Analysis of Crime in England and Wales in 2013/14

[redacted]

Executive Summary

Our analysis has been conducted on a subset of results from the Crime Survey for England and Wales, which collects data on respondent demographics, perceptions of crime, experience of crime and levels of antisocial behavior. Up-to-date analysis of crime data is pivotal for central policy changes and accurate budget allocation to law enforcement.

From our analysis, there seemed to be skewed perceptions of crime depending on respondent's sex. On average, we found women feel more afraid walking alone than men. Both sexes feel more worried about walking alone during the night than during the day, but the magnitude of worry appeared to increase more in women. Furthemore, women seemed to be significantly more worried about being a victim of a crime. Despite the imbalance of perceptions between sexes, both had similar proportions of victims and non-victims. This could be an effect of women reporting crimes less or be a result of the nature of the crimes they were subject to. Due to lack of data on the nature of crimes, we could not glean any more from this.

Country of residency seemed to have insignificant effect on perceptions of crime, as the levels of worry in both areas are of insignificant difference. Rural residency, however, did appear to report less crime than the national average. The level of deprivation in respondent's residency appeared to affect perceptions of crime, with people in more deprived areas feeling safer walking outside during both the day and night.

We found that younger respondents are of increased likelihood to be the victim of a crime and those aged over 75 proved approximately 4 times less likely to be a victim of a crime than those aged 16 to 24.

Lastly, we attempted to predict a respondent's probability of being a victim of a crime, based largely off demographic factors. We have been able to correctly identify victims and non-victims of crime 6 times out of 10.

Introduction

The British Crime Survey, as it was formerly known, was initially conducted in 1982[Home Office, 2012] and now is known as the Crime Survey for England and Wales (CSEW) as residents in Scotland are studied in a seperate survey. It is estimated that only 4 in 10 crimes are reported to the police[ONS, 2021], indicating the true picture of the prevalence of crime in the UK is not clear. Our study attempts to provide a more accurate estimate of the level of crime by surveying private residents across regions of England and Wales. The survey has a consistent methodology and so is unaffected by changes in reporting or recording practices that may happen in police departments.

It is important for the government to have an idea of crime levels to be able to evaluate future policies and assess specific social groups that could be subject to increased risk of certain crimes. Studies show crime had fallen from 110 per 1,000 people in 2002/03 to 71.2 per 1,000 people at the time of the sample we are looking at from 2013/14[Statista Research Department, 2021], although crime levels have been rising since then, possibly due to austerity measures. Crime bears large economic and social costs on governments. The UK Home Office estimates that the total cost of crime in the UK in 2015/16 was £50bn[Heeks et. al., 2018], with violent crimes making up the largest proportion of the total costs.

This report looks at a randomly selected subset of 8,843 responses to the 2013/14 CSEW survey, which is a quarter of the original CSEW data set that contained 35,371 observations. We will attempt to examine patterns in crime occurance, determine any relationships between demographics, social perception and degree of worry regarding being a victim of a crime. Lastly, we will explore whether predictive modelling can be used to determine the probability of a respondent being a victim of a crime based on their demographics.

Methodology

The cross-sectional data set contains 8,843 observations across 33 variables. [Buckley et. al., 2018]

Definition of variables

For a given observation i, the variables can be split into 4 categories, and we define them as follows:

- *Demographics of the respondent*
  Age group (A_i), sex (S_i), ethnic group, education level (Ed_i), deprivation quintile (split by England and Wales: D_i^E and D_i^W), marital status, years lived in the area, any paid work in past week (Wk_i), accomodation tenure type, area type (Ar_i).
- *Experience & perception of crime*
  Whether they have experienced a crime in the past year (V_i), what they consider to be the main cause of crime, how worried they feel about: walking alone during the day (W_i^D) and after dark (W_i^N), being alone at home at night, having their car stolen, being mugged, being raped, being physically attacked, being victim of a crime, etc.
- *Anti-social behaviour in their neighbourhood*
  How common in their immediate area are: litter and rubbish (Rb_i), vandalism and graffiti (Vd_i), run-down or poor homes (Po_i), and anti-social behaviour (As_i).
- *Metadata*
  Variables that relate directly to the data set: row label, index of the split (Sp_i), and the relative weight of the observation.

