ESRC SDAI Application

1 Intro

Domestic abuse is a complex phenomenon affecting people from all walks of life. It is increasingly recognised as a major public policy concern in many countries, including the UK (?, ?). While anyone can become a victim of domestic abuse, women are disproportionately affected, with more than 25% of women, and 15% of men in England and Wales reported to have experienced some form of domestic abuse since the age of 16 (?, ?). The current legal definition in the UK (?, ?) aims to capture the multifaceted nature of domestic abuse, by recognising that domestic abuse encompasses a wide range of behaviours, including emotional, sexual, and physical abuse, threatening, intimidating, coercive and controlling behaviour.

Domestic abuse has substantial mental health implications, with an estimated three-quarter of survivors experiencing posttraumatic stress disorder symptoms, and a significantly higher likelihood of reporting symptoms of anxiety and depression, compared to the general population (?, ?). The long lasting impacts of domestic abuse are not limited to the direct target of abuse. Witnessing domestic abuse at home can have severe developmental impacts on children, including an increased risk of experiencing mental and physical health problems and encountering difficulties in interpersonal relationships in later life, worse educational attainment, and increased likelihood of engaging in criminal behaviours (?, ?).

As the legal definition reflects, one characteristic of domestic abuse that differentiates it from other violent crimes is its repeated nature. Estimates show that on average, survivors live in the abusive relationship for 2.7 years, experiencing an estimated 50 cases of abuse before getting effective help (?, ?). The most reliable statistics on domestic abuse in the UK is the Crime Survey for England and Wales (CSEW; ?, ?), a victimisation survey including a self-completion module on domestic abuse. According to the CSEW, of those respondents who experienced domestic abuse between April 2017 and March 2018, only 17% reported it to the police (?, ?). This extremely high level of underreporting is another characteristic that is specific to domestic abuse.

In the most extreme cases, domestic abuse can culminate in domestic homicide. In the period between April, 2017 and March, 2018, 70 people in England and Wales were killed by their current or former partner, 90% of these victims were women (?, ?), demonstrating that domestic abuse is a fundamentally gendered phenomenon. While the pervasive problem of underreporting poses a significant obstacle to deriving reliable estimates of the true extent of the problem, the economic cost of domestic abuse in England and Wales between April, 2016 and March, 2017 was estimated to be £66 billion (?, ?). The largest component of this cost is represented by the physical and emotional consequences of abuse, reflected in a reduced expected quality of life for survivors. In addition, lost economic output resulting from missed workdays and reduced productivity, as well as costs to the health care system also significantly contribute to the overall figure.

2 Data

The first dataset is provided by the West Midlands Police, and includes all recorded crimes and incidents from the period between January, 2010 and October, 2018. The West Midlands Police is the third largest police force in England, serving a population of 2.9 million *(ref)*. Crimes that have a domestic abuse marker indicate cases of domestic abuse that meet the criteria for notifiable offences in the UK, whereas domestic abuse incidents refer to cases that do not qualify as a crime. About 31% of all crimes and incidents have a domestic abuse marker. For each record in this dataset, we have information about the exact location and time of the incident or crime, the gender, age and ethnicity of the offender and victim, and the severity of the injury sustained by the victim, if any. The first and last occurrence of the offence, as well as the exact time of reporting is also recorded. Each person in this dataset has a unique person identifier, allowing us to follow people over time. *Sentence about access to this dataset.*

The second dataset is the Crime Survey for England and Wales (CSEW), which is an annual, cross-sectional representative survey collecting information on victims of crime across England and Wales, including 30,000–40,000 households every year *(ref)*. The survey has a self-completion module containing questions about the respondent's experiences of domestic abuse throughout their life (since the age of 16), and in the past 12 months in particular *(ref)*. Given the serious problem of underreporting uniquely characteristic to domestic abuse, the CSEW is the most reliable source of information on the prevalence of domestic abuse in England and Wales. since when?

