![Crime Survey for England & Wales logo]

KANTAR PUBLIC

![Office for National Statistics logo]

Crime Survey for England and Wales

Technical Report 2020/21

Volume One

Contents

1.   Background

1.1   Introduction to the Crime Survey for England and Wales

The Crime Survey for England and Wales (CSEW) is a well-established study and one of the largest social research surveys conducted in England and Wales. The survey was first conducted in 1982 and ran at roughly two-yearly intervals until 2001, when it became a continuous survey[1]. Prior to this change, respondents were asked about their crime-related experiences in the previous calendar year; but when the CSEW changed to a continuous survey, respondents were asked about crime in the 12 months prior to interview (more information on the time periods covered can be found in section 2.4 of the user guide[2]).

Prior to April 2012, the survey was known as the British Crime Survey (BCS) and conducted on behalf of the Home Office. From April 2012 responsibility for the survey transferred to the Office for National Statistics (ONS) and it became known as the Crime Survey for England and Wales (CSEW). Since 2001, Kantar Public has been the sole contractor for the survey.

The CSEW is primarily a survey of victimisation in which respondents are asked about their experiences of both household crimes (e.g. burglary, vehicle crime) and personal crimes (e.g. robbery, snatch theft). Household crimes may have happened to anyone in the household, while personal crimes are only counted if they relate to the individual being interviewed. The traditional reference period for all interviews prior to 2020-21 relates to incidents that have happened in the last 12 months before the date of interview. Although there have been changes to the design of the survey over time, the wording of the screener questions that are asked to elicit respondents' experiences of victimisation have been consistent over the lifetime of the survey. In 2015-16 an additional set of screener questions was added to measure fraud and cybercrime.

Respondents are asked about their experience of crime, irrespective of whether they reported these incidents to the police. As such, the CSEW provides a record of peoples' experiences of crime which is unaffected by variations in reporting behaviour of victims or variations in police practices of recording crime.   The CSEW and police recorded figures are two complementary series, which together provide a better picture of crime than can be obtained from either series alone.

Since the survey became continuous in 2001 there have been a few significant changes to the design of the survey. Where changes have been incorporated these have been described in detail in the relevant technical reports. The most significant changes prior to 2020-21 include:

– Between 2004-05 and 2011-12, the core sample size was increased from 37,000 to 46,000, with a target of at least 1,000 interviews in each Police Force Area (PFA).

– Long-standing boost samples of Black and Asian respondents (3,000 sample boost per year) and 16 to 24 year olds (2,000 sample boost per year) were dropped in 2006-07 and 2008-09 respectively.

– In 2009-10, after an extensive development period, the survey was extended to cover young people aged 10 to 15 with a target sample size of 4,000 per year (reduced to

---

[1] Previous sweeps of the British Crime Surveys were carried out in 1982, 1984, 1988, 1992, 1994, 1996, 1998 and 2000.

[2] User guide to crime statistics for England and Wales:

https://www.ons.gov.uk/peoplepopulationandcommunity/crimeandjustice/methodologies/userguidetocrimestatisticsforenglandandwales#crime-survey-for-england-and-wales-csew

3,000 from 2012-13 onwards)[3]. The first results for this age group were published in June 2010[4] as experimental statistics and estimates of victimisation among children have traditionally been presented alongside the adult crime statistics.

– In 2012-13 the core adult sample size was reduced from 46,000 to 35,000. In the same year a new sampling approach was adopted based around a three-year un-clustered sample design.

In 2015-16 the questionnaire was updated to include measures of fraud and cybercrime following an extensive development phase, including a large-scale field test. A methodological note about the development of the fraud measures and the field trial was published in 2015 and the questions were put on the survey from October 2015[5].

In 2020-21 the Covid-19 pandemic necessitated the largest single change in the history of the CSEW when face-to-face interviewing was suspended on 17th March 2020, with no certainty about when it would resume. While this has little impact on the 2019-20 survey beyond stopping fieldwork slightly earlier than the normal 30th March end date, it meant the 2020-21 survey had to be re-designed to accommodate the suspension of face-to-face fieldwork.

1.2   Introduction to the Telephone Crime Survey for England and Wales (TCSEW)

When it became clear that Covid-19 would necessitate the indefinite suspension of all face-to-face fieldwork across the UK, work began to move the survey to a telephone approach (TCSEW), with the first telephone interviews being conducted on the 20th May 2020.

Details of the changes that this necessitated are covered in subsequent chapters, but the key changes are summarised below:

– It was clear that the switch to telephone would require a shorter average interview length both to help maximise the overall response rate and reduce respondent burden. The self-completion modules on drugs and intimate personal violence were removed as these were too sensitive to ask over the telephone. The rest of the questionnaire was streamlined to include only six modules – five existing modules and a new, Covid-19 specific module. The aim was to have an average interview length of 30 minutes compared with an average of 50 minutes on the face-to-face survey. This is discussed in more detail in Chapter 3.

– While most modules remained the same during 2020-21, the Covid and Demographics modules were more fluid, with both being updated as the impact of the pandemic changed across the year. Whereas questionnaire changes on the CSEW happened quarterly, the TCSEW was set up to enable monthly changes as and when needed.

– The face-to-face random sample design was set aside in favour of a re-contact sample based on those who had previously taken part in the face-to-face survey between May 2018 and February 2020 and had expressed a willingness to be re-contacted again. This is discussed in more detail in Chapter 2.

– Given that the available re-contact sample was finite, it was recognised that the sample would be insufficient to sustain the TCSEW throughout 2020-21. It was therefore agreed that all respondents who completed the initial telephone survey (and were willing to take

---

[3] A feasibility study was carried out before the survey was extended to this age group. See Pickering, K., Smith, P., Bryson, C. and Farmer, C. (2008) British Crime Survey: options for extending the coverage to children and people living in communal establishments. Home Office Research Report 06. London: Home Office.

[4] Millard, B. and Flatley, J. (2010) Experimental statistics on victimisation of children aged 10 to 15: Findings from the British Crime Survey for the year ending December 2009. Home Office Statistical Bulletin 11/10.

[5] CSEW Fraud and Cyber-crime Development: Field trial

part again) would be re-contacted at intervals of approximately three months, thereby creating a panel design somewhat similar to the Labour Force Survey[6].

– In follow-up interviews, respondents were asked about incidents that had happened since their last telephone interview, rather than the normal last 12 months. This is discussed in more detail in Chapter 3.

– Given the survey design it was agreed that the survey would not include any survey of 10-15 year olds, which has been a standard part of the CSEW since 2009-2010. It was also decided that the TCSEW would be a survey of adults aged 18 or over rather than of those aged 16 or over as is the case of the CSEW. Again, this is discussed in more detail in Chapter 2.

1.3   Outputs from the CSEW

Following the move of the processing and publication of crime statistics to ONS from the Home Office, the standard quarterly releases were extended to include more long-term trends and other data sources.

In addition to the regular quarterly publication, ONS publish additional thematic publications and articles on particular aspects of crime. Recent examples of thematic reports and articles based on CSEW data include:

- Domestic abuse in England and Wales: November 2020

- Sexual offences in England and Wales overview: year ending March 2020

- Online bullying in England and Wales: year ending March 2020

- Child abuse in England and Wales: year ending March 2020

- Drug misuse in England and Wales: year ending March 2020

The publications mentioned above are intended only to illustrate the types of reports and findings that are produced from the CSEW. Full details of all publications associated with the CSEW, and crime statistics more generally, can be found on the ONS website[7].

As well as published reports, anonymised CSEW data is made available through the UK Data Archive at the University of Essex[8] and through the ONS Secure Research Service[9]. The CSEW is a complex study with data organised at different levels (households, individuals, and incidents) and it includes numerous sub-samples who are asked specific questions. Accordingly, considerable effort and expertise is required to analyse the data and to interpret it in a valid manner. Some of the analysis routines that play a key role in the published estimates are implemented after the data have been supplied to the ONS and so are not documented in this report. Further information on how to use the data is available from the UK Data Service[10].

ONS also produces a User Guide for those interested in understanding CSEW data and outputs which contains further detail on the content and structure of the data[11].

1.4   Outputs from the TCSEW

---

[6] https://www.ons.gov.uk/surveys/informationforhouseholdsandindividuals/householdandindividualsurveys/labourforcesurvey

[7] https://www.ons.gov.uk/peoplepopulationandcommunity/crimeandjustice

[8] https://www.data-archive.ac.uk/

[9] https://www.ons.gov.uk/aboutus/whatwedo/statistics/requestingstatistics/approvedresearcherscheme/

[10] https://www.ukdataservice.ac.uk/

[11]

https://www.ons.gov.uk/peoplepopulationandcommunity/crimeandjustice/methodologies/userguidetocrimestatisticsforenglandandwales

While the switch to telephone-based interviewing has ensured that experiences of both household and personal crimes are still being captured during the pandemic, these estimates cannot be directly compared with those previously published from the face-to-face CSEW[12]. However, TCSEW annual estimates of total crime are similar to levels estimated by the face-to-face CSEW.

Furthermore, because of the change in data collection mode, and the time needed to instigate the revised approach, there is a break in the CSEW/ TCSEW data time series to reflect the lack of interviewing between March-May 2020.

1.5   Structure of the Technical Report

This report documents the technical aspects of the 2020-21 TCSEW. The analysis in this report relates to the total sample that was issued in the financial year 2020-21, irrespective of when interviews took place. The distinction between issued sample and achieved sample is explained in more detail in Chapters 4 and 5 of this report. Despite the fundamental differences between the TCSEW approach and the more traditional face-to-face survey, the basic structure of this technical report will mirror previous years.

As such, the sample design is set out in Chapter 2. Data collection is the major task for the organisation commissioned to conduct the TCSEW and forms the central part of this report. Chapter 3 covers the content and development of the questionnaire, while Chapter 4 details our fieldwork procedure (including response rates, documents and quality control) and Welsh fieldwork. Chapter 5 discusses response rate and reasons for non-response in the core sample. Chapter 6 gives details of the tasks that are involved in preparing the data for analysis, including the coding and offence classification and Chapter 7 covers the preparation and delivery of the TCSEW data files. Chapter 8 outlines the weighting required for analysis of the data. Chapter 9 provides the results of some checks on the profile of the TCSEW achieved sample against estimates for the population that the TCSEW aims to represent.

---

[12] An assessment into the comparability of estimates produced from the face-to-face CSEW and TCSEW:
https://www.ons.gov.uk/peoplepopulationandcommunity/crimeandjustice/methodologies/comparabilitybetweenthetelephoneope ratedcrimesurveyforenglandandwalesandthefacetofacecrimesurveyforenglandandwales

2.   Survey design

2.1   Introduction

In May 2020, the CSEW was replaced with the TCSEW, a simpler variant based on telephone interviews rather than face-to-face interviews. The original design for the TCSEW was intended to cover the nine month period May 2020 through January 2021 inclusive. However, in late 2020, ONS made the decision to extend it and the TCSEW will now run until March 2022. This technical report covers the period May 2020 through to March 2021.

2.2   Populations of inference

The TCSEW populations of inference are (i) private residential households in England and Wales, and (ii) adults aged 18+ living in these households. The CSEW also covered 10-17 year olds but this population was omitted from the TCSEW for reasons of practicality. The TCSEW sample frame - CSEW respondents who had given permission to be recontacted – did not include any 10-15 year olds and also (quite naturally) under-represented people who were aged 16-17 in 2020.

2.3   Sample frame

The sample frame comprised all CSEW respondents from May 2018 through February 2020 inclusive who had given permission to be recontacted by Kantar on behalf of ONS. This permission ran to 24 months so nobody interviewed before May 2018 could be included in the sample frame. This rule also affected *when* some CSEW respondents could be issued for the TCSEW. CSEW respondents from May 2018 could only be issued in May 2020; CSEW respondents from June 2018 had to be issued by June 2020 at the latest; and so on.

Some CSEW respondents had been aged 16 or 17 at the time of interview but the majority of these would be aged 18+ by the time of issue for TCSEW. Consequently, they were not excluded from the frame, although any that were aged 16 or 17 when interviewed for the TCSEW were given a survey weight of zero.

In total, the sample frame comprised 42,702 individuals from an original respondent set of 63,139. This original respondent set is called the Reference Sample and is a critical tool for weighting TCSEW data

2.4   Survey design

The TCSEW utilised a panel survey design similar to that used for the Labour Force Survey, albeit for reasons of sample conservation rather than to benefit measurement.

Three months after the first interview (W1), each TCSEW respondent was eligible for a second interview (W2). Three months after that interview, each respondent was eligible for a third interview (W3). Anyone who responded to W1 but did not respond to W2 was eligible for a W3 interview six months after the W1 interview. For cost efficiency reasons, non-respondents to W1 were not issued for W2 or W3.

Because the initial plan was to cover only the period May 2020 through January 2021, the maximum number of interviews was expected to be three. In the event, the plan was extended to cover through to the end of 2021 and the maximum number of interviews was revised upwards to a maximum of seven. However, the sample issue design was based on the *initial plan only*.

2.5   Sample issue design

The sample frame was divided into nine differently sized *replicates* of the whole. One replicate was issued to the field for the first time each month from May 2020 through to January 2021.

