	Dummy	Agression commise par un homme / uniquement des hommes	S16, S21 et Q03d, Q03g, Q21c, Q22bis, Q43d, Q43g, Q44d, Q44f
	At least one woman	Dummy	Agression commise par une femme / uniquement des femmes	
Link-victim offender	Known before	Dummy	Agresseur (ou au moins un) connu avant	S17, S22 (hors-ménage)
	Unknown before	Dummy	Agresseur(s) totalement inconnu(s) avant	
Location of the assault	In a public space	Dummy	Dans un transport en commun ou dans la rue	S9 (hors-ménage)
	Elsewhere	Dummy	A votre domicile, dans le logement de quelqu'un d'autre, dans votre immeuble, sur votre lieu de travail ou d'étude, ailleurs	
Victim's partner	Yes	Dummy = 1	Conjoint	Q03b, Q03c, Q03e, Q03f, Q21a, Q21b, Q22a, Q43b, Q43c, Q43e, Q43f, Q44b, Q44c, Q44e (intra-ménage)
	No	Dummy = 0	Père, mère, conjoint d'un de vos parents/enfants, un de vos enfants, autre membre de la famille, autre personne qui vit actuellement avec vous	
Physical violence/weapon use	Yes	Dummy = 1	Objet lancé contre vous, giflé(e), mordu(e), cheveux tirés ou bousculé(e) brutalement ; Frappé(e) avec les pieds/poings, brûlures infligées, jeté(e) au sol ou autres brutalités physiques ; Tenté de vous étrangler, de vous tuer ; Arme ou objet dangereux utilisé contre vous ou menacé de l'être	S13 -> S14b (hors-ménage) et Q03j -> Q03m, Q241 -> Q244, Q451 -> Q454 (intra-ménage)
	No	Dummy = 0	Aucun de ces faits	
Physical injury	Yes	Dummy = 1	Agression a causé une ou plusieurs fractures ou des blessures physiques visibles ou non-visibles	S28 et Q04, Q25, Q46
	No	Dummy = 0	Aucun de ces faits	
Offender drugged/drunk	Yes	Dummy = 1	Auteur(s) sous emprise de l'alcool ou d'une drogue	S20bis, S26bis et Q04ter, Q25ter, Q47bis
	No	Dummy = 0	Non	
OUTCOME VARIABLE				
Filed a complaint	Yes	Dummy = 1	S'est déplacé(e) à la police/gendarmerie pour déclarer cette agression ET a déposé plainte	S32, S33 et Q07, Q07bis, Q28, Q29, Q50, Q51
	No	Dummy = 0	Ne s'est pas déplacé(e) à la police/gendarmerie OU s'est déplacé mais n'a pas déposé plainte	

Note: The variable names and question numbers listed here refer to those used in the last CVS survey of 2019 (but may have changed from year to year and have been harmonised with the creation of new variables in this case).

Table A1.3: Victims of sexual violence in the population

		Victims of sexual violence					
		Outside current household		**Inside current household**		**Total**	
Total sample size	*N = 157 972*	*N*	Weighted prop.	*N*	Weighted prop.	*N*	Weighted prop.
Number of victims		1476	0.0085 (0.0003)	471	0.0037 (0.0002)	1854	0.0131 (0.0004)
Gender	Women	1201	0.0126 (0.0005)	408	0.0059 (0.0004)	1523	0.0198 (0.0006)
	Men	275	0.0040 (0.0003)	63	0.0014 (0.0002)	331	0.0059 (0.0004)
Age	18-29	491	0.0165 (0.0009)	80	0.0028 (0.0004)	545	0.0216 (0.0012)
	30-54	705	0.0072 (0.0003)	285	0.0041 (0.0003)	936	0.0116 (0.0005)
	55-76	280	0.0052 (0.0005)	106	0.0037 (0.0005)	373	0.0099 (0.0008)
Marital status	Single	960	0.0151 (0.0006)	160	0.0047 (0.0006)	1076	0.0309 (0.0012)
	In a couple	516	0.0055 (0.0003)	311	0.0034 (0.0002)	778	0.0086 (0.0004)
Place of birth	In France	1212	0.0077 (0.0003)	375	0.0033 (0.0002)	1515	0.0119 (0.0004)
	Abroad	264	0.0134 (0.0011)	96	0.0061 (0.0009)	339	0.0204 (0.0015)
Education level	High	392	0.0073 (0.0004)	106	0.0027 (0.0003)	481	0.0110 (0.0006)
	Medium	652	0.0082 (0.0004)	205	0.0033 (0.0003)	817	0.0124 (0.0006)
	Low	430	0.0100 (0.0007)	160	0.0054 (0.0006)	554	0.0163 (0.0009)
Socioeconomic category	Upper class	125	0.0035 (0.0004)	50	0.0019 (0.0003)	174	0.0060 (0.0005)
	Middle class	264	0.0067 (0.0005)	85	0.0031 (0.0004)	336	0.0105 (0.0007)
	Lower class	734	0.0086 (0.0004)	248	0.0045 (0.0004)	925	0.0138 (0.0006)
	Other	352	0.0203 (0.0014)	87	0.0047 (0.0007)	417	0.0269 (0.0017)
Activity status	Employed	699	0.0067 (0.0003)	243	0.0033 (0.0003)	901	0.0106 (0.0004)
	Unemployed	245	0.0168 (0.0014)	61	0.0051 (0.0008)	288	0.0236 (0.0018)
	Inactive	532	0.0103 (0.0006)	167	0.0044 (0.0004)	665	0.0160 (0.0008)
Household income	1st - 5th decile	455	0.0119 (0.0005)	155	0.0050 (0.0004)	594	0.0182 (0.0007)
	6th - 10th decile	1021	0.0053 (0.0003)	316	0.0025 (0.0002)	1260	0.0085 (0.0004)
City size	Countryside	169	0.0048 (0.0005)	88	0.0031 (0.0004)	249	0.0082 (0.0006)
	$<$ 20,000 inhab.	175	0.0064 (0.0006)	60	0.0027 (0.0004)	222	0.0096 (0.0008)
	20,000 - 100,000 inhab.	245	0.0101 (0.0009)	77	0.0038 (0.0005)	304	0.0151 (0.0011)
	$\geq$ 100,000 inhab.	646	0.0110 (0.0006)	181	0.0045 (0.0004)	786	0.0171 (0.0008)
	Parisian agglomeration	240	0.0099 (0.0008)	64	0.0041 (0.0007)	291	0.0156 (0.0012)
Priority district	Living in a QPV/ZUS	359	0.0144 (0.0010)	111	0.0058 (0.0007)	437	0.0215 (0.0014)
	Not living in a QPV/ZUS	1117	0.0080 (0.0003)	360	0.0035 (0.0002)	1417	0.0124 (0.0004)

Note: Individuals living in metropolitan France and aged between 18 and 75 years, who answered the module on sensitive violence. *N*: unweighted number of observations (victims). Weighted prop.: weighted proportion in the total sample and standard errors in parentheses.
Source: CVS surveys, 2007-2019 (pooled data).

Table A1.4: Descriptive statistics for variables relating to legal remedies undertaken by victims

		Victims of sexual violence		
		Extra-household	**Intra-household**	**Total**
Went to the police	Mean	0.156	0.161	0.161
	(Se)	(0.012)	(0.025)	(0.012)
	N	1400	423	1742
Filed a complaint	Mean	0.677	0.581	0.635
	(Se)	(0.037)	(0.047)	(0.032)
	N	212	80	292
"Main courante"	Mean	0.247	0.430	0.315
	(Se)	(0.036)	(0.040)	(0.035)
	N	212	38	252
Gave up making a report	Mean	0.075	-	0.088
	(Se)	(0.021)	-	(0.021)
	N	212	22	235
Went to the police and	Mean	0.102	0.099	0.100
filed a complaint	(Se)	(0.010)	(0.020)	(0.009)
	N	1386	423	1727

Note: Victims of at least one sexual assault committed during the two years preceding the survey, living in metropolitan France and aged between 18 and 75 years. *N*: unweighted number of observations (victims). Weighted proportion in the sample and standard errors in parentheses.

