Review into the criminal justice system response to adult rape and serious sexual offences across England and Wales

Appendix E: A statistical analysis of factors that influence the probability of charges for adult rape offences

Authors: Henryk Faas, Alex Dunnett, Rob Fulton and Andy Feist, Home Office

Contents

Acknowledgements

Executive summary

This report summarises quantitative analysis on trends and factors affecting the probability of a charge for rape offences recorded by police forces in England and Wales.

It is based on the analysis of a record-level dataset held by the Home Office covering approximately 214,000[1] rape offences recorded by the police between 2015/16 and 2019/20. The analysis used three forms of regression modelling to explore different aspects of the predictors of charge in adult rape offences.

The analysis focused on factors – selected based on their availability and completeness in administrative data records – likely to predict the outcome of rape cases recorded by the police.

Logistic regression[2] was used to identify the offence characteristics most likely to predict the probability of a charge in adult rape cases. The impact of the Crown Prosecution Service (CPS) and police force areas was estimated using a multilevel model.[3] Regression discontinuity analysis[4] was used to examine the shift in the charge probability around the time of the collapse of a high-profile Crown Court case in 2017 due to the non-disclosure of digital evidence.

Key findings

- The time between an offence taking place and being recorded by the police was a significant predictor of a charge. Charge probability was found to be highest for rape offences recorded in the first 48 hours after the offence was committed. After this period, charge probability falls but stabilises after seven days between the offence and recording. This stable pattern extended to offences that were recorded more than ten years after they took place.[5]

- Charge outcomes were least likely for offences where the victim and the suspect were identified as partners or ex-partners. These cases also have the highest probability to be closed due to a victim's unwillingness to support the police investigation.

- There is no evidence from the analysis to suggest that the long-term decline in rape charges in this period results from fundamental changes in case characteristics.

---

1 This figure relates to the full dataset. Most analyses were undertaken on a subset of the data.
2 Logistic regression measures the effect of independent variables on a binary outcome (in this analysis 'charge' versus 'other outcomes').
3 Multilevel regression extends conventional models to account for variation in the outcome i.e. charges across clusters (for example, police forces within CPS areas).
4 Regression discontinuity analysis measures level and/or trend shift at a specific event, cut-off or boundary.
5 Recording of an offence takes place at the earliest opportunity and at the most within 24 hours after it has been designated a recorded crime.

Comparing charge probabilities across different years shows that in-year patterns are broadly consistent over time. Other factors appear to have driven the fall in charge probability over the period.

- While charge rates vary between different CPS and police force areas, this can mainly be explained by differences in individual offence characteristics. Only a small part of the variation in charge probability can be attributed to CPS area or police force areas.

- At the start of 2018 there was a precipitous fall in charge volumes for adult rape offences. This took place soon after the high-profile collapse of a Crown Court rape case in December 2017. Charges against a defendant were dropped due to the failure of the prosecution to disclose mobile phone records. Regression discontinuity analysis was used to explore any relationship between the fall in rape charges and the response to this case.

- The charge probability for adult rape cases was compared with rape offences that might be expected to have lower levels of digital evidence: 'historical' adult rape cases, which occurred more than 10 years ago, and rape offences involving victims under the age of 13.

- The analysis shows that the probability of a charge for adult rape offences was statistically significantly lower for cases that concluded after 14 December 2017. When the equivalent analysis was undertaken on 'historical' offences committed before 2010, and for offences involving child victims under 13 years of age, the change in charge volumes was not found to be statistically significant.

- The results suggest that the criminal justice system's response to the case, alongside wider changes to the handling of disclosed material, contributed to the marked and sustained fall in charge probability of rape offences. However, the analysis does not allow the identification of the precise mechanism involved.

1. Background

Since 2013 there has been a continuous rise in recorded adult rape offences. While there is no strong indication that the prevalence of rape has changed fundamentally (see for instance the Crime Survey of England and Wales), the increase in police recorded rapes is likely to reflect, in part, an increase in the reporting and recording of these offences (Figure 1). At the same time, there has been a decrease in charge volumes for adult rape. Charge volumes peaked in the first quarter of financial year[6] 2015/16, fell in 2016/17 and then suffered a precipitous fall in the last quarter of 2017/18, from which they have not yet recovered.

Figure 1: Charge volumes and recorded rape offences, 2016–20

All recorded rape offences with a charge outcome ▬▬▬

All recorded rape offences ▬▬▬

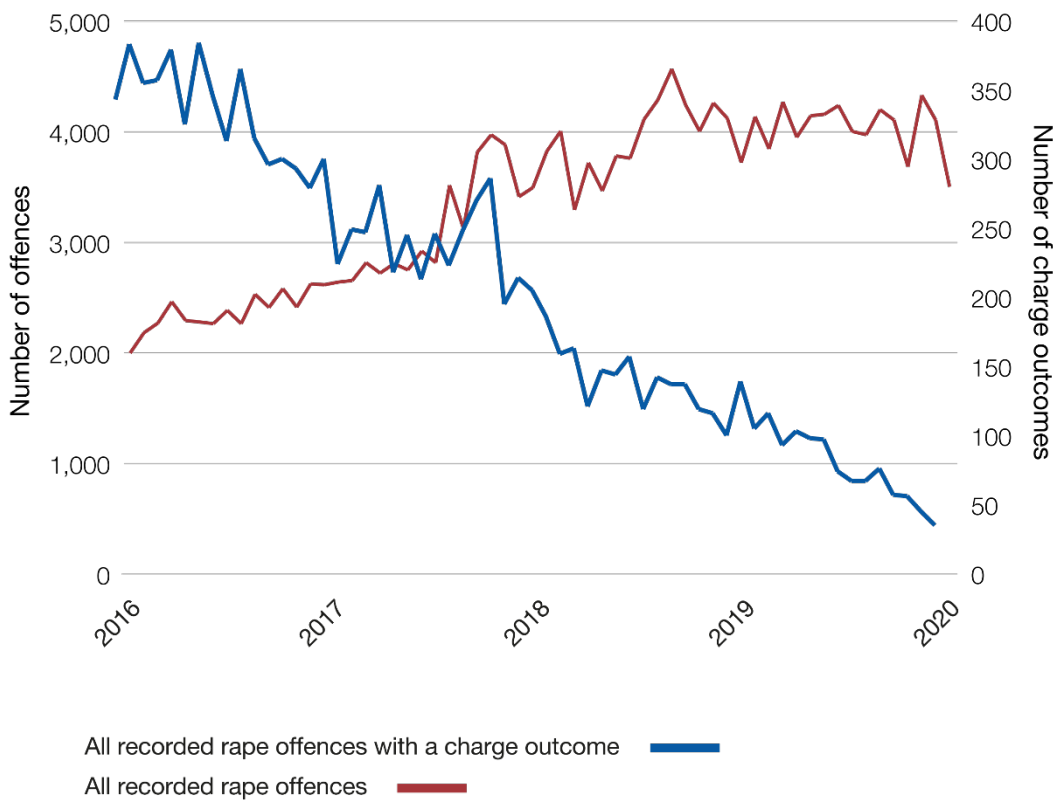**Figure 1: Rape offences in England and Wales: charge volumes and recorded crime trends, 2016-20**

Source: Home Office data

---

6   To correspond to data recording standards, the analysts use a financial year (FY) reporting cycle (1 April of the calendar year – 31 March of the following calendar year). Therefore, the last quarter of 2017/18 corresponds to the three months from January to March 2018.

The analysis summarised in this report was undertaken to support the cross-government end-to-end review into the criminal justice system's (CJS) response to adult rape offences across England and Wales.[7] The analysis sought to address three main research questions.

- What is the contribution of factors that influence the probability of a charge being brought for an adult rape offence?
- How much of the variation in charges can be attributed to differences in police force areas or CPS areas?
- Can the sharp reduction in charge volumes from January 2018 be linked to the impact of a high-profile Crown Court case in December 2017, which centred on the failure to disclose digital evidence?

---

[7] Review commissioned by the Violence Against Women and Girls (VAWG) inter-ministerial group and the National Criminal Justice Board (NCJB).

2.   Data and methodology

2.1 Data on the factors that predict charge outcomes in rape offences

The factors influencing whether a suspect is charged in recorded rape cases are generally well identified (Kelly et al., 2005)(Daly and Bouhours, 2009)(Hester and Lilley, 2017)(Morabito et al., 2019). These studies have generally been undertaken by extracting data from case file records. For instance, based on data from 8 police forces across England and Wales, Feist *et al.*, (2007) identified 21 variables as significant predictors of case outcome. These included:

- general offence characteristics (such as the age and employment characteristics of the victim, the relationship between the offender and the victim);
- specific details about the offence and offender characteristics; and
- details about the police investigation.

In an examination of 587 rape offences recorded by the London Metropolitan Police Service (MPS) over 2 months in 2012, Hohl and Stanko (2015) found a broadly similar range of factors – incident characteristics, evidence, reporting, police investigation, victim and suspect characteristics, victim–suspect relationship – as influential in determining charge rates. More recently, the Mayor of London's Office for Policing and Crime (MOPAC) examined 501 cases reported to the MPS and found that procedural characteristics were the strongest predictors of victim withdrawal, one of the most common outcomes for rape investigations (MOPAC, 2019).

The analysis presented here used a large dataset of administrative records from the Home Office data hub for almost 215,000 rape offences recorded over 5 years between April 2015 and March 2020 to assess the variation in charge volumes for rape offences across geographical regions and over time. These data cover 37 police forces within 13 Crown Prosecution Service (CPS) areas across England and Wales.[8] These data are limited by having fewer case variables than those derived from case-file analyses in the studies described above. However, the size and coverage of the dataset in this analysis is much larger than many earlier studies.

Because of the delay between the offence and recording, 62,522 offences took place before the sample period. In addition, due to the time it takes for some offences to be given an outcome, 29,767 offences were closed after the final offence recorded date in the sample period. Finally, there were also 17,168 'ongoing' cases, i.e. they were without an outcome recorded in the target period and had still not received an outcome at the time of data extraction.[9]

---

8   Offence level data and reference data (for example, lookup tables of offence codes and corresponding offence categories and descriptions) were accessed from the Home Office data hub with a database query in SQL and further processing in the R software environment accessed via RStudio (version 1.2.1335, R version 3.6.1.). Descriptive statistics were carried out in R or Excel. Logistic regression analysis was generally performed in R, with multilevel logistic regression performed in MLwiN (version 3.04).

9   Valid at the time the data were downloaded (7 September 2020).

2.2 Methodology – regression approaches

To model the probability of a charge outcome[10] for an adult recorded rape offence, logistic regression models were generated to describe patterns in the probability of a charge outcome and selected covariates. Incidents recorded in the reference period (2015/16 to 2019/20) that were yet to receive an outcome – 8% of the total sample – were excluded from the analysis.

The data for the independent variables selected come from police administrative data (Home Office data hub). The binary dependent variable for the logistic regression was 'charge' (coded as '1') versus all other outcomes (coded as '0').

Three types of regression approaches were used to explore the key research questions.

Approach 1: Logistic regression for probability of charge outcome

Predictors used in the analysis were selected from variables that were sufficiently well recorded in the dataset and reflected, where possible, significant predictors of outcome identified in previous studies. These were:
- the financial year (FY) in which the offence was recorded;
- the victim's age (categorised in 10-year age groups);
- gender (female, male);
- time before reporting the offence to the police (under 24 hours, 25 to 48 hours, 49 to 72 hours, 4 to 7 days, 8 days to 3 months, 4 months to 2 years, 3 to 10 years, or more than 10 years); and
- whether the offence occurred between 8pm and 5am on a Friday or Saturday night.

The victim's age (adult, under 16, under 13), based on the recorded offence code, was used to compare outcomes between adult and child victims. This variable had greater coverage than the specific victim age variable, which was affected by missing values.

Models were compared in terms of fit to the data using the Bayesian Information Criterion (BIC) to inform the choice of the most suitable model.

Modelling the effect of victim–offender relationship

In four police force areas – Hampshire, Norfolk, Staffordshire, Suffolk– across four years (2016/17 to 2019/20), information relating to the relationship between the victim and the suspect was sufficiently well recorded to allow the expansion of the logistic regression to include this predictor variable. Previous research has identified the victim–offender relationship to be a key variable in determining case outcome. This variable was grouped into five categories: partner; family; acquaintance; stranger; and unknown.[11] 'Partner'

---

[10] Crime outcomes are divided into 1 of 22 outcomes according to the current crime outcomes framework. For the purpose of this analysis, outcomes have been split into two broad categories: 'charges'; and 'all other outcomes'. *Home Office Crime Outcomes Bulletin*, 2020.