Prior to the preliminary data analysis, we have analysed the missing variables in the data. There is a substantial amount of missingness, which can be classified into two categories:

- *Systematically missing*
  These are the responses missing because a question has not been asked, e.g. only 2,074 people

(23%) were asked what they consider to be the main cause of crime.

- *Refused by the respondent*
  These occur when the respondent refuses to answer a question, e.g. 2 people (0.02%) refused to answer whether they have undertaken paid work in the past week.

The survey questions are split into 4 categories ($A-D$), and each respondent was only asked quesitons from that category. Some questions were asked in all categories, and some only in one category. This systematic missingness prevents us from using the variables from different disjoint categories in the same linear model, since the intersection of the data where the variables are not missing becomes empty.

Analysis

We immediately merge the variables D_i^E and D_i^W, since every observation has at least one of them not missing, and create the variable R_i representing the country of residence.

We aim to determine the factors that are associated with an increased risk of being involved in crime, and ultimately we aim to be able to predict the probability of being a victim of crime based on the given variables.

First, we would like to paint a picture of the types of people who are more likely to be victims of crime in the UK. For this we estimate the percentage of national crime occurance adjusted for demographic variables such as sex, age, ethnicity, etc. For this we calculate pair-wise proportion tables based on the victim variable (V_i). We do this both with the weights and without.

Then, we attempt to explore the main drivers of fear of walking alone during the day and during the night. We use proportion tables, box plots and bar plots to summarise the key differences in worrying about walking alone during the day (and night) between the deprivation quintile and the level of anti-social behaviour of the interviewee's neighbourhood. To statistically quantify the differences in the level of worry, we fit the following linear models:

$$W_i^D = \beta_0 + \beta_1 D_i + \beta_2 R_i + \varepsilon_i \quad (1)$$
$$W_i^N = \beta_0 + \beta_1 D_i + \beta_3 R_i + \varepsilon_i \quad (2)$$
$$As_i = \beta_0 + \beta_1 D_i + \beta_3 R_i + \varepsilon_i \quad (3)$$

where W_i^D and W_i^N are the integer variables quantifying the levels of worry about walking alone during the day and during the night, taking integer values from 1 to 4 where 1 represents *"Very safe"* and 4 represents *"Very unsafe"*. D_i is the numerical version of deprivation quintile (1 being 1-20% and 5 being 81-100%).

To determine gender-specific associations between worrying about being a victim of crime and the experience of being a victim of crime, we use $t-$tests comparing the mean worry levels of being a victim across male and female respondents.

To further examine the main drivers of worry, we use $t-$tests to compare the mean levels of worry about walking during the day and during the night across females and males, and also how the levels of worry change between walking alone during the day and during the night for each sex. For this, we use the numeric versions of the variables W_i^D and W_i^N.

We also determine whether one sex is more likely to be a victim of crime. For this, we use a $t-$test to compare the overall unweighted mean percentage of victims for male and female respondents.

Next, we aim to find a method for predicting whether a given person has been a victim of a crime in the past year based on their demographic and social characteristics. For this, we use a Binomial GLM with the logit link, where the respondent's victim status (V_i) is the response variable. We first split the data into a *training set* ($\approx$ 70% observations) and a *testing set* ($\approx$ 30% observations). Then we use the training set to build and estimate our GLM.

A key consideration here is that any of the variables relating to the experience of crime, that is those concerning ratings of how worried respondents are of certain forms of crime, are likely to be subject to reverse causality. In the current context, this means that the respondents who were a victim of a crime are more likely to report higher levels of worry. We omit these from modelling along with those variables with a high volume of missingness at random. Then, due to a relatively low number of total relevant variables (15), we are able to programmatically fit every possible GLM up to interaction terms (2^{15} = 32,768 models in total) and then rank them by lowest AIC values. For each of the 10 best models by AIC, we calculate: the cut-off probability for predicting a positive result and the brier score. Then, we use the 10 estimated models to predict the values in the testing set, and for each we calculate: sensitivity, specificity, misclassification rate, and Cohen's Kappa. The resulting output is shown in ***Table*** 2.