A2 Reasons for not reporting sexual violence

Figure A2.1: Reasons for not reporting sexual violence by victim's gender

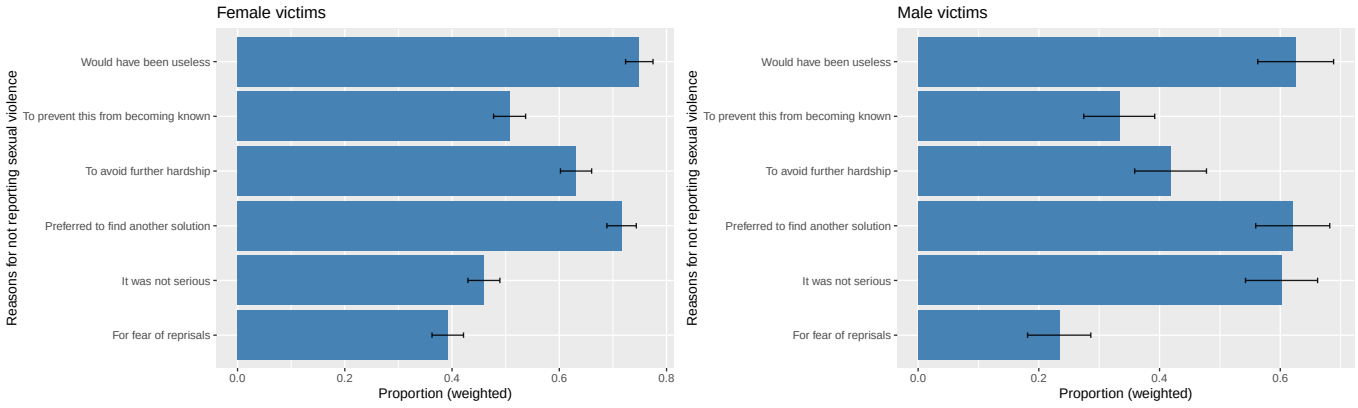
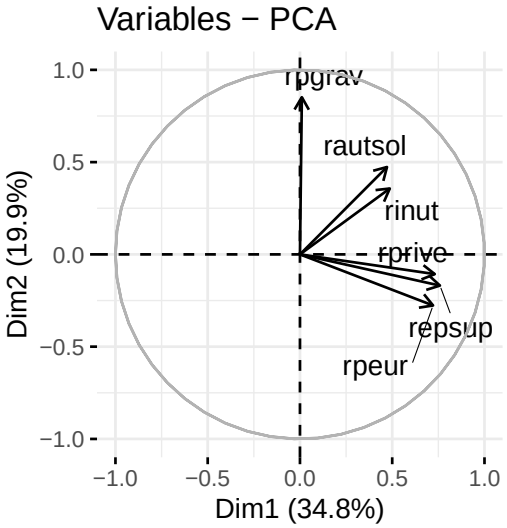


Table A2.1: Correlation Matrix - Reasons for not reporting sexual violence

	rprive	rautsol	rpeur	rinut	rpgrav	repsup
rprive	1	0.240	0.388	0.185	-0.028	0.413
rautsol	0.240	1	0.156	0.149	0.166	0.200
rpeur	0.388	0.156	1	0.200	-0.098	0.439
rinut	0.185	0.149	0.200	1	0.133	0.236
rpgrav	-0.028	0.166	-0.098	0.133	1	-0.054
repsup	0.413	0.200	0.439	0.236	-0.054	1

Note: The variables are "To prevent this from becoming known" (rprive), "Preferred to find another solution" (rautsol), "For fear of reprisals" (rpeur), "Would have been useless" (rinut), "It was not serious" (rpgrav), "To avoid further hardship" (repsup)

Figure A2.2: Principal Component Analysis - Reasons for not reporting sexual violence



Note: The variables are "To prevent this from becoming known" (rprive), "Preferred to find another solution" (rautsol), "For fear of reprisals" (rpeur), "Would have been useless" (rinut), "It was not serious" (rpgrav), "To avoid further hardship" (repsup)

A3 Determinants of the victim's reporting decision

Table A3.1: Determinants of the reporting propensity for victims of extra-household sexual violence - LPM

	Dependent variable: Filed a complaint (0/1)				
	(1)	(2)	(3)	(4)	(5)
Age	−0.002		−0.005	−0.004	−0.003
	(0.009)		(0.008)	(0.009)	(0.008)
Age-squared	0.0001		0.0001	0.0001	0.0001
	(0.0001)		(0.0001)	(0.0001)	(0.0001)
Female	0.041		0.004	0.005	0.008
	(0.027)		(0.031)	(0.031)	(0.030)
Born abroad	−0.019		−0.0004	−0.003	−0.004
	(0.033)		(0.036)	(0.035)	(0.035)
In a couple	0.007		0.028	0.028	0.024
	(0.026)		(0.027)	(0.027)	(0.027)
Education (Ref: High education)					
Medium education	0.006		−0.009	−0.011	−0.009
	(0.030)		(0.030)	(0.030)	(0.030)
Low education	0.033		0.007	0.007	0.011
	(0.037)		(0.040)	(0.040)	(0.040)
Employment status (Ref: Employed)					
Unemployed	0.042		0.019	0.018	0.025
	(0.038)		(0.041)	(0.041)	(0.041)
Inactive	−0.028		−0.043	−0.044	−0.038
	(0.057)		(0.057)	(0.059)	(0.058)
Socioeconomic category (Ref: Upper class)					
Middle class	0.002		0.022	0.023	0.024
	(0.047)		(0.045)	(0.045)	(0.042)
Lower class	0.063		0.080	0.081	0.085*

To be continued on next page

Table A3.1: Determinants of the reporting propensity for victims of extra-household sexual violence - LPM

	(0.054)		(0.051)	(0.051)	(0.048)
Other	0.081		0.107	0.102	0.104
	(0.069)		(0.070)	(0.072)	(0.069)
Income decile (Ref: 6th-10th deciles)					
1st-5th decile	0.013		0.004	0.005	0.006
	(0.023)		(0.023)	(0.024)	(0.024)
Number of household members	0.001		0.005	0.004	0.005
	(0.010)		(0.011)	(0.011)	(0.011)
Size of the urban unit (Ref: Rural areas)					
Less than 20,000	0.087**		0.103**	0.105**	0.098**
	(0.043)		(0.045)	(0.046)	(0.045)
20,000-100,000	0.022		0.007	0.010	−0.001
	(0.028)		(0.031)	(0.033)	(0.034)
More than 100,000	0.083***		0.080**	0.084***	0.079**
	(0.031)		(0.033)	(0.032)	(0.033)
Paris urban unit	0.072**		0.064	0.154**	0.138**
	(0.036)		(0.040)	(0.065)	(0.069)
Living in QP/ZUS	0.082*		0.064	0.058	0.042
	(0.045)		(0.045)	(0.044)	(0.043)
Assault type (Ref: Other sexual assault)					
Rape		0.104***	0.098***	0.099***	0.120***
		(0.031)	(0.030)	(0.030)	(0.031)
Attempted rape		0.103***	0.105***	0.106***	0.109***
		(0.035)	(0.036)	(0.036)	(0.038)
Offender (dummies)					
Several offenders		−0.063	−0.083*	−0.085*	−0.084*
		(0.042)	(0.047)	(0.048)	(0.047)
Known offender		−0.069**	−0.055	−0.058	−0.071*
		(0.035)	(0.038)	(0.038)	(0.038)
Occurrence (dummies)					
Occurred a public space		0.004	0.016	0.017	0.005
		(0.043)	(0.046)	(0.046)	(0.046)
Occurred twice or more		−0.048*	−0.055**	−0.055**	−0.049**
		(0.025)	(0.023)	(0.023)	(0.023)
Physical injuries		0.163***	0.140***	0.138***	0.131***
		(0.031)	(0.031)	(0.031)	(0.030)
Intercept	−0.077	0.088***	−0.032	−0.124	−0.171
	(0.141)	(0.033)	(0.142)	(0.159)	(0.169)
Region fixed effects	No	No	No	Yes	Yes
Year fixed effects	No	No	No	No	Yes
Observations	1,382	1,286	1,282	1,282	1,282
R^2	0.050	0.103	0.143	0.148	0.175
Adjusted R^2	0.037	0.098	0.126	0.125	0.145
F Statistic	3.784***	20.907***	8.084***	6.555***	5.837***
	(df = 19; 1362)	(df = 7; 1278)	(df = 26; 1255)	(df = 33; 1248)	(df = 45; 1236)

Note: Robust standard errors are reported in parentheses. * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$.

