

## Article

# Improving victimisation estimates derived from the Crime Survey for England and Wales

Improvements to the way in which victimisation is estimated using the Crime Survey for England and Wales (CSEW) and the impact they have on the survey's outputs.



Contact:  
Meghan Elkin  
[crimestatistics@ons.gov.uk](mailto:crimestatistics@ons.gov.uk)  
+44 (0)20 7592 8695

Release date:  
24 January 2019

Next release:  
To be announced

## Table of contents

1. [Main points](#)
2. [Overview](#)
3. [Adults: What are the 98th percentile values for the number of incidents in a series?](#)
4. [Adults: Impact on Crime Survey for England and Wales data](#)
5. [Children: What are the 98th percentile values for the number of incidents in a series?](#)
6. [Children: Impact on Crime Survey for England and Wales data](#)
7. [Implementation of new methodology for measuring repeat victimisation in published material](#)
8. [Appendix: Refinement to the weighting methodology](#)
9. [Acknowledgements](#)

# 1 . Main points

This article provides detail about the improvements made to the way in which victimisation incidents are estimated using the Crime Survey for England and Wales (CSEW). Data using this new methodology are published for the first time in the [Crime in England and Wales: year ending September 2018 release on 24 January 2019](#).

The new methodology changes how “repeat” incidents are estimated in the survey and includes a small refinement to the design weights. Repeat victimisation is defined as the same thing, done under the same circumstances, probably by the same people, against the same victim.

This methodological change was implemented in response to feedback we received. We announced the change in November 2016 in a [response to our user consultation](#). This was followed by a [methodological note](#) in October 2017, which outlined additional details of how we would be implementing this methodology.

Using this new methodology, there has been no impact on the long-term picture of total crime. However, the number of incidents for all CSEW crime are slightly higher across the entire time series than previously published. Since the year to March 2002 CSEW, the average increase in total CSEW crime (excluding fraud and computer misuse) was 2.8%.

For most crime types, the estimated number of incidents is unaffected. The increases to the number of incidents are seen primarily in violent offences, where since the year to March 2002, CSEW estimates have increased between 6.4% and 31.6%. This is due to repeat incidents being more common in violent offences.

Changes to the way in which repeat incidents are calculated does not affect the number of victims of crime. However, small changes made to the design weights had a marginal effect on all estimates calculated by the survey. For example, for the year to March 2018 CSEW, the estimated number of victims of violent crime increased by 0.4%.

The improvements of this new methodology include:

- removing the arbitrary limit of 5 on the number of repeat incidents of crime included in the survey estimates
- replacing this limit with a crime-specific imputation method based on the 98th percentile value, to track changes in repeat victimisation over time
- adjusting the design weights used on the survey to reduce the level of variance in the weights, which will in turn lessen the volatility in survey estimates

All releases of crime statistics using CSEW data will use this new methodology from 24 January 2019 and all historic data have been revised to the new methodology. Estimates based upon the previous methodology (incident numbers capped at 5) will no longer be published from January 2019 onwards. Users should not use releases published before January 2019 for data on the number of incidents from the CSEW.

## 2 . Overview

The Crime Survey for England and Wales (CSEW) was initially designed as a research tool. The aim was to answer questions that the police recorded crime series could not. These included, for example, which groups in society were at greater risk of victimisation and how much crime went unreported to the police. It was not conceived as being a source of [National Statistics](#). Therefore, those designing it took pragmatic decisions about weighting and processing on a relatively small sample of 10,000 respondents.

After becoming an established survey, the CSEW has since been used to estimate the number of times a person is a victim of crime. This allows estimation of the total number of crimes experienced by adults living in households in England and Wales. However, neither the survey weights, nor the processing of incident counts, were initially designed with this in mind.

In recent years the survey has received criticism<sup>1</sup> for the way it measures repeat victimisation and by extension, the resulting estimates of incidents of crime. Feedback from users suggests a continuing need for estimates of both the number of victims and the number of crimes experienced in the general population. However, there was also a need to improve the way we measure incidents of repeat victimisation.