To make the GLM as generalisable as possible, we do not make any assumptions about the respondent's subjective level of worry about being a victim of crime. This way we are also able to avoid the variables with very high systematic missingness, which means we are able to make use of the most of the training set.

The Binomial GLM with the lowest AIC value was determined to be:

$$\begin{aligned} V_i &\sim Bernoulli(\mu), \\ logit(\mu) &= \beta_0 + \alpha_j^A + \alpha_k^{Ed} + \alpha_l^{Ar} + \\ &\quad + \alpha_m^{Rb} + \alpha_n^{Po} + \alpha_p^D, \end{aligned} \quad (4)$$

where $j \in \{16\text{-}24, \ldots, 75+\}$, $k \in \{\text{None}, \ldots, \text{Other}\}$,
$l \in \{\text{Urban}, \text{Rural}\}$, $p \in \{1\text{-}20\%, \ldots, 81\text{-}100\%\}$,
$m, n \in \{\text{Very common}, \ldots, \text{Not at all common}\}$.

The first level of each of the categorical variables is the baseline, and is set to 0. An assumption of this model is that the selected variables are independent of each other, which might not be the case as age may have an effect of education level.

The estimated GLM was then assessed by its diagnostics and performance when applied on the test set. The binned residual plot indicated no problems: the residuals are randomly scattered around 0, and there is no sign of heteroskedasticity. So we chose to proceed with this model.

The final model is able to make use of 6,121 observations. Variables relating to the behavioural questions in the survey are missing systematically, but we didn't use those variables in the GLM for prediction. Of those variables that we used for prediction, the missing observations were completely at random, which meant that the result is likely unbiased.

Statistical methods

For the null-hypothesis tests, p-values are reported. An effect is deemed statistically significant if $p < 0.05$. Confidence intervals (CI) are reported at a 95% confidence level.

When training a Binomial GLM model, we choose the cut-off for predicting a positive value as the one which maximises sensitivity and specificity when the model is applied on the training set. This is the value at which sensitivity and specificity are equal.

Brier score was utilised to assess the model accuracy when fitted on the training data set. It is defined as $BS = \frac{1}{n}\sum_1^n (y_i - \hat{y}_i)^2$, where $y_i \in \{0,1\}$ are the true values and $\hat{y}_i$ are the predicted probabilities.

For evaluating predictions of the Binomial GLM model on the test data, we use misclassification rate (MR) defined as $MR := \frac{1}{n}\sum_1^n \mathbb{I}[\hat{y}_i \neq y_i]$, where $\mathbb{I}$ is an indicator function. We also use Cohen's kappa, which is a measure of how much better than pure chance the GLM is at predicting outcomes, and is defined as $\kappa = 1 - \frac{1-p_0}{1-p_e}$ where p_0 is the relative observed agreement between the predicted and the true values, and p_e is the probability of chance agreement.

The analysis was performed in R version 3.6.3. Unless stated otherwise, throughtout this report we use the `survey` package in R to create a sample based on the weights given in the data. The weights account for an unequal selection probabilities of each observation, and therefore help to reduce the selection bias.

Results

Out of the 8,843 respondents in the data set, 1,383 (15.64%, CI [14.88, 16.4]) have reported being a victim of crime in the past 12 months. The corresponding weighted mean percentage is **16.95%** (CI [15.99, 17.91]). The weigthed percentage is higher, suggesting the crime occurance may be underrepresented in our data, but the CI includes the unweighted mean, so the difference is not statistically significant.

