

# Background Quality Report

## Theft from a shop or stall data

### Section 1: Background to the statistics

The Sentencing Council was set up in 2010 and produces guidelines for use by all criminal courts in England and Wales. The *Theft Offences Definitive Guideline* came into force in February 2016 and replaced the Sentencing Guidelines Council (SGC) guideline *Theft and Burglary in a Building other than a Dwelling* (published in 2008) and guidance for theft offences in the *Magistrates' Court Sentencing Guidelines* (MCSG).

The Council has a statutory duty to monitor the impact of the sentencing guidelines it produces. The Council's *Theft Offences Definitive Guideline* included a new definitive guideline for the highest volume theft offence, theft from a shop or stall.<sup>1</sup> The aim of this guideline was to ensure consistency in sentencing practice and no changes in overall sentencing severity were predicted from the resource assessment.<sup>2</sup> In order to evaluate the actual impact once the guideline was in use, an assessment of the *Theft Offences Definitive Guideline* was conducted and then published in February 2019.<sup>3</sup>

To support this assessment, a data collection exercise was conducted pre- and post-guideline in a sample of 81 magistrates' courts, in order to gather detailed information from magistrates and district judges about how they sentenced offences using the *Theft from a shop or stall* definitive guideline.<sup>4</sup> The data were collected in two waves: 16<sup>th</sup> November 2015 – 5<sup>th</sup> February 2016 for the pre-guideline stage and 19<sup>th</sup> September – 16<sup>th</sup> December 2016 for the post-guideline stage. A total of 2,959 valid forms were returned from 81 courts in the pre-guideline stage and 2,417 valid cases involving theft from a shop or stall were returned by 80 courts in the post-guideline stage. Given that the guideline applied to adults only,<sup>5</sup> six cases were excluded from the published pre-guideline data as the date of birth and sentencing date indicated the offender was under 18. This means the total published number of records in the pre- and post-guideline datasets respectively are 2,953 and 2,417. The volume of cases published in the context of total theft from a shop or stall offending is discussed later on.

---

<sup>1</sup> <https://www.sentencingcouncil.org.uk/offences/magistrates-court/item/theft-from-a-shop-or-stall/>

<sup>2</sup> <https://www.sentencingcouncil.org.uk/publications/item/theft-offences-final-resource-assessment/>

<sup>3</sup> <https://www.sentencingcouncil.org.uk/publications/item/theft-offences-assessment-of-guideline/>

<sup>4</sup> Rand Europe administered the survey in the magistrates' courts, they then collated and cleaned the data before providing it to the Sentencing Council.

<sup>5</sup> Offenders under the age of 18 were not in scope of the *Theft Offences Definitive Guideline*.

Both the pre- and post- data collection stages involved asking magistrates and district judges to complete a paper form for every adult offender they sentenced for theft from a shop or stall (where it was their principal offence). The form asked sentencers to give detailed information on the: date of birth and gender of the offender, the sentencing date, the culpability of the offender (including the level of planning, use of force/threat, their role and other culpability factors); the value of goods stolen; other factors relating to harm; the overall levels of culpability and harm; the sentence starting point; aggravating and mitigating factors (including previous convictions); the sentence before any reduction for guilty plea; information on whether there was a guilty plea and if so when it was entered and the reduction applied; and detailed information on the final sentence outcome. The specific court returning the survey was also recorded against each form. They were also given an opportunity to state the single most important factor they took into account when sentencing the offender. This bespoke data collection was the first of several which have provided detailed information on sentencing factors taken into account by sentencers since the Crown Court Sentencing Survey (CCSS)<sup>6</sup> stopped.

The CCSS was a data collection exercise that ran from October 2010 to March 2015. Crown Court judges were asked to fill in a form every time they sentenced an offender, to record the details of the factors they took into account when determining the appropriate sentence. It was a rich source of detailed sentencing data, providing a wealth of information on sentencing for a wide range of offences, from arson to sexual offences, and including theft. For this reason, it was used as one of the key data sources for multiple guideline assessments for offences sentenced within the Crown Court. However, following an external review, the CCSS was ended at the end of March 2015, and the Council evolved its analytical approach to develop more focussed and targeted “guideline-specific” data collection in both magistrates’ courts and the Crown Court.