For each replicate, a three-month fieldwork period was planned, with the expectation that 70% of interviews would be achieved in month 1, 20% in month 2, and 10% in month 3[13]. The relative conversion rates for W1, W2 and W3 were also estimated: W1 = 40%; W2|W1 = 70%; W3|W1 = 57%.

Based on these assumptions, the size of each replicate was set to *minimise the expected variation between calendar months with respect to the total number of interviews* (as defined by data collection date). Table 2.1 shows the intended size of each replicate.

Table 2.1 TCSEW Intended issued sample sizes

Replicate: issue month	Number of cases in frame
1: May 2020	12,219
2: June 2020	8,728
3: July 2020	7,982
4: August 2020	2,707
5: September 2020	2,606
6: October 2020	2,533
7: November 2020	2,092
8: December 2020	1,955
9: January 2021	1,880
Total	**42,702**

To divide the sample frame into replicates, it was first stratified by (i) CSEW interview month (May 2018 through February 2020), (ii) NUTS1 region, (iii) a six category sex/age variable (male 16-29, male 30-59, male 60+, female 16-29, female 30-59, female 60+), and finally (iv) CSEW victimisation status (non-victim, victim of non-fraud crime, victim of fraud only).

Each case in each stratum was then allocated to a replicate (TCSEW issue month) using a systematic random sampling method that accounted for (i) the different intended size of each replicate, and (ii) the issue date constraints affecting CSEW respondents from May 2018 through December 2018.

2.6   Revisions to the sample design

Shortly after replicate 2 was issued to the field (June 2020), ONS requested that several thousand unissued CSEW respondents be redirected from the TCSEW to the Covid Infection Survey (CIS).

In total, 8,402 cases were systematically sampled for the CIS from across replicates 3 to 9. These cases were sampled from each replicate to minimise the variation in expected interview numbers between the calendar months July 2020 through January 2021.

Of the 8,402 cases sampled for the CIS, 4,305 were eligible for use in the CIS (agreed to be recontacted by a different organisation than Kantar). The remaining 4,097 cases were initially excluded from both the CIS and the TCSEW but were restored to the TCSEW sample frame in time for the issue of replicate 4 (August 2020). As before, this allocation was implemented to minimise the expected variation in the number of interviews per remaining calendar month.

---

[13] In the event, this was reduced to two months from the August 2020 (W1) issue onwards.

Table 2.2 shows the actual issued sample sizes for each replicate after revisions. In total, 38,397 CSEW respondents were issued for the TCSEW.

Table 2.2   TCSEW Actual issued sample sizes

Replicate: issue month	Number of cases in frame (+ initial allocation)
1: May 2020	12,219
2: June 2020	8,728
3: July 2020	6,985 (7,982)
4: August 2020	1,957 (2,707)
5: September 2020	1,928 (2,606)
6: October 2020	2,289 (2,533)
7: November 2020	1,579 (2,092)
8: December 2020	1,459 (1,955)
9: January 2021	1,253 (1,880)
Total	38,397 (42,702)

3.   Questionnaire content and development

3.1   Structure and coverage of the adult questionnaire

As discussed in the opening chapters, the TCSEW relies on a wave formation, with respondents taking part in a maximum of three waves between May 2020 and March 2021.

While the basic structure of the questionnaire was consistent across all these three waves, there were some structural differences between wave 1 and subsequent waves. Additionally, the content of the questionnaire varied throughout the year, in part to reflect the impact that the pandemic was having at specific periods of the year, such as the easing and subsequent tightening of lockdowns.

With this in mind, this chapter looks at the questionnaire content of wave 1 and waves 2-3 separately, as well as looking at key changes in the questionnaire that took place up to and including March 2021.

3.2   Overview of the adult questionnaire (wave 1)

The wave 1 TCSEW questionnaire was a cut-down version of the original CSEW version.

Only five of the existing modules were retained, although all of these were modified to some extent to reflect the switch from face-to-face, in-home interviewing to telephone. All self-completion modules were removed, as were any modules asked only of random-sub samples.

The abbreviated 2020-21 TCSEW questionnaire therefore consisted of the following modules:

- – Household box

- – Screener questionnaire

- – Victimisation module for non-fraud incidents identified at the screeners (up to a maximum of six)

- – Victimisation module for fraud incidents identified at the screeners (up to a maximum of six, including the non-fraud incidents)

- – Covid module, including questions related to the impact of Covid on children aged 10-15

- – Demographics module

The basic structure of the core questionnaire is shown in Table 3.1. The complete questionnaire is documented in Appendix H of Volume 2. The remainder of this chapter outlines the broad content of each module of the questionnaire.

Table 3.1 Modules of the 2020-21 TCSEW questionnaire and the sub-set of respondents who were asked each module

Questionnaire module	Core sample
Household grid	All
Perceptions of crime	All
Screener questions	All
Victimisation module	All victims of non-fraud
Fraud victimisation module	All victims of fraud
Covid	All
Covid & Children	All parents with at least one 10-15 year old in the household
Demographics	All

Almost every question in the survey included 'Don't know' and 'Refused' options that the interviewer could use. At most questions these options did not appear as part of the code frame, to try to ensure that interviewers did not overuse them. This largely replicated how these codes were presented in the face-to-face survey.

In the questionnaire in Appendix H of Volume 2, 'Don't know' and 'Refused' codes are only shown if they were explicit response categories and so actually appeared as an option on the screen.

3.3 Household grid

3.3.1 Wave 1

Basic socio-demographic details (age, sex and martial status) were collected in the household grid for every adult in the household and the age and sex of all children in the household under 16 years old were also collected. Additionally, some basic information was collected on length of time at the address and vehicle ownership.

3.3.2 Waves 2 and 3

While the basic structure of the household grid was consistent with wave 1 at subsequent waves, it was recognised that most household characteristics were unlikely to have changed in the intervening period. As such, updated household information was only collected if the respondent confirmed that their circumstances had changed since the date of the last interview.

3.4 Traditional (non-fraud) screener questions

3.4.1 Wave 1

All respondents were asked whether they had experienced certain types of non-fraud crimes or incidents within a specified reference period, namely the last 12 months from the date of interview.

Questions were designed to ensure that all incidents of crime within the scope of the TCSEW, including relatively minor ones, were mentioned. The screener questions deliberately avoided using terms such as 'burglary' or 'robbery' which have a precise definition that respondents might not know or not fully understand the precise meaning. The wording of these screener

questions has been kept consistent since the CSEW began to ensure comparability across years, apart from the minor updating of some terminology. The wording also remained broadly consistent on the TCSEW, although a small number of questions were modified to mitigate for the lack of showcards. The only screener question that was not asked as part of the TCSEW was the sexual assault screener as it was considered too sensitive to be asked over the phone.

One significant change was made to the wording of the threat screener question on the TCSEW compared with the CSEW. Whereas previously the question asked whether anyone had *'threatened you in any way'* the revised questions asked whether anyone had *'threatened, harassed or intimidated you in any way'*. This change was not connected to the change in mode but instead was due to ongoing work by ONS to consider how well the survey captures harassment. It was decided that given the natural break in the time series, and possible increased interest in levels of harassment during the pandemic, this would be an opportune time to introduce a slightly different question wording to try and capture more detail on harassment and intimidation which was not previously captured. Additional questions on harassment were also added to the victimisation module as part of this investigation.

Depending upon individual circumstances, a maximum of 24 screener questions were asked which can be grouped into four main categories:

– All respondents who owned vehicles or bicycles were asked about their experience of vehicle-related crimes (e.g. theft of vehicles, theft from vehicles, damage to vehicles, bicycle theft);

– All respondents were asked about experience of property-related crimes in their current residence (e.g. whether the property was broken into, whether anything was stolen from the property, whether the property was damaged);

– All respondents who had moved in the last 12 months were also asked about their experience of property-related crimes at their previous residence(s); and

– All respondents were asked about experience of personal crimes (e.g., whether any personal property was stolen, whether any personal property was damaged, whether they had been a victim of violence or threats)

The questions were designed to ensure that the respondent does not mention the same incident more than once. As a check, at the end of the screener questions, the interviewer is shown a list of all incidents recorded and asked to check with the respondent that all incidents have been recorded and nothing has been counted twice. If there is any evidence of double counting, the respondent has an opportunity to correct the information before proceeding.

Within the screener questions there is a crucial distinction between **household** incidents and **personal** incidents.

All vehicle-related and property-related crimes are counted as household incidents. Respondents are asked whether anyone currently residing in their household has experienced any relevant incidents within the reference period. A typical example of a household incident is criminal damage to a car. It is assumed that the respondent will be able to recall these incidents and provide information even in cases where he/she was not the owner or user of the car.

Personal incidents refer to all crimes against the individual and so only relate to things that have happened to the respondent personally, but not to other people in the household. This is often a difficult concept for respondents to understand as their natural inclination is to tell the interviewer about incidents affecting other members of their household. An example of a personal incident would be an assault. An assault against other household members (no matter how serious) are not recorded, unless the respondent was also assaulted as part of the same incident.

3.4.2 Waves 2 and 3

The screener questions asked in subsequent waves were identical apart from the reference period covered. Unlike wave 1, the screener questions in the follow-up interviews asked about the time that had elapsed since the last interview rather than the last 12 months. This was a minimum of three months but could be longer (potentially up to six months) depending on when the follow-up interview took place.

3.5   Fraud screener questions

3.5.1 Wave 1

The fraud screener questions were asked to all respondents and were administered in the same way as the traditional non-fraud screeners.

The six main topic areas covered by the fraud screeners were:

– Incidents which occurred as a direct result of a previous non-fraud incident

– Personal information or account details been used to obtain money, or buy goods or services without permission

– Being tricked or deceived out of money or goods

– Attempts to trick or deceive the respondent out of money or goods

– Theft of personal information or details held on your computer or in on-line accounts

– Computer or other internet-enabled device being infected or interfered with by a virus

As with the non-fraud screener questions, the wording remained consistent during the transition to telephone interviewing, although a small number of questions had to be modified to overcome the inability to rely on showcards. The most significant change was in relation to the screener about incidents which occurred as a direct result of a non-fraud incident which involved a show card on the face-to-face survey. For the telephone survey this single screener question was split into five separate questions as it was felt the original format could not be replicated on the telephone as a single question.

3.5.2 Waves 2 and 3

Again, the fraud screener questions asked in subsequent waves were identical to wave 1 apart from the time period covered. Unlike wave 1, the fraud screener questions in the follow-up interviews asked about the time that had elapsed since the last interview rather than the last 12 months. This was a minimum of three months but could be longer (potentially up to six months) depending on exactly when the follow-up interview took place.

3.6   Victimisation modules (All waves)

All incidents identified at the screener questions (up to a maximum of six) were followed through in more detail in the victimisation module. Incidents are covered in a specific priority order which has been consistent since the start of the CSEW and was maintained in the move to the TCSEW. However, to accommodate the shorter interview length the number of questions asked in the victimisation module was significantly reduced for the telephone survey compared with the face-to-face survey. Priority was given to those questions which were considered critical for classifying offences.

3.6.1   Identification and ordering of incidents for victimisation modules

In 2020-21, 81% of all adult respondents interviewed did not report any incidents of crime during the last 12 months[14], and therefore did not complete any victimisation modules as part of the interview.

Where a respondent had experienced one or more incidents in the reference period, the questionnaire script automatically identified the order in which the modules were to be asked. Fraud crimes were given a lower priority than the existing non-fraud crime types. The automatic selection meant that the interviewer had no discretion about the selection or order of the modules[15]. The priority ordering used by the script was as follows:

– **According to the type of crime.** Non-fraud victimisation modules were asked first, in reverse order to the screener questions. Broadly speaking this means that all personal incidents were asked before property-related incidents, which were asked before vehicle-related incidents. Fraud victimisation modules were asked but in the same order as the fraud screener questions. Overall, across both non-fraud and fraud crimes a maximum of six victimisation modules were completed, with non-fraud incidents taking priority.

– **Chronologically within each type of crime.** If a respondent reported more than one incident of the same type of crime, modules were asked about the most recent incident first and worked backwards chronologically.

In the 2020-21 survey, a total of 9,682 victimisation modules were completed by 6,941 individual victims, with 18.9% of all respondents interviewed reporting at least one incident (see Table 3.2).

---

14 Respondents could be interviewed up to three times in this time period.

[15] In the case of the incidents of domestic violence, the interviewer had an option to suspend the victimisation module, as this might make the respondents feel uncomfortable or endanger the respondent in some situations.

Table 3.2 Core sample respondents who completed victimisation modules, 2020-21 TCSEW

	N	% of all respondents interviewed	% of victims
Non victims	29,860	81.1	
Victims[16]	6,941	18.9	
No. of victim modules completed			
1	5,154	14.0	74.3
2	1,233	3.4	17.8
3	333	0.9	4.8
4	111	0.3	1.6
5	41	0.1	0.6
6	69	0.2	1.0
Total	**9,682**		
Bases:		*36,801*	*6,941*

3.6.2 Defining a series of incidents

Most incidents reported represent one-off crimes or single incidents. However, in a minority of cases a respondent may have been victimised a number of times in succession. At each screener question where a respondent reported an incident, they were asked how many incidents of the given type had occurred during the reference period. If more than one incident was reported, the respondent was asked whether they thought that these incidents represented a 'series' or not. A series was defined as "the same thing, done under the same circumstances and probably by the same people". Where this was the case, only one victimisation module was completed in relation to the most recent incident in the series. Again, this was done to minimise respondent burden.