Out of the 16-24 year olds, 24.24% (CI [20.55, 27.92]) were victims of crime in the past year. The CI does not include the national average, so this age group is statistically more likely to experience crime. The unweighted mean is very similar (23.65%, CI [20.47, 26.83]). This is summarised in ***Table*** 1.

Older age groups are progressively less likely to experience crime, and the difference from the national average is statistically significant. The weighted mean percentages are 14.09% (CI [12.04, 16.15]) for 55-64 year olds, 9.81% (CI [7.99, 11.63]) for 65-74 year olds, and 6.52% (CI [4.81, 8.22]).

None of the ethnic groups have a statistically significant difference in weighted percentage mean from the national average. Mixed race, Asian and Chinese respondents have numerically higher average percentages: 28.42% (CI [16.42, 40.41]), 20.57% (CI [15.66, 25.48]) and 20.78% (CI [11.31, 30.25]) respectively. The table for this is omitted for brevity.

Respondents who have a degree are statistically significantly more likely to be victim of crime then the national average (18.59%, CI [16.99, 20.2]). On the other hand, respondents who have no education or list their education as *Other* are less likely to be a victim, with averages 11.89% (CI [10.04, 13.75]) and 12.11% (CI [8.2, 16.01]) respectively.

Respondents living in rural areas are statistically significantly less likely to be victims of crime (12.41%, CI [10.74, 14.08]) than the national average.

In England, respondents living in the bottom deprivation quintile (1-20%) are more likely to be a victim of crime (20.32%, CI [17.87, 22.76]) than average, and people in the 61-80% and 81-100% quintiles are less likely to be victims: 14.75% (CI [12.74, 16.76]) and 13.43% (CI [11.45, 15.41]) respectively. There was not enough data for Wales to establish a relationship.

We found that the higher the deprivation quintile, the lower the level of worry about walking outside during the day and the lower the perceived anti-social behaviour level of the respondent's neighbourhood. 85.6% of the respondents in the top deprivation quintile report feeling very safe walking outside, compared to 61.4% in the lowest quintile. This is seen in ***Figure*** 1. Quantified statistically, an increase in deprivation quintile by 1 level (which is 20%) decreases the level of worry by 0.07 levels on average ((1): $\hat{\beta}_1 = -0.07$, CI $[-0.09, -0.05]$, $p<0.001$). This association holds for England, but there is not enough evidence for Wales.

Each increase in deprivation quintile decreases level of worry about walking alone during the night by 0.14 levels on average ((2): $\hat{\beta}_1 = -0.14$, CI $[-0.17, -0.11]$, $p<0.001$), and the perceived level of antisocial behaviour by 0.21 levels ((3): $\hat{\beta}_1 = -0.21$, CI $[-0.24, -0.18]$, $p<0.001$). These associations hold for both England and Wales.

There is no evidence that the respondents in England have a higher level of worry than those in Wales about walking alone during the day (means: 1.28 vs. 1.25, p=0.506) and during the night (means: 2.12 vs. 2.09, p=0.711). Numerically, 77.1% of respondents from the lowest quintile in Wales report feeling very safe walking alone during the day compared to 60.16% in England.

We found that females have a higher level of worry about being victims of crime on average than males (means: 0.18 vs. −0.27, $p<0.001$). This is also visible in ***Figure*** 2.

Females on average have a higher level of worry about walking alone during the day than males (means: 1.34 vs. 1.19, $p<0.001$). This is also true and even more pronounced for walking during the night (means: 2.40 vs. 1.72, $p<0.001$).

Both sexes feel more worried about walking alone during the night compared to during the day: males (means: 1.72 vs. 1.20, $p<0.001$), but the difference for females are more pronounced (means: 2.40 vs. 1.35, $p<0.001$). This is visualised in ***Figure*** 3.

We found no evidence that female respondents are more likely to have been victims of crime in the past 12 months than males (means: 15% vs. 16%, $p=0.226$). This is seen in ***Figure*** 5.