Table A3.2: Determinants of the reporting propensity for victims of intra-household sexual violence - LPM

	Dependent variable: Filed a complaint (0/1)				
	(1)	(2)	(3)	(4)	(5)
Age	0.004		0.004	0.005	0.012
	(0.011)		(0.010)	(0.010)	(0.010)
Age-squared	−0.00005		−0.0001	−0.0001	−0.0002
	(0.0001)		(0.0001)	(0.0001)	(0.0001)
Female	−0.026		−0.088	−0.101	−0.108
	(0.061)		(0.070)	(0.068)	(0.067)
Born abroad	0.086		0.069	0.076	0.075
	(0.094)		(0.084)	(0.082)	(0.084)
In a couple	−0.091*		−0.116*	−0.109	−0.084
	(0.052)		(0.067)	(0.068)	(0.074)
Education (Ref: High education)					
Medium education	0.020		0.003	−0.004	0.005
	(0.040)		(0.044)	(0.046)	(0.050)
Low education	0.061		0.073	0.067	0.065
	(0.052)		(0.065)	(0.071)	(0.067)
Employment status (Ref: Employed)					
Unemployed	−0.013		−0.016	−0.014	−0.019
	(0.042)		(0.044)	(0.048)	(0.055)
Inactive	0.125		0.193	0.201	0.207
	(0.133)		(0.156)	(0.151)	(0.144)
Socioeconomic category (Ref: Upper class)					
Middle class	0.113		0.118	0.123*	0.125*
	(0.069)		(0.074)	(0.073)	(0.073)
Lower class	0.024		0.034	0.045	0.037
	(0.038)		(0.055)	(0.053)	(0.058)
Other	−0.033		−0.103	−0.093	−0.095
	(0.154)		(0.178)	(0.169)	(0.172)
Income decile (Ref: 6th-10th decile))					
1st-5th decile	−0.027		−0.016	−0.022	−0.027
	(0.057)		(0.053)	(0.054)	(0.051)
Number of household members	−0.009		−0.022	−0.024	−0.022
	(0.022)		(0.021)	(0.021)	(0.020)
Size of the urban unit (Ref: Rural areas)					
Less than 20,000	−0.027		−0.035	−0.012	−0.014
	(0.042)		(0.047)	(0.052)	(0.052)
20,000-100,000	0.028		0.034	0.053	0.043
	(0.067)		(0.065)	(0.064)	(0.064)
More than 100,000	0.019		0.017	0.023	0.017
	(0.049)		(0.050)	(0.053)	(0.053)
Paris urban unit	−0.030		−0.037	0.005	0.057
	(0.061)		(0.058)	(0.128)	(0.227)
Living in QP/ZUS	0.176**		0.127	0.133*	0.110
	(0.085)		(0.079)	(0.076)	(0.077)
Assault type (not exclusive cat.)					
Rape		0.013	0.018	0.033	0.030
		(0.050)	(0.060)	(0.055)	(0.064)
Attempted rape		−0.075	−0.081	−0.077	−0.069
		(0.051)	(0.052)	(0.049)	(0.049)
Other sexual assault		0.068	0.075	0.069	0.042
		(0.047)	(0.052)	(0.049)	(0.042)

To be continued on next page

Table A3.2: Determinants of the reporting propensity for victims of intra-household sexual violence - LPM

Offender (dummies)					
Several offenders		−0.094	−0.085	−0.113*	−0.101
		(0.059)	(0.052)	(0.061)	(0.070)
Victim's partner		−0.005	0.062	0.059	0.040
		(0.046)	(0.049)	(0.051)	(0.052)
Occurred twice or more		−0.006	−0.008	−0.004	−0.0005
		(0.048)	(0.045)	(0.043)	(0.045)
Physical injuries		0.178***	0.175***	0.185***	0.183***
		(0.049)	(0.047)	(0.047)	(0.050)
Intercept	0.026	0.032	0.123	0.053	−0.180
	(0.273)	(0.039)	(0.272)	(0.309)	(0.373)
Region fixed effects	No	No	No	Yes	Yes
Year fixed effects	No	No	No	No	Yes
Observations	421	375	373	373	373
R^2	0.115	0.129	0.242	0.282	0.327
Adjusted R^2	0.073	0.112	0.185	0.212	0.235
F Statistic	2.749***	7.759***	4.256***	4.025***	3.533***
	(df = 19; 401)	(df = 7; 367)	(df = 26; 346)	(df = 33; 339)	(df = 45; 327)

Note: Robust standard errors are reported in parentheses. $^{*}p < 0.1$; $^{**}p < 0.05$; $^{***}p < 0.01$.

Table A3.3: Determinants of the reporting propensity for female victims - LPM

	Dependent variable: Filed a complaint (0/1)				
	(1)	(2)	(3)	(4)	(5)
Age	0.008		0.007	0.008	0.010*
	(0.006)		(0.005)	(0.005)	(0.005)
Age-squared	−0.0001		−0.0001	−0.0001	−0.0001
	(0.0001)		(0.0001)	(0.0001)	(0.0001)
Born abroad	−0.001		0.013	0.017	0.013
	(0.038)		(0.041)	(0.042)	(0.041)
In a couple	−0.029		−0.007	−0.009	−0.011
	(0.024)		(0.027)	(0.027)	(0.027)
Education (Ref: High education)					
Medium education	0.009		−0.006	−0.006	0.002
	(0.027)		(0.028)	(0.028)	(0.029)
Low education	0.057*		0.029	0.031	0.042
	(0.034)		(0.037)	(0.038)	(0.039)
Employment status (Ref: Employed)					
Unemployed	0.048		0.037	0.039	0.049
	(0.034)		(0.036)	(0.037)	(0.037)
Inactive	0.068		0.065	0.067	0.071
	(0.055)		(0.057)	(0.058)	(0.058)
Socioeconomic category (Ref: Upper class)					
Middle class	0.045		0.036	0.037	0.040
	(0.033)		(0.035)	(0.035)	(0.036)
Lower class	0.059*		0.051	0.049	0.052
	(0.031)		(0.034)	(0.033)	(0.034)

To be continued on next page

Table A3.3: Determinants of the reporting propensity for female victims of sexual violence - LPM

Other	0.028		0.027	0.026	0.030
	(0.059)		(0.063)	(0.065)	(0.064)
Income decile (Ref: 6th-10th decile)					
1st-5th decile	0.007		0.005	−0.003	−0.006
	(0.024)		(0.025)	(0.026)	(0.025)
Number of household members	−0.002		−0.005	−0.007	−0.007
	(0.010)		(0.011)	(0.011)	(0.011)
Size of the urban unit (Ref: Rural areas)					
Less than 20,000	0.067*		0.073*	0.079**	0.074*
	(0.037)		(0.038)	(0.039)	(0.038)
20,000-100,000	0.018		0.018	0.024	0.016
	(0.030)		(0.032)	(0.032)	(0.032)
More than 100,000	0.060**		0.054*	0.062**	0.061*
	(0.030)		(0.032)	(0.032)	(0.032)
Paris urban unit	0.052		0.056	0.137**	0.131*
	(0.034)		(0.038)	(0.063)	(0.067)
Living in QP/ZUS	0.073*		0.064	0.064	0.052
	(0.039)		(0.041)	(0.041)	(0.040)
Assault type (not exclusive cat.)					
Rape		−0.017	−0.015	−0.014	−0.006
		(0.031)	(0.033)	(0.033)	(0.034)
Attempted rape		0.015	0.009	0.011	0.010
		(0.026)	(0.028)	(0.028)	(0.029)
Other sexual assault		−0.041	−0.034	−0.034	−0.043
		(0.026)	(0.027)	(0.027)	(0.027)
Offender (dummies)					
Several offenders		0.036	0.026	0.025	0.035
		(0.050)	(0.052)	(0.051)	(0.049)
Victim's partner or ex-partner		−0.014	−0.008	−0.011	−0.004
		(0.029)	(0.028)	(0.028)	(0.029)
Occurred twice or more		−0.006	−0.018	−0.018	−0.015
		(0.024)	(0.023)	(0.023)	(0.023)
Physical injuries		0.139***	0.118***	0.119***	0.114***
		(0.028)	(0.027)	(0.027)	(0.026)
Intercept	−0.178	0.084***	−0.160	−0.246**	−0.311**
	(0.110)	(0.027)	(0.107)	(0.121)	(0.141)
Region fixed effects	No	No	No	Yes	Yes
Year fixed effects	No	No	No	No	Yes
Observations	1,441	1,346	1,341	1,341	1,341
R^2	0.039	0.055	0.078	0.084	0.104
Adjusted R^2	0.026	0.050	0.061	0.062	0.074
F Statistic	3.170***	11.044***	4.455***	3.745***	3.429***
	(df = 18; 1422)	(df = 7; 1338)	(df = 25; 1315)	(df = 32; 1308)	(df = 44; 1296)

Note: Robust standard errors are reported in parentheses. * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$.