While there was some detailed analysis of the magistrates’ courts data collection exercise undertaken specifically for the guideline assessment, it is hoped that publication of the raw underlying data will be useful, adding to the knowledge base to better understand magistrates’ courts sentencing factors in relation to outcomes. The intention is also that the Sentencing Council will, in the near future, publish the data from further data collection exercises, collected for a similar purpose from a range of courts for different offences.

The aim of this document is for it to be read alongside the raw data in order to better understand the overall quality of it, in the context of its further use.

When considering the data, it is important to keep in mind that every case is unique and there are many factors, both relating to the offence and the offender’s personal circumstances that will be taken into account when deciding on the appropriate sentence. Therefore, there may be factors other than those collected on the form and detailed in the data that impact on the final sentence. Furthermore, while the same factors may be present in more than one case, the specific circumstances of each case may mean that the factors are not given the same importance in both cases which may, in turn, be reflected in the decision regarding an appropriate sentence for the offender in question.

---

<sup>6</sup> <https://www.sentencingcouncil.org.uk/research-and-resources/data-collections/crowncourt-sentencing-survey/>

## Section 2: Assessment of quality

### i. Relevance

Relevance is about making sure that users of statistics and data are at the centre of statistical production: that their needs should be understood, their views sought and acted on, and their use of statistics supported. Relevance to the user is one of the key principles under the pillar of 'Value' in the Code of Practice for Statistics,<sup>7</sup> so the usefulness of this data has been considered from this user-perspective.

The datasets contain detailed information on the variety of sentencing factors sentencers are asked to consider when using the *Theft from a shop or stall* guideline. These factors may be relevant in determining the type of sentence, the sentence length, any requirements attached to the sentence or the level of any fine imposed. The factors taken into account will vary depending upon the facts of each individual case. Sentencers were also asked to state the 'single most important factor' that they took into account, with regards to the sentencing outcome. A cleaned version of this field has been published alongside the main dataset, and as far as we are aware, this is the first time that data like this will be available in the public domain. This should provide a new insight into the key factors affecting sentence outcomes.

The data also contains some basic demographic data about the offenders (their age and gender)<sup>8</sup>, which could be used to examine how different groups are represented within the data and how factors and sentencing outcomes may vary from one group to another.

It is intended that this data will be useful for any user who wants to better understand magistrates' courts sentencing factors and outcomes for this specific offence. Alongside the publication of this data, the Council has launched a user feedback exercise to gather the views of users on the data and the way the datasets and accompanying documents have been published. This should help the Council to identify the needs of users and make any improvements for future publications of data if it is possible to do so.

Publishing this data contributes to fulfilling one of the Council's responsibilities, of '*promoting awareness amongst the public regarding the realities of sentencing and publishing information regarding sentencing practice in magistrates' and the Crown Court*' as well as one of the Council's additional functions which says it must '*play a greater part in promoting understanding of, and increasing public confidence in, sentencing and the criminal justice system.*'

### ii. Accuracy and Reliability

Accuracy is the proximity between an estimate and the (unknown) true value. Reliability is the closeness of early estimates to subsequent estimated values. This

---

<sup>7</sup> The Code of Practice for Statistics is a document that sets out the standards that producers of official statistics should commit to. The framework for the Code is based on three pillars: Trustworthiness, Quality and Value. This section, 'Assessment of quality', covers elements which span all three of these pillars. The Code can be found here on the UK Statistics Authority website: <https://code.statisticsauthority.gov.uk/wp-content/uploads/2018/02/Code-of-Practice-for-Statistics.pdf>

<sup>8</sup> It was not possible to directly collect ethnicity data in this data collection.

section will provide users with an overview of how accurate and reliable the data are thought to be, by considering possible sources of error and bias.

### ***Sources of error and bias***

There are several types of error that can arise within data such as these, including coverage error, sampling error, non-response error and measurement error. Each of these, including how they may have occurred within the published data and how they have been dealt with (where possible) are described in detail below.