[11] Forces submit information about the victim–suspect relationship to the Home Office in 96 categories, which were summarised for this analysis in 5 categories (partner, family, acquaintance, stranger and unknown).

offences were coded to include both current and ex-partners, and 'family' offences to include blood relatives, in-laws and step relationships.

Approach 2: Multilevel logistic regression for probability of charge outcome allowing for variation across police forces and CPS areas

To simultaneously explore the probability of charge outcome, individual case characteristics *and* the pattern of variation in police force areas and CPS areas, a three-level random-intercept logistic regression mode was used (Charlton *et al.*, 2020).[12] These offences were grouped into 37 police force areas at the second level, and 13 CPS areas at the third level.[13] The estimation method used were a first order marginal quasi-probability (MQL1) method followed by a second order predictive quasi-probability (PQL2). Given the low degree of clustering in the data, this was considered suitable and the more complex Markov Chain Monte Carlo (MCMC) methods were not required.

The first level predictors considered in the final model were the year that the offence was recorded (during 2015/16 and 2018/19), and the time between when the offence occurred and when it was recorded by the police. In order to reduce the complexity of the multilevel analysis, it was restricted to adult female victims.

To establish the relative influence of each level of the model hierarchy, the proportion of observed response variation was determined (variance partition coefficients) (Snijders, 2011). The potential impact of possible outlying CPS areas and police forces on these statistics was investigated.

Approach 3: Regression discontinuity design

A regression discontinuity method (Imbens, G., 2007) (Mueller-Smith and Schnepel, 2016) was used to assess the effect of the collapse of a high-profile Crown Court case in 2017 (*R.* v. *Allan*) (*The Guardian*, Grierson, J., 2017) on charges for rape offences. Specifically, it was hypothesised whether a decline in charge volumes after this date could be linked to the case and its impact on the disclosure of digital evidence.

A cut-off date in the data was assigned to distinguish whether an offender was charged before or after this case collapsed, and the probability of an adult rape charge was modelled either side of the decision (14 December 2017). A discontinuity logistic regression model was used for cases with an outcome between 1 April 2017 and 31 March 2019. A 'sharp' disruption was modelled because the analysts would expect a non-linear relationship between probability of charge and time across this disruption. Predictors included in the model were the year of the outcome of the police investigation, the age of victim, and the time between the offence and reporting to the police.

As the main analysis was undertaken for adult rape offences, control models were built for several sub-categories of rape where the presence of digital evidence was likely to be lower (historical offences, and offences involving victims under the age of 13). Two other

---

12 Bristol University Centre for Multilevel Modelling (2020). The model was fitted to the data using MLwiN version 3.04.

13 Because, from a multilevel modelling perspective, there are relatively few clusters at the second and especially the third level, second and third level predictors were not included.

non-rape control models were also constructed, for burglary offences and indecent image offences.

Descriptive statistics and model outputs relating to data and regression approaches are given in the following sections (3 to 5). The full regression results for all models are reported in Appendix A.

3. Results of the logistic regression for probability of charge outcome

3.1 Victim gender and age

The full dataset covers 214,832 recorded crimes of rape between 2015/16 and 2019/20. More than half of these offences involved female adult victims, with females aged between 13 and 15 making up the second largest group (Table 1). Ongoing investigations, i.e. those still pending an outcome, totalled 17,168. These were excluded from the dataset, so the effective sample size of adult rape offences involving male and female victims was 140,384 offences, of which around 1 in 20 resulted in a charge.

Table 1: Victim gender and age, by outcome, 2015/16 to 2019/20

Victim category	Charges	All other outcomes	Ongoing	Total
Adult female	6,200	125,612	9,521	**141,333**
Adult male	169	8,403	600	**9,172**
13- – 15-year-old female	1,966	23,632	3,228	**28,826**
13- – 15-year-old male	280	3,008	443	**3,731**
Under 13-year-old female	2,557	16,779	2,454	**21,790**
Under 13-year-old male	900	8,158	922	**9,980**
Total	**12,072**	**185,592**	**17,168**	**214,832**

Source: Home Office data

Figure 2 gives a more detailed breakdown of the victim's age. The age distribution across rape offences shows a peak for victims at 16 years of age, with a much smaller secondary peak for young victims at 8 years old. The mode, i.e. the most common, age of an adult victim is 16 years old.

Figure 2: Recorded rape offences, by age of victim, 2015/16 to 2019/20
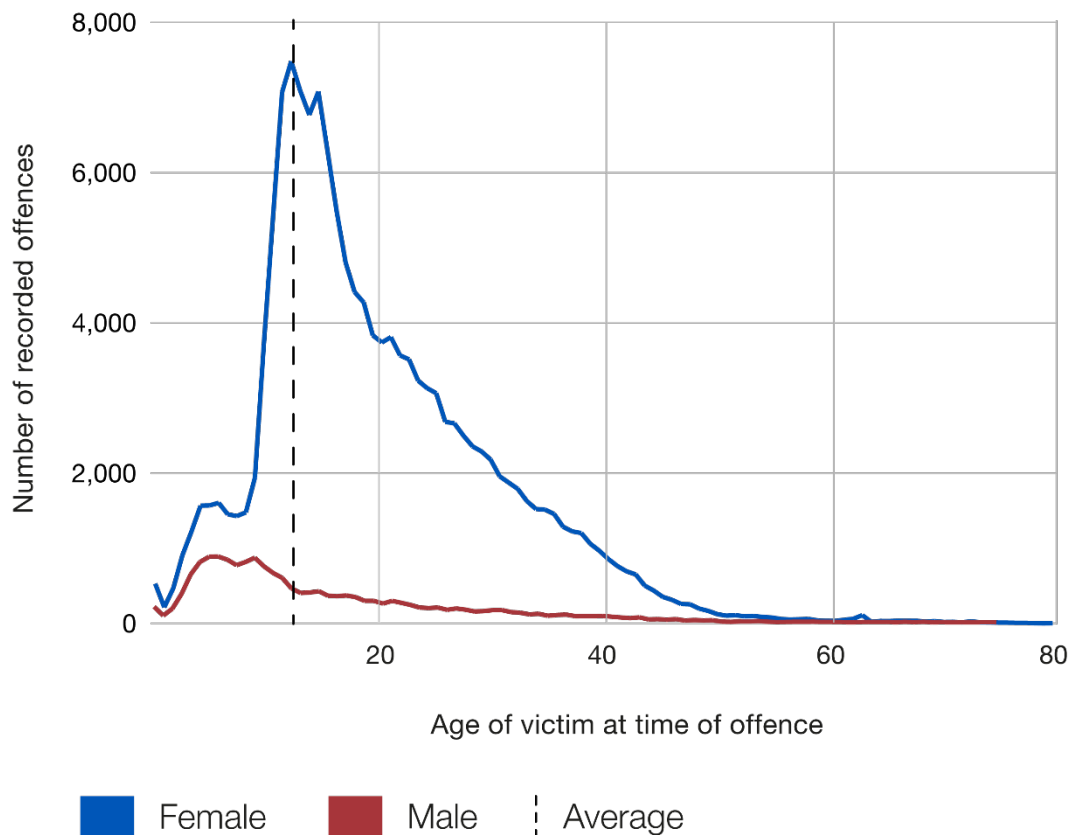
Figure 2: Rape offences by age of victim

Source: Home Office data

The regression model showed that the following variables were significant predictors of whether the case resulted in a charge:

- the age of the adult victim (by age band);
- victim gender;
- the time between when the offence was committed and when it was recorded by the police;
- whether the offence was committed on a weekend night (a Friday or Saturday night between 8pm and 5am); and
- the year of recording.

These findings tend to echo those from previous studies, which have found victim characteristics, incident and reporting characteristics, as key predictors of a charge or conviction.

Table 2 shows the distribution of cases by victim age for adults, categorised by outcome.

Table 2: Victims, by age band and outcome, 2015/16 to 2019/20

Age category	Charges	All other outcomes	Ongoing	Total
16 – 24	2,622	46,794	3,803	**53,219**
25 – 34	1,491	30,312	2,234	**34,037**
35 – 44	713	18,067	1,083	**19,863**
45 – 54	362	9,015	575	**9,952**
55 and over	142	3,669	215	**4,026**
Under 16	4,294	38,671	4,894	**47,859**
Missing data	2,448	39,064	4,364	**45,876**
Total	**12,072**	**185,592**	**17,168**	**214,832**

Source: Home Office data

A statistically significant difference in charge probability was identified between age groups within the adult category (Figure 3). Offences with victims aged 16 to 24 years old have a significantly higher probability of a charge outcome than older victims.

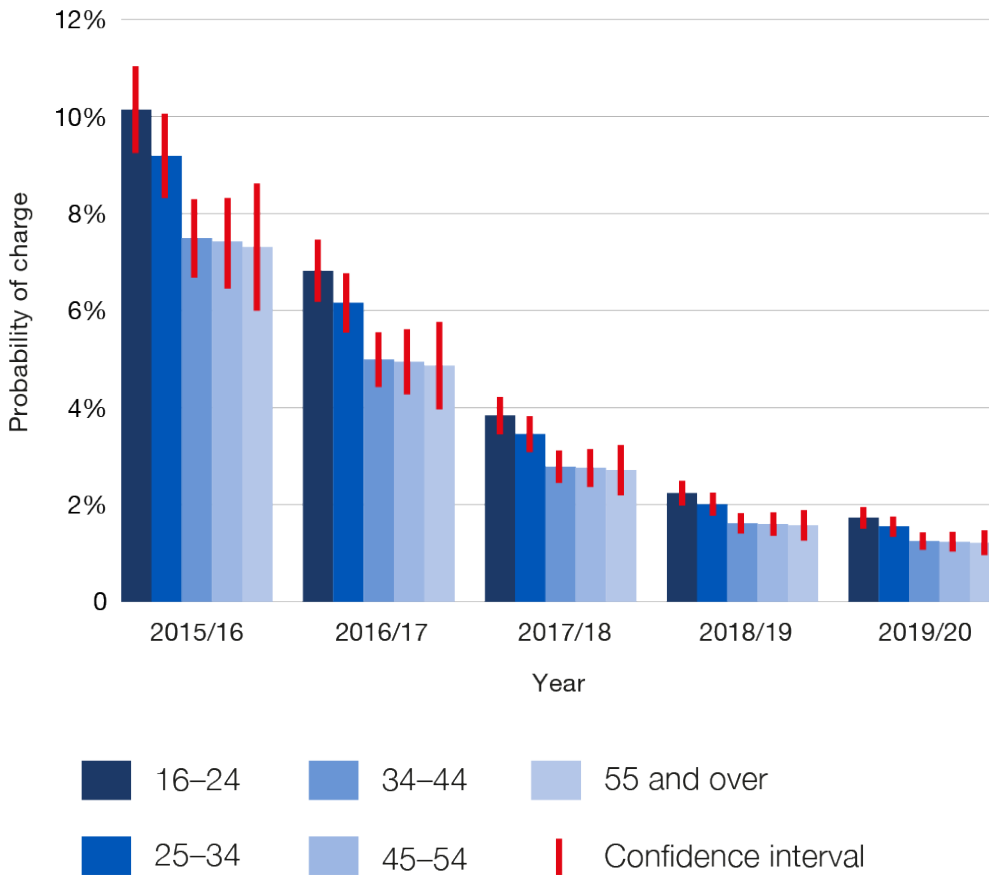**Figure 3: Age of victim and charge probability, 2015/16 to 2019/20**

Figure 3: Charge probability by victim age

Source: Home Office data

Cases involving female victims had a statistically significant higher probability of a charge than those involving male victims (Figure 4).