A4 Probit models

Figure A4.1: Predicted probabilities for LPM - Determinants of complaint filing for victims of sexual violence

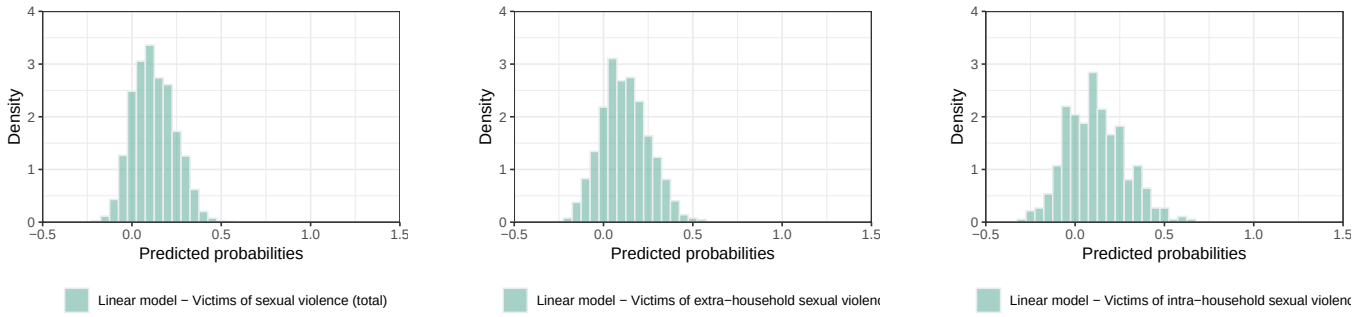


Figure A4.2: Predicted probabilities for LPM - Determinants of reasons for not reporting sexual violence

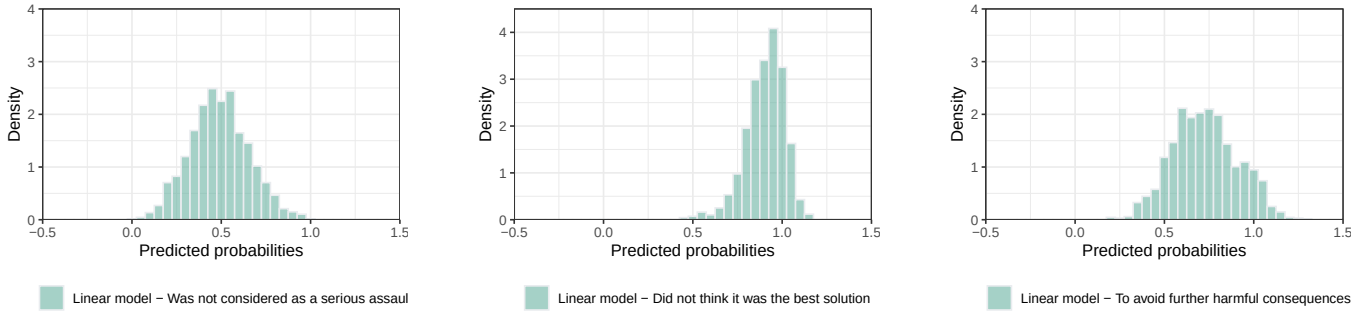


Table A4.1: Determinants of the reporting propensity for victims of sexual violence (total) - Probit (ME)

	Dependent variable: Filed a complaint (0/1)				
	(1)	(2)	(3)	(4)	(5)
Age	−0.0002		−0.0008	−0.0006	0.0022
	0.0051		0.0040	0.0040	0.0040
Age-squared	0.0000		0.0000	0.0000	−0.0000
	0.0001		0.0001	0.0001	0.0001
Female	0.0442		0.0154	0.0166	0.0203
	0.0267		0.0259	0.0255	0.0249
Born abroad	0.0027		0.0134	0.0163	0.0063
	0.0254		0.0253	0.0254	0.0243
In a couple	−0.0078		0.0152	0.0155	0.0191
	0.0206		0.0219	0.0215	0.0212
Education (Ref: High education)					
Medium education	0.0257		0.0014	0.0014	0.0070
	0.0306		0.0311	0.0306	0.0296
Low education	0.0493		0.0156	0.0173	0.0224
	0.0318		0.0339	0.0335	0.0334
Employment status (Ref: Employed)					
Unemployed	0.0430		0.0354	0.0351	0.0488
	0.0263		0.0289	0.0288	0.0280

To be continued on next page

Table A4.1: Determinants of the reporting propensity for victims of sexual violence (total) - Probit Models (ME)

Inactive	0.0086		0.0198	0.0209	0.0310
	0.0398		0.0373	0.0366	0.0347
Socioeconomic category (Ref: Upper class)					
Middle class	0.0242		0.0289	0.0282	0.0279
	0.0579		0.0571	0.0557	0.0491
Lower class	0.0502		0.0470	0.0455	0.0452
	0.0556		0.0544	0.0524	0.0466
Other	0.0607		0.0535	0.0518	0.0443
	0.0622		0.0604	0.0599	0.0562
Income decile (Ref: 6th-10th decile)					
1st-5th decile	0.0063		0.0017	−0.0004	0.0012
	0.0233		0.0236	0.0236	0.0227
Number of household members	−0.0022		−0.0030	−0.0040	−0.0040
	0.0079		0.0086	0.0082	0.0080
Size of the urban unit (Ref: Rural areas)					
Less than 20,000	0.0823		0.1022	0.1047	0.0917
	0.0409		0.0423	0.0427	0.0390
20,000-100,000	0.0383		0.0426	0.0467	0.0350
	0.0374		0.0393	0.0403	0.0394
More than 100,000	0.0856		0.0907	0.0950	0.0819
	0.0359		0.0376	0.0374	0.0361
Paris urban unit	0.0678		0.0758	0.2436	0.2219
	0.0391		0.0426	0.1124	0.1083
Living in QP/ZUS	0.0721		0.0673	0.0659	0.0477
	0.0229		0.0215	0.0211	0.0212
Assault type (not exclusive cat.)					
Rape		0.0228	0.0181	0.0196	0.0309
		0.0279	0.0257	0.0239	0.0235
Attempted rape		0.0215	0.0209	0.0219	0.0191
		0.0235	0.0216	0.0212	0.0206
Other sexual assault		−0.0294	−0.0205	−0.0221	−0.0318
		0.0244	0.0224	0.0218	0.0204
Offender (dummies)					
Several offenders		0.0011	0.0027	0.0040	0.0204
		0.0315	0.0287	0.0291	0.0263
Victim's partner or ex-partner		−0.0177	−0.0180	−0.0228	−0.0206
		0.0260	0.0232	0.0230	0.0222
Occurred twice or more		−0.0283	−0.0326	−0.0301	−0.0254
		0.0218	0.0192	0.0188	0.0183
Physical injuries		0.1461	0.1297	0.1278	0.1221
		0.0254	0.0226	0.0221	0.0198
Intercept	−4.627***	−2.733***	−4.659***	−6.593***	−7.810***
	(1.211)	(0.292)	(1.116)	(1.620)	(1.815)
Region fixed effects	No	No	No	Yes	Yes
Year fixed effects	No	No	No	No	Yes
Observations	1,721	1,584	1,578	1,578	1,578
Log Likelihood	−509.104	−479.742	−449.785	−444.917	−420.001
Akaike Inf. Crit.	1,058.208	975.483	953.571	957.835	932.001

Note: Standard errors are reported in parentheses. * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$.