#### ***Coverage error***

Coverage error occurs when the list used to select a sample (the ‘sampling frame’) does not have a one-to-one correspondence with the target population (the total group of units or people that we want to sample from). The Council is confident that the list of courts used to select the sample for this data collection was accurate and included all of the magistrates’ courts open at the time that the sample was selected. The contractors that ran the data collection exercise had direct correspondence with the courts and the lists were checked against multiple sources. The Council is therefore confident that there is no coverage error within the data.

#### ***Sampling error***

Sampling error occurs when a sample is taken, instead of observing the whole population, and where there are differences between estimates generated using the sample and the actual unknown true value for the population.

Theft from a shop or stall is a triable either way offence, meaning that it can be dealt with at either magistrates’ courts or the Crown Court, but this data collection exercise only took place in magistrates’ courts. The vast majority of these offenders are sentenced in magistrates’ courts (99% of 62,000 in 2015 and 98% of 55,000 in 2016<sup>9</sup>). However, the data may be slightly skewed towards lower level offending and sentencing outcomes.

Additionally, these data were collected from a *sample* of magistrates’ courts instead of all of them. The sample of 81 courts represented around half of the courts open at the time of the data collection. Since the data collection did not achieve a 100% response rate from all courts and there are no comparable published sources of data on the key factors used in magistrates’ courts to sentence this offence, there is a risk of the data either being biased or not being representative of all theft from a shop or stall offences and offending. Furthermore, the courts were selected in a specific way to target those with a higher volume of specific drug offences,<sup>10</sup> for which data were being collected at the same time as the theft data collection.<sup>11</sup> This was a deliberate choice to maximise the efficacy of the data collection and volume of expected form returns. However, since the courts were not randomly sampled, there is a possibility

---

<sup>9</sup> See Criminal Justice Statistics quarterly, sentencing data tool: [https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment\\_data/file/895527/sentencing-tool-2019.xlsx](https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/895527/sentencing-tool-2019.xlsx)

<sup>10</sup> There was a specific focus on production of a class B drug offences. However, 75 per cent of the courts in the sample also saw above-average volumes of Shoplifting offences.

<sup>11</sup> The data collected on drug offences will also be published in due course.

that the data are not representative of all magistrates' courts sentencing practice for sentencing theft from a shop or stall.

To check whether analysis of these data could lead to sampling error, a comparison was made with data on theft from a shop or stall from the Court Proceedings Database (CPD), an administrative database of court outcomes for both Crown Court and magistrates' courts held by the Ministry of Justice.<sup>12</sup> Data on a principal offence basis<sup>13</sup> from the equivalent time periods to the pre- and post-guideline data collection exercises were examined, and it was found that the data were broadly representative in terms of both all courts and specifically magistrates' courts sentencing outcomes, with only small differences between the two sources. However, it is worth noting the small differences, as they may affect interpretation of the data.

- The data collection captured slightly smaller proportions of offenders sentenced to immediate custody than there were overall (according to the CPD). In both the pre and post data collections, 19 per cent of offenders were given an immediate custodial sentence, compared to 22 per cent and 23 per cent (respectively) in all magistrates' courts over the same periods. Consequently, the data collection captured slightly higher proportions of offenders sentenced to lower-level outcomes, such as fines and discharges.
- The data collection also captured smaller proportions of offenders committed to the Crown Court for sentence. In the CPD, 2 per cent of offenders found guilty in the magistrates' courts during the same months and years as each pre and post data collection were committed for sentencing to the Crown Court. By comparison, only 0.5 per cent of the sentencing outcomes in both the pre- and post- data collections were 'Committal to the Crown Court for sentence'.<sup>14</sup>

As a result of these small differences, it is possible that the factors ticked in these cases are not wholly representative of all offenders sentenced for theft from a shop or stall. However, as the differences are small, it is expected that the data are largely representative and still useful in identifying, for example, the most and least common factors taken into account and the sentences imposed.

An assessment was also made of how representative the demographics of the offenders contained in the data collection were of the total population of offenders sentenced for theft from a shop or stall. The proportion of offenders of each gender and age group were compared with the equivalent proportions from the CPD. The samples from the data collection were found to be highly representative of the CPD data for both age and gender, which means that users can be confident when using these variables in examining how factors and sentencing outcomes may vary from one group to another.