Figure 4: Sex of victim and charge probability, 2015/16 to 2019/20
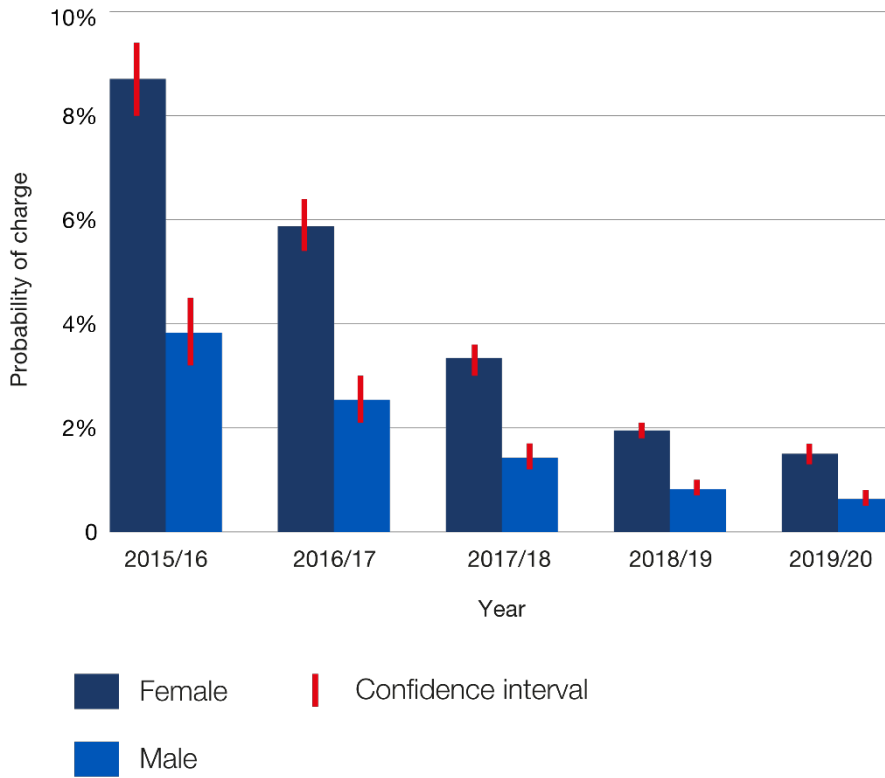
Figure 4: Charge probability by victim sex

Source: Home Office data

3.2 Time between offence and recording by the police

Table 3 shows the distribution of cases for adult victims by the time between the offence and when it was recorded by the police.

Table 3: Time between offence and recording by the police, by outcome, adult victims, 2015/16 to 2019/20

Time between offence and recording	Charges	All other outcomes
Up to 1 day	2,661	31,525
1 to 2 days	1,408	19,454
2 to 7 days	1,052	22,345
7 days to 3 months	1,210	30,717
3 months to 2 years	1,407	29,079
2 years to 10 years	1,557	22,156
Over 10 years	2,777	30,316
Total	**12,072**	**185,592**

Source: Home Office data

Figure 5 shows that the probability of a charge outcome is significantly higher if recorded in the first 24 to 48 hours after the offence. After the first week the probability of a suspect being charged remains stable, irrespective of elapsed time between the offence taking place and recording. This pattern is consistent across the reference period (2015/16 to 2019/20), however with a marked decrease in probability year-on-year.

Figure 5: Time between offence and recording, and charge probability, 2015/16 to 2019/20
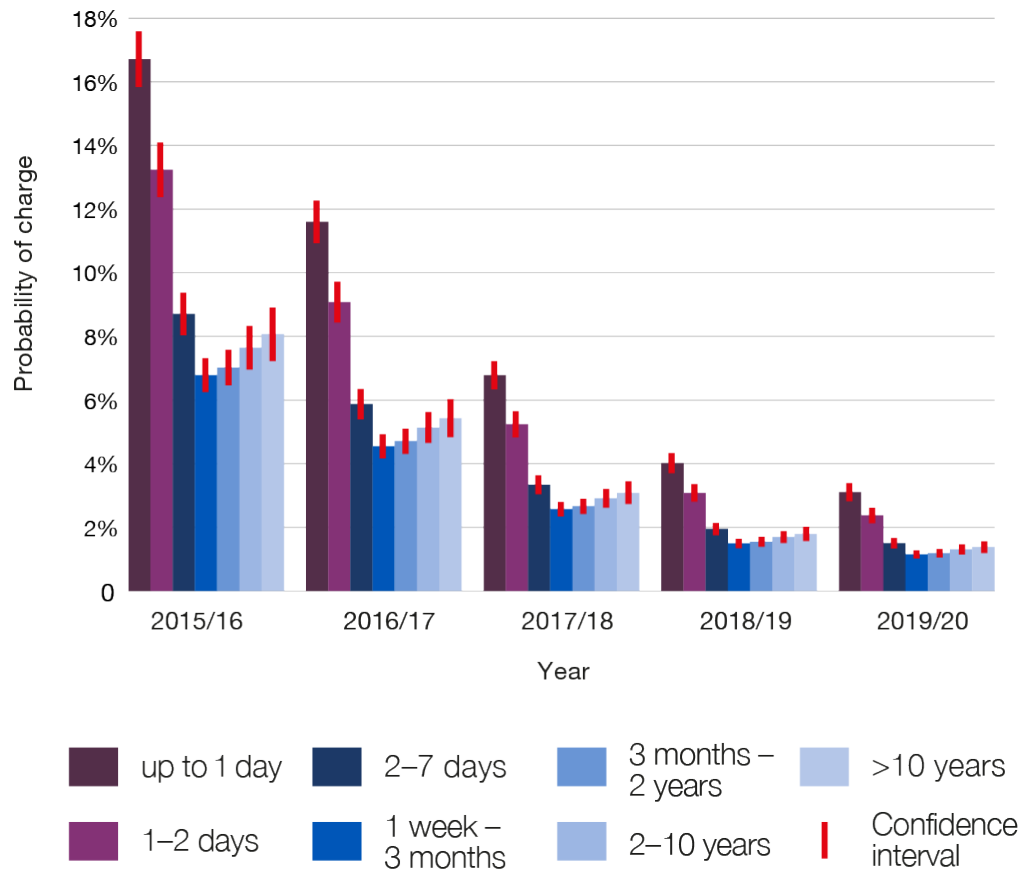
Figure 5: Charge probability by time between offence and recording

Source: Home Office data

3.3 Victim–suspect relationship

Sufficiently comprehensive data on victim–suspect relationship was available for 4 of the 37 police forces. Table 4 and Figure 6 give the overall breakdown for these – acquaintance, familial, partner, stranger and unknown – by outcome, for the four forces. In 11% of cases the relationship was not known. Of those offences where the relationship was known (8,279), just under half were 'partner' offences (including ex-partners) (4,016). 'Stranger' offences accounted for just under 5% of all offences where the relationship was known.

Table 4: Victim–suspect relationship, by outcome,[1] 2016/17 to 2019/20

	Charges	All other outcomes	Total
Acquaintance	172	2,645	**2,817**
Family	175	876	**1,051**
Partner	103	3,913	**4,016**
Stranger	53	342	**395**
Unknown	74	972	**1,046**
Total	**577**	**8,748**	**9,325**

[1] Based on four forces, Hampshire, Norfolk, Staffordshire, Suffolk, 2016/17 – 2019/20.
  Source: Home Office data

The relationship between the victim and the suspect had a significant impact on both charge probability and the outcome 'victim does not support the investigation'.

Figure 6: Predicted probability of charge outcomes for adult offences, by victim–suspect relationship, 2015/16 to 2019/20
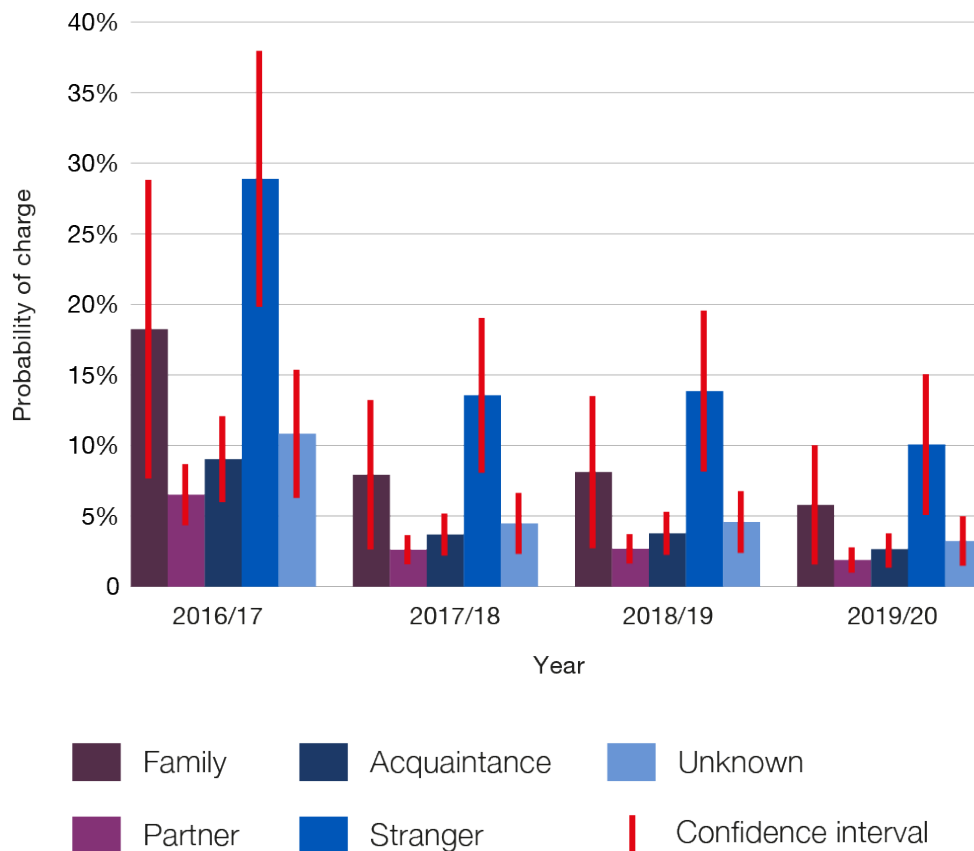
Figure 6: Charge probability by victim-suspect relationship

Source: Home Office data

Among all offences where a suspect was identified, and information about the relationship between the victim and the suspect was recorded, the probability of the identified suspect being charged was found to be lower for partner offences than for other relationship categories. This aligns with findings from earlier studies, for example, Feist *et al.* (2007).

A general observation from reviewing the predicted probabilities across different years in Figures 3 to 6, is that the patterns are broadly consistent over time. There is no evidence to suggest any systematic year-on-year changes in case characteristics from the analysis. Other factors appear to have driven the fall in charge probability over the period.

3.4 Investigations that close with the outcome 'evidential difficulties – victim does not support investigation'

Table 5 shows the distribution of those cases that were closed with the outcome 'evidential difficulties – victim does not support the investigation' (30% of all cases).

Table 5: Distribution of cases closed 'evidential difficulties – victim does not support the investigation', by offence type, 2015/16 to 2019/20

Category	Evidential difficulties – victim does not support investigation	All other outcomes	Missing data
Adult female	69,696	64,075	7,562
Adult male	4,088	4,651	433
13- – 15-year-old female	11,513	14,891	2,422
13- – 15-year-old male	1,108	2,257	366
Under 13-year-old female	6,308	13,523	1,959
Under 13-year-old male	2,417	6,809	754
Total	**95,130**	**106,206**	**13,496**

Source: Home Office data

Logistic regression was used to identify the factors that predict the outcome 'victim does not support' compared with all other outcomes. Figure 7 shows that the probability of an adult rape case receiving the outcome of 'evidential difficulties – victim does not support investigation' was significantly higher for victims aged between 35 and 44.

Figure 7: Probability of case closing 'victim does not support', by age group, 2015/16 to 2019/20

Figure 7: Probability of case closing with evidential difficulties, 'victim does not support the investigation', by age group

Source: Home Office data

The probability of an outcome 'victim does not support the investigation' was also higher for rape offences that were recorded more than one week after the offence, and highest for offences dating back between two and ten years (Figure 8). However, in those cases where the time between the offence and the police recording was in excess of ten years, the probability of victims not supporting the investigation was significantly reduced.

Figure 8: Probability of case closing 'victim does not support investigation', by time to recording, 2015/16 to 2019/20

Figure 8: Probability of case closing with evidential difficulties, 'victim does not support the investigation', by time between offence and recording

Source: Home Office data

Using the victim–suspect relationship dataset, logistic regression was also used to measure the impact of the relationship on the probability of victim withdrawal. The probability of the case closing with 'evidential difficulties – victim does not support the investigation' was highest for 'partner' offences and lowest for 'stranger' offences, mirroring the pattern for the outcome 'charged' (Figure 9).