In fraud cases the definition of a series is more complex, as the survey is intended to replicate the way in which the police would record fraud incidents as close as possible. The key measures for identifying a series with fraud offences is whether all the incidents are identified at the same time, and whether the victim responded in the same way. This is designed to ensure that cases of fraud involving multiple transactions on a single account are counted as a single incident rather than multiple incidents. For example, if someone discovers four

---

[16] Victims refers to the number of respondents who started at least one victimisation module. This is slightly different to the number of respondents who reported at least one incident at the screener questions (n=7,049). This is due to respondent drop out after the screener questions or victimisation modules being skipped under certain circumstances.

separate transactions on their bank account these will be recorded as a single incident rather than four separate incidents or a series. However, if they later discover more transactions on their account then this would be recorded as a separate incident or as the second incident in a series.

There are two practical advantages to the approach of only asking about the most recent incident where a series of similar incidents has occurred. First, since some (although not all) incidents classified as a series can be petty or minor incidents (e.g. vandalism) it avoids the need to ask the same questions to a respondent several times over. And second, it avoids using up the limit of six victimisation modules on incidents which <u>may</u> be fairly trivial, while missing out potentially more serious incidents.

In 2020-21, 86% of all victimisation modules related to single incidents and 14% related to a series of incidents. This split between single and series incidents was broadly the same as on previous surveys.

In the rare cases where a respondent has experienced a mixture of single incidents and a series of incidents the interview program has a complex routine which handles the sequence of individual and series incidents and allows the priority ordering of the victimisation modules to be decided.

In terms of estimating the victimisation rates, series incidents receive a weight corresponding to the number of incidents in the series that fall within the reference period, subject to a maximum limit that is specific to the offence code group (see section 8.5). This is a relatively recent change to how the data is weighted as previously all offence types were capped at a limit of five.

3.6.3   Content of victimisation module

The victimisation module collects the key information needed to classify each incident to a particular offence type, which is the basis for calculating the prevalence and incidence rates. It contains three types of information:

– **The exact month(s) in which the incident or series of incidents occurred**. In a few cases, respondents may have reported an incident which later turns out to have been outside the reference period. In such cases, the victimisation module is simply by-passed. If respondents were unsure about the exact month in which something happened, they were asked to narrow it down to a specific quarter. For incidents that were part of a series, respondents were asked how many incidents occurred in each quarter and the month in which the most recent incident had occurred.

– **An open-ended description of the incident where the respondent describes exactly what happened in their own words**. The open-ended description is vital to the accurate coding of offences that takes place in the office. Short, ambiguous or inconsistent descriptions can often make offence coding difficult. In fraud victimisation modules a second open-ended description is included to collect information about the action the respondent took following the fraud or attempted fraud, as this is a key aspect of the fraud offence coding. At the end of each victimisation module, the original open-ended description that the interviewer had entered at the start is reshown to the interviewer along with the answers to some of the key pre-coded questions. By presenting this information on a single screen, interviewers have the chance to confirm with respondents that the information is correct and consistent. If the respondent and/or interviewer wish to add or clarify any information they can do this.

– **A series of key questions used to establish important characteristics about the incident.** These include where and when the incident took place; whether anything was stolen or damaged and, if so, what; the costs of things stolen or damaged; any details of the offenders (if known); whether force or violence was used and, if so, the nature of the force used and any injuries sustained; and whether the police were informed or not. While

many of the questions in the fraud victimisation module reflect the non-fraud module there are also other questions which are more relevant for these specific types of crime.

3.7 Reference dates

3.7.1 Wave 1

In the questionnaire script, reference dates were automatically calculated based on the date of interview and appropriate text substitution was used to ensure that the questions always referred to the correct reference period.

Because the 12-month reference period changed each month throughout the fieldwork year, some date-related questions in the victimisation module had different text each month to reflect this changing reference period. Thus, for example, any interviews conducted in May 2020 would use the reference period "*since the first of May 2019*". This means that in practice the 12-month reference period consisted of the last 12 full calendar months, plus the current month (i.e., slightly more than 12 months). This is taken into account when the victimisation rates are estimated.

In the previous section it was noted that for each incident the respondent is asked which month of the year the incident happened in. At these questions the code frame presented to the interviewer and respondent always displays the last 13 months counting back from the date of interview.

If respondents are unable to narrow it down to a particular month, they are then asked for the quarter of the year it happened in. Additionally, where respondents have reported a series of incidents in the last 12 months, they are asked how many incidents happened in each quarter. The time period used for both these questions is not 'rolling quarters' but rather are fixed to match the standard quarters used in both the survey design and in terms of how the estimates are reported (i.e., January – March, April – June, July – September, October – December). However, the first quarter would only include the months pertinent to that quarter, e.g. an interview conducted in June 2020 would only show June 2019 in quarter 1 as the two previous months would be outside the 12 month period.

Since the reference period is based on a rolling 12 months based on the month of interview it is important in cases where only the quarter is recorded to be able to establish whether the incident is in scope (within the last 12 months) or out of scope (more than 12 months ago). This requires some questions within the victimisation module to have an adjusted code frame which differs based on the exact month of interview.

Table 3.3 illustrates this for the full year 2020-21. In each case the first code is always out of scope (more than 12 months ago) and the other codes are in scope.

Table 3.3 Code frame by month of interview at the victimisation module (wave 1)

Interview month=April 2020

No interviews conducted in April 2020

Interview month=May 2020

1. Before the 1st of May 2019 (Out of scope)
2. In May or June 2019
3. Between July and September 2019
4. Between October and December 2019
5. Between January and March 2020
6. Between the 1st of April 2020 and present

Interview month=June 2020

1. Before the first of June 2019 (Out of scope)
2. In June 2019
3. Between July and September 2019
4. Between October and December 2019
5. Between January and March 2020
6. Between the 1st of April 2020 and present

Interview month =July 2020

1. Before the 1st of July 2019 (Out of scope)
2. Between July and September 2019
3. Between October and December 2019
4. Between January and March 2020
5. Between April and June 2020
6. Between the 1st of July 2020 and present

Interview month=August 2020

1. Before the 1st of August 2019 (Out of scope)
2. In August or September 2019
3. Between October and December 2019
4. Between January and March 2020
5. Between April and June 2020
6. Between the 1st of July 2020 and present

Interview month=September 2020

1. Before the 1st of September 2019 (Out of scope)
2. In September 2019
3. Between October and December 2019
4. Between January and March 2020
5. Between April and June 2020
6. Between the 1st of July 2020 and present

Interview month=October 2020

1. Before the 1st of October 2019 (Out of scope)
2. Between October and December 2019
3. Between January and March 2020
4. Between April and June 2020
5. Between July and September 2020
6. Between the 1st October 2020 and present

Interview month=October 2020

7. Before the 1st of October 2019 (Out of scope)
8. Between October and December 2019
9. Between January and March 2020
10. Between April and June 2020
11. Between July and September 2020
12. Between the 1st October 2020 and present

Interview month=November 2020

1. Before the 1st of November 2019 (Out of scope)
2. In November or December 2019
3. Between January and March 2020
4. Between April and June 2020
5. Between July and September 2020
6. Between the 1st of October 2020 and present

Interview month=December 2020

1. Before the 1st of December 2019 (Out of scope)
2. In December 2019
3. Between January and March 2020
4. Between April and June 2020
5. Between July and September 2020
6. Between the 1st of October 2020 and present

Interview month=January 2021

1. Before the 1st of January 2020 (Out of scope)
2. Between January and March 2020
3. Between April and June 2020
4. Between July and September 2020
5. Between October and December 2020
6. Between the 1st of January 2021 and present

Interview month=February 2021

1. Before the 1st of February 2020 (Out of scope)
2. In February or March 2020
3. Between April and June 2020
4. Between July and September 2020
5. Between October and December 2020
6. Between the 1st of January 2021 and present

Interview month =March 2021

1. Before the 1st of March 2020 (Out of scope)
2. In March 2020
3. Between April and June 2020
4. Between July and September 2020
5. Between October and December 2020
6. Between 1st of January 2021 and present

3.7.2 Waves 2 and 3

As with wave 1, the questionnaire script reference dates were automatically calculated based on the date of the last interview and the date of the follow-up interview (e.g., between the wave 1 interview date and the wave 2 interview date) and appropriate text substitution was used to ensure that the questions always referred to the correct reference period.

Because the reference period varied by respondent, the date-related questions in the victimisation module had different text to reflect this changing reference period. Thus, for example, if the first interview was 20th June 2020, a wave 2 interview conducted on 10th October 2020 would cover the month prior to the previous interview (May 2020), up to and including the current month (October 2020). This means that in practice the script consisted of a longer period of time than strictly needed for the wave 2 interview, i.e., May 2020 and the 1-19th June 2020. This is again taken into account when the victimisation rates are estimated.

Respondents were asked in which month the incident happened, and at these questions the code frame presented to the interviewer always displayed the months counting back from the date of the last interview plus the month prior to the last interview. In this way it was possible to establish a rolling 12-month period by utilising part of the date from the original wave 1 interview combined with the wave 2 information.

As a result of the much shorter reference dates, respondents in waves 2-3 were not asked for the quarter of the year an incident happened if they did not know the exact month.

Where respondents reported a series of incidents since the last interview, they were asked how many incidents happened in each month rather than in each quarter.

Having established our core (Wave 1) universe, it was agreed that Wave 1 respondents who did not complete Wave 2 would still be invited to take part in one further way (i.e., re-invited at the start of Wave 3). For any respondents who fell into this category, the reference dates were automatically re-calculated to be based on their first interview and the wave 3 interview date. Thus, if their first interview was 31st May 2020 and they skipped their second interview, but completed when next invited in December 2020, this interview covered the month prior to the last interview (April 2020), up to and including the current month (December 2020) to ensure the 'catch-up' interview covered the full period since the last interview.

3.8 Covid module

The Covid module was the only new module and was designed to be more fluid than the other modules. As such, while other modules remained relatively consistent throughout 2020-21 (changes between wave 1 and waves 2-3 notwithstanding), the Covid module was updated on multiple occasions throughout the year. A full breakdown of the changes are included in Appendix H of Volume 2 of the Technical report.

Topics covered in the module at the outset included:

– Worry about crime
– Anti-social behaviour
– Harassment
– Reporting crimes/ lockdown breaches
– Perception and satisfaction of the police response to Covid
– Children and Covid (see next section)

The topics were reviewed monthly and modified where relevant to reflect the changing needs of TCSEW and to reflect the fact that as the initial lockdown eased then some questions became less relevant. Similarly, some questions were modified so that they were only asked once a year (i.e. in wave 1 only), rather than being included in every subsequent interview.

3.9 Covid module – questions about children aged 10-15

Having agreed that the interviews with 10-15 year olds would not be part of the TCSEW, a small number of questions were asked of parents/ guardians who had children aged 10-15 in the household. These questions looked at awareness of what the child was doing online and the child's online experiences, as well as awareness of what the child was doing when they went out by themselves.

Where multiple children aged 10-15 were in residence, children were selected based on most recent birthday or chosen at random in the case of multiple births.

Again, the questions were modified to reflect the changing needs over the course of the year, with a sub-set of questions only being asked once (wave 1 only).

3.10 Demographics

This section collected additional information on the respondent.

Question topics included:

– health and disability
– employment details
– housing tenure
– well-being

Although an established module, the Demographics module was also designed to be more fluid over time. The topics were reviewed monthly and again modified where relevant, either to remove questions completely or to ensure they were only asked once a year. The single biggest change to the Demographics module was the addition of three new topic areas in September 2020:

– alcohol
– drugs
– mental health

The first two of these were an abridged version of the questions previously asked in the CSEW self-completion module and as such were only asked of those aged 18-74.

3.11 Question development and testing

In most survey years, question testing is a standard component of the CSEW questionnaire development process. This usually takes the form of cognitive testing conducted via face-to-face interviews. However, given the unprecedented nature of the changes forced upon the survey, no such development work was feasible before the switch to telephone. Instead, the modified questionnaire was reviewed by multiple (internal) teams, with various feedback loops in place at the time of launch to review the updated questionnaire. As a result of the extensive pre-testing of the questionnaire very little feedback was received post-launch, although a small number of questions were modified in the first few days of fieldwork, following initial feedback from the interviewers. This feedback loop was maintained throughout fieldwork as/ when additional questions were introduced to the survey.

4. Fieldwork

This chapter documents all aspects of the data collection process, focusing on fieldwork procedures, the management of fieldwork across the survey year, quality control procedures and aspects of how the interview was conducted.

4.1 Briefing of interviewers

Previously on the CSEW a certain number of new interviewers are brought onto the CSEW interviewer panel each year either to replace those who have left the panel or to boost the overall size of the panel. For the 2020/21 TCSEW it was decided only to use interviewers who were already on the CSEW interviewer panel and who had previously worked as face-to-face interviewers on the survey, in some cases for many years. This was critical in terms of being able to launch the TCSEW in the field quickly by ensuring that all interviewers were familiar with the complexities of the survey, such as avoiding double counting or understanding what constitutes a series of incidents.