---

<sup>12</sup> Some of the differences identified may be as a result of non-response error instead of (or as well as) sampling error. Non-response error is discussed later.

<sup>13</sup> When a defendant has been found guilty of two or more offences, the principal offence is the offence for which the heaviest penalty is imposed. Where the same disposal is imposed for two or more offences, the offence selected is the offence for which the statutory maximum penalty is the most severe. Although the offender will receive a sentence for each of the offences that they are convicted of, which will all appear in the CPD, it is only the sentence for the principal offence from the CPD that has been used for comparison here.

<sup>14</sup> This low response rate for committals was picked up during data collection; the research team did clarify with SPOCs that committals should be included, but this communication happened part-way through collection.

### *Non-response error*

There are two types of non-response: in the context of this data collection, ‘unit non-response’ is where a form was not filled in for an offender sentenced for this offence during the data collection period, and ‘item non-response’ is where a form was filled in, but a question or box that should have been completed was left blank, so the non-response was specific to a certain set of items on the form. Where these types of non-response occur, this can lead to error (or bias) in the data.

When the volume of forms returned was compared to the total number of adult offenders sentenced within the same dates as the data collection, this equated to an approximate response rate of 30 per cent in the pre-guideline phase and 26 per cent in the post-guideline phase. If certain types of courts were more or less likely to respond, then this may have affected the data. For example, given that the survey was not supervised at the court level, there is a chance that the administration of the paper forms may have differed between courts, which could introduce bias into the data. If the administration of the forms differed across courts depending on the amount of resource available to distribute and collect the forms, or on the resource of the sentencers to fill in the survey, relating to how busy they were, then response rates may have differed across courts. If the data lead to biased estimates as a result of a form not being completed when an offender was sentenced for this offence during the running of this data collection, then unit non-response error may occur.

Item non-response is another type of non-response which occurred across many of the variables, although it affected some more than others. The variable with the highest proportion of unknown or missing values in the pre-guideline data is sentence ‘starting point’ (21 per cent unknown/missing), closely followed by ‘outcome before guilty plea’ (20 per cent). For the post-guideline data, starting point also has the highest proportion of unknowns, at 45 per cent unknown/missing. Unfortunately, it is not known why the proportion of missing records is so much higher for this variable in the post-guideline data. If the records with unknown or missing data are systematically different to those where clear data have been provided, this could lead to item non-response error.

Aside from the comparison with the CPD data discussed in the ‘Coverage error and sampling error’ section earlier, there is no other relevant source to compare the data collection with and in particular, there is no other source of evidence on the factors taken into account in magistrates’ courts for this offence. It is therefore not possible to measure the extent to which these data may be affected by non-response error. However, there are several reasons why it is thought that non-response error may not be substantial within any analysis of the data:

- The sentencing outcomes were found to be broadly representative of all outcomes imposed for this offence at the time (as detailed earlier), so it could also be assumed that the factors indicated on the forms are also representative,
- A high volume of data was collected, so users do not need to rely on only a small number of offenders to conduct any analysis, and,
- There is no explicit evidence of sentencers being more likely to fill in data collection forms for some types of cases more than for others, so it is assumed that this does not happen.

### *Measurement error*

We have assumed that sentencers have interpreted the form correctly<sup>15</sup> and accurately recorded all the case details, that these have then been accurately interpreted and inputted by the external contractors and accurately cleaned in preparation for publication. However, there is always the chance of human error at each of these stages, and any differences between the true values related to the sentence imposed and the final published dataset are known as measurement error. Furthermore, given the wording of the instructions in the form,<sup>16</sup> if a sentencer did not tick a particular factor then it has been assumed that this particular factor was not taken into account during sentencing. Similarly, if a factor was ticked then it has been assumed it was taken into account. However, this may not be the case and omission as a mistake may have been conflated with omission due to lack of relevance.

There are three variables – ‘No additional harm factors’, ‘No relevant aggravating factors’ and ‘No relevant mitigating factors’ – that have been removed from both the pre- and post-guideline datasets due to concerns that the factors were misinterpreted. In the guideline, harm is categorised into the value of the goods stolen and then all other harm factors are considered to be additional. Sentencers were asked to tick a box when there were ‘no additional harm factors’ they considered during sentencing, which was ticked in 1,359 pre- and 1,309 post-guideline forms. However, in a greater number of forms, the multiple tick boxes for sentencers to indicate that additional harm factors (e.g. a particularly vulnerable victim) had been considered, were left blank (1,671 pre- and 1,353 post-guideline). This indicates that more sentencers had considered no additional harm factors than this variable alone suggests.