By estimating (4), we found that age groups 55-64, 65-74 and 75+ are less likely to be victims of crime than the age of group 16-24. Their respective odds ratios are: 0.62 (CI [0.46, 0.83], $p<0.001$), 0.43 (CI [0.31, 0.59], $p<0.001$) and 0.28 (CI [0.19, 0.40], $p<0.001$). This suggests that older age groups are progressively less likely to experience crime than the 16-24 age group.

We also found that the respondents who are educated up to A-level and degree level are more likely to experience crime than those with no education. Their odds ratios are 1.40 (CI [1.10, 1.80], $p=0.007$) and 1.27 (CI [1.01, 1.60], $p=0.042$). This suggests that the odds of being a victim of crime are 40% higher if a respondent is educated up to A-level compared to having no education. Similarly, respondents with a degree-level eduaction have 27% higher odds of being a victim than those with no education.

The respondents who live in rural areas have 19% (CI [0.67, 0.97], $p=0.025$) lower odds of being a victim of crime than those living in urban areas.

Those who live in the top two deprivation quintiles are also less likely to be affected by crime. Those who live in the 61-80% quintile have 24% (CI [0.60, 0.97], $p=0.026$) lower odds of being a victim, and those living in the 81-100% quintile have 27% (CI [0.57, 0.94], $p=0.015$) lower odds.

Overall, the model performs fairly. As seen in ***Table*** 2, its sensitivity and specificity are 0.607 and 0.600, giving a misclassification rate of 0.399. This suggests that the model is able to correctly identify true victims with probability 60.7%, and will correctly identify those who aren't a victim with probability 60.0%, giving a total misclassification rate of 39.9%.

Since the prevalence of victims in the population is around 16.95%, a 39.9% misclassification will result in a large number of false positives in the population. There is a trade-off between sensitivity and specificity of a model, so depending on the context, having a larger number of false-positives or false-negatives may be less undesirable.

All of the models in ***Table*** 2 have nearly identical performance in terms of prediction, which might suggest that a 60% sensitivity and specificity is the best we can hope for, and there is no more juice left to be squeezed out of the data set.

Conclusion

We first sought to depict the typical victim of a crime in the UK, and then to predict whether a given person will be a victim of a crime based on their demographics and their socio-economical characteristics.

Unsurprisingly, we found that people living in wealthier neighbourhoods are less likely to experience crime, especially those living in the top 40% quintile of deprivation. Similarly, so do the people who report not having many run-down houses in their immediate neighbourhood.

Younger people and people who live in urban areas are more likely to be victims of crime. Even though urban areas tend to have younger populations[Champion, 2014], the associations hold, even adjusting for age and area type: younger people are more likely to be victims of crime even in rural areas.

Despite female respondents worrying more about being affected by crime, we did not find evidence within the data that they are at a greater risk of being a victim.

Our logistic model suggested some respondent groups vary significantly in probability of being a victim of crime, namely: age group, level of deprivation, scarcity of poor housing, level of education and area type (urban or rural). The moderate sensitivity and specificity of our model is perhaps somewhat expected, we would not necessarily expect to be able to predict, with high accuracy, those susceptible to a generic crime based predominantly on demographic and background factors. Nonetheless, we are able to predict victims with a 60% accuracy.

If we had the nature of the specific crimes the respondents were victims of, we may perhaps be able to identify respondents of certain demographic variables to be more susceptible to certain crimes and hence our model would have increased accuracy. For example, females are perhaps more likely to be victims of sexual assault crimes, with one study, conducted in the United States, finding nearly 1 in 5 women, compared to 1 in 71 men, having been raped at some time in their lives[Black et. al., 2011]. Hence, predictions for models based on demographic variables for sexual assault related crimes could thus be of increased sensitivity and specificity.

Due to the data being cross-sectional in nature we are unable to comment on trends in the rates or perceptions of crime over time. Furthermore, without the nature of the crimes victims were the subject of, this analysis may not be specific enough for some real-world applications. Nonetheless, the analyses performed can provide an overview of the perceptions of crime and rates of victimisation in specific societal groups in England and Wales.