The two datasets have different merits, and combining them allows us to gain a deeper understanding the characteristics and dynamics of domestic abuse, and deliver policy-relevant insights. The CSEW is more likely to give a better estimate of the true prevalence of domestic abuse, and contains rich demographic, socio-economic and geographic (LSOA level) information on the respondent (survivor). However, owing to the fact that it is a victimisation survey, it does not contain any information on the perpetrators and has limited longitudinal information. In contrast, the crime dataset has information (age, gender, ethnicity) on both the victim and the offender, and records all reported incidents for the same victim-offender pair, allowing us to conduct a quantitative analysis of the dynamics of abusive relationships and identify the predictors of escalation. Information on the exact timing of the abuse and the time of reporting will help us identify high-risk times of the year and understand the factors influencing the decision to report. Information on alcohol involvement is more likely to be accurate in the crime dataset than in the self-completion survey. Furthermore, the crime dataset provides us with information on other criminal behaviours of the offender and victim. We do not intend to link the two datasets.

Research Programme

Our research programme involves four research projects exploring different aspects of domestic abuse. In outlining each of these research projects, for the sake of clarity, we will specify a set of core regression analyses we plan to conduct. We will preregister all of our analysis plans using the Open Science Framework (OSF) ahead of beginning research. We will also aim to complement all analyses with robustness checks and further exploratory regressions as the data allows (see, *Trendl Football*, for an example of this approach).

1. Understanding domestic abuse: survivor characteristics and police mis-recording

Short intro about aims, highlighting added value. In this research project, we will first use the CSEW to investigate the demographic and socio-economic characteristics of survivors of domestic abuse. Exploring the risk factors in domestic abuse victimisation is key in designing effectively targeted policy measures. We will also use the WMP data to explore the extent of police mis-recording of domestic abuse. Increasing victim's trust in the police through improving law enforcement response to reported incidents is key in encouraging victims to report domestic abuse and prevent further harm.

Detailed plan First, using the CSEW data, we will extend our understanding of the demographic and socio-economic factors predicting domestic abuse victimisation. The CSEW provides a broad range of information on the characteristics of the respondent, including age, sex, marital status, number of children, ethnicity, education, employment, income, benefit history, physical and mental health, frequency of going out, house and car ownership, self-reported well-being, and frequency of drug and alcohol use. Previous research has identified a number of risk factors predicting domestic abuse victimisation (being female, young, unemployed, separated or divorced, living in single parent household and earning in the lowest income bracket), based on descriptive statistics derived from the CSEW data. Building on what we know so far, we will explore the predictive power of a broader range of victim characteristics in statistical model, using a logistic regression approach. Investigating the explanatory power of these characteristics within one statistical model will provide us with a deeper understanding of the importance of these factors in predicting domestic abuse victimisation, compared to a purely descriptive statistics approach. We are also interested in how the predictive power of victim characteristics vary with the type of abuse suffered (physical abuse, threats, sexual abuse).

However, given the complex nature of domestic abuse, it is likely that there are important interactions between various victim characteristics in predicting the risk of domestic abuse victimisation (for example, we would expect that not having stable employment is a more significant risk factor of domestic abuse victimisation for those who have young children). To uncover the structure of these interactions (without having to a priori specify them, as is required in a logistic regression model), we will also use a random forest classification algorithm, a machine learning method that has the ability to detect non-linear relationships between variables, allowing us to identify particularly vulnerable subgroups of the population. Insights from both methods will provide us with the first extensive characterisation of the risk factors predicting domestic abuse victimisation in England, using a national-level, large representative sample like the CSEW.

Police forces in England, and particularly the West Midlands Police, have previously been criticised for mis-recording violent cases of violent crime, including domestic abuse. Mis-recording these already heavily underreported crimes decrease victim's trust in the Criminal Justice System and their willingness to report subsequent incidents. *Why don't we use the aggregate number of police-recorded domestic abuse cases for each police force to do this? But if we don't particularly rely on the WMP data is there a point in this? They use CSEW and police data in this report, but they don't seem to compare the two*

Core outcomes

The first extensive investigation of the demographic and socio-economic predictors of domestic abuse victimisation.