Producing estimates for incidents of crime is unproblematic for most crime types, since most crimes are unaffected by high levels of repeat victimisation. However, for some crime types, a relatively small number of victims yield a high number of victimisations. The inclusion or exclusion of these individuals in the sample for any given year can make estimates volatile and difficult to use, particularly when understanding changes in crime over time.

Since the survey began in 1981, “repeat” incidents have been limited to a total of 5. Historically, including a maximum of 5 repeat incidents for any individual victim had proven to be an effective way of reducing the effects of sample variability from year to year. This approach enabled the publication of incident rates that were not subject to large fluctuation between survey years. This approach yields a more reliable picture of changes in victimisations over time once high order repeat victimisations were treated in this way.

However, for some crime types, such as violence, this resulted in point estimates being consistently lower than estimates if all high order repeat victimisations were included. It may also have introduced additional measurement error where high order repeat victimisation disproportionately affected a sub-group within the population, for example, women suffering from sustained repeat victimisation by a violent partner or family member.

## Decision-making

In July 2016 we commissioned an [independent review](#) and ran a [user consultation](#) on the issues associated with measuring high order repeat victimisation in the CSEW. The results of the consultation were presented at the National Statistician's Crime Statistics Advisory Committee in September 2016. Based on advice from the Committee and the consultation responses received, in November 2016 we published our [response](#).

In our response, we recognised that removing the cap of 5 was essential to improving crime statistics for many of our users. We summarised the feedback we received and the decisions made, agreeing to:

- change the current methodology of arbitrarily capping repeat incidents at 5
- adopt a lighter cap of the 98th percentile of victim incident counts for each crime type (calculated over several years)
- not use annualised multiple-year aggregations of data
- revise the time series back as far as possible
- make available uncapped estimates (with appropriate caveats) as part of our methodology information

In making these decisions, different methodological approaches were assessed to consider factors important to our users. These included: the level of volatility introduced into time series data; the sensitivity of different approaches to measuring changes in repeat victimisation over time; and the level of transparency surrounding the methodology. The review of existing methodology also revealed an issue with the production of dwelling weights, which led us to refine our weighting procedures.

Based on this work and advice from the National Statistician's Crime Statistics Advisory Committee, we published a further [methodological note](#) in October 2017. This note outlined some specifics to adopting the 98th percentile value (of the number of incidents within series of each headline crime type) as a maximum value imposed on incident counts. This included:

- adopting the use of three-year rolling datasets to calculate 98th percentile values for the number of incidents in a series, which enabled us to obtain 98th percentiles that balance the need for stability with the ability to respond to changes in repeat victimisation over time
- not lowering the cap of 5 for crime types with 98th percentile values lower than this, avoiding introducing additional bias when there is very little volatility
- removing (difficult to interpret) “too many to remember” responses from the data when calculating the 98th percentile values and subsequently imputing the 98th percentile value in their place
- adjusting our design weights to better suit the inclusion of count data by trimming component weights prior to calibration
- publishing uncapped data available as part of our methodology information to give users choice over which estimates to use, while recognising these estimates of incidents will be subject to considerable volatility from year to year and are not the preferred trend measure

The decision to trim the component weights prior to calibration was added to reduce large variability between weights that had the potential to interact with count data and produce misleading analysis. Making this amendment prior to calibration ensured non-response weights were untouched and that all weights were still calibrated to the equivalent of the resident household population for England and Wales. The result is a negligible change in some prevalence estimates and smoother, less volatile time series for numbers of incidents of crime.

Except for a final decision on adjusting CSEW weights, our [methodological note](#) published in October 2017 outlined each of these decisions in detail and indicated the impact these changes would have on the CSEW data. Further details on the changes we have made to our weighting procedures are available in the [Appendix: Refinement to the weighting methodology section](#).