In previous years, interviewers already working on the survey have attended a half-day face to face refresher briefing annually. This half-day refresher briefing was replaced in 2020/21 by a remote two and a half hour briefing which took place over Microsoft Teams. Over the year, five of these briefings took place and were attended by a total of 83 interviewers before they began to work on the wave 1 sample. The purpose of these briefings was to give interviewers an overview of the changes made to the survey to accommodate the switch of mode from CAPI to CATI. Interviewers were also briefed on the differences between face-to-face and telephone interviewing and given tips on best practice for telephone interviewing.

A follow-up video briefing was uploaded to Kantar's interviewer portal on 17th August 2020. All interviewers were given access to this briefing and asked to watch the follow-up video before staring work on the wave 2 sample. The video briefing re-capped the topics covered in the initial May briefings and informed interviewers of additional changes to the questionnaire implemented at wave 2.

4.2 Supervision and quality control

Several methods were used to ensure the quality and validity of the data collection operation.

Each individual shift (lasting 3.5 hours) was assigned a dedicated Team Leader or Senior Interviewer to supervise and oversee quality control during the shift. At the start of each shift, the Team Leader or Senior Interviewer would organise a communications call to confirm that all interviewers were logged in, and to provide any additional briefing instructions which interviewers would require. Interviewers were also informed on this call which wave of sample they were being allocated to for that shift.

During each shift, a chat group on Microsoft Teams was made available to all interviewers. In this chat group interviewers were able to ask any questions or queries that came up during the course of the shift, and these could be immediately addressed by the Team Leader or Senior Interviewer.

As is standard on all telephone projects, a certain proportion of interviews are listened to for quality control purposes[17]. On the TCSEW, 7% of all interviews were listened to for at least 75% of the interview in order to meet standard quality control requirements (amounting to approximately 2,576 surveys). Beyond this standard requirement, further quality control measures were put in place for the TCSEW specifically, with a supervisor within the telephone unit responsible for quality control listening to at least one completed survey per interviewer

---

[17] Both live interviews and recordings are listened to

each week. As a result of this additional quality control process, the quality of every interviewer's work was checked frequently throughout the year.

4.3 Fieldwork dates and fieldwork management

During the 2020/21 survey, sample was grouped into three linked waves which were released in batches each month.

Wave 1 issued sample ran from May 2020 to January 2021, with one batch of sample being released per month. The first three wave 1 batches, those released in May-July 2020 remained open for three months. All subsequent wave 1 batches remained open for just two months. In total, 38,397 pieces of sample were released as part of wave 1 across the year.

The wave 2 issued sample (September 2020 to March 2021) and wave 3 issued sample (December 2020 to March 2021) were all released in two batches per month. The only exception to this was the wave 3 February sample, which was released as one batch. All Wave 2 and Wave 3 batches remained open for approximately two months. In total, wave 2 comprised of 15,717 pieces of sample and wave 3 consisted of 9,764 pieces up to the end of March 2021. Over the course of the whole year willingness to be re-contacted remained at 98% across each wave of the survey.

How sample was worked in the field differed from the face-to-face survey in several ways. First, the TCSEW sample was not interviewer specific: that is each interviewer did not have their own assignment of sample which they worked through to a final outcome. Rather, sample was managed through an automatic dialler, with each piece of sample being allocated to the next available interviewer. The dialler is able to prioritise some batches of sample over others. Another difference compared to the face-to-face survey was that no sample was re-issued. Rather, sample stayed live for the full fieldwork period with the time between calls being automatically set based on previous outcomes. This meant that the maximum amount of time for a batch of sample to remain in field was around three months. It should be noted that, due to this set-up, some issued sample was still being worked into May 2021[18]. This is similar to the management of the CSEW fieldwork where sample batches issued later in the survey year are worked beyond the 31st March.

4.4 Advance letter

All selected respondents were sent a letter from the Office for National Statistics in advance of an interviewer calling them to administer the survey. This explained a little about the survey, why they had been selected and informed them that an interviewer from Kantar would be calling in the next few weeks. The letter also provided a telephone number and an email address for people to contact to find out more about the survey, to make an appointment for an interviewer to call, or to opt out of the survey.

As well as an advance letter, an advance email was also sent to those respondents with an email address included in the sample. This email was also sent from the Office for National Statistics and it covered similar information to that provided in the letter. It also included a direct link through which respondents could email the team at Kantar Public to find out more about the survey.

Both the advance letters and advance emails varied slightly by wave. The most notable difference between waves was that, for those respondents being offered an unconditional incentive at Waves 2 and 3, this was included with their advance letter or email (depending on whether it was a postal or online voucher).

Examples of the advance letters and emails used can be found in Appendix A-F of Volume 2.

---

[18] The final batches of Wave 2 and Wave 3 sample (released in mid-March 2021) closed on the 15th May 2021.

Respondents living in Wales received a bilingual version of the advance letters and emails. The bilingual versions of the letters and emails included the same information as the English versions but displayed this in both English and Welsh. Again, examples of the Welsh advance letters and emails can be found in Appendix A-F of Volume 2.

4.5 Respondent website

A website with information about the survey was set up, with the style and content of information updated regularly. Respondents could be directed to this website by the interviewer and the website was also referenced in all respondent-facing survey materials.

Information displayed on this website includes what the survey is about and what types of questions are asked, survey results, confidentiality and data security, as well as a section on frequently asked questions. The website is available in both English and Welsh.

The URL for the website is: http://www.crimesurvey.co.uk

4.6 Incentives

Since 2005, a booklet of six first class stamps has been sent with the advance letter as a 'thank you' to people for taking part in the survey and this was the case for Wave 1 of the 2020/21 survey.

Due to the re-contact nature of the survey in 2020/21, additional incentives were offered to encourage respondents to continue to participate in future waves. At wave 1, alongside the book of six stamps which was sent with the advance letter, respondents were also offered a £10 incentive conditional on their completion of the survey. This £10 incentive was a voucher which could be provided by post or by email and respondents were given a choice between Asda, Tesco, Sainsbury's, Marks & Spencer, John Lewis/ Waitrose and Amazon as providers.

At wave 2, an unconditional £10 incentive was given to all respondents, this was included either in their advance letter or advance email depending on how they had requested their incentive upon completion of their wave 1 survey. The voucher sent as part of the wave 2 advance materials was for the same provider requested by the respondent upon completion of their wave 1 survey.

For those respondents who completed a wave 2 survey, another unconditional £10 incentive was sent in their advance letter for wave 3. For those respondents who did not complete a survey at wave 2, the wave 3 advance letter included a book of stamps and the offer of a £10 incentive conditional on their completing the wave 3 survey.

4.7 Presence of others during the interview

During the interviewer briefing sessions emphasis was placed on the importance of trying, wherever possible, to conduct the interview in private. This generally helped to make the interview run more smoothly but was also felt likely to encourage respondents to mention certain incidents or events, which they might be embarrassed or unwilling to talk about in front of others. However, given that the survey was conducted by telephone there was clearly a limitation on what interviewers could do. Interviewers were instructed to flag to respondents at the outset that they may want to be alone for some, or all, of the questionnaire. Interviewers were also reminded to restate this prior to asking certain sensitive questions.

Privacy during the interview is a particular concern for respondents who had been attacked, hurt, harassed, or intimidated (answering Yes at DELIBVIO, THREVIOL2, HHLDVIOL). Where respondents had experienced such incidents in the last 12 months[19], interviewers asked respondents whether they were happy to be asked more detailed questions about the incident. If the respondent said no, interviewers had the option of suspending the victim form. This procedure meant that the interviewer could complete the rest of the questionnaire, rather than

---

[19] Or since the last interview in the case of Wave 2 or Wave 3 interviews

having to abandon the whole interview. During 2020/21, a total of 206 victimisation modules were suspended by interviewers for this reason. The proportion of suspended victimisation modules was higher than levels seen on the face-to-face survey in previous years.

4.8 Length of interview

Timing stamps were placed throughout the questionnaire to allow timing of individual sections. In a small number of cases, the time stamps were invalid due to technical issues or interviews conducted over multiple days although valid times were available for over 95% of interviews.

Wave 1 interviews took longer to complete on average (mean: 31minutes, median: 28 minutes) compared with wave 2 and 3 interviews (mean: 24 minutes, median: 22 minutes). At wave 1, just 61% of interviews took 30 minutes or less to complete, compared to 85% at wave 2 and 86% at wave 3. At waves 2 and 3 only 4% of interviews lasted over 45 minutes, compared to 11% at wave 1.

The main influence on core interview length was whether the respondent had been a victim of crime. The average interview length for victims of crime was 43 minutes compared with 24 minutes for non-victims.

The average length of interview by number of victimisation modules completed is shown in Table 4.1. Not unexpectedly, interview length was strongly related to the number of victimisation modules completed, with those completing four or more modules (0.6% of victims) having an average interview length of around 70 minutes.

Table 4.1   Average (mean) time of interview by number of victimisation modules, 2020/21 TCSEW

Number of victimisation modules	Average time (minutes)			
	Total	Wave 1	Wave 2	Wave 3
Non victims	**24**	**27**	**22**	**22**
All victims	**43**	**44**	**39**	**39**
1	39	40	37	37
2	50	51	48	48
3	59	60	57	59
4 or more	70	70	72	67
All adult respondents	**28**	**31**	**24**	**24**

5. Response rates

5.1 Wave 1: survey response rate and non-response

The full response and non-response breakdown for the 2020-21 Wave 1 sample is shown in Table 5.1.

For wave 1, 10.1% of issued cases were identified as not being an eligible case (known as deadwood). The most common type of deadwood was dead/ invalid number, which accounted for 7.3% of all issued cases.

Interviewers made contact with either the selected respondent or another responsible adult in the household at 75% of eligible addresses, meaning a non-contact rate of 25%. The most common reason for non-contact (11.9% of eligible cases) was where the respondent had set up a caller ID block or call barring system.

For eligible cases where contact was made, the most common reason for not getting an interview was due to a respondent refusal, which accounted for 9.5% of all eligible cases. Proxy refusals (someone refusing on behalf of the named respondent) were less common (1%).

Half of eligible cases (50.3%) were categorised as unproductive for other reasons including broken appointments, people who were ill/ in hospital during the period of the survey and people who had inadequate English to complete the survey.

Overall, 17,167 Wave 1 interviews were achieved in 2020-21, representing a response rate of 49.7% (44.7% across all wave 1 sample issued). The overall sample conversion rate (achieved interviews/issued sample) was 44.7% against the target conversion rate of 40%.

Table 5.1   Wave 1 sample response rate and non-response outcomes, 2020-21 TCSEW

	N	% of issued	% of eligible
TOTAL ISSUED ADDRESSES	38,397	100%	
Deadwood			
Business number	111	0.3%	
Dead / Invalid number	2,794	7.3%	
Modem/ Fax number	326	0.8%	
Respondent has moved	126	0.3%	
Respondent unknown at number	517	1.3%	
TOTAL DEADWOOD	**3,874**	**10.1%**	
TOTAL ELIGIBLE ADDRESSES	**34,523**	**89.9%**	**100%**
Non-contact			
Caller ID Block/ Call Barring Message	4,099	10.7%	11.9%
General call back (not arranged with respondent)	1,256	3.3%	3.6%
No answer/ Answer Machine/ Number Busy	3,289	8.6%	9.5%
Total non-contact	*8,644*	*22.5%*	*25.0%*
Refusal			
Office refusal	233	0.6%	0.7%
Respondent refusal	3,265	8.5%	9.5%
Proxy refusal	333	0.9%	1.0%
Quit mid interview, refused to finish	90	0.2%	0.3%
Total refusal	*3,921*	*10.2%*	*11.4%*
Other unproductive			
Broken Appointment	2,956	7.7%	8.6%
Dialler Error[20]	989	2.6%	2.9%
Inadequate English	97	0.3%	0.3%
Physically or mentally unable	254	0.7%	0.7%
Respondent has died	112	0.3%	0.3%
Respondent too ill/ in hospital	266	0.7%	0.8%
Other unsuccessful	117	0.3%	0.3%
Total other unsuccessful	*4,791*	*12.5%*	*13.9%*
TOTAL UNPRODUCTIVE	**17,356**	**45.2%**	**50.3%**
TOTAL INTERVIEWS	**17,167**	**44.7%**	**49.7%**

---

[20] Dialler Error refers to calls that fail due to dial tone irregularities

5.1.1 Wave 1 response rates by Government Office Region

Table 5.2 shows the different response rates and reasons for non-response achieved by region for 2020-21 Wave 1. This shows that across regions the response rate ranged from 52.9% in London to 44.9% in the North West.

Table 5.2   **Wave 1 sample response rates and non-response by Government Office Region, 2020-21 TCSEW**

	North East	North West	Yorkshire & The Humber	East Midlands	West Midlands	East of England	London	South East	South West	Wales
Percentage of eligible addresses (%):										
Non-contact	27.2	28.4	25.9	26.7	25.6	23.7	18.9	24.4	26.3	25.3
Refusal	10.1	11.5	10.6	10.6	10.7	10.2	12.3	10.9	9.6	10.9
Other unproductive	15.5	15.2	14.0	12.7	15.3	13.3	15.9	12.3	12.9	14.0
Achieved interview	47.2	44.9	49.5	50.0	48.4	52.8	52.9	52.4	51.2	49.8

5.1.2 Wave 1 response rate by Police Force Area

Table 5.3 overleaf shows the number of Wave 1 interviews achieved in each PFA and the response rates.