Table 1: Understanding domestic abuse: survivor characteristics and police mis-recording, regressions

Research question	Dataset	Unit of analysis	Outcome variables	Explanatory variables	Model
What are the characteristics of domestic abuse victims?	CSEW	Individual (respondent-level)	Domestic abuse victimisation; Victimisation by type (physical, sexual abuse, threats)	demographic and socio-economic characteristics	Logistic regression/random forest
What is the extent of police mis-recording of domestic abuse?	ONS data				

2. Predicting serious harm and understanding the decision to report

Short intro about aims, highlighting added value. In this project, we will explore the victim and offender characteristics that are predict serious harm and the factors determining the decision to report the abuse. Identifying high-risk victim-offender pairs is key in preventing serious harm. Understanding what affects the decision to report is particularly important since underreporting represents the biggest obstacle to effectively tackling domestic abuse. Insights from these analyses can inform decisions about the optimal timing and target audience for domestic abuse awareness campaigns.

Detailed plan Drawing on the relative strengths of the CSEW and the WMP datasets, we will explore the relative importance of the risk factors predicting serious harm. Previous studies investigating this question mostly relied on police data and focused on identifying high-risk offenders. We will complement and extend this research by simultaneously considering the victim characteristics that are predictive of suffering serious injuries using the CSEW. Specifically, using a multinomial logistic approach, we will explore to which financial independence and resilience (income bracket), mental health (self-reported well-being), drug and alcohol dependence (frequency of usage), social isolation (not going out, living alone or with the perpetrator), recent separation, and feeling frightened (increased home security, reporting feeling scared) are predictors of various levels of harm (threats, minor or serious injury) *causality issue?*. We will know the level of harm suffered, because the CSEW asks whether the respondent needed to see a doctor, nurse or other health worker, or needed to take time off work because of the abuse.

Using the WMP data, we will also explore the risk factors associated with the perpetrator, using each perpetrator-victim pair as a unit of analysis. The strength of the WMP data is that it has a temporal dimension, and we can follow the same perpetrator-victim pair over time. Our detailed crime dataset will allow to investigate the extent to which previous recorded self-harm and suicide attempts, previous violent and drug offences, breaching court orders and separation will predict causing serious harm. From the detailed description of the incident, we will know if the perpetrator and the victim live together. We will explore this question using a multinomial logistic regression analysis with perpetrator-victim fixed effects.

We will also use both datasets to explore the predictors of the decision to report the abuse. Using the CSEW, and a logistic regression approach to understand how the decision to report to the police (or the decision to seek any form of external help) depends on the individual-level characteristics of the survivor , controlling for the severity and type of the abuse suffered.

Second, we can use the WMP data to investigate drivers of the decision to report. For this analysis, we will focus on the subset of incidents (6% of all incidents), where the victim reported a long history of abuse (at least 3 days between first and last reported occurrence). We are interested in whether these reports are more likely to occur after certain days of the year (e.g., birthday of the victim or perpetrator, Christmas, Halloween, Easter, etc.), potentially reflecting cases where the abuse has suddenly escalated, or became known to outsiders.

Core outcomes

Table 2: Predicting serious harm and understanding the decision to report, regressions

Research question	Dataset	Unit of analysis	Outcome variables	Explanatory variables	Model
What victim characteristics predict serious harm?	CSEW	Individual (respondent-level)	Level of harm resulting from domestic abuse victimisation (no physical harm, minor, serious)	financial resources, mental health, dependence, social isolation, separation, feeling scared	Multinomial logistic regression
What perpetrator characteristics predict serious harm?	WMP	Perpetrator-victim pairs	Violent offences resulting in injury	previous self-harm and suicide attempts, previous violent and drug offences, breaching court orders, separation	Multinomial logistic regression with perpetrator-victim fixed effects

3. The long-lasting effects of domestic abuse

it's impossible to establish causality

Short intro about aims, highlighting added value. Domestic abuse has long-lasting adverse effects on victims and those close to them. In this research project, our is aim to explore some of these consequences on the direct victims of domestic abuse and the children who live in the same household. Gaining a deeper understanding of the tangible, far-reaching consequences of domestic abuse will help to design survivor support programmes and quantify the societal harm caused by it.