## Impact of the new method

Since October 2017, we have been working on the production of revised Crime Survey estimates (back to 1981<sup>2</sup>) in line with this approach, maintaining valuable trend data for our users. This includes estimates of crimes against children aged 10 to 15 years from the year ending March 2010, when these estimates were first produced by the survey. We have also taken this opportunity to make other improvements to the data and have now implemented consistent weighting and calibration techniques back as far as 1991 for the first time. Data are now available in the [Crime in England and Wales: year ending September 2018](#) release published on 24 January 2019.

Under both the new and old methodologies the way in which crime has changed over time, as measured by the CSEW, is almost identical. For example, total CSEW crime excluding fraud and computer misuse for both series peaked in 1995. This was followed by a series of falls that were sustained throughout the early parts of the new millennium with most recent estimates reflecting a more stable picture. The number of victims of crime is also unaffected by changes to the way in which incidents are counted<sup>3</sup>.

Although most crime types are minimally affected by the new methodology, reweighting has meant that all data are affected to some extent. For the year ending March 2018 CSEW, the new methodology has increased the total number of incidents for all CSEW crime (including fraud and computer misuse) by 158,000 from 10.6 million to 10.7 million incidents, an increase of 1.5%.

The most substantial effect was found in relation to violent incidents. For the year ending March 2018 CSEW, the new methodology has increased the total number of violent incidents by 167,000 from 1.3 million to 1.4 million incidents, an increase of 13.3%.

The increase in violence being greater than the increase in total CSEW crime occurs because estimates for other crime types have decreased. This is caused by a combination of no change in the capped number of incidents for these crime types (with the 98th percentile values being no larger than 5) and the trimming of the component survey weights. The [Appendix: Refinement to the weighting methodology section](#) has further details on our weighting refinements.

For the purposes of this article, we have focused on incident numbers since this is where the greatest impact can be seen. However, new time series for estimates of incident rates, prevalence rates and the number of victims have also been published (see the [Appendix Tables](#) in the [Crime in England and Wales: year ending September 2018](#) release).

## Notes for: Overview

1. There is a long history of criticism of the way repeat victimisation is measured by the CSEW, both [historical criticism](#) (PDF, 782KB) and [more recent criticism](#) (PDF, 256KB).
2. The calendar years 1981, 1983 and 1987 have not been reweighted or recalibrated since the available datasets do not lend themselves to this. However, all annual datasets from 1991 onwards have been reweighted and recalibrated and are now consistently weighted for the first time, enhancing the comparability of our time series.
3. A small revision to the weighting procedure whereby the component weights are now trimmed does mean that the estimated number of victims of crime are marginally affected, but differences are practically indistinguishable from previous estimates.

## 3 . Adults: What are the 98th percentile values for the number of incidents in a series?

[Table 1](#) details the three-year rolling 98th percentiles for the number of incidents in a series by crime type within the Crime Survey for England and Wales (CSEW). These 98th percentiles have been applied to the dataset associated with the latest year included in the calculations. For example, 98th percentiles for the three years ending March 2018 will be applied to the year ending March 2018 estimates. The 98th percentiles act as a new maximum number of incidents that can be included within a series.

For most crime types, in the majority of instances, the 98th percentile value is lower than 5, indicating a low level of repeat victimisation for these crimes. Where this is the case, we have not lowered the maximum number of incidents counted within a series below 5 (and thus, included numbers of incidents above the 98th percentile). Hence, the impact on most estimates is minimal and any changes will only result from the minor adjustments made to some components of the design weights<sup>1</sup>. The estimates for criminal damage and violence are the only categories to be noticeably impacted by implementing the 98th percentile methodology.

For criminal damage, the 98th percentile values ([Table 1](#)) vary slightly over time, with a small peak in the three years to December 1997 and 1999 (a 98th percentile value of 9 in both periods). For many years the value of the 98th percentile is only slightly higher (with a value of 6) than the previous cap of 5. As a result, the number of criminal damage incidents estimated by the survey is marginally higher under the new methodology compared with the old one. The increase is no more than 4% in any single year, with one exception of 9% in the year ending December 1997.