Table A4.2: Determinants of the reasons for not reporting sexual violence - Full specification (Probit - ME)

	Dependent variable: Reason (0/1)		
	Was not considered as a serious assault	Did not think it was the best solution	To avoid further harmful consequences
	(1)	(2)	(3)
Age	−0.001	−0.002	0.017
	0.008	0.005	0.007
Age-squared	0.000	−0.000	−0.000
	0.000	0.000	0.000
Female	−0.170	0.050	0.103
	0.044	0.025	0.038
In a couple	0.015	−0.004	0.020
	0.037	0.022	0.036
Born abroad	0.050	−0.025	0.003
	0.050	0.028	0.049
Education (Ref: High education)			
Medium education	0.050	−0.002	0.134
	0.041	0.031	0.037
Low education	0.013	−0.035	0.150
	0.052	0.035	0.048
Employment status (Ref: Employed)			
Unemployed	−0.082	0.024	−0.048
	0.054	0.034	0.046
Inactive	−0.040	0.030	0.034
	0.066	0.041	0.059
Socioeconomic category (Ref: Upper class)			
Middle class	−0.118	0.033	0.013
	0.063	0.040	0.053
Lower class	−0.046	0.010	0.046
	0.062	0.040	0.053
Other	−0.008	−0.013	−0.018
	0.084	0.053	0.074
Income decile (Ref: 6th-10th decile)			
1st-5th decile	−0.008	0.002	−0.027
	0.038	0.023	0.033
Number of household members	0.017	−0.011	−0.032
	0.015	0.010	0.016
Size of the urban unit (Ref: Rural areas)			
Less than 20,000	0.037	−0.022	0.064
	0.060	0.031	0.056
20,000-100,000	0.150	0.048	−0.018
	0.058	0.036	0.054
More than 100,000	0.084	0.082	0.036
	0.051	0.031	0.045
Paris urban unit	0.198	0.005	0.020
	0.166	0.102	0.123
Living in QP/ZUS	0.027	0.034	0.078
	0.049	0.036	0.047
Assault type (not exclusive cat.)			
Rape	−0.057	0.010	0.076
	0.045	0.031	0.051
Attempted rape	−0.042	−0.004	0.091
	0.043	0.033	0.050

To be continued on next page

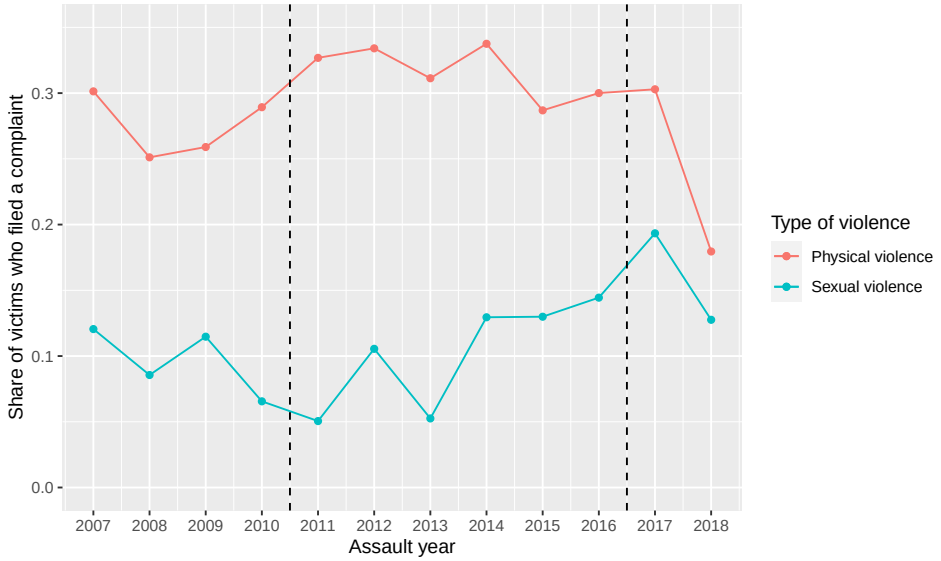
Table A4.2: Determinants of the reasons for not reporting sexual violence - Full specification (Probit Models - ME)

Other sexual assault	0.079	0.071	0.023
	0.041	0.033	0.048
Offender (dummies)			
Several offenders	−0.139	−0.097	−0.065
	0.063	0.027	0.054
Victim's partner or ex-partner	0.011	0.060	0.009
	0.042	0.031	0.040
Occurred twice or more	−0.030	0.027	0.059
	0.035	0.020	0.031
Physical injuries	−0.133	0.015	0.209
	0.035	0.026	0.034
Intercept	0.119	1.796	−2.941**
	(1.097)	(1.788)	(1.186)
Region fixed effects	Yes	Yes	Yes
Year fixed effects	Yes	Yes	Yes
Observations	1,335	1,352	1,352
Log Likelihood	−824.149	−395.459	−702.202
Akaike Inf. Crit.	1,740.297	882.919	1,496.404

Note: Standard errors are reported in parentheses. $^{*}p < 0.1$; $^{**}p < 0.05$; $^{***}p < 0.01$.

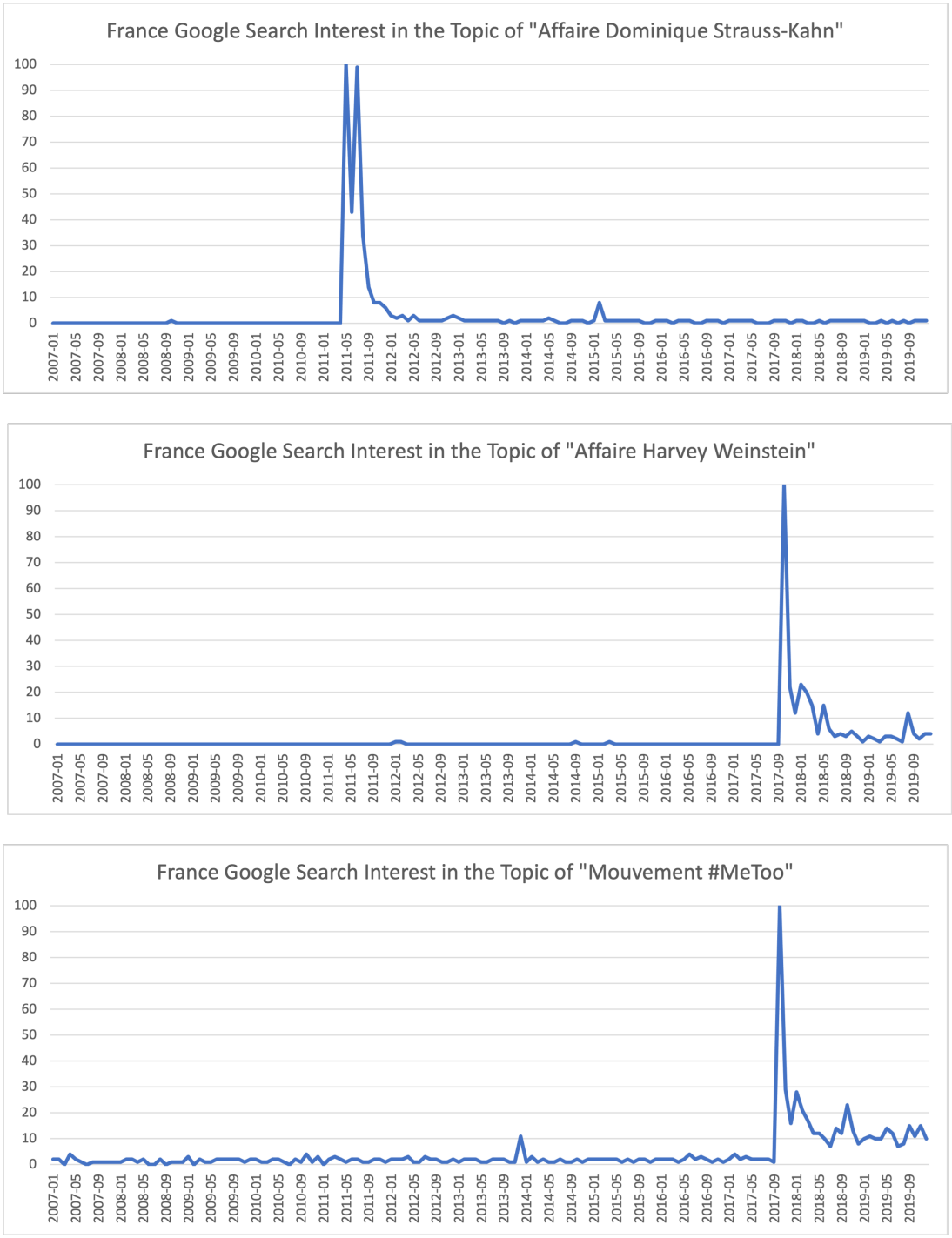
A5 High-profile sexual assault cases

Figure A5.1: Pre-trends for filing a complaint for female victims - raw data



Population: Female victims of sexual violence or physical violence in the two years preceding the survey. Weighted proportions.
Source: CVS surveys, 2007-2019.