* * *

Figures

Deprivation quintile and fear of walking alone

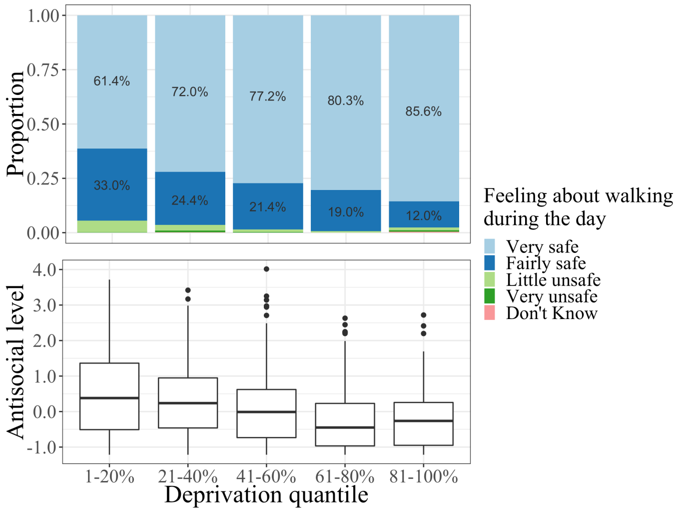

Figure 1: *(weighted) The higher the deprivation quintile, the safer people feel walking alone during the day, and the lower the perceived level of antisocial behaviour in the neighbourhood.*

Sex and level of worry about being a victim of crime

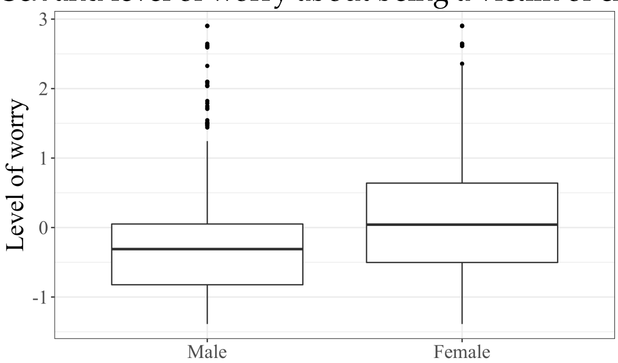

Figure 2: *Female respondents have a higher level of worry of being a victim of a crime on average than male respondents.*

Sex and feeling about walking alone

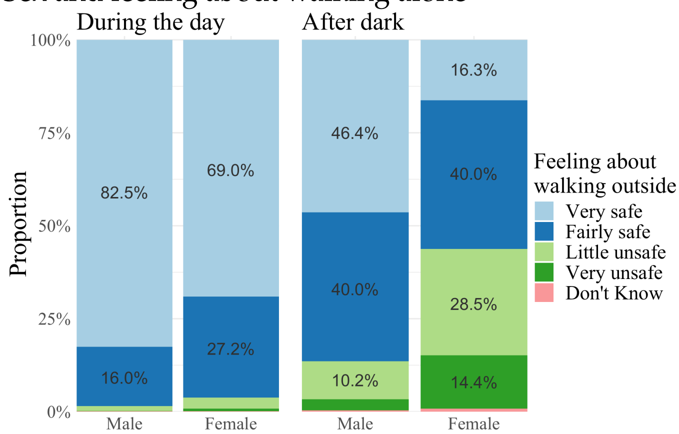

Figure 3: *(weighted) Both sexes feel more worried about walking alone during the night than during the day, but the difference among female respondents is more pronounced.*