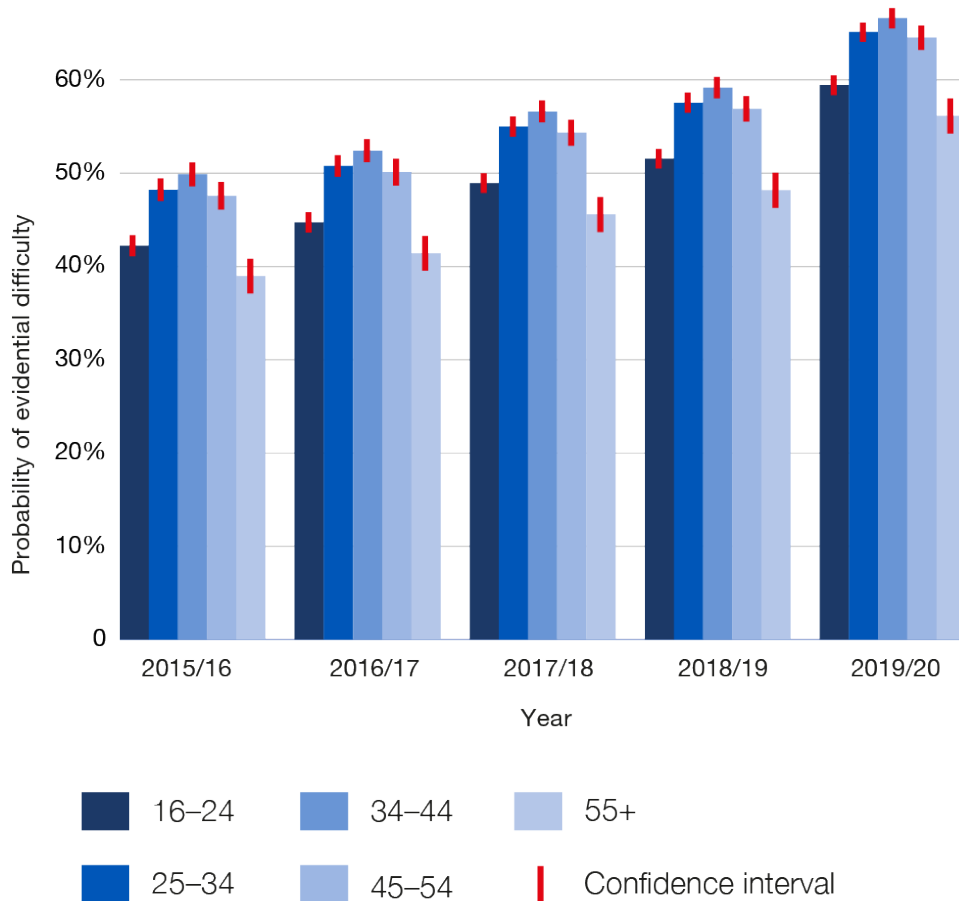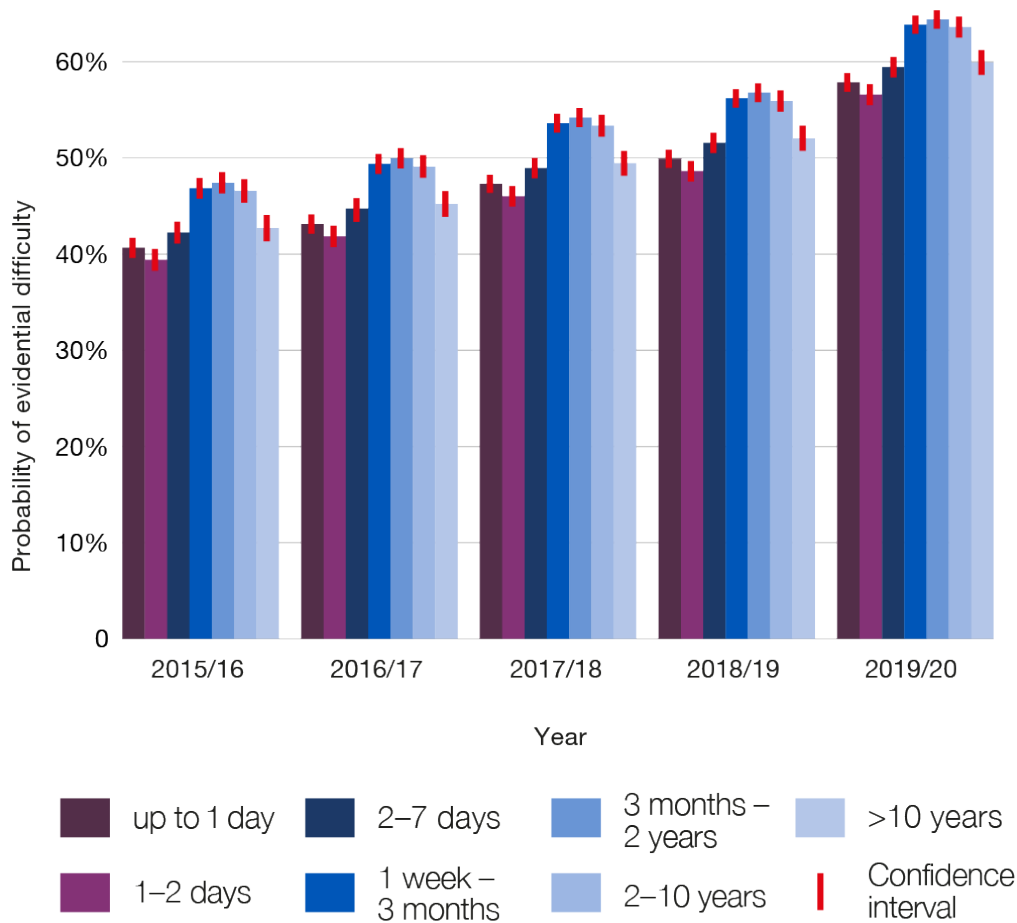**Figure 9: Probability of case closing 'victim does not support investigation', by victim–suspect relationship, 2015/16 to 2019/20**
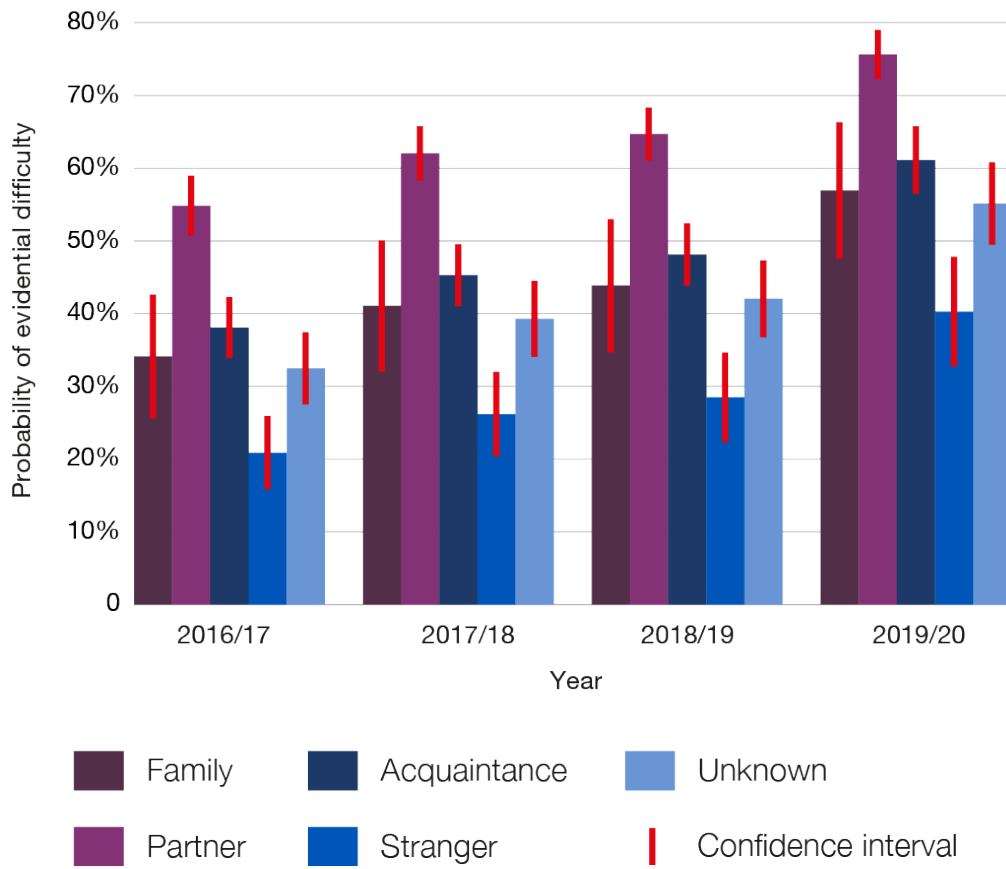
Figure 9: Probability of case closing with evidential difficulties, 'victim does not support the investigation', by victim – suspect relationship

Source: Home Office data

3.5 Results of the multilevel logistic regression exploring the variation of charge outcome across police forces and Crown Prosecution Service areas

Multilevel logistic regression was used to assess how much the charge rate varied across police forces and Crown Prosecution Service (CPS) areas and how much of this variability can be attributed to CPS areas or police forces, rather than individual offence characteristics.

Figure 10: Charge rate, by police force grouped by CPS areas, 2016/17

1 Bedfordshire	11 Surrey	21 Northamptonshire	31 West Yorkshire
2 Hertfordshire	12 Metropolitan Police	22 Derbyshire	32 Durham
3 Hampshire	13 West Mercia	23 Leicestershire	33 Northumbria
4 Dorset	14 West Midlands	24 Nottinghamshire	34 Cleveland
5 Cambridgeshire	15 Staffordshire	25 Lincolnshire	35 Cumbria
6 Essex	16 Avon & Somerset	26 South Wales	36 Lancashire
7 Suffolk	17 Gloucestershire	27 Gwent	- - - mean charge ratio
8 Norfolk	18 Devon & Cornwall	28 North Wales	
9 Thames Valley	19 Merseyside	29 Dyfed Powys	
10 Sussex	20 Cheshire	30 North Yorkshire	

Figure 10: Charge ratio by CPS area in 2016-17

Source: Home Office data

Figure 10 shows how charge rates differed between CPS areas and between police forces *within* CPS areas in 2016/17. However, it is hard to meaningfully interpret these differences without considering the variation in offence characteristics. The multilevel model allows this to be addressed.

To estimate the contribution that CPS areas and police forces make to the overall variation in charge probability, a three-level logistic regression model was fitted for adult female victims. The predictor variables were:
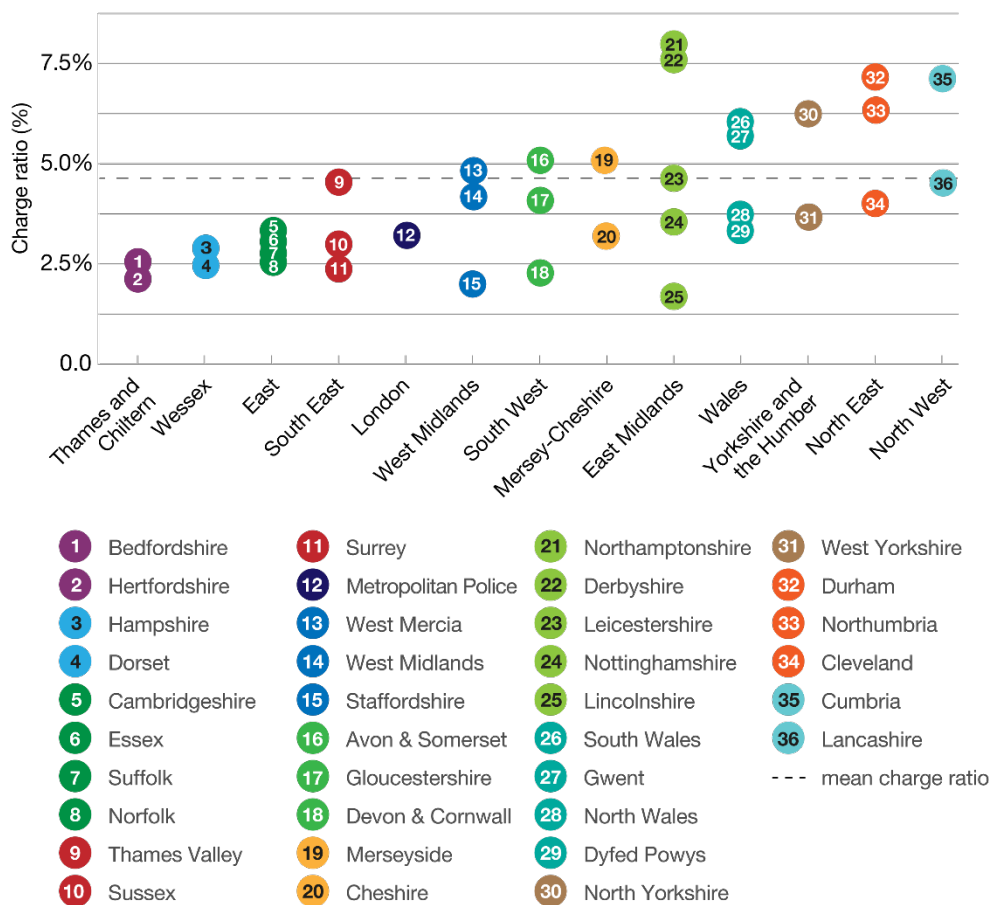- the year the offence was recorded by the police; and
- the time from offence to recording by the police.