Furthermore, there were quite a few instances of direct contradictions, either from sentencers ticking to say ‘no additional harm factors’ were relevant while also indicating that an additional harm factor (e.g. a particularly vulnerable victim) was relevant in their sentencing decision (90 pre- and 143 post-guideline), or from sentencers not ticking any of these additional harm factors but also not ticking the box to indicate that ‘no additional harm factors’ were considered (402 pre- and 187 post-guideline). Similar contradictions were also seen with the variables relating to ‘no relevant aggravating factors’ and ‘no relevant mitigating factors’. Given that users will be able to re-create these factors themselves from the data using the absence of the other factors involved, it was decided that the removal of the potentially misleading data would be the appropriate approach with regards to the accuracy and quality of the overall data.

In recording the custodial sentence outcome, sentencers were asked to record the duration in either days or weeks, and then to indicate which of the two measures they used. It is possible that sentencers used a different unit of measurement to that which they indicated, and therefore our impression of the sentence imposed may be considerably different to the real value. Aside from making sure that the original paper forms were as clear as possible, several steps were taken during the cleaning of the data to mitigate errors of this nature, where possible, which are detailed in the

---

<sup>15</sup> The survey was initially piloted with a handful of sentencers who provided detailed feedback on the structure and content of the surveys, which then fed into revisions to the survey forms.

<sup>16</sup> ‘Please tick all the factors that were relevant, leaving blank any factors that were not relevant or where you have insufficient information to say’.



metadata document. Additionally, we checked that there were no custodial sentence lengths longer than the sentencers had the powers to impose,<sup>17</sup> as this would have also indicated a data error.

The style of questions and the format of the survey may also have contributed to the level of error. The data were collected using paper forms as opposed to being collected digitally (online), so there was not the option to add in any internal assurance processes to flag inconsistent answering within the same form, for example, where a sentencer may have ticked the mitigating factor of the offender not having any previous convictions but then also provided a number for the volume of previous convictions taken into account during sentencing. To improve the data quality, we have applied these types of internal consistency checks prior to publication and amended some records where there were obvious discrepancies. For more details, please see the metadata file.

While free text fields are useful for gathering detailed individualised comments, these take a lot of resource to process and are potentially more prone to misinterpretation, introducing error in the data. To minimise the effect of this, tick-box options were used for most questions and free text fields were only used where necessary. As mentioned above, sentencers were asked to state the ‘single most important factor’ influencing their sentence, and they were given a box to provide a free text answer. The answers to this question varied widely, both in the nature of the answers (the factors mentioned) and in how they were worded/structured. It was not possible to publish the raw answers as they included very 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, and to make the data easier to analyse, the data have been cleaned and provided in a different dataset to the rest of the data. See Annex A for more detail.

### **iii. Timeliness and Punctuality**

The pre- and post- guideline data collections were undertaken between November 2015 and December 2016, either side of the definitive guideline being published and then effective from February 2016. Thus, with regards to the original intention for collecting the data (to monitor the impact of the guideline on sentencing), it captured the change in guideline in a timely way. The same guideline is still in place at the time of publishing, and the Council is not aware of any other policies, legislation or other changes that are likely to have had an impact on sentencing practice since the evaluation of the impact of the guideline was published in February 2019.<sup>18</sup> Thus, the data should still be wholly relevant and useful.

Nevertheless, we recognise that the nature of theft offending and other external factors may have changed since the data collection exercises were undertaken and so the factors that sentencers considered in 2015 and 2016 may not be entirely representative of the current year.

---

<sup>17</sup> This was assumed to be 12 months for immediate custody and 24 months for suspended sentence orders.

<sup>18</sup> The Council decided that the relevant guideline assessment should be published before the underlying data.