Table 5.3   Wave 1 sample achieved interviews and response rates by PFA, 2020-21 TCSEW

PFA	Achieved	Response rate
	N	*%*
Avon & Somerset	399	52.2%
Bedfordshire	331	55.5%
Cambridgeshire	354	51.2%
Cheshire	287	46.7%
Cleveland	240	45.8%
Cumbria	311	46.3%
Derbyshire	337	50.0%
Devon & Cornwall	570	52.8%
Dorset	359	48.4%
Durham	214	43.4%
Dyfed Powys	372	51.6%
Essex	484	52.1%
Gloucestershire	295	51.4%
Greater Manchester	637	43.5%
Gwent	333	47.4%
Hampshire	536	49.5%
Hertfordshire	331	55.3%
Humberside	340	50.1%
Kent	473	51.9%
Lancashire	346	44.4%
Leicestershire	303	49.0%
Lincolnshire	336	49.3%
Merseyside	404	45.2%
Metropolitan and City of London	1,672	52.9%
Norfolk	393	51.2%
North Wales	324	50.0%
North Yorkshire	322	54.6%
Northamptonshire	312	50.2%
Northumbria	279	52.1%
Nottinghamshire	333	51.4%
South Wales	325	50.3%
South Yorkshire	325	47.8%
Staffordshire	258	48.7%
Suffolk	376	52.3%
Surrey	416	55.9%
Sussex	388	50.6%
Thames Valley	635	54.3%
Warwickshire	365	53.1%
West Mercia	318	49.7%
West Midlands	585	45.1%
West Yorkshire	601	47.7%
Wiltshire	348	50.4%

5.2   Wave 2: survey response rate and non-response

The full response and non-response breakdown for the 2020-21 wave 2 sample is shown in Table 5.5[21]. For wave 2, 2.1% of issued cases were identified as not being an eligible case (known as deadwood). The most common type of deadwood was dead/ invalid number, which accounted for 1.5% of all issued cases.

Interviewers made contact with either the selected respondent or a responsible adult at 90.9% of eligible addresses, meaning a non-contact rate of 9.1%. The most common reason for non-contact (4.8% of eligible cases) was where the respondent had set up a caller ID block or call barring system.

For eligible cases where contact was made, the most common reason for not getting an interview was due to broken appointment, which accounted for 5.8% of all eligible cases. Respondent refusals accounted for 3.6% of all eligible cases, while proxy refusals (someone refusing on behalf of the named respondent) were less common (0.3%).

Overall, 10,623 wave 2 interviews were achieved in 2020-21, representing a response rate of 79.4% and a conversion rate of 77.7%.

---

[21] Response rates are based only on wave 2 sample batches which were closed in the field by 31st March 2021. A total of 15,717 wave 2 cases were issued during the year but some batches remained open until as late as May 2021.

Table 5.5   Wave 2 sample response rate and non-response outcomes, 2020-21 TCSEW

	N	% of issued	% of eligible
TOTAL ISSUED ADDRESSES	13,677	100%	
Deadwood			
Business number	11	0.1%	
Dead / Invalid number	206	1.5%	
Modem/ Fax number	28	0.2%	
Respondent has moved	18	0.1%	
Respondent unknown at number	27	0.2%	
TOTAL DEADWOOD	**290**	**2.1%**	
TOTAL ELIGIBLE ADDRESSES	**13,387**	**97.9%**	**100%**
Non-contact			
Caller ID Block/ Call Barring Message	647	4.7%	4.8%
General call back (not arranged with respondent)	257	1.9%	1.9%
No answer/ Answer Machine/ Number Busy	313	2.3%	2.3%
Total non-contact	*1,217*	*8.9%*	*9.1%*
Refusal			
Office refusal	64	0.5%	0.5%
Respondent refusal	483	3.5%	3.6%
Proxy refusal	34	0.2%	0.3%
Quit mid interview, refused to finish	11	0.1%	0.1%
Total refusal	*592*	*4.3%*	*4.4%*
Other unproductive			
Broken Appointment	779	5.7%	5.8%
Dialler Error[22]	82	0.6%	0.6%
Inadequate English	4	0.0%	0.0%
Physically or mentally unable	24	0.2%	0.2%
Respondent has died	15	0.1%	0.1%
Respondent too ill/ in hospital	49	0.4%	0.4%
Other unsuccessful	2	0.0%	0.0%
Total other unsuccessful	*955*	*7.0%*	*7.1%*
TOTAL UNPRODUCTIVE	**2,764**	**20.2%**	**20.6%**
TOTAL INTERVIEWS	**10,623**	**77.7%**	**79.4%**

---

[22] Dialler Error refers to calls that fail due to dial tone irregularities

5.2.1 Wave 2 response rates by Government Office Region

Table 5.6 shows the different response rates and reasons for non-response achieved by region for 2020-21 Wave 2. This shows that across regions the response rate ranged from 81.5% in Wales to 77.1% in the West Midlands.

Table 5.6   Wave 1 sample response rates and non-response by Government Office Region,   2020-21   TCSEW

	North East	North West	Yorkshire & The Humber	East Midlands	West Midlands	East of England	London	South East	South West	Wales
Percentage of eligible addresses (%):										
Non-contact	10.9	10.1	9.3	9.3	10.4	7.2	8.2	10.0	8.9	8.1
Refusal	3.4	4.2	4.0	3.8	4.3	3.6	5.3	3.6	4.1	3.1
Other unproductive	6.3	6.5	5.8	7.7	8.2	7.9	8.3	6.4	7.3	7.3
Achieved interview	79.3	79.3	81.0	79.3	77.1	81.3	78.1	80.0	79.7	81.5

5.2.2 Wave 2 response rate by Police Force Area

Table 5.7 overleaf shows the number of Wave 2 interviews achieved in each PFA and the response rates.

Table 5.7   Wave 2 sample achieved interviews and response rates by PFA, 2020-21 TCSEW

PFA	Achieved	Response rate
	N	*%*
Avon & Somerset	252	79.2%
Bedfordshire	214	80.8%
Cambridgeshire	225	80.6%
Cheshire	187	80.3%
Cleveland	147	78.2%
Cumbria	174	77.0%
Derbyshire	199	76.2%
Devon & Cornwall	318	75.2%
Dorset	242	84.6%
Durham	140	80.0%
Dyfed Powys	232	84.4%
Essex	294	78.2%
Gloucestershire	172	79.3%
Greater Manchester	409	79.4%
Gwent	189	79.7%
Hampshire	323	78.6%
Hertfordshire	211	82.7%
Humberside	212	79.7%
Kent	290	78.6%
Lancashire	224	79.7%
Leicestershire	193	79.4%
Lincolnshire	208	79.1%
Merseyside	259	79.4%
Metropolitan and City of London	1005	78.1%
Norfolk	244	82.2%
North Wales	208	81.9%
North Yorkshire	217	83.5%
Northamptonshire	200	79.7%
Northumbria	177	79.7%
Nottinghamshire	211	82.1%
South Wales	185	79.4%
South Yorkshire	209	81.0%
Staffordshire	145	74.0%
Suffolk	249	84.1%
Surrey	260	82.3%
Sussex	240	78.4%
Thames Valley	407	81.9%
Warwickshire	221	79.5%
West Mercia	203	82.2%
West Midlands	323	74.1%
West Yorkshire	386	80.4%
Wiltshire	219	82.6%

5.3   Wave 3: survey response rate and non-response

The full response and non-response breakdown for the 2020-21 Wave 3 sample is shown in Table 5.9[23].

For Wave 3, 2% of issued cases were identified as not being an eligible case (known as deadwood). The most common type of deadwood was dead/ invalid number, which accounted for 1.6% of all issued cases.

Interviewers made contact with either the selected respondent or a responsible adult at 91% of eligible addresses, meaning a non-contact rate of 9%. The most common (4.6% of eligible cases) was where the contact was consistently going to answer phone or there was no answer.

For eligible cases where contact was made, the most common reason for not getting an interview was due to broken appointment, which accounted for 5.5% of all eligible cases. Respondent refusals accounted for 3% of all eligible cases, while proxy refusals (someone refusing on behalf of the named respondent) were less common (0.3%).

Overall, 2,913 Wave 3 interviews were achieved in 2020-21, representing a response rate of 81.1% and a conversion rate of 79.5%.

---

[23] Response rates are based only on wave 3 sample batches which were closed in the field by 31st March 2021. A total of 9,764 wave 3 cases were issued during the year but most batches were still open in the field after the end of the survey year.

Table 5.9   Wave 3 sample response rate and non-response outcomes, 2020-21 TCSEW

	N	% of issued	% of eligible
TOTAL ISSUED ADDRESSES	3,666	100%	
Deadwood			
Business number	2	0.1%	
Dead / Invalid number	60	1.6%	
Modem/ Fax number	2	0.1%	
Respondent has moved	3	0.1%	
Respondent unknown at number	8	0.2%	
TOTAL DEADWOOD	**75**	**2.0%**	
TOTAL ELIGIBLE ADDRESSES	**3,591**	**98.0%**	**100%**
Non-contact			
Caller ID Block/ Call Barring Message	93	2.5%	2.6%
General call back (not arranged with respondent)	66	1.8%	1.8%
No answer/ Answer Machine/ Number Busy	164	4.5%	4.6%
Total non-contact	*323*	*8.8%*	*9.0%*
Refusal			
Office refusal	21	0.6%	0.6%
Respondent refusal	106	2.9%	3.0%
Proxy refusal	9	0.2%	0.3%
Quit mid interview, refused to finish	1	0.0%	0.0%
Total refusal	*137*	*3.7%*	*3.8%*
Other unproductive			
Broken Appointment	196	5.3%	5.5%
Dialler Error[24]	7	0.2%	0.2%
Inadequate English	0	0.0%	0.0%
Physically or mentally unable	7	0.2%	0.2%
Respondent has died	3	0.1%	0.1%
Respondent too ill/ in hospital	4	0.1%	0.1%
Other unsuccessful	1	0.0%	0.0%
Total other unsuccessful	*218*	*5.9%*	*6.1%*
TOTAL UNPRODUCTIVE	678	18.5%	18.9%
TOTAL INTERVIEWS	2,913	79.5%	81.1%

---

[24] Dialler Error refers to calls that fail due to dial tone irregularities

5.3.1 Wave 3 response rates by Government Office Region

Table 5.10 shows the different response rates and reasons for non-response achieved by region for 2020-21 Wave 3. This shows that across regions the response rate ranged from 84.4% in the South East to 77.4% in Wales.

Table 5.10   **Wave 3 sample response rates and non-response by Government Office Region, 2020-21 TCSEW**

	North East	North West	Yorkshire & The Humber	East Midlands	West Midlands	East of England	London	South East	South West	Wales
Percentage of eligible addresses (%):										
Non-contact	11.8	10.8	11.7	7.2	11.5	9.8	9.0	6.1	7.9	6.8
Refusal	2.8	3.7	4.1	4.3	0.7	2.0	2.8	3.4	3.2	6.4
Other unproductive	2.8	6.4	5.1	7.5	5.9	5.3	7.7	6.1	4.9	9.4
Achieved interview	82.6	79.1	79.1	81.0	82.0	82.9	80.6	84.4	84.0	77.4

5.3.2 Wave 3 response rate by Police Force Area

Table 5.11 overleaf shows the number of Wave 3 interviews achieved in each PFA and the response rates.

Table 5.11   Wave 3 sample achieved interviews and response rates by PFA, 2020-21 TCSEW

PFA	Achieved	Response rate
	N	*%*
Avon & Somerset	80	82.5%
Bedfordshire	60	78.9%
Cambridgeshire	54	76.1%
Cheshire	52	80.0%
Cleveland	38	80.9%
Cumbria	53	76.8%
Derbyshire	52	86.7%
Devon & Cornwall	100	88.5%
Dorset	76	86.4%
Durham	55	83.3%
Dyfed Powys	54	81.8%
Essex	89	87.3%
Gloucestershire	51	77.3%
Greater Manchester	107	81.7%
Gwent	49	81.7%
Hampshire	93	85.3%
Hertfordshire	60	85.7%
Humberside	45	78.9%
Kent	90	81.1%
Lancashire	63	76.8%
Leicestershire	65	84.4%
Lincolnshire	54	80.6%
Merseyside	69	78.4%
Metropolitan and City of London	261	80.6%
Norfolk	69	83.1%
North Wales	42	73.7%
North Yorkshire	57	86.4%
Northamptonshire	49	74.2%
Northumbria	54	83.1%
Nottinghamshire	63	79.7%
South Wales	37	71.2%
South Yorkshire	46	82.1%
Staffordshire	50	87.7%
Suffolk	74	84.1%
Surrey	73	86.9%
Sussex	73	88.0%
Thames Valley	98	82.4%
Warwickshire	68	89.5%
West Mercia	47	83.9%
West Midlands	85	73.3%
West Yorkshire	102	74.5%
Wiltshire	56	82.4%

6.   Offence Coding

This chapter outlines the offence coding process that takes place on the survey. Although changes were made to the victimisation module for the TCSEW compared with the CSEW the aim was to retain all the questions that were critical for offence classification. As such the offence coding processes carried out on the TCSEW were largely consistent with the CSEW.