Sex and mean percentage victim status

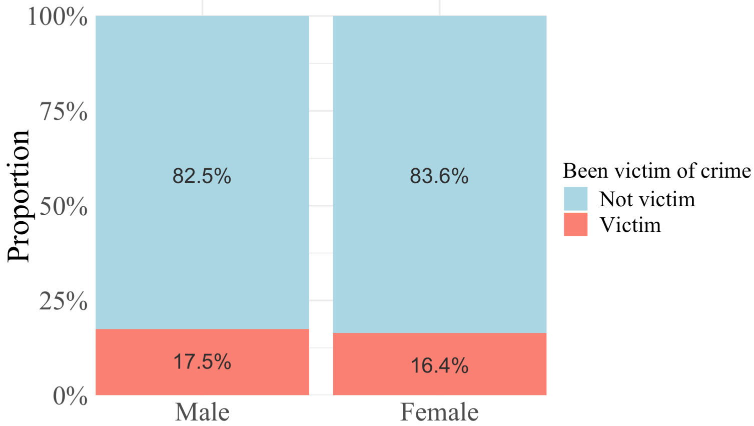

Figure 4: *(weighted) Both sexes have have similar mean percentages of victims.*

Forest plot of the Odds Ratios for the victim model (4)

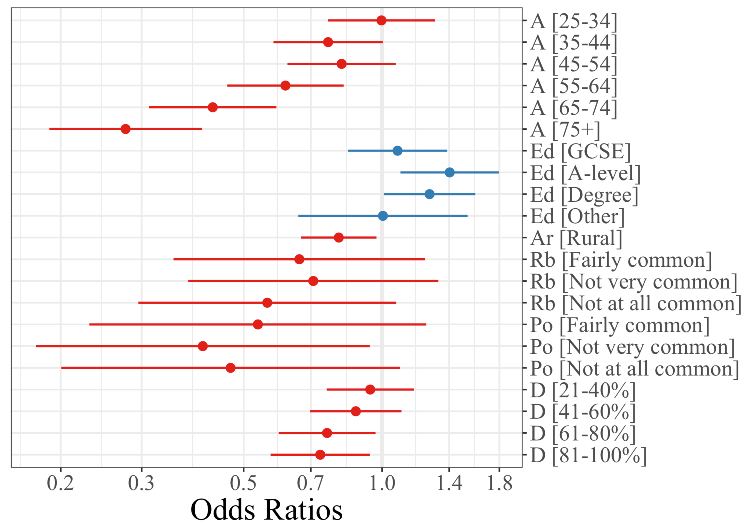

Figure 5: *Lower odds mean lower probability of being a victim of a crime. Red: odds* ≤ 1*, blue: odds* > 1.

Age group	% Victim unweighted	% Victim weighted	Intersecting CIs
16-24	24.24 (20.55–27.92)	23.65 (20.47–26.83)	Yes
25-34	22.82 (20.19–25.44)	22.81 (20.58–25.04)	Yes
35-44	17.34 (15.14–19.54)	17.20 (15.25–19.15)	Yes
45-54	17.21 (15.06–19.36)	18.02 (16.08–19.96)	Yes
55-64	14.09 (12.04–16.15)	14.15 (12.34–15.95)	Yes
65-74	9.81 (7.99–11.63)	9.2 (7.64–10.76)	Yes
75+	6.52 (4.81–8.22)	6.1 (4.68–7.51)	Yes

Table 1: *National victim % by age groups. 95% confidence interval in brackets.*

Table of the top 10 best GLM models by AIC for predicting whether a person has been a victim of a crime in the past year. Here the response variable is $V_i \sim Binomial(\mu_i)$.