#### iv. Accessibility and Clarity

Publishing this information means that the data are made free and equally available to all users. It is thought that this data might be of most interest to an expert user comfortable with processing and manipulating raw datasets. Alongside the raw datasets, we have also assembled a metadata document. This is intended to be read alongside the dataset to understand, for each variable in the data, what the range of values mean and if there any limitations of using this variable to draw conclusions. Examples of the forms completed by the sentencers have also been published, to aid users' understanding of the way the questions were asked, the layout, etc, which may be useful in any analysis.

For the user who still wants to understand the impact of the *Theft from a shop or stall* guideline but is not comfortable analysing data themselves, the Theft Offences guideline assessment discussed previously fulfils this purpose, by utilising the same data source but providing additional narrative around the analysis comparing the pre- and post-guideline periods.

One of the challenges has been ensuring that the data are published at a sufficient level of detail to enable users to sufficiently delve into the individual factors behind magistrates' court sentencing decisions, while still taking steps to reduce the risk of disclosure for the individual offenders as much as possible. There is a disclosure statement published alongside the data itself, and further details can be found in the metadata document.

#### v. Coherence and Comparability

Coherence and comparability are the degrees to which data derived from different sources or methods, but that refer to the same topic, are similar, and the degrees to which data can be compared over time and domain.

##### *Comparability of the pre- and post-guideline data*

The pre- and post- data collections were both administered in the same way, using paper forms which were filled in by magistrates and district judges. The forms were designed to capture the change in guideline so there are some necessary small differences between the two forms (for example in the offence categorisation),<sup>19</sup> but generally speaking, the two forms are highly comparable. It is therefore unlikely that any changes in sentences observed between the two datasets are related to differences in the data collection methods.

The variable 'Outcome before Guilty Plea' was a free text box on the pre-guideline data form whereas the post-guideline form had preset tick boxes. This meant that both variables had to be treated differently in preparation for publication. The pre-guideline field underwent thematic analysis to manually yet systematically interpret the contents and recode them into similar categories to the final sentencing outcome and to the categories in the post-guideline data. Nevertheless, the response rate across both forms was similar, with the proportion of unknown/missing values 16%

---

<sup>19</sup> The pre-guideline form asked sentencers to select an offence category from 1 (least serious) to 4 (most serious) whereas the post-guideline form separated out Culpability (A, high to C, lesser) and Harm (category 1 to 3).

pre- and 20% post-guideline. For more details, please see the metadata file for this variable.

One of the aims of the *Theft Offences Definitive Guideline* was to emphasise the impact on the victim in Step one. It also contained specific monetary values within the harm model whereas the previous guideline only referenced value as ‘low’ (under offence seriousness) and ‘high’ (as an aggravating factor). Despite this difference, the pre- and post- guideline data forms both contained tick boxes asking sentencers to indicate the total value of the goods stolen. The only difference was that the post-guideline form split ‘Up to £10’ and ‘£11 to £50’ into separate categories, whereas the pre-guideline form had a single category of ‘Up to £50’.

Also included in the data are some demographic details of the offenders being sentenced for an offence of theft from a shop or stall. The proportion of records with a female offender recorded are comparable across both pre- and post- data (27 and 28 per cent respectively). Additionally, the recorded ages of the offenders<sup>20</sup> are very similar across both datasets. The highest proportion of offenders (40 and 42 per cent respectively) were recorded in the 30 to 39 age bracket.

In the data there are some variables that we have had to remove due to disclosure concerns from either the pre-, post- or both datasets. We recognise that this may mean that a small number of variables cannot be compared across the two datasets, but the decision was made to limit undue disclosure risk in counts of fewer than 10 records. Given that small counts are unlikely to yield strong conclusions anyway, it is not believed this decision will overall impede the usefulness of the data. For more information about disclosure control, please see the disclosure statement.

Any differences between the pre- and post- datasets in terms of the factors ticked by sentencers and the implications this has in terms of interpreting the true impact of the *Theft from a shop or stall* guideline are discussed and analysed in the Theft Offences guideline assessment, published on the Council’s website.