Figure A5.2: Google Search Interest in France



Reading: Monthly time series for the relative search volume in the topic of the DSK case, the Weinstein case and the #MeToo movement in France.
Source: Google Trends data, extracted on 24/07/2021. This data creates an index value between 0 and 100 based on the relative search rate in each time period, where 100 is the time where this search rate is maximised.

A6 Heterogeneity analysis

Table A6.1: DiD estimation – Effect by age (female victims)

	Dependent variable: Filed a complaint (0/1)			
	Age above median (>40)		Age below median (<40)	
	(1)	(2)	(3)	(4)
Sexual violence	−0.263***	−0.171***	−0.140***	−0.218***
	(0.047)	(0.050)	(0.043)	(0.034)
Assault post 2011	0.075		0.037	
	(0.097)		(0.068)	
Sexual violence*Assault post 2011	−0.110		−0.058	
	(0.105)		(0.083)	
Assault post 2017		−0.101*		−0.064
		(0.057)		(0.055)
Sexual violence*Assault post 2017		0.128		0.163**
		(0.093)		(0.079)
Intercept	0.511	0.593**	0.134	−0.113
	(0.333)	(0.254)	(0.280)	(0.147)
Socio-demographic controls	Yes	Yes	Yes	Yes
Region fixed effects	Yes	Yes	Yes	Yes
Period	2007-2011	2012-2018	2007-2011	2012-2018
Observations	561	685	757	904
Adjusted R^2	0.143	0.089	0.072	0.116

Note: OLS estimates from LPM. Robust standard errors in parentheses *p<0.1; **p<0.05; ***p<0.01

Table A6.2: DiD estimation – Effect by level of education (female victims)

	Dependent variable: Filed a complaint (0/1)					
	DSK			Weinstein		
	(1) High	(2) Medium	(3) Low	(4) High	(5) Medium	(6) Low
Sexual violence	−0.068	−0.176***	−0.275***	−0.168***	−0.255***	−0.174**
	(0.067)	(0.042)	(0.065)	(0.037)	(0.045)	(0.069)
Assault post 2011	0.042	0.006	0.140			
	(0.113)	(0.079)	(0.105)			
Sexual violence*Assault post 2011	−0.121	−0.027	−0.188			
	(0.122)	(0.093)	(0.128)			
Assault post 2017				−0.101*	−0.133**	0.028
				(0.062)	(0.063)	(0.100)
Sexual violence*Assault post 2017				0.111	0.274**	0.066
				(0.077)	(0.113)	(0.151)
Intercept	−0.011	0.742***	−0.118	0.017	0.354*	0.404
	(0.342)	(0.281)	(0.292)	(0.127)	(0.204)	(0.457)
Socio-demographic controls	Yes	Yes	Yes	Yes	Yes	Yes
Region fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Period	2007-2011	2007-2011	2007-2011	2012-2018	2012-2018	2012-2018
Observations	360	563	395	489	699	401
Adjusted R^2	0.031	0.082	0.149	0.045	0.072	0.089

Note: OLS estimates from LPM. Robust standard errors in parentheses *p<0.1; **p<0.05; ***p<0.01

Table A6.3: DiD estimation – Effect by socio-economic category (female victims)

	Dependent variable: Filed a complaint (0/1)					
	DSK			Weinstein		
	(1) Upper class	(2) Middle class	(3) Lower class	(4) Upper class	(5) Middle class	(6) Lower class
Sexual violence	−0.120*	−0.233***	−0.227***	−0.192**	−0.201***	−0.209***
	(0.072)	(0.070)	(0.042)	(0.086)	(0.055)	(0.043)
Assault post 2011	0.246	−0.102	−0.003			
	(0.246)	(0.112)	(0.079)			
Sexual violence*Assault post 2011	−0.297	0.015	−0.020			
	(0.263)	(0.140)	(0.091)			
Assault post 2017				−0.078	−0.158**	0.009
				(0.177)	(0.066)	(0.070)
Sexual violence*Assault post 2017				0.113	0.210**	0.091
				(0.223)	(0.104)	(0.110)
Intercept	−0.137	0.662*	0.431	0.006	−0.191	0.279
	(0.485)	(0.342)	(0.363)	(0.222)	(0.161)	(0.180)
Socio-demographic controls	Yes	Yes	Yes	Yes	Yes	Yes
Region fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Period	2007-2011	2007-2011	2007-2011	2012-2018	2012-2018	2012-2018
Observations	133	228	636	158	348	738
Adjusted R^2	0.090	0.152	0.093	0.037	0.112	0.096

Note: OLS estimates from LPM. Robust standard errors in parentheses *p<0.1; **p<0.05; ***p<0.01

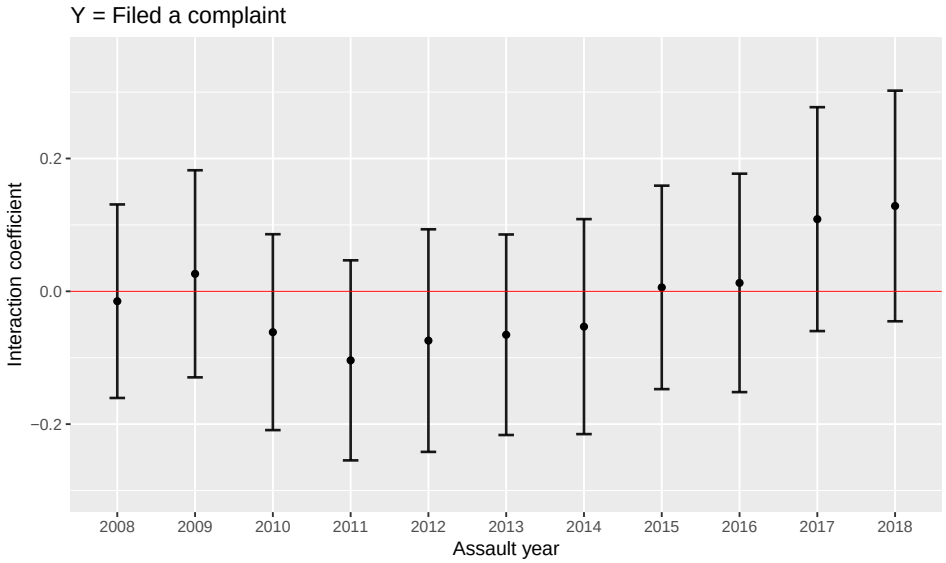
Table A6.4: DiD estimation – Effect by type of territory (female victims)

	Dependent variable: Filed a complaint (0/1)			
	Less than 100,000 inhabitants		More than 100,000 inhabitants	
	(1)	(2)	(3)	(4)
Sexual violence	−0.259***	−0.269***	−0.117***	−0.167***
	(0.041)	(0.044)	(0.042)	(0.035)
Assault post 2011	0.027		0.074	
	(0.090)		(0.071)	
Sexual violence*Assault post 2011	−0.007		−0.163**	
	(0.102)		(0.080)	
Assault post 2017		−0.041		−0.110**
		(0.066)		(0.054)
Sexual violence*Assault post 2017		0.076		0.214***
		(0.096)		(0.079)
Intercept	0.294	0.418***	0.021	0.055
	(0.236)	(0.145)	(0.089)	(0.089)
Socio-demographic controls	Yes	Yes	Yes	Yes
Region fixed effects	Yes	Yes	Yes	Yes
Period	2007-2011	2012-2018	2007-2011	2012-2018
Observations	555	677	763	912
Adjusted R^2	0.123	0.114	0.073	0.066