What we know so far Women's aid, effect on victims increased likelihood of mental health problems (isolation, low self-esteem), alcohol and drug dependency, financial consequences due to financial abuse or losing a job as a result of the abuse, fighting back?

Effect on children I witnessing violence is major predictor for PTSD and depressive symptoms in children; behavioural problems, bullying (either victims or perpetrators)

Effect on children II cognitive development is affected, learning difficulties are more common, difficulties in social relationships, withdrawal, antisocial behaviours, they are more likely to become both perpetrators and victims of domestic abuse; protective factors: literacy, intelligence, social competence, supportive relationship with at least one influential adult

Detailed plan

Aim 1. First, using the CSEW, we will explore the long-lasting effects of experiencing domestic abuse. In the CSEW, in two separate questions, respondents are asked whether they have experienced any form of domestic abuse (use of violence, sexual abuse, threatening behaviour, stalking) since the age of 16 or in the last 12 months. This way we can identify respondents who have suffered specific forms of abuse in the past, but not in the last 12 months. Using propensity score matching, we can then compare their physical and mental health-related outcomes, and alcohol and drug consumption with those, who have never experienced any form of abuse, but otherwise have similar socio-economic characteristics. This will allow us to estimate the effect of past abuse on present outcomes whilst controlling for individual- and macro-level characteristics.

Aim 2. Using data from the 10-15 year old questionnaire of the CSEW, which can be linked to the adult questionnaire, we can examine the effect of domestic abuse on children living in such households. The questionnaire records information on whether the child had been physically or verbally abused by someone (and if that person lives in the household), as well as their experiences with bullying, carrying knives and

information on gang membership, school truancy and health. We can explore the effect of domestic and child abuse on these outcomes through propensity score matching.

Domestic abuse often co-occurs with child abuse. We can use the WMP data to estimate this co-morbidity, and identify under 18s who had been a survivor of a domestic abuse offender. No control group.

Core outcomes

Environmental factors

Short intro about aims, highlighting added value. In this research project, we will be using the CSEW and WMP datasets, as well as external sources of data to investigate some of the environmental predictors of domestic abuse. First, using data from the CSEW and drawing on previous research on the topic, we will be looking at the association between neighbourhood-level characteristics, and the prevalence of domestic abuse as well the decision to report. We will complement this with an additional analysis on WMP and census data. We will be able to account for the dynamic nature of these neighbourhood characteristics, by using longitudinal and repeated cross-sectional datasets. Understanding how neighbourhood-level characteristics affect the prevalence of domestic abuse and reporting behaviour is crucial for the most effective distribution of resources to help victims and targeted awareness campaigns.

Second, we will be looking at how time-varying, exogenous factors affect domestic abuse. We will first explore the effect of local events that increase alcohol-consumption, since our previous study have found a significant relationship between alcohol-related domestic abuse and football. Second, we will conduct two studies to understand the relationship between financial stress and the prevalence of domestic abuse. First, we will explore the association between gambling accessibility in the local area, and reported numbers of domestic abuse. Second, we will examine how the roll-out of the Universal Credit system affected levels of domestic abuse across the West Midlands. Apart from delivering important policy-relevant insights, these investigations will also aim to inform the local police about the optimal allocation of police resources.

What we know so far interestingly, there are no studies about this using UK data; neighbourhood-level predictors of domestic abuse: social disorganization theory (socio-economic disadvantage and residential instability disrupt social bonds a limit collective ability to maintain control and recognise IPV); cultural norms through social learning process (high IPV neighbourhoods where people observe influential others being rewarded for engaging in IPV, whereas in low IPV neighbourhoods perpetrators get socially ostracized); women's empowerment reduces IPV, especially financial independence; abuser's lack of employment and alcohol consumption can be a risk factor; alcohol outlet density; churches, playgrounds, community networks might foster community cohesion and reduces IPV; however, in deprived neighbourhoods, where IPV is accepted social interactions can increase IPV

The relationship between electronic gaming machine accessibility and police-recorded domestic violence: A spatio-temporal analysis of 654 postcodes in Victoria, Australia, 2005âĂŞ2014