As expected, violent offences have seen the most significant impact from the change to the new methodology. The 98th percentile values ([Table 1](#)) increase substantially and vary between 9 and 20 across all years of the survey back to 1981. The highest 98th percentile value of 20 in the time series has been applied to the year ending March 2015 dataset. This was a result of high levels of repeat victimisation being recorded by the survey in two of the three years on which the 98th percentile value was based: the year ending March 2013 and the year ending March 2015. Following this peak, the value of the 98th percentile decreases, tracking the fall in repeat victimisation to more usual levels. The impact of moving to the 98th percentile values on the data, particularly over time, is investigated in the remaining sections of this article.

It should be noted that the CSEW is not well-suited to measuring trends in crimes that occur in relatively low volumes, such as robbery. This is particularly true in the early years of the survey prior to the year ending March 2002 where the sample sizes were particularly small, varying between 10,000 and 19,000 each year. While the 98th percentile values for robbery changed only slightly with most years unaffected, a value of 10 was recorded in the three years to March 2002. However, even in this year the impact was minimal and from the year ending March 2002 onwards, incident estimates do not vary by more than 2% under the new methodology compared with the old one.

### **Notes for: Adults: What are the 98th percentile values for the number of incidents in a series?**

1. It is important to note that as a result of changes to the weighting, both prevalence and incidence estimates have changed somewhat for all crime types, regardless of whether the level at which we trim counts of repeat incidents is increased.

## **4 . Adults: Impact on Crime Survey for England and Wales data**

### **Trends in CSEW crime**

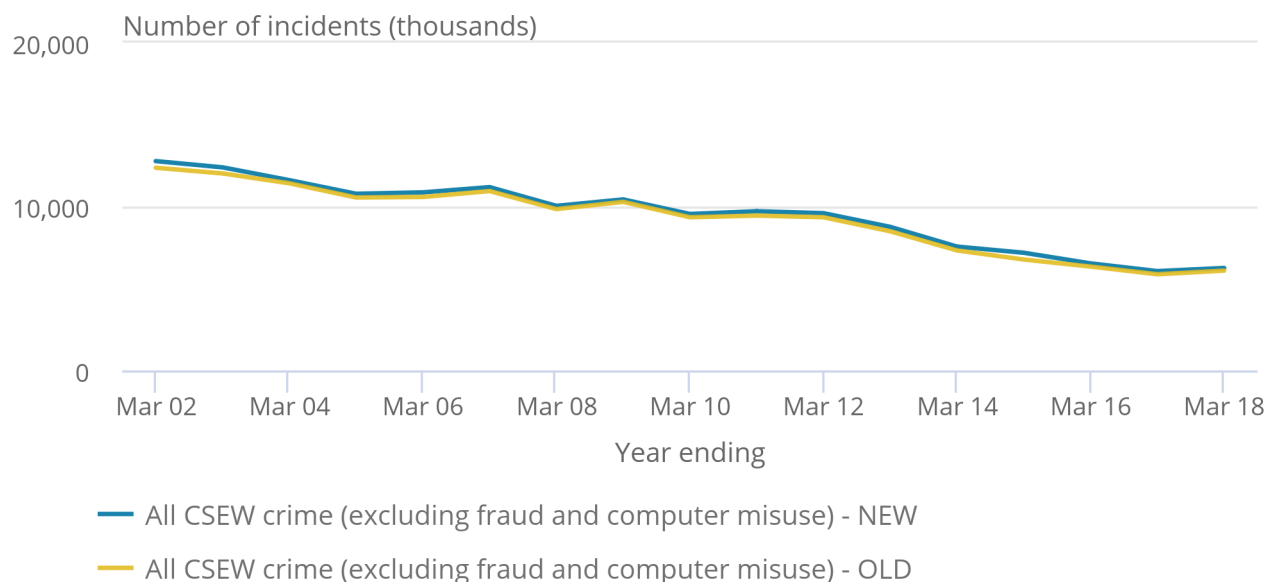
As seen in Figure 1, trends in total crime<sup>1</sup> estimated by the Crime Survey for England and Wales (CSEW) are not greatly impacted by the methodological change. This is owing to violence being the only crime type substantially impacted by the new methodology, typically equating to around 20% to 25% of all CSEW crime. It is therefore not surprising that the changes we have implemented have minimal impact on the trends in all CSEW crime, or that the increased volume we have reported is consistent over time.