#### *Comparability with existing analysis using the same data*

The data being published were used as one of the sources for the Theft Offences guideline assessment; however, they have undergone further cleaning and internal quality assurance in preparation ready for publication. As a result, while we believe they should still be analogous, some very small differences may exist. Specifically, in the pre-guideline data only, we are aware that the culpability factor ‘Little or no planning’ was recorded as ticked in 67 per cent of forms in the evaluation but is now recorded as relevant in 66 per cent of cases in the published data. Similarly, the culpability factor ‘Leading role’ was recorded in the evaluation as relevant for 16 per cent of forms but is present in 17 per cent of cases in the published data.

#### *Comparability with other data sources*

As far as the Council is aware, there are no other data sources available on magistrates’ courts sentencing practice for the offence of theft from a shop or stall that contain both the factors taken into account by sentencers and details of the sentences imposed. However, the Council collected similar data at the Crown Court

---

<sup>20</sup> Calculated using offender date of birth and date of sentencing.

as part of the CCSS between October 2010 and the end of March 2015, and this is published on the Council's website.

Although the types of factors taken into account may generally be comparable across the two data sources, there are several reasons why differences between the two would be expected. The CCSS data were collected during a different time period, were only collected at the Crown Court (when most offenders sentenced for this offence are sentenced at magistrates' courts) and the forms themselves were different. Users who intend to compare the two sources should bear these differences in mind when interpreting any results.

For further information about this data, please contact the Analysis and Research team at [Research@sentencingcouncil.gov.uk](mailto:Research@sentencingcouncil.gov.uk).

## Annex A: Single Factor Analysis

As mentioned above, sentencers were asked to note the ‘single most important’ factor considered while passing the sentence. This information was provided in a free text field which has been cleaned and coded internally within the OSC<sup>21</sup> using thematic analysis and provided in a separate dataset, with a unique identifier allowing it to be mapped onto the main pre- and post-guideline datasets.

The data were studied manually to draw out key themes (for example, it was noticed that previous convictions were mentioned regularly, and so this was noted as a key theme). These key themes were then coded, so that they could be picked out automatically by analytical software, including variations in the way factors were worded. For example, where the phrases “precons”, “pre cons”, “previous offences”, “antecedents”, “significant previous” and other similar phrases were mentioned within the free text box, these were all grouped together to form a new factor called ‘Precons’. Each new variable is binary, with a value of 1 indicating the phrase in question was mentioned and 0 indicating it was not.

Almost 90 per cent of records pre- and post-guideline had the ‘Single most important factor’ variable populated and of these, 97 per cent have been assigned at least one key theme. In some cases, sentencers highlighted more than one factor, where this was the case, the separate factors have been identified and individually coded. Checks were put in place to identify possible mistakes, e.g. to try to ensure that where the sentencer stated that there were no previous convictions, that this was coded as ‘No precons’ and not as ‘Precons’, or that where the offender had failed to engage with probation services, that this was coded as ‘Lack of engagement’ rather than ‘Engagement’. However, it is possible that the accuracy of the analysis has been affected by this semi-automated process, since although a sample of records were manually checked to make sure that they were being coded correctly, it was not possible to do this for every record. Any differences between the single most important factor as intended by the sentencer and the final published dataset is a source of measurement error.

There were 45 key themes identified within the dataset which have been broadly attributed to the key stages of the sentencing process: culpability, harm and aggravating and mitigating factors. Where key themes have been identified but cannot be attributed to a specific stage, they have been identified as ‘Other’. These themes are not an all-inclusive list but represent the main codes identified by the analyst; therefore, these themes and groupings should be interpreted with some caution. Further details on the methods used for assigning and grouping the themes and the possible issues associated with the method used are given within the metadata document.

It was not possible to apply the coding methodology to all records. As such, around three per cent of both pre- and post- data records have not been assigned a key theme due to the nuanced nature of these records.

This analysis has been conducted as an iterative process, by engaging fully with the data, generating and searching for themes, reviewing the themes and logic behind

---

<sup>21</sup> OSC analysts, policy and legal staff with expert knowledge of the criminal justice system all contributed to this work; feeding in key analysis and conducting peer review and quality assurance of the process.

the coding and conducting peer reviews and quality assurance. However, due to the subjective nature of thematic analysis and the semi-automated process used, care should be taken when interpreting this data. We hope that this methodology can provide a basis for further research whilst considering the potential limitations of using thematic analysis.