6.1   History of offence classification on the CSEW

The CSEW Offence Coding System, which was originally developed in 1982 as part of the first Crime Survey, is designed to replicate as far as possible how incidents are classified by the police. The survey counts crime according to the victim's account of events, rather than requiring criminal intent to be proven. This is reflected in how the police record crimes under the National Crime Recording Standard using the Counting Rules[25]. It should be noted, however, that the Counting Rules evolve and change over time, and while efforts are made to reflect these changes in the survey, there are always likely to be some discrepancies between the two systems.

To classify offences, detailed information is collected about the incidents reported by respondents in the victimisation modules. Once the data is returned to the office, all victimisation modules are reviewed by specially trained coders to determine whether what has been reported represents a crime or not and, if so, what offence code should be assigned to the crime.

Apart from some minor changes, the code frame and the instructions to coders on the core survey (see Volume 2 for a copy of the Coding Manual) have remained largely unchanged since 1982. The current operational procedures used for assigning codes have been in place since 2001. In 2010 the coding process was updated to include the coding of offences against 10 to 15 year olds, while in 2015 it was updated to include the classification of fraud and cyber offences. Neither of these changes affected the way in which non-fraud incidents affecting adults were coded.

The coding manual itself is reviewed annually. Most updates are minor modifications to account for new scenarios that evolve over time and to reflect changes in the Counting Rules. However, in October 2018, a more significant update was incorporated to change the classification of offences related to identity theft. Prior to the change these incidents were recorded as computer misuse offences due to unauthorised access to the victim's personal details. After the change was applied these offences were recorded as 'other fraud' offences, reflecting the fraudulent use of a victim's details to apply for a loan or another type of credit agreement. Despite the changes that were being applied to the TCSEW the approach to offence coding remained consistent with the CSEW.

The current Offence Coding System consists of the following steps:

–   For each victimisation module a summary (called an RTF) is produced drawing together the key information from the module into a single easy reference document. This allows the coders to review each incident as a whole and make a judgement on the most appropriate code to allocate based on the totality of the information.

–   In addition to these summaries, the coders use a specially developed computer assisted questionnaire to help them arrive at a final offence code for each incident.

–   As well as recording an offence code for all fraud crimes, coders record whether the offence meets the criteria for being a cybercrime or not.

---

[25] https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/877783/count-general-apr-2020.pdf

- A supervisor checks any codes that the original coder is uncertain about. Additionally, 5% of codes where the coder is certain of the outcome are also checked by a supervisor as a further quality check. These are systematically selected from all cases that have been coded (i.e. every n*th* case) in a particular period.
- A further quality check is carried out by a team at the Office for National Statistics who examine:
  - Any codes that Kantar is uncertain about.
  - Certain types of incident that are automatically referred (e.g. arson).
  - A proportion (minimum of 5%) of certain codes, as part of a general quality control check. Again, these cases are systematically selected from all cases that have been coded.

The result of this process is that every victimisation module has a final offence code assigned to it. Although the coding rules are broadly similar, separate instructions exist for the coding of traditional (non-fraud) incidents and fraud and computer misuse incidents.

A flow chart of the Offence Coding process is shown in Figure 6.1 and the offence coding system is explained in more detail below.

Figure 6.1 TCSEW Offence Coding Flowchart
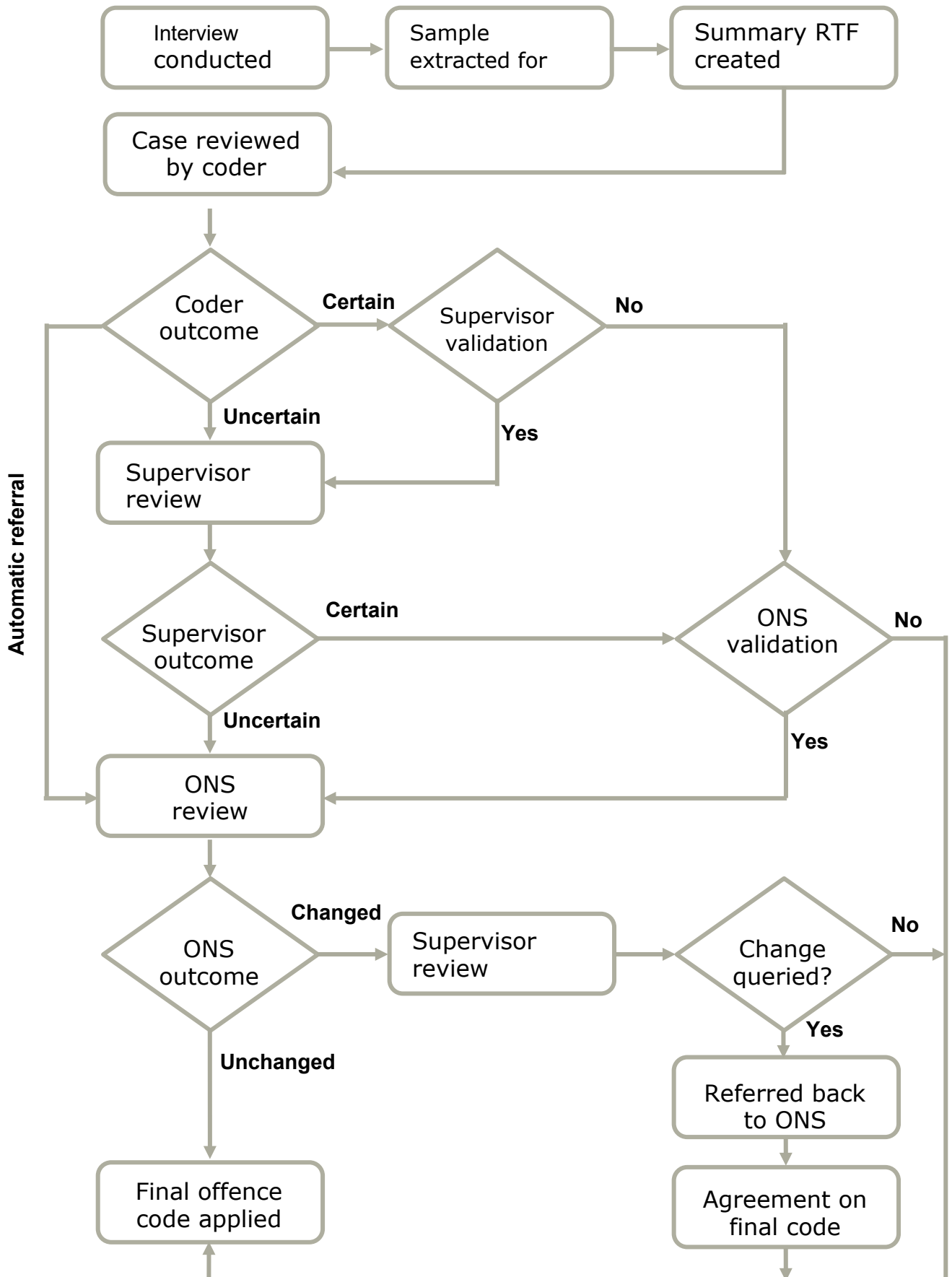
6.2 The offence coding task

Coders are provided with a summary sheet (called an RTF) of the key variables from each victimisation module and this information forms the basis of the coding. This summary sheet includes the open-ended description collected during the interview, as well as some of the key pre-coded questions in the survey which feed into the classifying of offences. It is important that the coders can consider all the information in its totality because sometimes the information collected may not be entirely clear or some of the information may appear contradictory or inconsistent. While a lot of emphasis is placed on the training and briefing of interviewers about collecting comprehensive and accurate data, inevitably there are cases where coders must make judgements about which bits of information to prioritise.

To assist with their task, coders use a specially designed computer assisted questionnaire to carry out the coding. This questionnaire consists of several different modules each of which relate to a high-level offence category (assault, burglary, theft, criminal damage, fraud, etc.). For each case coders must select an offence module to start with. Once in a module the questionnaire programme asks the coders a series of questions about the incident, and they are able to use the information from the RTF to record an answer. The questionnaire is structured like a flow chart to take account of the major rules that apply to offence coding (such as the priority of codes). By answering the sequence of questions based on the information provided in the victimisation module, the coder either reaches an offence code or is directed to another module to repeat the process.

The coders are also provided with a coding manual. The manual contains all the rules that govern offence coding plus further guidance by using specific examples. The manual also provides flow-charts that show how the coding questionnaire works, so that coders can see how they reach a particular offence code on the basis of the answers that they input. The coding manual is kept updated both in terms of major changes to the survey (such as the incorporation of coding guidelines for the 10 to 15 year olds survey in 2010 and the incorporation of fraud and cybercrimes in 2015), as well as being updated to add additional detail and guidance based on the experience of the coders and other feedback.

The current Offence Coding Manual can be found in Appendix I in Volume 2 of the 2020-21 Technical Report.

Once a coder arrives at an offence code using all the resources outlined above, they also record whether they are certain or uncertain that it is the right code. Any case where the coder is uncertain is automatically referred to a supervisor for checking. In addition, supervisors check a minimum of 5% of codes which coders are certain about as part of the quality assurance process.

6.3 Quality assurance by ONS coders

All cases where coders are uncertain about the correct code to assign are automatically referred to ONS. In addition to this, a minimum of 5% of all codes which coders are certain about are selected to be sent to ONS for quality control checking. These are selected in a systematic fashion by selecting every n*th* case in each two-week period.

All quality assurance checks carried out by researchers at ONS take place through an online offence coding portal. Victimisation modules to be checked by ONS staff are uploaded to the portal every week. The offence coding portal contains the unique serial number of each victimisation module, the code that the coder (and supervisor if applicable) has given the incident, how certain the coder (and supervisor) is about the code, and any notes that the coder has added about why they are uncertain. The RTF summary document providing the key variables from the victimisation module are also available from the portal for ONS staff.

Researchers at ONS review each of the victimisation modules sent to them via the portal and add any comments they have on each case. For all cases they either accept the code given by the coder or suggest a different code. These codes then appear on the offence coding portal so that the coders can see the changes that have been made. Apart from making the process more efficient the portal also ensures a complete audit trail for every case.

Once all cases have been reviewed by ONS staff the coding team at Kantar review all cases where a code has been changed. Particular attention is paid to cases where ONS has changed a code that Kantar coders had marked as "certain". If the Kantar coders disagree with the ONS coding decision, it is flagged up in the coding portal to both Kantar researchers and ONS researchers for further consideration and discussion. This approach of iterative review is continued until everyone is agreed on the final outcome code.

As part of the 2020-21 survey, a total of 1,749 cases were sent to ONS for checking, which represented about 18% of all adult victimisation modules (both traditional and fraud cases). Overall, 1,051 traditional (non-fraud) cases were sent for checking (18% of all cases) and 698 fraud cases were sent (18% of all cases).

6.3.1   Traditional (non-fraud) cases referred to ONS

Of the 1,051 traditional (non-fraud) modules sent to ONS:

– 111 cases were automatically referred. This covers cases of aggravated burglary, duplicate cases and cases where the victimisation module was invalid;

– 228 cases were sent because the Kantar coders were uncertain about the code; all uncertain codes are automatically referred;

– 388 cases were sent as part of the systematic quality control check; and

– 324 cases were related victimisation modules. To ensure that those checking offence codes have complete information, all the victimisation modules related to an individual respondent are sent to ONS, rather than just the single module under consideration.

Of the 1,051 non-fraud modules referred to ONS, only 8 cases initially had their code changed by ONS, representing less than 1% of all cases sent. In all cases where ONS changed a code that Kantar coders or supervisors had been certain about, the change was reviewed by a coding supervisor and if there was still disagreement over the final code it was referred back to ONS for further review based on providing additional information on the reasons for reaching a particular code. At the end of this iterative process, only 4 codes were changed from the code originally allocated by the coder or supervisor.

6.3.2   Fraud cases referred to ONS

Of the 698 fraud cases sent to ONS for checking as part of the 2020-21 survey:

– 98 cases were automatically referred to ONS. This covers duplicate cases and cases where the victimisation module was invalid;

– 163 cases were where the Kantar coders were uncertain about the code; all uncertain codes are automatically referred;

– 265 cases were sent as part of the systematic quality control check; and

– 172 cases were related victimisation modules.

Of the 698 fraud modules sent to ONS, 5 cases initially had their code changed by ONS staff, representing less than 1% of all cases sent. However, following further review and discussion only 2 cases were changed from the original code.

6.4   Final Offence Code

Unlike the CSEW, the TCSEW SPSS data sets were delivered to the ONS on a monthly basis. These include all the offence codes that have been given to each victimisation module at each stage of the coding process. This ensures an audit trail exists for each case. The final offence code is derived using a priority ordering system, whereby the ONS code takes priority over the supervisor code, which takes priority over the original code assigned by the coder. The variables on the data file are:

(T)VOFFENCE	Code assigned by the original coder
(T)SOFFENCE	Code assigned by the supervisor (if coded)
(T)FINLOFFC	Code assigned by the ONS team (if coded)
(T)OFFENCE	Final offence code

6.5   Checks on final offence code

Once the SPSS data sets are run some further consistency checks are applied to the final offence codes, checking the offence codes against key pre-coded variables in the victimisation module. The purpose of this is to highlight cases where some of the pre-coded data seems potentially anomalous with the final offence code. Such anomalies can arise because occasionally the information reported by the respondent is not consistent, or even seems contradictory. In particular, there can be inconsistencies between the verbatim description of the incident and subsequent pre-coded questions. While interviewers are carefully briefed to try and be aware of such inconsistencies arising during the interview it is inevitable that some will be missed. Consistency checks within the actual questionnaire script to try and pick up anomalies are not possible when a verbatim description is involved.