**Figure 1: Trends in all Crime Survey for England and Wales crime (excluding fraud and computer misuse), by old and new methodology**

England and Wales, year ending March 2002 to year ending March 2018

## Figure 1: Trends in all Crime Survey for England and Wales crime (excluding fraud and computer misuse), by old and new methodology

England and Wales, year ending March 2002 to year ending March 2018



Source: Office for National Statistics - Crime Survey for England and Wales

### Notes:

1. Data relate to adults aged 16 years and over or to households.
2. New victimisation questions on fraud and computer misuse were only introduced into the CSEW from October 2015; in order to present a consistent time-series, fraud and computer misuse offences have been excluded.

[Table 2](#) shows that the new methodology increases the number of incidents of all CSEW crime by around 2% to 3% for each year of the survey. The one exception is a 6% increase in the year ending March 2015, which was owing to particularly high order repeat victimisation for violence recorded by the survey around this time.

## Trends in violent crime

Repeat victimisation is more pronounced in violent crime than any other crime type. This is particularly the case for violence in a domestic setting, where victims are more likely to suffer multiple incidents under similar circumstances. We have now refined our methodology to be more sensitive to repeat victimisation.

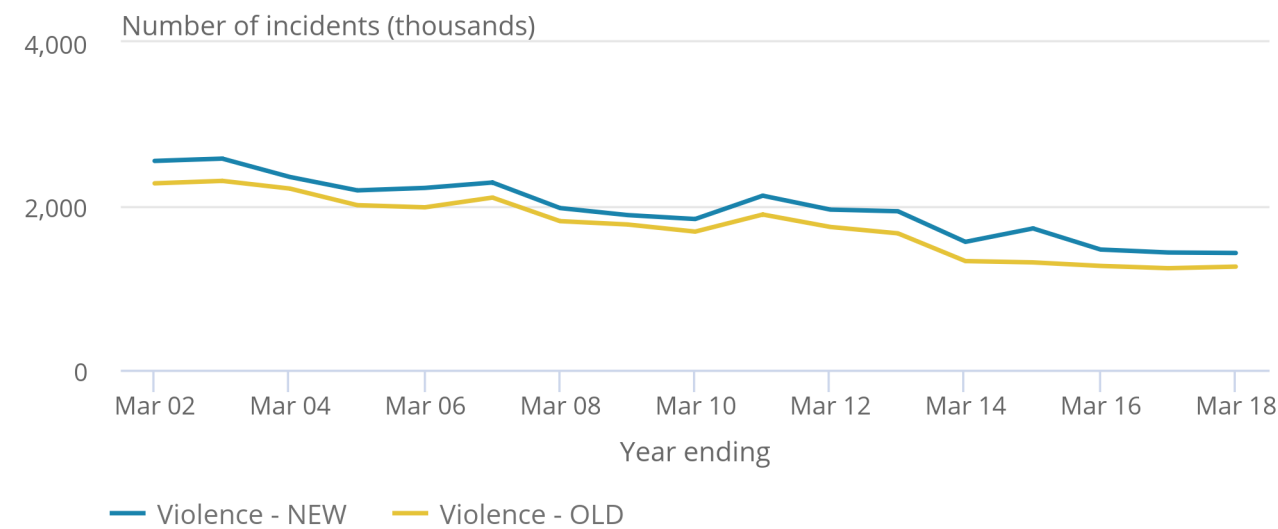
The new methodology is picking up anywhere between 6% (year to March 2009) to 32% (year to March 2015) more violent incidences, depending on the levels of repeat victimisation experienced by respondents in any given year ([Table 3](#)). At the same time, we have balanced the need for this improvement with the requirement for a usable time series of violent crime. As can be seen in Figure 2, our new estimates for all CSEW violence follows a similar trend to that seen in our already published data. Our new methodology has balanced out the extreme volatility that we saw in trialling other approaches and in looking at uncapped data.