Individual offences comprised the first level of the multilevel model, police force areas the second, and CPS areas the third level.

Only a small part of the variation in charge probability can be attributed to the CPS area or the police forces within a CPS area; 3% of the overall variation lies *between* CPS areas and only 1% lies *between* police forces *within* CPS areas.[14] Most of the variation in charge probability – some 96% – lies *within* police forces, between individual cases.

Allowing for case characteristics, the models also show that it is not possible to statistically separate the majority of CPS areas from one another in terms of their charge rates. Likewise, a similar picture exists when comparing the model outputs for different police force areas.

---

[14] Estimated from the variance partition coefficient

4. Regression discontinuity analysis on the impact of the Crown Court case *R.* v. *Allan*, 2017

Analysis was conducted to explore whether a specific event – the high-profile collapse of the Crown Court case *R.* v. *Allan* (MPS, 2018), due to the failure of the prosecution to disclose full mobile phone records to the defence – could explain the subsequent sharp decline in the number of charges.

Figure 11a shows charge volumes against all other outcomes for different categories of rape offences to examine the hypothesis that the court judgment impacted disproportionately on adult rape offences. For adult rape victims – 16 years or older – charge volumes fall markedly after the Crown Court case, December 2017 (Figure 11a). Comparing the six-monthly volumes for the second half of 2017 with the first half of 2018 shows a sharp drop in adult rape charges, down from an average of 251 to 147 per month, a fall of 42%.

However, for offences where at least 10 years had passed between offence and recording ('historical' offences) (Figure 11b) and for rape offences involving victims under 13 years old (Figure 11c), the drop is much less pronounced. For adult 'historical' offences, charge volumes fell by 28% – from an average of 19 to 14 charges per month – in the first half of 2018 compared with the second half of 2017. There was a fall in charges of 33% for cases involving under 13-year-old victims across the same period, down from an average of 69 to 46 offences per month. Arguably, both these offence types are less likely to have a digital component. 'Historical' cases will have taken place before the widespread use of mobile phones and the advent of smartphones. Cases with victims under the age of 13 are also less likely to have digital material relevant to the case.[15] The trend data provide some support for the hypothesis, by revealing a sharp fall in charges for adult rape cases in early 2018, but a more limited fall in charges for adult 'historical' rape offences and offences with victims under 13 years of age.

---

[15] There are two reasons why cases with under 13-year-old victims are likely to have a lower level of digital evidence. First, by their very nature, the issue of consent in these cases is not relevant so it is unlikely that mobile phone evidence would be pertinent. Secondly, ownership of mobile phones is expected to be lower for this age group (the mode for recorded rape victims under 13 years old is 8). An Ofcom report in 2019 found that fewer than 4 in 10 children aged between 8 and 11 years had their own smartphone. The comparable figure for 12- to 15-year-olds was over 8 in 10.

Figure 11a: Outcomes for rape offences 2016–20, by month of outcome
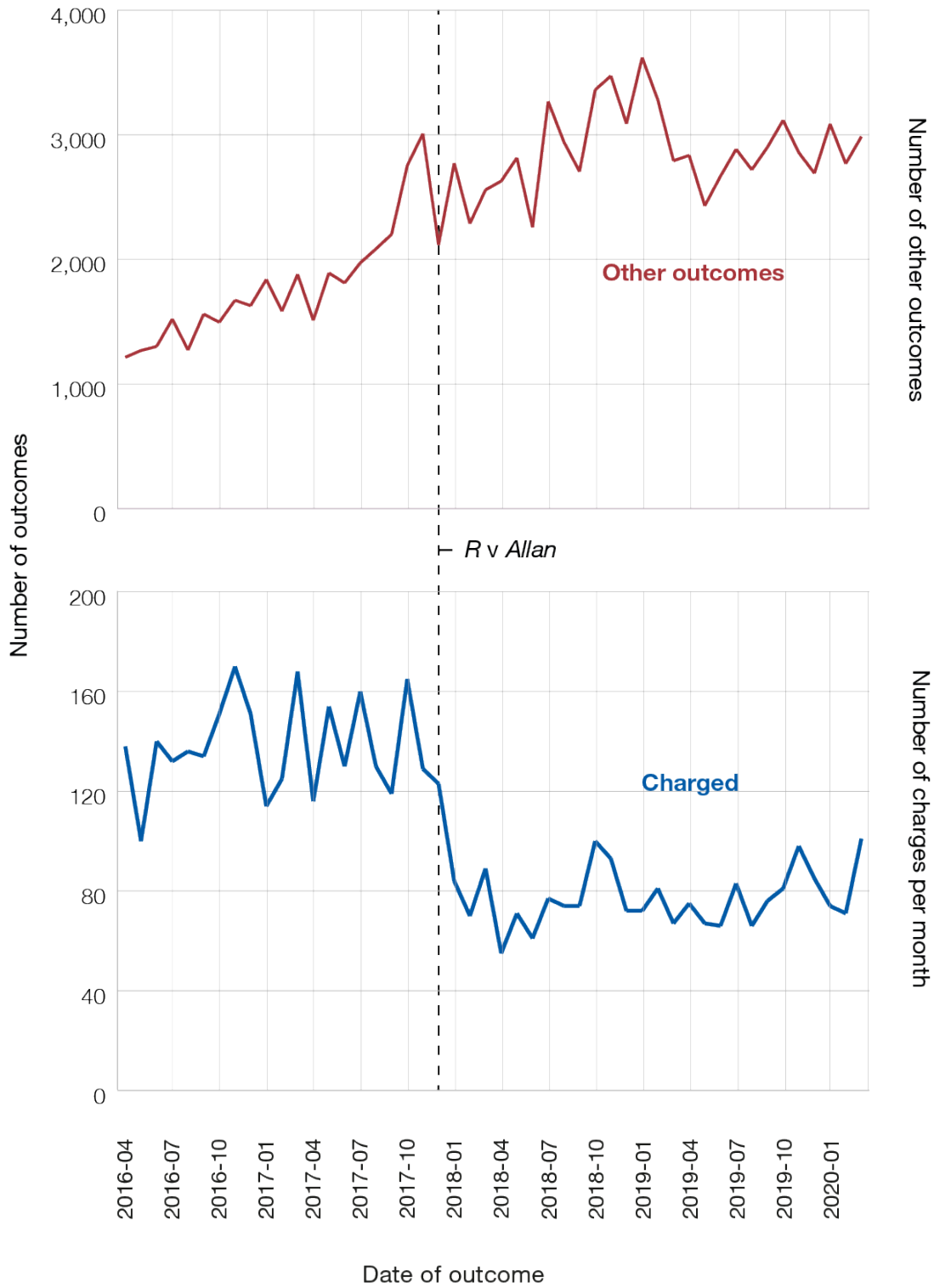
Figure 11a: Outcomes for rape offences with adult victims, by month of outcome

Source: Home Office data

Historical offences

Number of offences

Number of other outcomes

Other outcomes

R v Allan

Charged

Number of charges per month

Date of outcome

Figure 11b: Number of offences for adult rape victims, where offence occurred more than 10 years before it was recorded
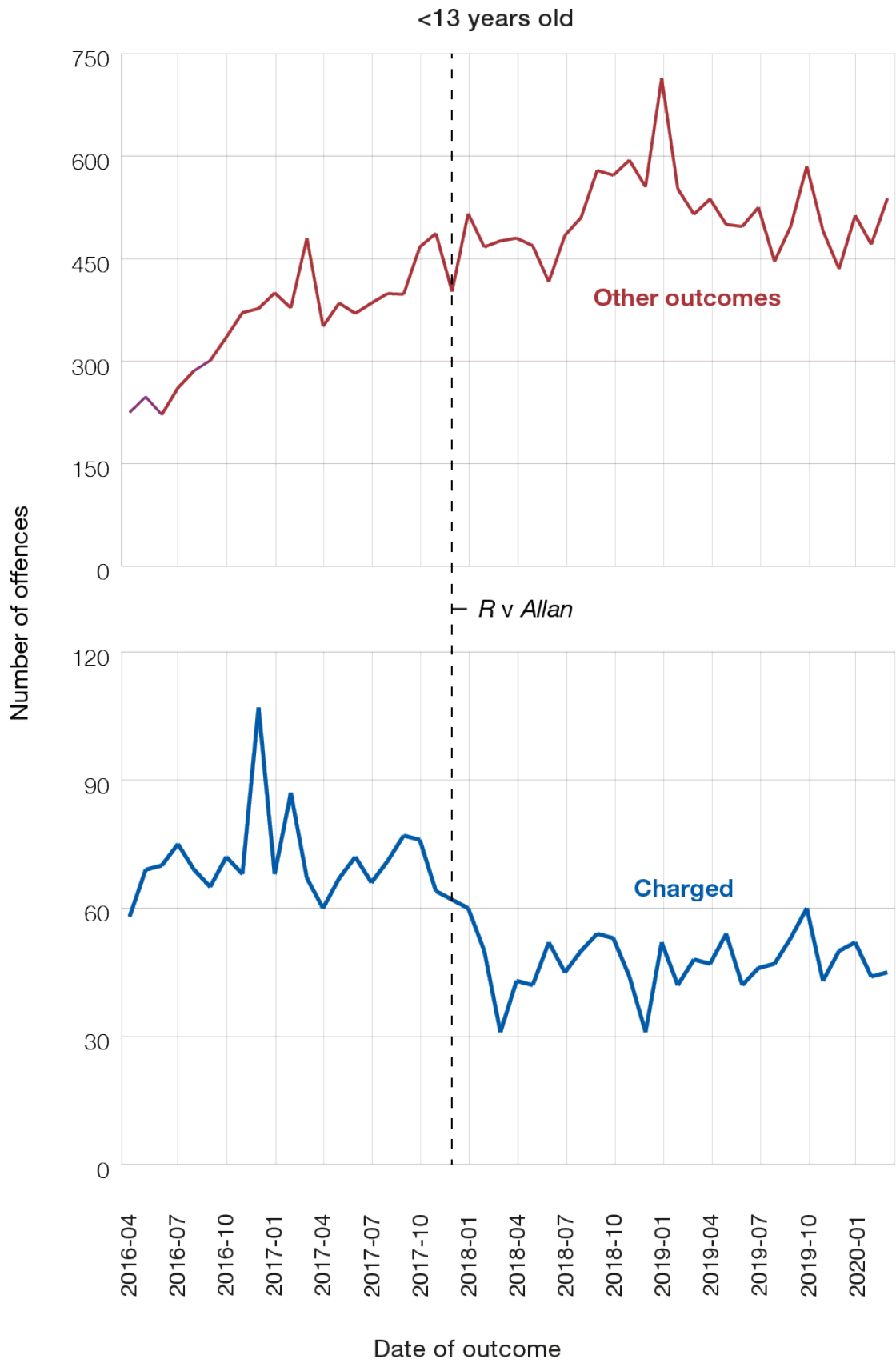
Figure 11c: Number of offences for rape victims under 13 years old

The hypothesis that the fall in adult charge volumes in early 2018 was associated with changes around digital evidence is supported by analysis that models this disruption. As Figure 12 shows, the charge probability for adult rape offences was significantly lower for cases that concluded after 14 December 2017 (also see Annex A).[16] However, when the equivalent analysis was undertaken on 'historical' offences committed before 2010, and for offences involving child victims under 13 years of age, the 'disruption' in charge volumes in December 2017 was not found to be statistically significant.