The consistency checks carried out are as follows:

- Assaults where no force or violence is recorded as having been used
- Burglary where entry to the property is recorded as being authorised
- Car thefts where no car is recorded as being stolen, or where the police were not informed
- Sexual assaults where there is no sexual element to the assault recorded
- Snatch thefts where the item stolen is not recorded as being held or carried
- Other thefts where the item stolen is recorded as being held or carried
- Wounding where no injury is recorded as being sustained
- In scope offences where the offender is perceived by the victim to be mentally ill
- Thefts where nothing is recorded as having been stolen
- Vandalism where no damage is recorded
- Threats where no threat is recorded

Further checks were added in 2015-16 to check the consistency of the fraud coding:

- Computer virus reported where the offence is not classified as a computer virus
- Computer virus where no virus is reported
- Unauthorised access to personal information with loss of money reported
- Fraud with no loss but a loss has been reported
- Checks that the respondent has been correctly identified as a specific intended victim
- Cyber flag checks if inconsistent reporting is evident:
    - o Computer virus but no cyber element is reported
    - o Classified as a cybercrime but no cyber element is reported
    - o Not classified as a cybercrime but a cyber element is reported.

All cases that fail these checks are examined individually by a researcher and, if changes are required the revised code is reviewed by a coding supervisor. Where clear anomalies in the

data do exist, it is up to the judgment of the researchers to decide which bits of information should be prioritised in arriving at the final agreed offence code. In such cases, greater credence tends to be given to a good verbatim description of the incident over the answers to specific pre-coded questions where, for example, anomalies may be a result of interviewer mis-keying, or respondent misreporting.

Experience of running these checks shows that most flagged cases do have the correct offence codes, but a few may be amended each month as a result of these additional checks.

6.6   Other coding

In addition to the offence coding, coders also looked at all questions where an "other –specify" had been given as an answer. The aim of this exercise, commonly known as back coding, is to see whether the answer given can actually be coded into one of the original pre-coded response options. This coding is done in Ascribe, a Windows based coding package.

Coders are provided with the code frames used in the questionnaire as a starting point for coding each year. Since most of the questions have been used in previous years of the survey, the code frames are already well developed and there is little need to add new codes to the frames. However, if the coding supervisor feels an extra code is needed, this is flagged up to researchers who approved any changes before they are implemented.

As with the offence coding a minimum of 5% of all cases are checked by supervisors as part of the standard quality assurance process.

7.   Data outputs

7.1   Overview

The main outputs provided to ONS on the TCSEW were SPSS data files that were delivered on a monthly basis one month after the end of each fieldwork month. Two data files were provided each month: The Non-Victim File and the Victim File.

The **Non-Victim File (NVF)** is produced at the level of the individual interview and contains all questionnaire data and associated variables, except for information that is collected in the victimisation modules. Because of the TCSEW wave formation a single respondent can have multiple interviews from different waves in this file. Data for both victims and non-victims are included on the Non-Victim File.

The **Victim File (VF)** is produced at the level of the individual incident and contains all the data collected in the victimisation modules. Thus, an individual respondent who reported three crimes and completed three victimisation modules would have three separate records in the Victim File. Because of the TCSEW wave formation reported incidents from a single respondent could be from different reference periods based on the interview wave. All generated victimisation modules were included on the file, including cases where the module either had been suspended or where the reference period was out of scope. Although such records contain no information and are not used for analysis, it is useful to keep these on the file to monitor the number of modules that fall into these categories.

7.2   Delivery of data output

During 2020-21 survey, ten data files (May 2020 to March 2021) were supplied to ONS on a monthly basis, with the initial data file containing both May and June data. Given the clear distinction between the CSEW and the TCSEW, data was supplied on a cumulative basis, meaning each new data delivery was updated by adding the newest month of data.

In addition to the achieved sample, a data file of the entire 2020-21 issued sample was supplied to ONS. This contained information on every issued respondent such as the final outcome, number of calls, call pattern and geo-demographic variables at each wave of the survey.

With the exception of the May-June file, data was delivered a month after the end of each monthly fieldwork period. Each monthly data delivery included interviews that were **achieved** in each specific month, rather than those that were **issued** in a specific month.

7.3   Content of SPSS data file

The SPSS data files delivered to the Office for National Statistics contain various types of variables. The main types of variables contained on the files are:

– Questionnaire variables (NVF and VF).

– Geo-demographic variables (NVF only). All interviews had a set of pre-specified geo-demographic variables attached to them.

– Coding variables (VF). On the Victim File, a full set of offence codes are attached as outlined in Chapter 6.

– Derived variables (NVF and VF). Many derived variables are also added to the file. These consisted primarily of two types: flag variables and classificatory variables

    o Flag variables (NVF and VF) that identify, for example, the date of interview, the month of issue, date of previous interview, whether a partial or full interview, whether a victim or non-victim, etc. On the Victim File, flag variables include whether the record was a long or short victimisation module, whether it was a series or a single incident, and whether it was inside or outside the reference period.

- o Classificatory variables (NVF only) derived from the data. These included standard classifications such as ONS harmonised variables, banded age groups, ethnic groups etc.
  - Weighting variables (NVF only). These are at an individual and household level.
  - Wave information (NVF and VF)

Both the Non-Victim and Victim files include variables that identify the wave of interview and any interview waves that have been missed by the respondent.

7.4   Case identifier

The case identifier is designed to meet the requirements of a continuous survey.

On the Non-Victim File, where each individual case or record represents an interview, the unique interview identifier (TNVFID) is a 10-digit number constructed as shown below

	Column position	Values
Year of issue	1-2	20-21
Area point number	3-6	1000-9999
Address number	7-8	1-99
Screen number[26]	9	0
Wave number	10	1-4

To identify a single respondent across their multiple interviews the respondent identifier (TSERIAL) is a 6-digit number constructed as shown below.

	Column position	Values
Area point number	1-4	1000-9999
Address number	5-6	1-99

On the Victim File, where each individual case or record represents a victimisation module, the unique case identifier (TVFID) is an 11-digit number, which is identical to TNVFID with the addition of the victimisation module number.

---

[36] Taken from original CAPI sample, kept here for ease of maintenance

[26] Screen numbers are used to identify the type of sample. '0' indicates a core sample case.

	Column position	Values
Year of issue	1-2	1-99
Area point number	3-6	1000-9999
Address number	7-8	1-99
Screen number	9	0 or 8
Wave number	10	1-4
Victimisation module number	11	1-6

7.5   Naming conventions

In creating the 2020-21 data files attention was paid to ensuring as much consistency as possible between the face-to-face survey and the telephone survey. Variable names on the TCSEW data files were kept the same as the previous CSEW wherever possible, but with the addition of a 'T' at the start of each variable to signify the switch in survey mode. While it is not the intention that data from the TCSEW and CSEW should even be combined it still made sense to ensure that equivalent variables from the two datasets could be easily linked by users.

One specific requirement arising from the panel approach was that data from one survey wave needed to be combined with data from one or more later waves during the course of the year. This meant it was especially important to systematically document and account for changes to questions over the course of the survey year to avoid confusion among users. For example, small changes to a question from one month to the next (such as adding an extra code to the code frame) could lead to data from different waves being wrongly merged because they appear similar even although they are not. To avoid such situations, the variable names on the 2020-21 data file were changed as and when any changes were made during the year.

While the CSEW was a relatively static questionnaire year-on-year, the unprecedented nature of the pandemic meant that there were questionnaire updates on five different occasions between July 2020 and February 2021. Any variables that were changed during the period are outlined in Table 7.1, overleaf:

Table 7.1   Changes in variables during 2020-21

Module	July 2020 variable	May-June 2020 variable	Reason for change
Covid-19 Module	Tworkcov2B	Tworkcov2	Change to question wording
Covid-19 Module	Tworkcov3Ba-o	Tworkcov3a-o	Change to question wording
Covid-19 Module	TcvkeyworkB	Tcvkeywork	Change to question wording
Covid-19 Module	TcovinchB	Tcovinch	Change to question wording
Module	**November 2020 variable**	**October 2020 variable**	**Reason for change**
Covid-19 Module	TcvrephowBa - j	Tcvrephowa - h	Change to code frame
Covid-19 Module	TcvrepfutB	Tcvrepfut	Change to code frame
Covid-19 Module	Tcvcage3B	Tcvcage3	Refused now allowed
Covid-19 Module	TcvbadexpBa - j	Tcvbadexpa - g	Change to code frame
Demographics and Media Module	Ttypinc3a - o	Ttypinc3a - o	Change to code frame
Module	**December 2020 variable**	**November 2020 variable**	**Reason for change**
Demographics and Media Module	Tanyalcof	Tanyalcof	No change to question text but routing changed
Demographics and Media Module	Tdrugmona - f	Tdrugmona - f	No change to question text but routing changed
Core Victim File			
No Change to existing questions in Victim File during 2020-21			

Table 7.2   Geo-demographic variables added to the survey in 2020-21

Variable	Comments
Tatyp2021	Added
Tagrp2021	Added
Tacat2021	Added
Tmtyp2020	Added
Tmgrp2020	Added

7.6   Don't Know and Refused values

The convention for Don't Know and Refusal codes used in the most recent surveys was maintained on the 2020-21 data. This meant that on the SPSS file the code for Don't Know was '9' for code frames up to 7, '99' for code frames up to 97, and so on. The code for Refused was 8, 98, and so on. Since these are standard codes used throughout the SPSS files, Don't Know and Refused codes are not labelled.

8. Weighting

8.1 Overview of weighting

There are two main reasons for computing weights on the TCSEW:

– To compensate for unequal selection probabilities. In the TCSEW, different units of analysis (households, individuals, instances of victimisation) have different probabilities of inclusion in the sample due to factors such as over sampling of smaller police force areas, the selection of one dwelling unit at multi-household addresses, the selection of one adult in each household, and the inclusion of a single victimisation module to represent a series of similar incidents.

To compensate for differential response. Differential response rates can arise both between different geographic units (e.g. differences in response between regions or between different types of neighbourhood) and between different age and gender sub-groups.

The TCSEW data was weighted after the end of each calendar month of fieldwork. The master dataset included all TCSEW interviews and was updated each month. By the end of March 2021, it contained 36,857interviews from 17,211 individuals (a mean of 2.14 interviews per responding individual, up to a maximum of three)[27].

Kantar produced both individual-level and household-level weights for different subsets of the master dataset, each defined by calendar month:

– 1. All interviews so far from May 2020 onwards (equal to the master dataset)
– 2. All interviews so far from July 2020 onwards (if any)
– 3. All interviews so far from October 2020 onwards (if any)
– 4. All interviews so far from January 2021 onwards (if any)
– 5. All interviews in the latest three calendar months (if at least three available)
– 6. All interviews in the latest calendar month

There are several observations to note about these subsets:

– Subsets of types 1-4 could include more than one TCSEW interview from the same individual.
– From May 2021 onwards, subset 1 will be replaced by 'the latest twelve calendar months' (i.e. June 2020 through May 2021 for that edition, then July 2020 through June 2021 for the next).
– A weight of zero was given to any interview from outside the defined calendar months for subset s.
– Interview data was accumulated each month, almost all of which came from the same (latest) calendar month. However, a handful of cases were left over from earlier calendar months (mainly interrupted interviews). These cases were added to the master dataset, but no revision was made to weights produced the previous month that did not include these cases.
– For each case in the master dataset (one per interview), standard questionnaire data was limited to that collected in the relevant TCSEW interview but *victim form* data could come from an earlier interview, whether TCSEW or CSEW. Victim form data attached to each case in the master dataset always covered the last twelve complete months,

---

[27] Number of interviews and individuals is slightly different from that quoted in previous chapters due to inclusion of under 18s (not included in the sent data) and a small number of cases that were resolved after March 2021 cut-off.

regardless of the source interview (CSEW, TCSEW W1, TCSEW W2, or TCSEW W3).28

8.2   Weighting procedure, stages 1-3

The weights for each subset s were produced in several stages and followed broadly the same pattern.

Stage 1 was to take the full CSEW Reference Sample (all CSEW cases from May 2018 through to February 2020, including those that were not part of the TCSEW sample frame) and, for each case, compute the mean of its 'rolling-12 month' CSEW individual-level calibration weights **C11IndivWgt**.29 This mean weight is called the Base Weight. Applying this weight to the Reference Sample produces an individual-level dataset that is maximally representative of the CSEW target population: individuals aged 16+ living in private residential accommodation in England or Wales, over the period May 2018 through to February 2020.

Stage 2 was to estimate a logistic regression model of the probability that case r in the Base-Weighted Reference Sample is also present in TCSEW subset s. For subsets with multiple cases for the same respondent (e.g. both TCSEW W1 and TCSEW W3), the logistic regression model was extended to estimate the probability that case r appears y number of times in TCSEW subset s.