Note: OLS estimates from LPM. Robust standard errors in parentheses *p<0.1; **p<0.05; ***p<0.01

A7 Robustness checks

Figure A7.1: Placebo tests for DiD analysis - female victims



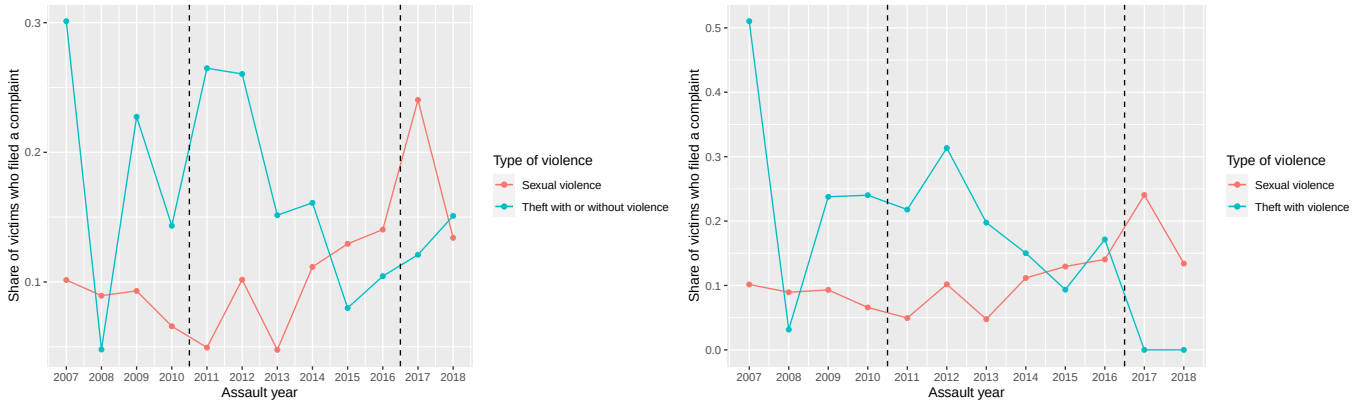
Note: Estimated coefficients β_3 from the equation $Y_{itr} = \beta_0 + \beta_1 SV_i + \beta_2 Post_t + \beta_3 SV_i \times Year_t + \gamma X_i + \phi_r + \epsilon_i$ with 10% confidence intervals. Sample: female victims of sexual violence or of physical violence in the two years preceding the survey.

Table A7.1: DiD estimation using other control groups – victims of theft/attempted theft

	Dependent variable: Filed a complaint (0/1)			
	Control: Theft/attempted theft with or without violence		Control: Theft/attempted theft with violence or threats	
	(1)	(2)	(3)	(4)
Sexual violence	−0.099**	−0.119***	−0.109	−0.165**
	(0.048)	(0.042)	(0.067)	(0.068)
Assault post 2011	0.118		0.024	
	(0.090)		(0.122)	
Sexual violence*Assault post 2011	−0.169*		−0.074	
	(0.094)		(0.125)	
Assault post 2017		−0.022		−0.220***
		(0.072)		(0.068)
Sexual violence*Assault post 2017		0.123		0.317***
		(0.085)		(0.081)
Intercept	−0.142	−0.201**	−0.044	−0.195*
	(0.127)	(0.092)	(0.097)	(0.117)
Socio-demographic controls	Yes	Yes	Yes	Yes
Region fixed effects	Yes	Yes	Yes	Yes
Period	2007-2011	2012-2018	2007-2011	2012-2018
Observations	786	961	688	852
Adjusted R^2	0.060	0.073	0.032	0.089

Note: OLS estimates from LPM. Robust standard errors in parentheses *p<0.1; **p<0.05; ***p<0.01

Figure A7.2: Pre-trends for filing a complaint with other control groups - raw data



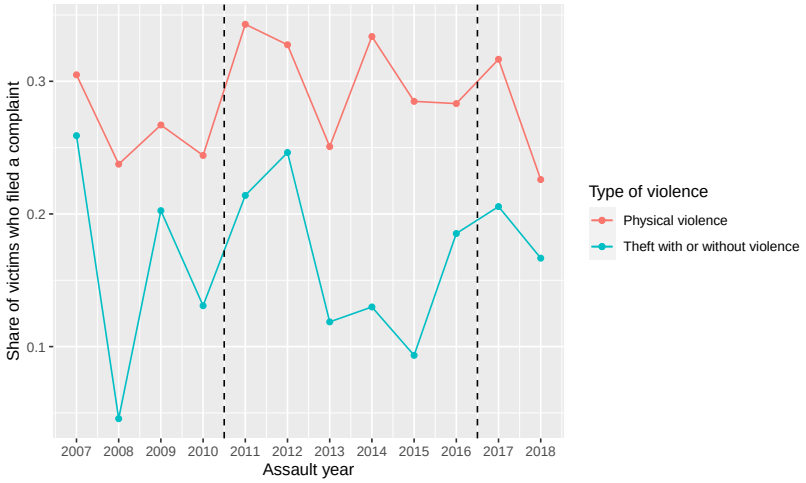
Population: Victims of sexual violence or theft/attempted theft in the two years preceding the survey. Weighted proportions.
Source: CVS surveys, 2007-2019.

Table A7.2: DiD estimation using other treatment group – victims of theft/attempted theft

	Dependent variable: Filed a complaint (0/1)	
	(1)	(2)
Theft or attempted theft	−0.076**	−0.111***
	(0.038)	(0.033)
Assault post 2011	0.077**	
	(0.038)	
Theft/attempted theft*Assault post 2011	−0.009	
	(0.081)	
Assault post 2017		−0.032
		(0.030)
Theft/attempted theft*Assault post 2017		0.088
		(0.075)
Intercept	0.051	0.065
	(0.159)	(0.110)
Socio-demographic controls	Yes	Yes
Region fixed effects	Yes	Yes
Period	2007-2011	2012-2018
Observations	1,702	2,125
Adjusted R^2	0.042	0.059

Note: OLS estimates from LPM. Robust standard errors in parentheses *p<0.1; **p<0.05; ***p<0.01

Figure A7.3: Pre-trends for filing a complaint with other treatment group - raw data



Population: Victims of physical violence or theft/attempted theft in the two years preceding the survey. Weighted proportions.
Source: CVS surveys, 2007-2019.

Table A7.3: DiD estimation using Probit models

	Dependent variable: Filed a complaint (0/1)			
	All victims	Women only	All victims	Women only
	(1)	(2)	(3)	(4)
Sexual violence	−0.230		−0.234	
	0.034		0.030	
Assault post 2011	0.071	0.030		
	0.026	0.036		
Sexual violence*Assault post 2011	−0.156	−0.134		
	0.072	0.075		
Assault post 2017			−0.029	−0.071
			0.025	0.035
Sexual violence*Assault post 2017			0.167	0.172
			0.057	0.064
Intercept	−2.104***	−1.271	−2.623***	−1.799***
	(0.750)	(0.981)	(0.518)	(0.674)
Socio-demographic controls	Yes	Yes	Yes	Yes
Region fixed effects	Yes	Yes	Yes	Yes
Period	2007-2011	2007-2011	2012-2018	2012-2018
Observations	2,158	1,318	2,692	1,589
Log Likelihood	−1,041.520	−583.325	−1,353.295	−763.212

Note: Marginal effects from Probit models (except the intercept). Standard errors below coefficients.