Figure 12: Model results: predicted versus observed charge rate for adult offences, quarterly results, 2016–18
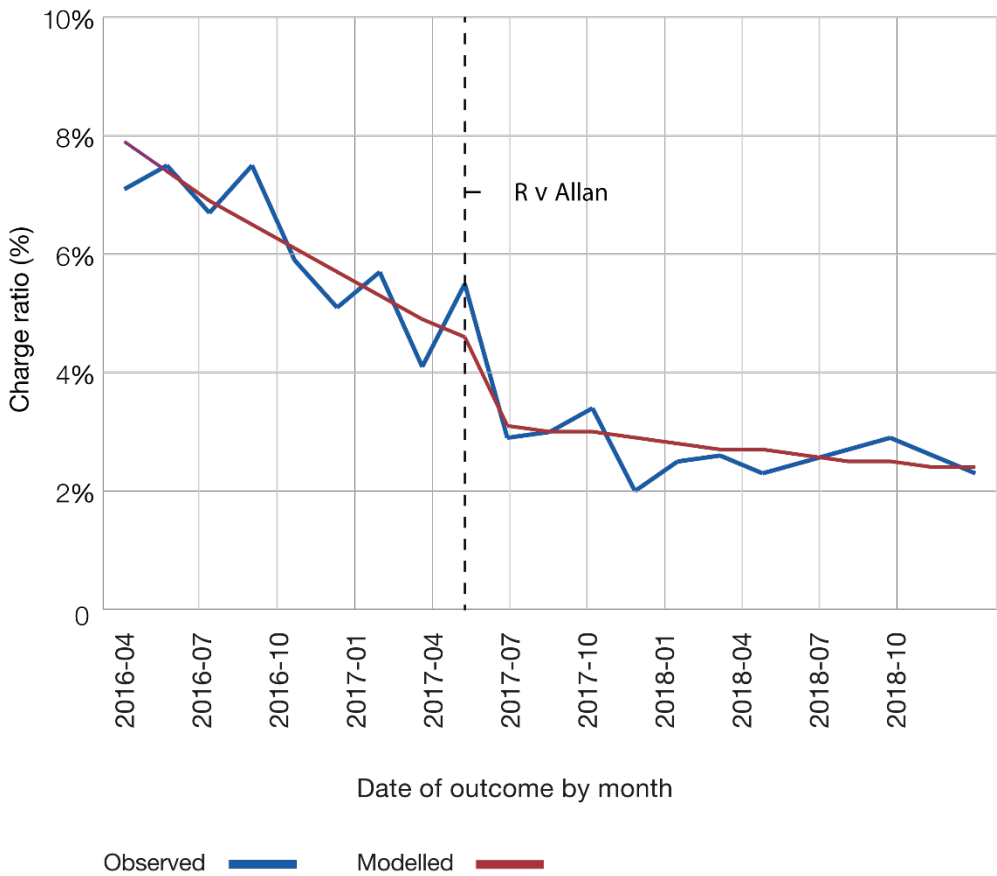
Figure 12: Predicted vs observed charge ratio for rape offences with adult victims

Two non-contact offences (burglary and indecent imagery), where the expected impact of digital evidence is less significant, serve as further controls. Specifically, burglary, an offence category with no – or limited – digital component, does not show signs of a 'disruption' (see Annex B, Figure B1). Possessi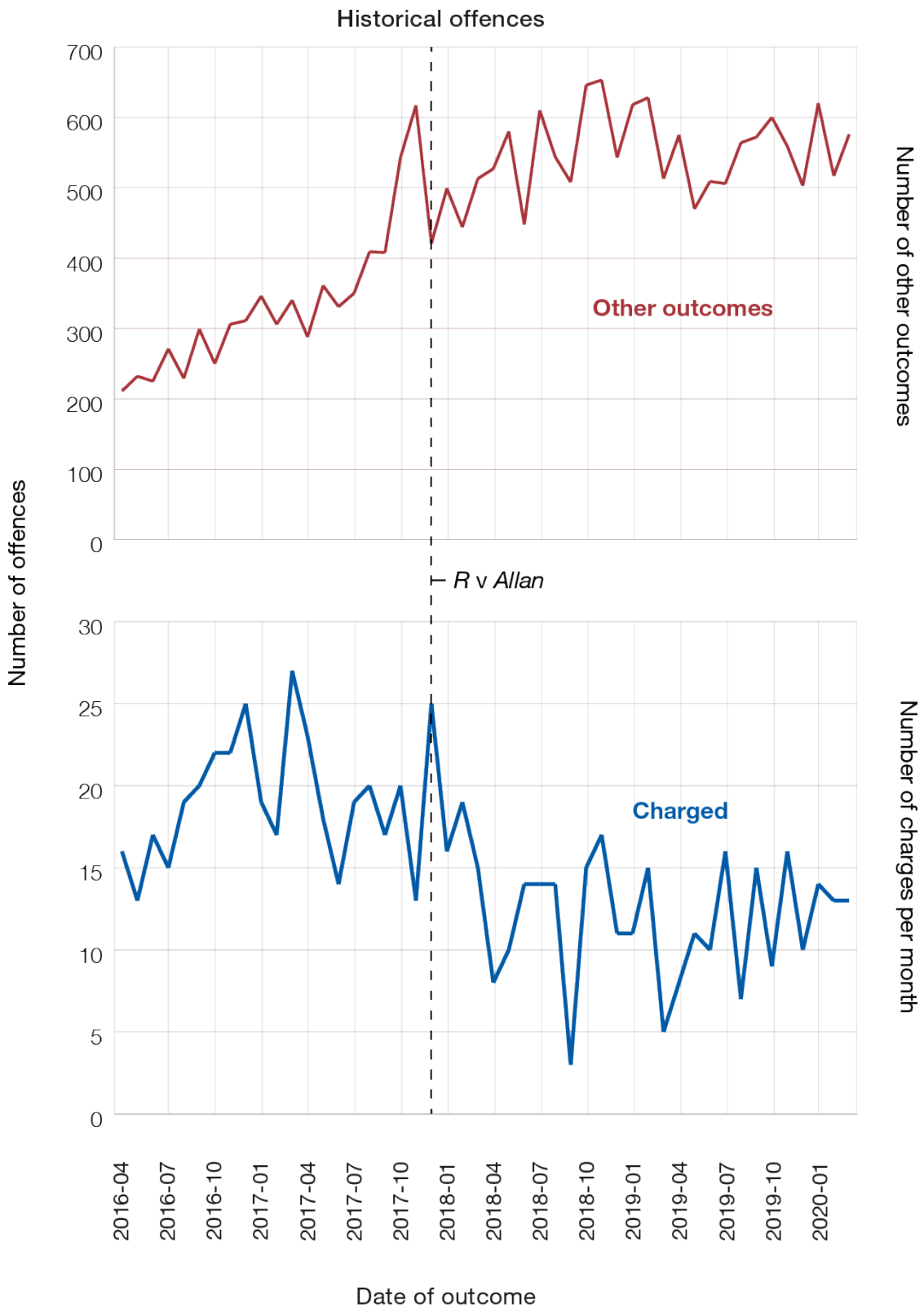on of indecent images, an offence category with an image-technology component, appears to enter a period of mild fluctuation in the six months after the 2017 case without showing a sharp and persistent decline (Figure B2).

---

16   The following predictor variables were used in this model: the date the offence received an outcome (centred around the date of the disruption); the disruption (D); time between offence and recording; CPS area.

5. Discussion

This analysis has explored, using a large administrative dataset of recorded adult rape offences in England and Wales, a range of factors that predict the probability of charge outcomes. Many of the findings echo those from earlier studies on critical factors in determining charge probability (for example, victim–suspect relationship and time between offence and recording). This analysis does, however, address some less well explored areas. The size of the dataset has permitted a more granular breakdown of outputs, particularly around the impact of timeliness in more historic cases. It has also explored predictors for non-charge outcomes, specifically those cases that are closed with the outcome 'victim does not support investigation'. These predictors have shown a high probability of this outcome for partner offences.

The analysis also highlights the importance of examining sudden shifts in charging patterns through granular time series data. In this case, the regression discontinuity analysis adds weight to the suggestion that something changed markedly in the months immediately after a high-profile court outcome involving the failure to disclose digital evidence. The pattern of falling and depressed charge volumes for adult rape offences from the start of 2018 was by no means common across all sub-categories of rape, nor non-contact offences.

The hypothesis – that charges for adult rape offences were adversely affected by this case – seems plausible given the evidence presented here. However, the precise theory of change is far from clear. Early in 2018 – and in response to this and other high-profile cases – the Crown Prosecution Service (CPS) announced that its senior prosecutors were assessing all cases in England and Wales in which someone had been charged with rape or serious sexual assault to ensure that disclosure was being managed effectively, 3,637 cases in total.[17]

This additional scrutiny involved specialist rape and serious sexual offence prosecutors assessing each case to be satisfied that the police had pursued all reasonable lines of inquiry, and that there was a clear strategy for disclosure to be carried out. Prosecutors identified where additional work was required, either to strengthen the prosecution case or to be satisfied that the evidence continued to support the decision to prosecute. Police were frequently asked to conduct further investigations.

This exercise identified 47 prosecutions for rape or serious sexual offences that were stopped in that period and were found to have issues with the disclosure of unused material (1% of cases reviewed). This appears to be a small, but direct impact on the volume of rape cases. However, it may have had wider impacts – in terms of the attitudes and behaviour of police, prosecutors, and victims – around the disclosure and handling of digital evidence during, and after, the exercise. The review itself would also have diverted substantial specialist investigator and prosecutor time away from ongoing rape cases.

The review did not make any new recommendations but did point to existing work to 'support police and prosecutors with their disclosure duties' on the back of the publication

---

[17] See https://www.cps.gov.uk/cps/news/cps-publishes-outcome-sexual-offences-review

of the joint National Police Chiefs' Council (NPCC), CPS and College of Policing (CoP) *National Disclosure Improvement Plan* in January 2018 (NPCC *et al.*, 2018). This sought to address several earlier recommendations around the handling and disclosure of digital and third-party material. Together, all these developments may have contributed to the fall in adult rape charges in 2018. It is also possible that these factors may have simply collided to accelerate issues around the disclosure of digital evidence, which had their roots somewhat earlier (HMCPSI and HMIC, 2017).

Limitations of this analysis

While the dataset contains information on a large number of recorded rape offences in England and Wales, it does not cover all recorded offences. Offence details – including time to outcome, time to recording, victim gender, and victim age – were available for many offences, but almost all variables had some level of missing values. Use of the detailed offence code could sometimes be used to yield proxy information on age band and gender to mitigate this issue. Information on victim–suspect relationships was only available in a consistent and comparable format for 4 out of 37 police forces, although the overall sample size for this 'constrained' dataset is still sizeable compared to many previous studies.

There were some changes in recording practices during the period. For example, a decision was made to ignore an accounting rule change brought in 2016, which labelled a small number of offence outcomes with alternate, lesser offences rather than 'charge'. However, robustness checks suggest that this change would not affect the main conclusions of the analysis.

The dataset only includes a select number of independent variables, covering those that had low levels of missing data. These corresponded with factors that previous studies had identified as significant predictors of charge probability. Other data covering more detailed characteristics on the offence and investigation (for example, the presence of injuries, use of forensics and medical records, availability of specialist sexual assault services) were not covered in the source data.

A limitation of the multilevel analysis was that it did not extend to investigating how the lower probability of a victim being charged in a CPS or police force area was ultimately reflected in the outcome at court.

References

Charlton, C., Rasbash, J., Browne, W. J., Healy, M. and Cameron, B. (2020) *MLwiN Version 3.05*. Centre for Multilevel Modelling, University of Bristol.

Daly, K. and Bouhours, B. (2010) 'Rape and Attrition in the Legal Process: A Comparative Analysis of Five Countries', *Crime and Justice*, 39, pp 565–650.

Feist, A., Ashe, J., Lawrence, J., McPhee, D. and Wilson, R. (2007) *Investigating and detecting recorded offences of rape*, Project Report. London: Home Office.

Hester, M. and Lilley, S. J. (2017) 'Rape investigation and attrition in acquaintance, domestic violence and historical rape cases', *J. Investig. Psychol. Offender Profil.*, 14, pp 175–188.

HMCPSI and HMIC (2017) *Making it fair: A joint inspection of the disclosure of unused material in Crown Court cases*. HM Crown Prosecution Service Inspectorate and HM Inspectorate of Constabulary.

Hohl, K. and Stanko, E. (2015). 'Complaints of rape and the criminal justice system: Fresh evidence on the attrition problem in England and Wales', *European Journal of Criminology*, 12 (3), pp 324–341.