A fixed set of 66 CSEW variables (individual and household level) was included in the Reference Sample dataset and can be found in Appendix K of Volume 2. These 66 variables were used as candidate predictors for each regression model. The subset of variables used in each model was selected from this list of 66, using an iterative filtering method. Variable v was included in the model if it passed all four of the following filters:

– A bivariate chi square p value of <=0.01 for the hypothesis of zero systematic difference between members of subset s and non-members (or between members with differing numbers of cases in subset s), with respect to variable v

   o A Wald F p value of <=0.2 for the hypothesis that a model containing all the variables passing the first filter has no more predictive power than the same one but excluding variable v.

      ▪ A Wald F p value of <=0.1 for the hypothesis that a model containing all the variables passing the second filter has no more predictive power than the same one but excluding variable v.

         • A Wald F p value of <=0.05 for the hypothesis that a model containing all the variables passing the third filter has no more predictive power than the same one but excluding variable v.

Stage 3 was to produce an interim individual level weight for case r that was equal to the product of (i) its Reference Sample Base Weight, and (ii) one divided by the model-based prediction of the number of times case r appears in subset s.30 Where case r appeared multiple times in subset s, each instance was given the same interim weight.

These interim weights were trimmed to reduce the influence of outliers. First, element (ii) was limited to no more than three times the median value. Second, the product of element (i) and the trimmed element (ii) was limited to no more than five times the median value. The trimmed weights were then scaled to sum to an estimate of the TCSEW target population (individuals

---

28 The TCSEW victim form was restricted to the period since the last interview so TCSEW W1 victim forms either covered the whole of the previous twelve months (if the CSEW interview had been carried out more than twelve months before) or covered all the months since the end of the CSEW interview reference period.

29 The majority of cases in the Reference Sample appear in four rolling-12 month datasets but the later ones appear in fewer.

30 For a very small number of cases in some subsets, this value was imputed because one or more predictor variables in the model had missing data.

aged 18+ living in private residential accommodation in England or Wales). These trimmed and scaled weights were called the stage 3 weights.

8.3 Stage 4: calibration

Kantar carried out stages 1-3 but stage 4 of the weighting procedure was carried out by ONS. Using the stage 3 weights as the baseline, ONS calibrated subset *s* to sex, age and region target population totals, themselves derived from a combination of the contemporary Labour Force Survey and other sources. The method used for calibration ensured that each case in subset *s* from the same individual was given the same stage 4 'calibration' weight just as each case from the same individual had been given the same stage 3 weight. Each individual level calibration weight for each subset *s* has a name of the form **C11Indivwgt**.

8.4 Household level weights

Kantar also produced a stage 3 household level weight for each case in subset *s*. This was equal to the stage 3 individual level weight divided by the most recently recorded total number of people aged 16+ in the individual's household. These household level weights were then scaled to sum to an estimate of the target population (private residential households in England or Wales). This approach treats the household's inclusion in subset *s* as dependent on the interviewed individual's inclusion in subset *s*.

As with the individual level weights, ONS carried out stage 4 of the weighting procedure for households. Taking the stage 3 household level weights as the baseline, ONS calibrated to population totals the set of individuals aged 16+ reported to be resident in households in subset *s*. ONS worked within the constraint that (i) the same calibration weight should be given to each individual in the same household, and (ii) each case from the same household in subset *s* must be given the same calibration weight. Each household level calibration weight for each subset *s* has a name of the form **C11HhdWgt**.

8.5 Victim form weights

Most victim forms cover one incident but some are representative of a 'series' of very similar victimisations, probably perpetrated by the same people. In the CSEW, these incidents are divided up and each is allocated to a specific three month period (calendar quarter) with the respondent's help. Because the TCSEW has a *monthly* data format, series incidents were instead divided up and allocated to specific calendar months. CSEW incidents that had been allocated only to a calendar quarter were randomly allocated to specific calendar months for the purpose of constructing a 'last twelve months' timeline of victimisations.

The base weight for each victim form was equal to either the individual level calibration weight or the household level calibration weight, depending on the type of victimisation. To obtain the final victim form weight, this base weight was multiplied by the number of incidents covered by the victim form that fell within the target reference period, subject to a maximum limit that is specific to the offence code group[31].

---

31 Although the number of incidents is capped for weighting purposes, the actual number of reported incidents in each series (uncapped) is also supplied on the data file.

Table 8.1 shows the maximum limits used for TCSEW. These limits are equal to *either* (i) the 98th percentile series incident count over the period April 2017 to Mar 2020, *or* (ii) 5, whichever is the higher value.

Table 8.1 Limits to 2020-21 victim form weights for each offence code group

Offence code group	98th percentile incident cap
INDIVIDUAL LEVEL OFFENCES	
Violence excepting sex offences, threats and robbery (codes 11,12,13,21,32,33)	10
Sex offences (codes 31,34,35)	5
Threats (codes 91,92,93,94)	9
Robbery (codes 41, 42)	5
Personal theft (codes 43,44,45)	5
Other personal theft (codes 67, 73)	5
Fraud (codes 200,201,202,203,204,205, 206,207,208,210,211,212)	5
Computer misuse (codes 320,321,322,323,324)	5
HOUSEHOLD LEVEL OFFENCES	
Burglary (codes 50,51,52,53,57,58)	5
Other household theft (codes 55,56,65)	5
Motor vehicle crime (codes 60,61,62,63,71,72)	5
Bike theft (code 64)	5
Vandalism (codes 80,81,82,83,84,85,86)	5

9.   Comparing key survey variables with the population

In order to assess the representativeness of the final achieved sample this chapter compares the profile of the 2020-21 survey against population estimates for a range of socio-demographic variables. In addition to comparing the age and sex profile of the survey with the latest population estimates, comparisons are also made with data from the 2011 Census. The tables presented below show the survey profile with the appropriate design weights applied (either household or individual weight) but without the application of the calibration weighting. Comparisons are made based on the 2020-21 achieved sample (i.e., from May 2020 to March 2021) rather than on the 2020-21 issued sample.

9.1   Regional distribution of the sample

Table 9.1 shows the distribution of households by region in the 2020-21 survey compared with the 2011 Census[32]. This shows that the regional profile of the weighted sample was broadly in line with the population distribution.

Table 9.1 Distribution of households by region in the 2020-21 survey compared with the 2011 Census

	2020-21 TCSEW	2011 Census	Difference
	%	%	%
North East	5.0	4.8	0.2
North West	12.5	12.9	-0.4
Yorkshire and The Humber	9.5	9.5	0.0
East Midlands	8.4	8.1	0.3
West Midlands	9.6	9.8	-0.2
East of England	10.9	10.4	0.5
London	12.9	14.0	-1.1
South East	15.5	15.2	0.3
South West	10.2	9.7	0.5
Wales	5.3	5.6	-0.3

9.2   Age and sex profile of the sample

Table 9.2 shows a comparison between the achieved 2020-21 core adult sample and the mid-2019 population estimates for England and Wales by sex and age. This shows that the survey slightly under-represented men and those aged under 35. The profile of the survey by sex and age was similar to previous years. These patterns are fairly typical of large-scale surveys and reflect the lower co-response rates generally achieved among these particular groups.

---

[32] All Census figures presented in the tables are sourced from http://www.nomisweb.co.uk/census/2011

Table 9.2 Age and sex profile of adult sample against mid-2019 population estimates

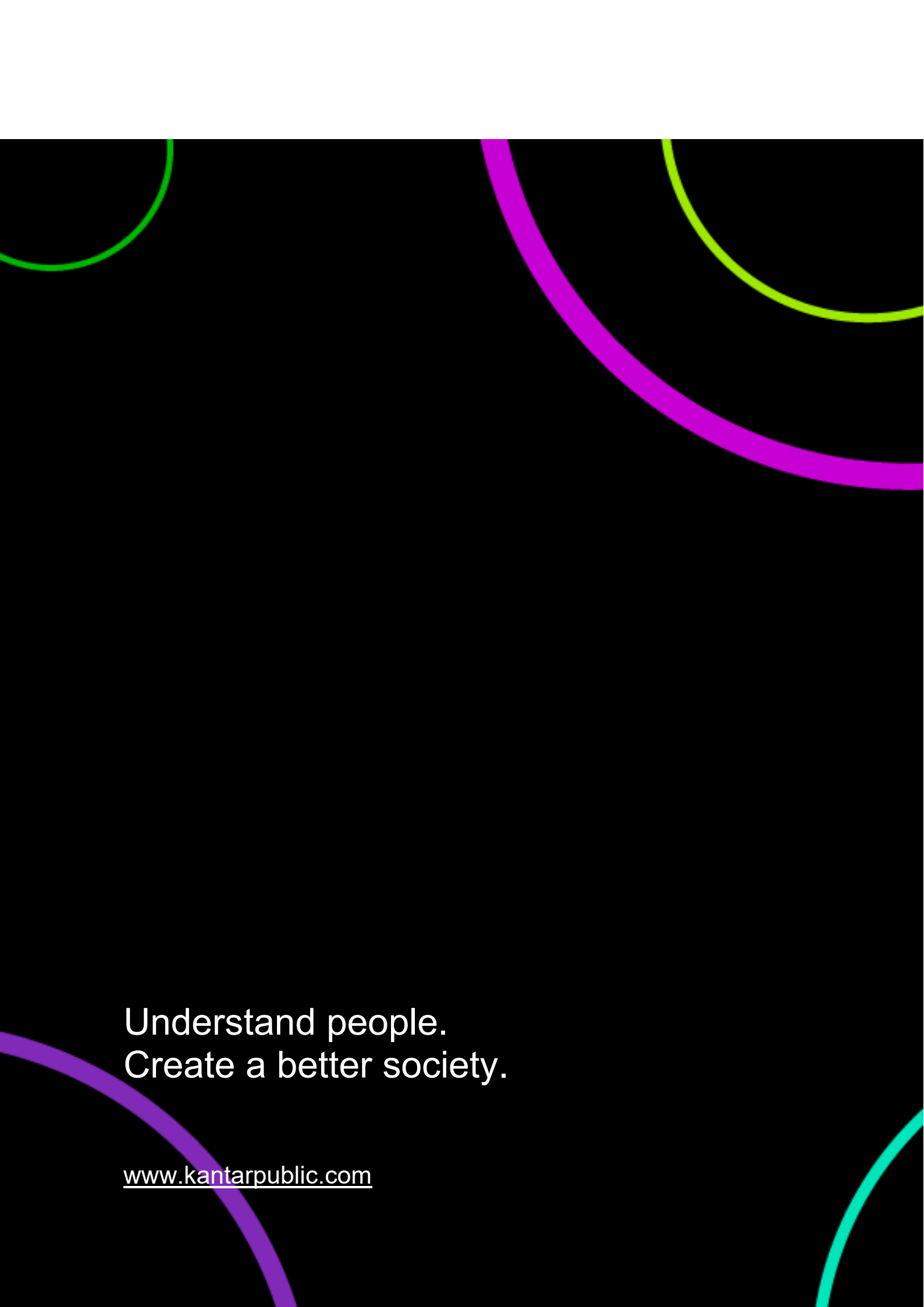|	2020-21 TCSEW	Mid-2019 population estimates	Difference
	%	%	%
Sex			
Male	48.4	49.0	-0.6
Female	51.6	51.0	0.6
Men			
16-19	1.8	5.7	-3.9
20-24	6.2	8.1	-1.9
25-34	16.5	17.1	-0.6
35-44	16.7	15.8	0.9
45-54	17.0	16.9	0.1
55-64	16.5	15.0	1.5
65-74	14.3	12.1	2.2
75-84	8.5	6.9	1.6
85 and over	2.5	2.3	0.2
Women			
16-19	2.0	5.2	-3.2
20-24	5.3	7.3	-2.0
25-34	15.9	16.2	-0.3
35-44	16.3	15.4	0.9
45-54	17.0	16.6	0.4
55-64	16.6	14.9	1.7
65-74	14.3	12.6	1.7
75-84	9.2	8.0	1.2
85 and over	3.4	3.8	-0.4

9.3   Other household characteristics

Table 9.3 shows the profile of the 2020-21 survey compared with some key household characteristics from the 2011 Census. This shows that the survey over-represented two person households. Although housing tenure was broadly in line with the Census there was a noticeable under representation of people living in flats. Those who do not own a car or van are also slightly under-represented.

Table 9.3 Household characteristic of the core adult sample against 2011 Census

	2020-21 TCSEW	2011 Census	Difference
	%	%	%
Tenure			
Owned	65.6	64.3	1.3
Social renting	16.9	17.5	-0.6
Private renting	16.9	18.2	-1.3
Accommodation type			
Whole house or bungalow	81.7	78.6	3.1
Flat, maisonette or apartment	17.7	20.7	-2.8
Household size			
1 person household	29.8	30.2	-0.4
2 person household	36.5	34.2	2.3
3 person household	14.4	15.6	-1.2
4 or more person household	19.3	19.9	-0.6
Car ownership			
No cars or vans	21.3	25.6	-4.3
1 car or van	41.3	42.2	-0.9
2+ cars or vans	37.3	32.1	5.2

Understand people.
Create a better society.