Disclosure statement

Robbery data

Background

This bespoke data collection exercise was conducted to gather detailed information from sentencers in the Crown Court about how they sentenced robbery, between 1 November 2016 and 28 April 2017. It was undertaken to support the assessment of the Robbery definitive guideline and involved administering surveys to all locations of the Crown Court several months after the guideline came into force.

The data collected information regarding the date and court location of the sentencing decision, some demographic details of the offenders, the relevant factors taken into account when each offender was sentenced for a principal offence of robbery, along with details regarding the offence category, subsequent sentence starting point and the outcome imposed after any reduction for guilty plea. The potential disclosure issues surrounding these data have been carefully considered and the subsequent actions that were taken to ensure that the offenders' confidentiality has been protected are outlined in this document.

Assessment of risk

In accordance with the Government Statistical Service (GSS) guidance, an assessment of the risk of disclosure was undertaken in order to determine which features of the data may help to protect individuals' identities and which aspects of the data may lead to disclosure risks. This is to help prevent identification of individuals and/or prevent additional knowledge of any previously unknown characteristics for those offenders sentenced for these robbery offences in the time period of the data collection.

The following information was used to assess the level of risk:

- the sample size of the survey
- the response rate of the survey
- the likelihood that the data could be linked or matched to other data sources
- the nature and level of detail of the information collected

Applying disclosure control

The following changes were made to the data to reduce the chance of the identification of individuals and thus maintain disclosure control. This included aggregation or banding of variables, where appropriate, in order to reduce unique counts and/or the removal of any variables that could make it easier to identify an individual:

- removal of court location
- original unique case identifiers removed and a new unique defendant identifier created instead
- grouping of offender age and removal of date of birth variable
- removal of the sentencing date
- cleaning and coding of the 'Single most important factor' variable and removal of the original free text answers from the final dataset
- removal of other free text fields where the raw data were deemed too disclosive and were not deemed high enough quality to clean and recode
- removal of any factors with counts fewer than 10

Comments on these data

- The data collection covered all locations of the Crown Court, but the response rate was 48 per cent. This means that over half of the offenders sentenced during the period of the data collection did not have a form filled in for them and are not captured in the data. Therefore, if someone knew some details about an offender sentenced for robbery during the period of the data collection and found a unique record in the data that seemed to match that individual, they could not be certain that the record in the data relates to that specific person. As such, the Council considers that the data are sufficiently anonymous and feels that any further disclosure control would be detrimental to the usefulness of the data.
- The likelihood has been judged as very low that individuals can be identified in the published data through linking to another dataset, given the lack of other publicly available data sources of the same information and that the sample cannot be replicated.
- The names of the courts and the exact date of the sentence have been removed and, as such, it is not possible to identify exactly where the sentencing took place or when any particular form was filled in.
- Where the counts of records were extremely low (fewer than 10), these variables have been removed from the dataset to reduce the risk of positive identification (see the metadata file for further information).
- Banding of specific numeric variables was considered and was conducted for the age variable. However, due to high demand for specific sentence lengths from academic researchers, the decision was made to include the actual custodial sentence lengths given to offenders, rather than to band them, after

balancing the very small risk of identifying individuals with the high usefulness of the data.

- In the raw data, the 'Single most important factor' free text field sometimes included specific details about the offender, offence, location and other details that may have risked being disclosive. To mitigate the risk of any offender being identifiable within the data, the data have been cleaned and recoded into a series of less disclosive factors, with the original answers removed from the published data.

- It is also acknowledged that the cases about which the robbery data relate would have been heard in open court. Although this may increase the risk of identification due to some of the defendant's details being more accessible, the impact is anticipated to be negligible due to the factors listed above. However, it also means that the impact of any potential disclosure issues is minimal as much of this information is already in the public domain.

If users would like any further details of the disclosure controls applied to the data, please contact the Analysis and Research team at Research@sentencingcouncil.gov.uk.