Home Office (2020) 'Crime outcomes in England and Wales 2019 to 2020', *Home Office Statistical Bulletin.* 17/20. London: Home Office.

Imbens G. (2007) *Regression Discontinuity Designs: A Guide to Practice*. National Bureau of Economic Research, Technical Working Paper 337.

Kelly, L., Lovett J. and Regan L. (2005) *A gap or chasm? Attrition in reported rape cases*. Home Office Research Study 293. London: Home Office.

MOPAC (2019) *The London Rape Review: A review of cases from 2016*. London: Mayor's Office for Policing and Crime.

MPS and CPS (2018) *A joint review of the disclosure process in the case of* R. *v.* Allen. London: Metropolitan Police Service and Crown Prosecution Service.

Morabito, M. S., Williams, L. M. and Pattavina, A. (2019) *Decision Making in Sexual Assault Cases: Replication Research on Sexual Violence Case Attrition in the U.S.* Office of Justice Programs' National Criminal Justice Reference Service. Document Number 252689.

Mueller-Smith, M. and Schnepel, K. (2016) *Avoiding Convictions: Regression Discontinuity Evidence on Court Deferrals for First-Time Drug Offenders*. IZA Discussion Paper No. 10409

NPCC (2020) *Police replace processing notice to obtain agreement from victims to search digital devices*. news.npcc.police.uk/releases. National Police Chiefs' Council.

NPCC, CoP and CPS (2018) *National Police Chiefs' Council National Disclosure Improvement Plan*. National Police Chiefs' Council; College of Policing and Crown Prosecution Service.

Ofcom (2020) *Children and parents: Media use and attitudes report*.

Snijders T. A. B. and Bosker, R. J. (2011) *Multilevel analysis. An introduction to basic and advanced multilevel modelling*, second edition. Sage Publishers.

***The Guardian* (Grierson J)** (2017) 'Scotland Yard carrying out "urgent review" after rape trial collapses, *The Guardian*, 15 Dec 2017.

Appendix A: Details of logistic regression modelling results

- Binomial regression for adult rape victims

- Victim–suspect relationships

- Regression with investigation outcome 'evidential difficulties'

- Multilevel model assessing the role of the Crown Prosecution Service (CPS) and police forces

- Regression discontinuity analysis on the impact of the 2017 Crown Court case *R.* v. *Allan*

Binomial regression, all adult victims

Model 1a (including predictor: time from offence to police recording)

Outcome ~ year of offence + victim gender + weekend + time to recording

Outcome: 'Charged' or 'all other outcomes' (reference)

Predictor	Coefficients	Estimate	Std. Error	z value	Pr(>|z|)		Sign.		Reference category
	Intercept	-2.77422	0.044	-63.053	<	2.00E-16	***		
Year	FY1516	0.42427	0.03481	12.189	<	2.00E-16	***	Year offence recorded	FY2016/17
	FY1718	-0.59035	0.03879	-15.218	<	2.00E-16	***		
	FY1819	-1.14287	0.04456	-25.646	<	2.00E-16	***		
	FY1920	-1.40964	0.05107	-27.601	<	2.00E-16	***		
Gender	mf1	-0.87359	0.08009	-10.907	<	2.00E-16	***	Victim gender	Female
Weekend evenings	event_late_wknd	0.20073	0.03047	6.588		4.45E-11	***	Weekend evenings	All other times
Time to recording	time_rec0	0.74366	0.04388	16.947	<	2.00E-16	***	up to 1 day	2 days – 1 week
	time_rec1	0.46995	0.04933	9.527	<	2.00E-16	***	1 – 2 days,	
	time_rec3	-0.27036	0.05298	-5.103		3.35E-07	***	1 week – 3 months,	
	time_rec4	-0.23287	0.05378	-4.33		1.49E-05	***	3 months – 2 years	
	time_rec5	-0.1411	0.0584	-2.416		0.0157	*	2 – 10 years	
	time_rec6	-0.08305	0.06541	-1.27		0.2042		10 years +	

Sign.: Statistical significance: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Model 1b (model including predictors: time to recording and adult age category)

Outcome ~ year of offence + victim gender + weekend + time to recording + age group

Outcome: 'Charged' or 'all other outcomes' (reference)

Predictor	Coefficients	Estimate	Std. Error	z value	Pr(>|z|)		Sign.		Reference category
	Intercept	-2.72333	0.05393	-50.498	<	2.00E-16	***		
Year	FY1516	0.43283	0.03866	11.196	<	2.00E-16	***	Year offence recorded	FY2016/17
	FY1718	-0.60691	0.04267	-14.224	<	2.00E-16	***		
	FY1819	-1.16232	0.04917	-23.638	<	2.00E-16	***		
	FY1920	-1.42428	0.05784	-24.624	<	2.00E-16	***		
Gender	mf1	-0.82563	0.08747	-9.439	<	2.00E-16	***	Victim gender	Female
Weekend evenings	event_late_wknd	0.1821	0.0341	5.34		9.29E-08	***	Weekend evenings	All other times
Time to recording	time_rec0	0.73931	0.04866	15.194	<	2.00E-16	***	up to 1 day	2 days – 1 week
	time_rec1	0.46996	0.05446	8.63	<	2.00E-16	***	1 – 2 days	
	time_rec3	-0.31185	0.05909	-5.278		1.31E-07	***	1 week – 3 months	
	time_rec4	-0.24504	0.05942	-4.124		3.72E-05	***	3 months – 2 years	
	time_rec5	-0.2006	0.06544	-3.066		0.002172	**	2 – 10 years	
	time_rec6	-0.28538	0.07715	-3.699		0.000216	***	10 years +	
Age group	age_grp16–24	0.10903	0.03491	3.123		0.001789	**	16 – 24	24 – 35
	age_grp35–44	-0.22304	0.04871	-4.579		4.68E-06	***	35 – 44	
	age_grp45–54	-0.23247	0.06354	-3.659		0.000253	***	45 – 54	
	age_grp55–100	-0.24956	0.0927	-2.692		0.0071	**	55 +	
	age_grpother	0.60593	0.11176	5.422		5.90E-08	***		

Victim–suspect relationship

Outcome ~ year of offence + victim gender + weekend + time to recording + relationship

Outcome: 'Charged' or 'all other outcomes' (reference)

Predictor	Coefficients	Estimate	Std. Error	z value	Pr(>|z|)	Sign.		Reference category
	Intercept	-2.66362	0.18177	-14.654	< 2.00E-16	***		
Year	FY1718	-0.95232	0.17539	-5.43	5.65E-08	***	Year offence recorded	FY2016/17
	FY1819	-0.92733	0.17581	-5.275	1.33E-07	***		
	FY1920	-1.29043	0.21475	-6.009	1.87E-09	***		
Gender	mf1	-0.66916	0.39654	-1.687	0.09151	.	Victim gender	Female
Weekend evenings	event_late_wknd	0.068	0.16825	0.404	0.68609		Weekend evenings	All other times
Time to recording	time_rec0	0.42986	0.20521	2.095	0.0362	*	up to 1 day	2 days – 3 weeks
	time_rec1	0.05885	0.23646	0.249	0.80345		1 – 2 days	
	time_rec3	-0.36344	0.21566	-1.685	0.09195	.	1 week – 3 months	
	time_rec4	-0.18143	0.22835	-0.795	0.4269		3 months – 2 years	
	time_rec5	-0.71793	0.31554	-2.275	0.02289	*	2 – 10 years	
	time_rec6	0.19802	0.28571	0.693	0.48826		10 years +	
Relationship victim–suspect	rel_catgAquaintance	0.35281	0.17335	2.035	0.04183	*	Acquaintance	Partner
	rel_catgFamily	1.16389	0.34006	3.423	0.00062	***	Family	
	rel_catgStranger	1.76314	0.20609	8.555	< 2.00E-16	***	Stranger	
	rel_catgUnknown	0.55449	0.21643	2.562	0.01041	*	unknown	

Evidential difficulties

Outcome ~ year of offence + victim gender + weekend + time to recording + age group

Outcome: 'Evidential difficulties' or 'all other outcomes' (reference)

Predictor	Coefficients	Estimate	Std. Error	z value		Pr(>|z|)	Sign.		Reference category
	Intercept	0.031	0.02341	1.324		0.185468			
Year	FY1516	-0.10166	0.02168	-4.69		2.74E-06	***	Year offence recorded	FY2016/17
	FY1718	0.16948	0.01922	8.819	<	2.00E-16	***		
	FY1819	0.27373	0.019	14.405	<	2.00E-16	***		
	FY1920	0.59382	0.02017	29.434	<	2.00E-16	***		
Gender	mf1	-0.20278	0.02578	-7.867		3.63E-15	***	Victim gender	Female
Weekend	event_late_wknd	-0.12869	0.01523	-8.451	<	2.00E-16	***	Weekend evenings	All other times
Age group	age_grp16–24	-0.24266	0.01496	-16.222	<	2.00E-16	***	16 – 24	
	age_grp35–44	0.06616	0.01925	3.437		0.000587	***	35 – 44	25 – 34
	age_grp45–54	-0.02631	0.02463	-1.068		0.285497		45 – 54	
	age_grp55–100	-0.37802	0.03563	-10.61	<	2.00E-16	***	55 +	
	age_grpother	-0.21733	0.05055	-4.299		1.71E-05	***		
Time to recording	time_rec0	-0.06516	0.02127	-3.063		0.002188	**	up to 1 day	
	time_rec1	-0.11738	0.02366	-4.961		7.02E-07	***	1 – 2 days	
	time_rec3	0.18679	0.02189	8.534	<	2.00E-16	***	1 week – 3 months	2 days – 3 weeks
	time_rec4	0.2104	0.02242	9.386	<	2.00E-16	***	3 months – 2 years	
	time_rec5	0.17606	0.02484	7.087		1.37E-12	***	2 – 10 years	
	time_rec6	0.01986	0.02852	0.696		0.486292		10 years +	

Multilevel model with CPS and police forces

Multilevel model with predictors time from offence to recording and year.

Outcome: 'Charged' or 'all other outcomes' (reference)

$$\text{outc}_{ijk} \sim \text{Binomial}(\text{denom}_{ijk}, \pi_{ijk})$$

$$\text{logit}(\pi_{ijk}) = \beta_{0jk}\text{cons} + -0.055(0.071)\text{"6"}_{ijk} + 0.790(0.048)\text{"0"}_{ijk} + 0.464(0.053)\text{"1"}_{ijk} +$$
$$-0.267(0.056)\text{"3"}_{ijk} + -0.175(0.057)\text{"4"}_{ijk} + -0.092(0.062)\text{"5"}_{ijk} +$$
$$-1.606(0.044)\text{"1819"}_{ijk} + -0.426(0.035)\text{"1617"}_{ijk} + -1.036(0.038)\text{"1718"}_{ijk}$$

$$\beta_{0jk} = -2.323(0.106) + v_{0k} + u_{0jk}$$

$$\begin{bmatrix} v_{0k} \end{bmatrix} \sim N(0, \Omega_v) : \Omega_v = \begin{bmatrix} 0.103(0.048) \end{bmatrix}$$

$$\begin{bmatrix} u_{0jk} \end{bmatrix} \sim N(0, \Omega_u) : \Omega_u = \begin{bmatrix} 0.036(0.014) \end{bmatrix}$$

$$\text{var}(\text{outc}_{ijk}|\pi_{ijk}) = \pi_{ijk}(1 - \pi_{ijk})/\text{denom}_{ijk}$$

(100632 of 100632 cases in use)

UNITS:

      cps: 13 (of 13) in use

      pfa: 37 (of 37) in use

Regression discontinuity model of the 2017 Crown Court case *R.* v. *Allan*

All adult victims

Outcome ~ time + disruption + time to recording

Outcome: 'Charged' or 'all other outcomes' (reference)

	Coefficients	Estimate	Std. Error	z value		Pr(>|z|)	Sign.		Reference category
	Intercept	-3.17532	0.08384	-37.873	<	2.00E-16	***		
Time	day	-0.00229	0.000406	-5.637		1.73E-08	***	Times in days from disruption	
Disruption	D	-0.38821	0.088593	-4.382		1.18E-05	***	Disruption, 0 before, 1 after 2017 Crown Court case	0
Time to recording	time_rec0	0.664327	0.074198	8.953	<	2.00E-16	***	up to 1 day	
	time_rec1	0.454715	0.083518	5.445		5.19E-08	***	1 – 2 days	
	time_rec3	-0.32014	0.090441	-3.54		0.000401	***	1 week – 3 months	2 days – 3 weeks
	time_rec4	-0.22252	0.090797	-2.451		0.014255	*	3 months – 2 years	
	time_rec5	-0.04228	0.096147	-0.44		0.660146		2 – 10 years	
	time_rec6	-0.15062	0.110207	-1.367		0.171712		10 years +	
Interaction	day:D	0.001432	0.000501	2.855		0.004303	**	Interaction time – disruption	