Figure A7.4: Predicted probabilities for LPM - Main DiD estimation

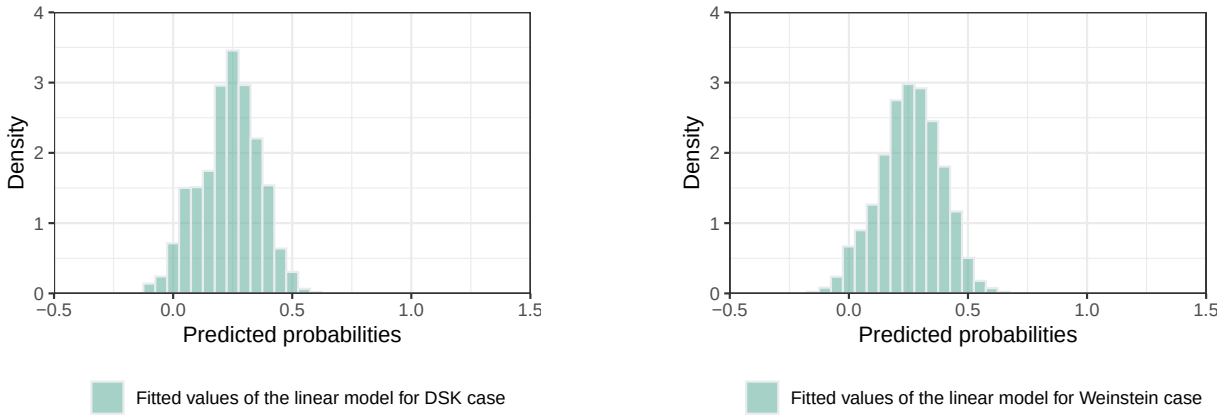
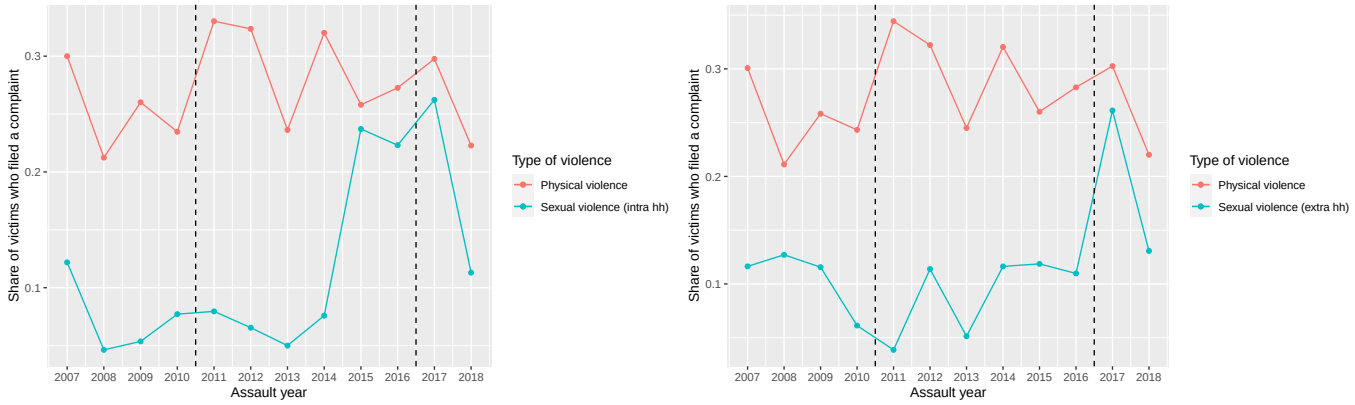


Table A7.4: DiD estimation restricting the treatment group – intra- or extra-household sexual violence

	Dependent variable: Filed a complaint (0/1)			
	Treatment: Intra-household violence		Treatment: Extra-household violence	
	(1)	(2)	(3)	(4)
Sexual violence	−0.248***	−0.219***	−0.166***	−0.193***
	(0.030)	(0.046)	(0.027)	(0.024)
Assault post 2011	0.079**		0.091**	
	(0.035)		(0.036)	
Sexual violence*Assault post 2011	−0.076		−0.144***	
	(0.067)		(0.047)	
Assault post 2017		−0.027		−0.034
		(0.028)		(0.029)
Sexual violence*Assault post 2017		0.145		0.146***
		(0.107)		(0.056)
Intercept	0.045	0.035	0.066	0.009
	(0.157)	(0.106)	(0.148)	(0.090)
Socio-demographic controls	Yes	Yes	Yes	Yes
Region fixed effects	Yes	Yes	Yes	Yes
Period	2007-2011	2012-2018	2007-2011	2012-2018
Observations	1,806	2,194	2,039	2,576
Adjusted R^2	0.057	0.058	0.062	0.076

Note: OLS estimates from LPM. Robust standard errors in parentheses *p<0.1; **p<0.05; ***p<0.01

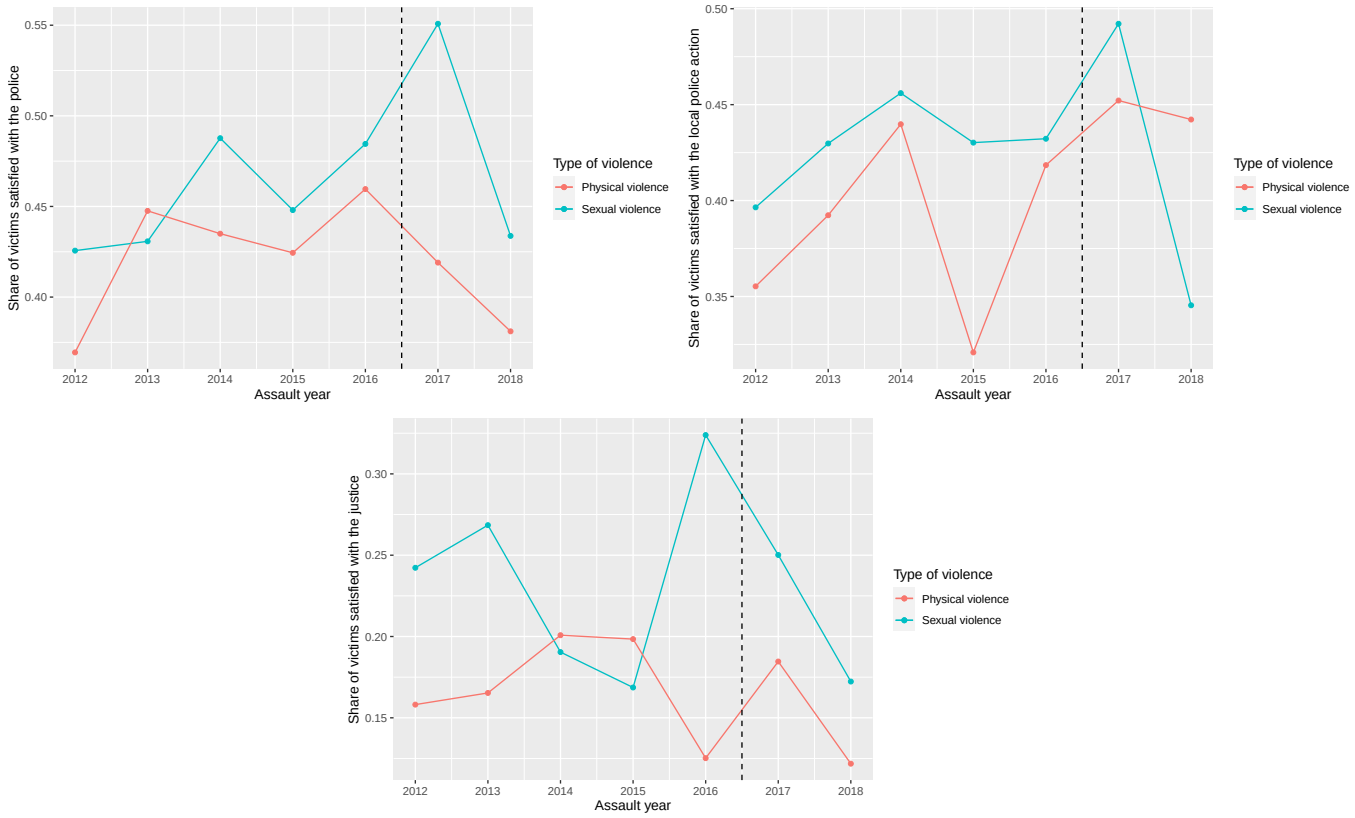
Figure A7.5: Pre-trends for filing a complaint when restricting the treatment group - raw data



Population: Victims of sexual violence (two subgroups) or physical violence in the two years preceding the survey. Weighted proportions.
Source: CVS surveys, 2007-2019.

A8 Potential mechanisms

Figure A8.1: Pre-trends for satisfaction with police and justice - raw data



Population: Victims of sexual violence or physical violence in the two years preceding the survey. Weighted proportions.
Source: CVS surveys, 2012-2019.