**Figure 2: Trends in Crime Survey for England and Wales violence, by old and new methodology**

England and Wales, year ending March 2002 to year ending March 2018

Figure 2: Trends in Crime Survey for England and Wales violence, by old and new methodology

England and Wales, year ending March 2002 to year ending March 2018



Source: Office for National Statistics - Crime Survey for England and Wales

Notes:

1. Data relate to adults aged 16 years and over.

## Types of violent crime

[Appendix Tables 1 and 6](#) within the [Crime in England and Wales, year ending September 2018](#) release present the new time series for the breakdowns of violent crime. Under the new methodology, the volumes have increased for every year (compared with the previously published estimates). For example, “Violence with injury” increases by between 5% and 26% over the time series and “Violence without injury” increases by between 2% and 40% compared with previously published figures. Trends for violence with and without injury from the year ending March 2002 to March 2018 are displayed in Figures 3 and 4.

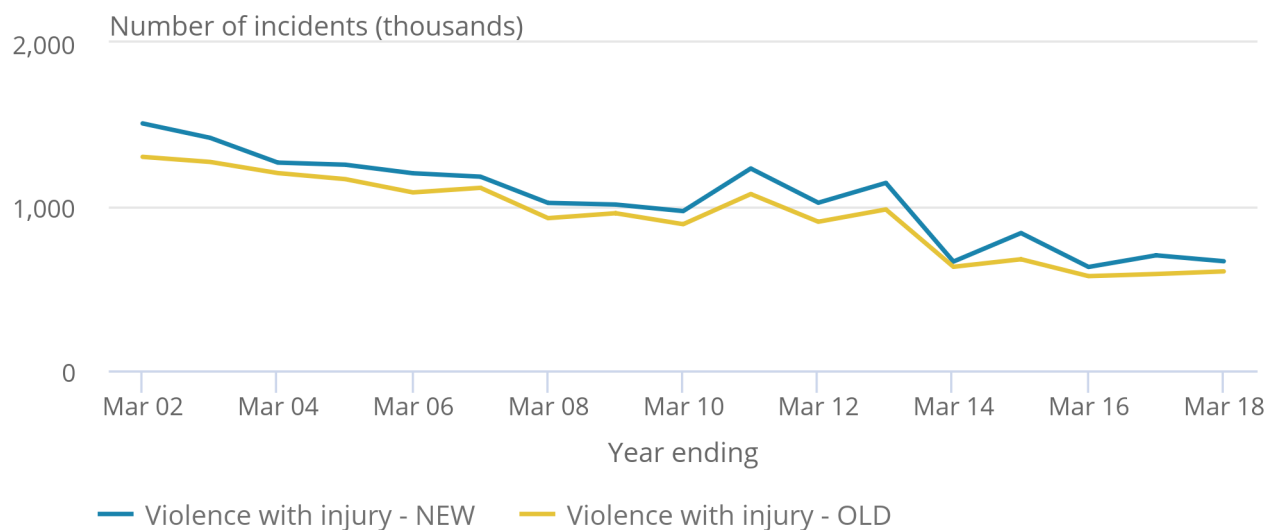


**Figure 3: Trends in Crime Survey for England and Wales violence with injury, by old and new methodology**

England and Wales, year ending March 2002 to year ending March 2018

### Figure 3: Trends in Crime Survey for England and Wales violence with injury, by old and new methodology

England and Wales, year ending March 2002 to year ending March 2018



Source: Office for National Statistics - Crime Survey for England and Wales

**Notes:**

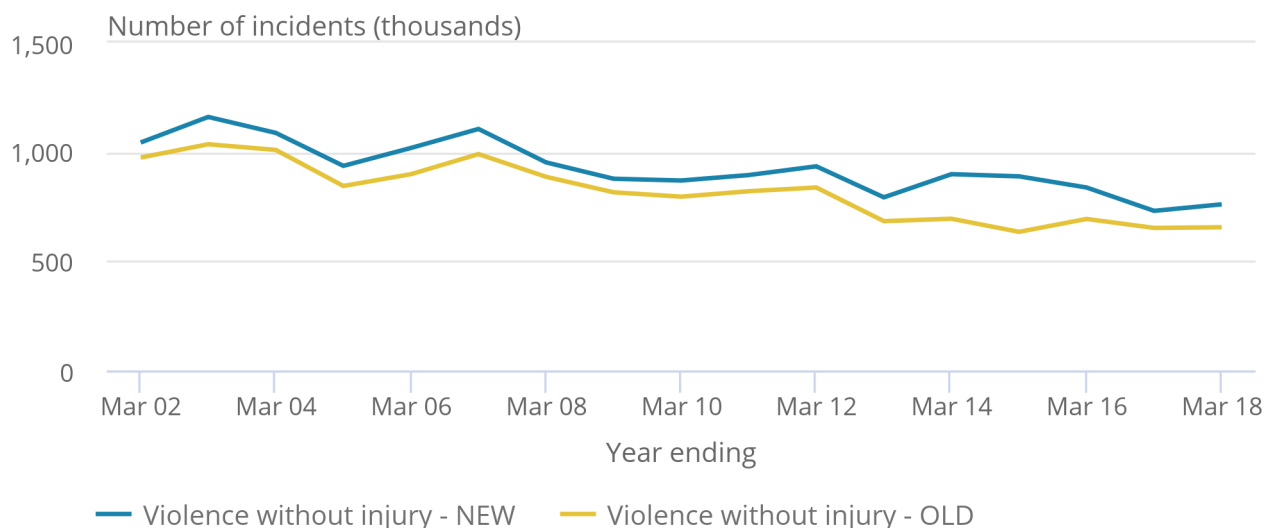
1. Data relate to adults aged 16 years and over.

**Figure 4: Trends in Crime Survey for England and Wales violence without injury, by old and new methodology**

England and Wales, year ending March 2002 to year ending March 2018

## Figure 4: Trends in Crime Survey for England and Wales violence without injury, by old and new methodology

England and Wales, year ending March 2002 to year ending March 2018



Source: Office for National Statistics - Crime Survey for England and Wales

### Notes:

1. Data relate to adults aged 16 years and over.