Historical offences

Outcome ~ time + disruption + time to recording + interaction time-disruption

Outcome: 'Charged' or 'all other outcomes' (reference)

	Coefficients	Estimate	Std. Error	z value	Pr(>|z|)	Sign.		Reference category
	Intercept	-3.44991	0.16571	-20.819	< 2.00E-16	***		
Time	day	-0.00316	0.001071	-2.955	0.00313	**		Day of Crown Court case
Disruption	D	0.180822	0.220636	0.82	0.41247		Disruption, 0 before, 1 after 2017 Crown Court case	0
Time to recording	time_rec6	-0.11286	0.115692	-0.976	0.32928		10 years +	2 days – 3 weeks
Interaction	day:D	0.001196	0.001294	0.924	0.35549		Interaction time – disruption	

Victims under 13 years of age

Outcome ~ time + disruption + time to recording + CPS area + interaction CPS-disruption

Outcome: 'Charged' or 'all other outcomes' (reference)

	Coefficients	Estimate	Std. Error	z value	Pr(>|z|)	Sign.		Reference category
	Intercept	-2.29E+00	2.44E-01	-9.363	< 2.00E-16	***		
Time	day	-7.48E-04	6.02E-04	-1.241	0.21454			Day of Crown Court case
Disruption	D	-1.24E-01	1.98E-01	-0.627	0.53057			0
Time to recording	time_rec0	5.23E-01	2.26E-01	2.315	0.02059	*	up to 1 day	
	time_rec1	4.01E-01	2.50E-01	1.602	0.10908		1 – 2 days	
	time_rec3	6.16E-01	2.29E-01	2.686	0.00722	**	1 week – 3 months	2 days – 3 weeks
	time_rec4	8.90E-01	2.22E-01	4.007	6.15E-05	***	3 months – 2 years	
	time_rec5	8.23E-01	2.10E-01	3.913	9.13E-05	***	2 – 10 years	
	time_rec6	4.36E-01	2.02E-01	2.153	0.03135	*	10 years +	
CPS area	cps0	-9.19E-01	3.90E-01	-2.357	0.01841	*		
	cps1	-3.96E-01	2.21E-01	-1.794	0.07277	.		
	cps3	6.27E-01	1.77E-01	3.547	0.00039	***		
	cps4	-4.11E-01	2.83E-01	-1.452	0.14655			
	cps5	-2.52E-01	2.09E-01	-1.204	0.22865			
	cps6	1.86E-02	2.25E-01	0.083	0.93414			
	cps7	-3.17E-01	2.07E-01	-1.536	0.12463			
	cps8	-3.27E-01	2.20E-01	-1.487	0.1369			
	cps9	-4.94E-02	2.12E-01	-0.233	0.81539			
	cps10	-6.38E-01	2.65E-01	-2.413	0.01582	*		
	cps11	-1.34E-01	1.75E-01	-0.769	0.44188			
	cps12	-1.51E-03	1.99E-01	-0.008	0.99393			
Interaction	day:D	-8.36E-05	7.22E-04	-0.116	0.90781		Interaction time – disruption	

	Coefficients	Estimate	Std. Error	z value	Pr(>|z|)	Sign.		Reference category
Interaction	D:cps0	4.76E-01	4.97E-01	0.958	0.33818		Interaction CPS area with disruption	
	D:cps1	-8.63E-01	3.28E-01	-2.636	0.0084	**		
	D:cps3	-1.56E+00	2.55E-01	-6.118	9.46E-10	***		
	D:cps4	3.55E-01	3.72E-01	0.953	0.34043			
	D:cps5	2.46E-01	2.72E-01	0.904	0.36599			
	D:cps6	-2.50E-01	2.98E-01	-0.838	0.40207			
	D:cps7	-5.08E-01	2.99E-01	-1.7	0.08921	.		
	D:cps8	1.04E-01	3.00E-01	0.348	0.72763			
	D:cps9	8.14E-02	2.75E-01	0.296	0.76724			
	D:cps10	-2.48E-01	3.63E-01	-0.682	0.495			
	D:cps11	-2.73E-01	2.39E-01	-1.141	0.25383			
	D:cps12	1.54E-01	2.73E-01	0.567	0.57091			

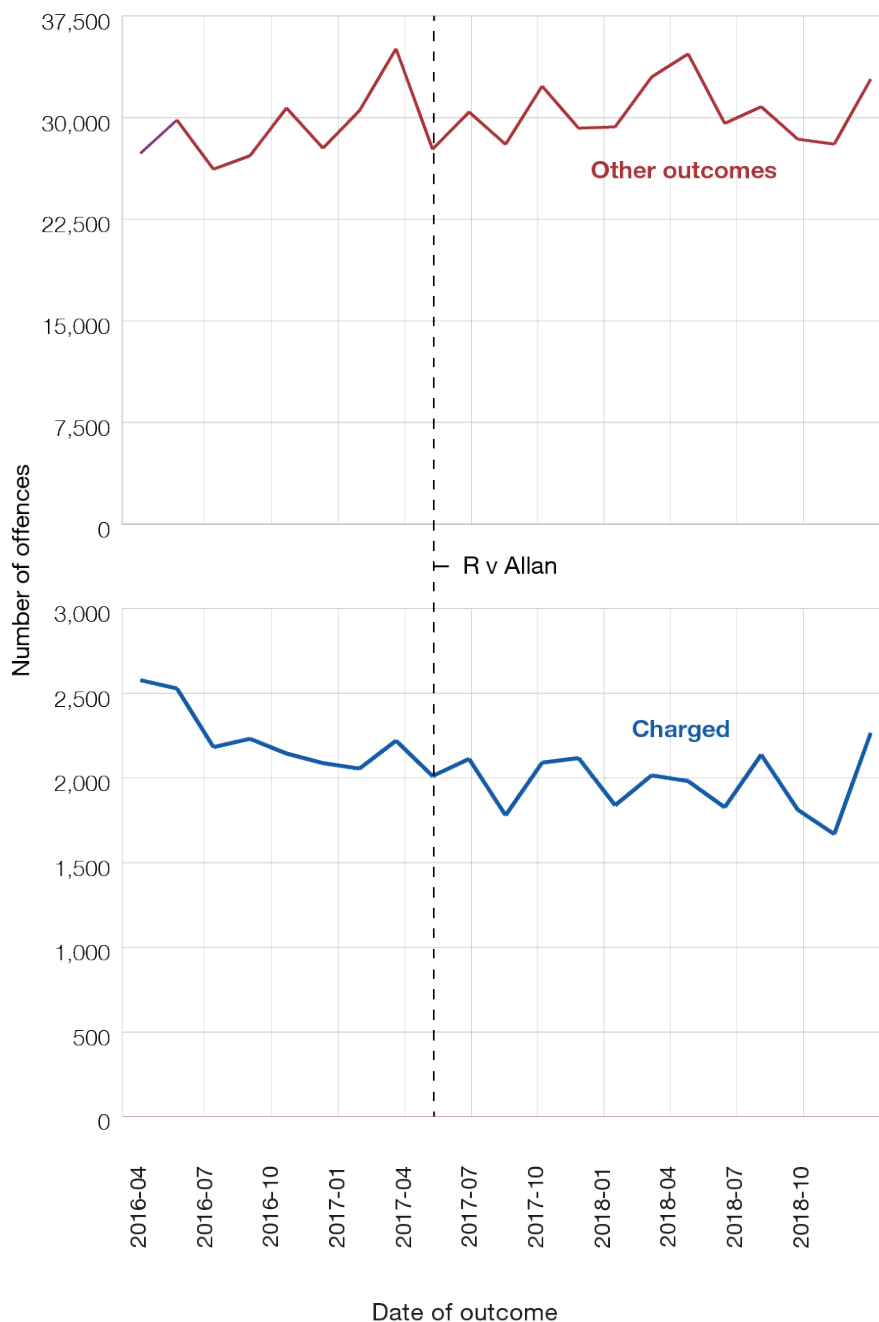# Appendix B: Outcomes for selected comparison offences 2016 – 2020, by month of outcome

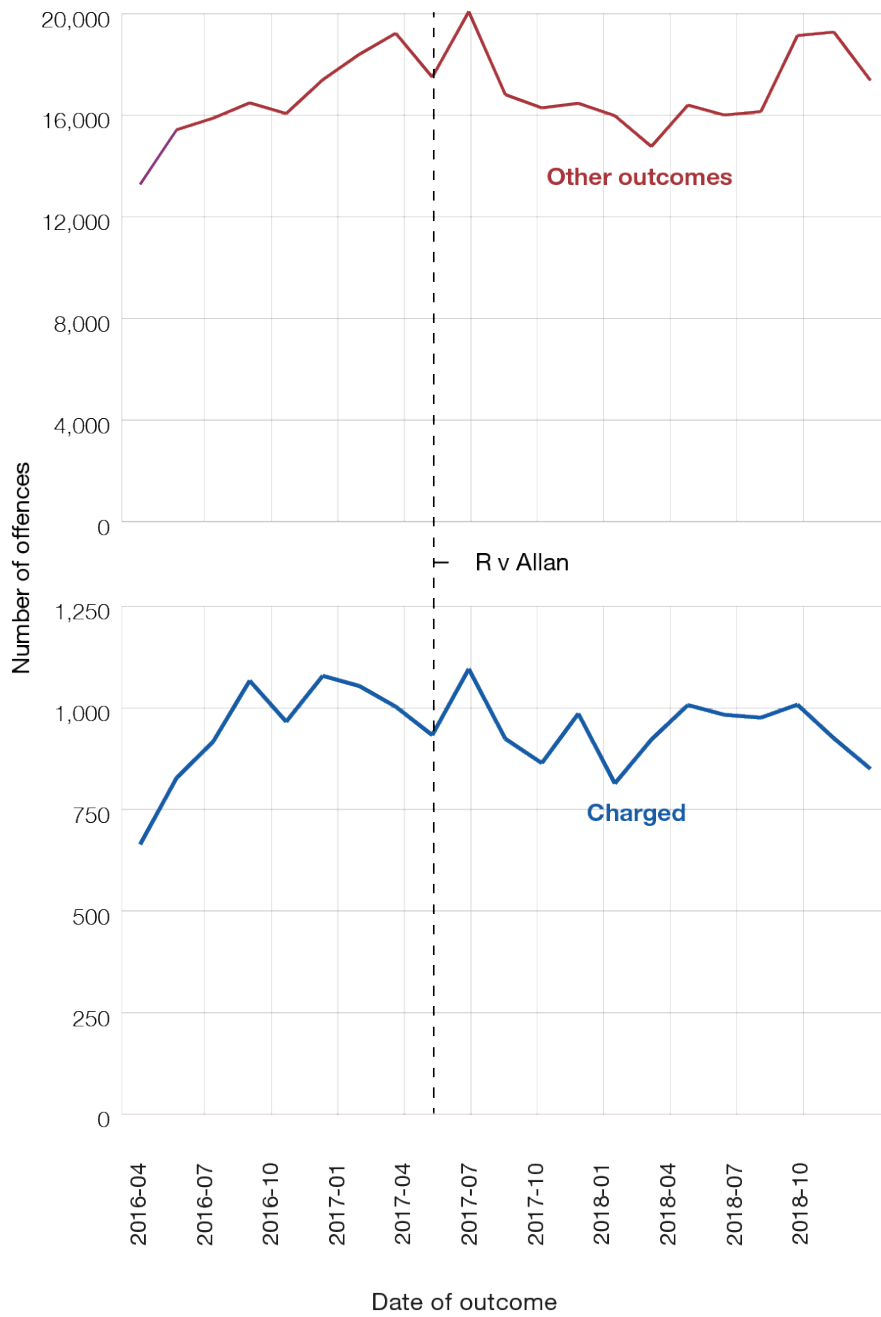Figure B1: Outcomes for burglary offences, by month of outcome

Figure B2: Outcomes for indecent imagery offences, by month of outcome