Economic Stress and Domestic Violence: Examining the Impact of Mortgage Foreclosures on Incidents Reported to the Police

The Impact of Neighborhoods on Intimate Partner Violence and Victimization

the type sof social tie matters: friendship ties reduce the risk of IPV, family ties have no effect

Neighborhood Environment and Intimate Partner Violence: A Systematic Review

Detailed plan

Aim 1. First, we will use the CSEW to investigate if there are significant associations between the prevalence of domestic abuse, the willingness to report, and the environmental characteristics of the neighbourhood they live in. Previous literature has suggested various pathways of how neighbourhood characteristics affect domestic abuse, but the question has not been explored in a UK context yet. Our rich data allows us to determine the relative explanatory power of these. For example, social disorganization theory posits that neighbourhood disadvantage is associated with an increased risk of domestic abuse through reduced collective efficacy. Another potential pathway is social norms, suggesting that acceptance and normalisation of violence within a community can encourage domestic abuse. Using a logistic regression, we will explore how the interviewer's perception of the street (signs of rubbish, vandalism, and the general condition of houses, *member of Neighbourhood Watch?*) as well as the respondent's connection and attitude towards the neighbourhood (length of time living in the local area, noisy neighbours, rubbish lying around, teenagers hanging around, vandalism, drunken and anti-social behaviour, drug trafficking, abandoned cars, speeding traffic, police presence in the area, worries about crime levels) predict domestic abuse victimisation and trust in CJS, reflected in the propensity to report abuse. Second, using the WMP and the census, we can compliment this with an LSOA-level spatial regression, where we use the socio-economic characteristics (average income, education levels and benefit dependency) of the area to predict the number of reported abuse cases. Using a spatial regression will account for the spatial

dependency between our observations and improve the reliability of our estimates. *Not sure whether we can do the same with the CSEW, probably not on the LSOA level, but maybe on the MSOA level?* This analysis, contrasted with that outlined in project 1 about the individual-level predictors of domestic abuse, will provide us with the most comprehensive understanding on the risk factors of domestic abuse within the context of the UK.

Aim 2. In this section, we will be looking at time-varying environmental factors that may influence domestic abuse. This part of the project will benefit from a dataset containing credit card and current account spending of x customers across the UK, to which we will have access as part of a data sharing agreement with a very large UK bank. *I'm not sure about the specifics of this, would we combine this with the WMP or CSEW? I guess we won't know the card owner's address?* First, by using the WMP and credit card data, we will identify time-specific changes in alcohol consumption in parts of the West Midlands by identifying externals events that may affect it (e.g., local festivals, weather, bank holidays, sport tournaments), and investigate their effect on alcohol-related domestic abuse. Previously, we have found a 60% increase in alcohol-related domestic abuse when the England national football team won, highlighting the profound effect exogenous events can have on the propensity for violence.

The recent reforms of the UK benefit system, in the form of the introduction of the Universal Credit (UC) had been criticised widely, due to the temporal financial strain it imposes on the least financially resilient people in society. Drawing on previous findings about the link between financial stress and domestic abuse, we will explore what effect the roll-out of the UC across the West Midlands had on the reported number of domestic abuse cases. The fact that the exact date of the roll-out varied across the seven metropolitan boroughs within the county allows for a more precise estimation of the effect of UC on the reported number of domestic abuse incidents in the West Midlands. For this estimation, we will use data on the proportion of benefits claimants in each LSOA. *I know we were talking about using the credit card data for this one, but I am not sure what was the exact plan*. Gambling is increasingly recognised as a serious health concern across the UK, and can have an adverse impact on family finances. To assess the link between gambling accessibility and the prevalence of domestic abuse, we will use data on the number of licensed betting shops in the West Midlands within a spatial regression approach.

List of licensed gambling businesses We can also contact the local authorities, as they should have an up-to-date list

Core outcomes The first extensive analysis of the complex individual– and neighbourhood level predictors of domestic abuse within the UK. An exploration of the exogenous, time-varying factors affecting levels of reported domestic abuse (mostly through alcohol). An examination of how gambling accessibility and temporal financial stress affect the reported number of domestic abuse cases in the West Midlands.