Model	Formula	AIC (train)	Obs (train)	Brier (train)	Sens. (test)	Spec. (test)	MR (test)	Kappa (test)	Cut-off (train)
Model 1	$logit(\mu_i) \sim A_i + Ed_i + Ar_i + Rb_i + Po_i + D_i$	5,156.4	6,121	0.128	0.607	0.600	0.399	0.120	0.167
Model 2	$logit(\mu_i) \sim A_i + Ed_i + Ar_i + Rb_i + Po_i$	5,157.0	6,121	0.128	0.600	0.611	0.391	0.124	0.170
Model 3	$logit(\mu_i) \sim A_i + Ed_i + Ar_i + Rb_i + Vd_i + D_i$	5,157.0	6,121	0.128	0.604	0.601	0.398	0.120	0.167
Model 4	$logit(\mu_i) \sim A_i + Ed_i + Ar_i + Po_i + D_i$	5,157.1	6,121	0.128	0.595	0.599	0.402	0.113	0.168
Model 5	$logit(\mu_i) \sim A_i + Ed_i + Ar_i + Rb_i + D_i$	5,157.4	6,121	0.128	0.604	0.596	0.403	0.116	0.166
Model 6	$logit(\mu_i) \sim A_i + Ed_i + Ar_i + Rb_i + Vd_i + Po_i + D_i$	5,157.5	6,121	0.128	0.607	0.596	0.403	0.117	0.166
Model 7	$logit(\mu_i) \sim Sp_i + A_i + Ed_i + Ar_i + Rb_i + Po_i + D_i$	5,157.6	6,121	0.128	0.590	0.606	0.397	0.115	0.168
Model 8	$logit(\mu_i) \sim A_i + Ar_i + Rb_i + Po_i$	5,157.6	6,121	0.128	0.597	0.611	0.392	0.123	0.170
Model 9	$logit(\mu_i) \sim S_i + A_i + Ed_i + Ar_i + Rb_i + Po_i + D_i$	5,157.7	6,121	0.128	0.595	0.603	0.398	0.116	0.167
Model 10	$logit(\mu_i) \sim Wk_i + A_i + Ed_i + Ar_i + Rb_i + Po_i + D_i$	5,157.9	6,121	0.128	0.602	0.599	0.400	0.117	0.167

Table 2: *Obs* is the number of observations, *Brier* is the Brier score, *Sens.* is the sensitivity, *Spec.* is the specificity, *MR* is the misclassification rate, *Kappa* is Cohen's kappa, *Cut-off* is the cut-off which maximises sensitivity and specificity of a given model's predictions on the training set.

Bibliography

[Black et. al., 2011] Black, M. C., Basile, K. C., Breiding, M. J., Smith, S .G., Walters, M. L., Merrick, M. T., Stevens, M. R. 2011. The National Intimate Partner and Sexual Violence Survey: 2010 summary report. Retrieved from the Centers for Disease Control and Prevention, National Center for Injury Prevention and Control. [online] Available at: http://www.cdc.gov/ViolencePrevention/pdf/NISVS_Report2010-a.pdf [Accessed 27 April 2021]

[Botchkarev, 2018] Botchkarev, A. (2018). Performance Metrics (Error Measures) in Machine Learning Regression, Forecasting and Prognostics: Properties and Typology. ArXiv, abs/1809.03006. Available from: https://arxiv.org/pdf/1809.03006.pdf

[Buckley et. al., 2018] Buckley, J. and Ceraolo, M. 2016. Crime Survey for England and Wales 2013- 2014: Unrestricted Access Teaching Dataset, version 1.0.

[Champion, 2014] Champion, T. 2014. People in cities: the numbers. [online] Available from: https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/321814/14-802-people-in-cities-numbers.pdf

[Heeks et. al., 2018] Heeks, M., Reed, S., Tafsiri, M, Prince, S. (2018). The economic and social costs of crime. Second edition. Available from: https://www.gov.uk/government/publications/the-economic-and-social-costs-of-crime [Accessed 21 Apr 2021]

[Home Office, 2012] Home Office (UK). 2012. British Crime Survey: methodology. [online] Available at: https://www.gov.uk/government/statistics/british-crime-survey-methodology [Accessed 27 April 2021].

[ONS, 2021] Office For National Statistics. 2021. Crime in England and Wales: year ending September 2020. [online] Available at: https://www.ons.gov.uk/peoplepopulationandcommunity/crimeandjustice/bulletins/crimeinenglandandwales/yearendingseptember2020 [Accessed 27 April 2021].

[Statista Research Department, 2021] Statista Research Department. 2021. UK crime rate 2020 | Statista. [online] Available at: https://www.statista.com/statistics/1030625/crime-rate-uk/#:~:text=The%20crime%20rate%20of%20the,with%20110%20in%202002%2F03 [Accessed 27 April 2021].