While in the main these trends do not change, despite resulting in slightly larger volumes, there are some exceptions. For example, we see a slightly different trend in violence without injury from the year ending March 2012 onwards. The new series shows a (non-significant) 13% rise compared with a (non-significant) 2% rise in the old series between the year ending March 2013 and year ending March 2014.

The new series also shows a (non-significant) 6% fall compared with a (non-significant) 9% rise in the old series between the year ending March 2015 and year ending March 2016. This is the result of the 98th percentile cap falling from 20 in the year to March 2015 to 14 in the year to March 2016, whereas previously, incident numbers for both years were capped at the same amount. This demonstrates that the new series is accounting for higher levels of repeat victimisation and better tracks these changes over time.

The other breakdown of violent crime that is of particular interest is based on the relationship between the perpetrator and the victim. We regularly publish estimates of violent incidents broken down by whether the perpetrator was an acquaintance, a stranger, or someone who lived in the same household as the victim (categorised as "domestic"). These estimates were cited most in critiques of the way we previously estimated levels of repeat victimisation and there was a call to produce uncapped estimates of these types of violence.

It is important to note that estimates of domestic violence reported here are from the main section of the Crime Survey. These estimates are collected from face-to-face interviews with respondents and are separate from estimates of physical domestic abuse reported in the self-completion section of the survey. The self-completion module employs a broader definition of physical domestic abuse (it includes threats or force) and is unaffected by the cap. In contrast, the face-to-face module has been affected by the cap of 5 and includes only incidents of physical violence.

Due to these definitional differences, it is difficult to make any direct comparison between those who reported physical domestic abuse in the self-completion module with those who report the similar category of domestic violence in the face-to-face interview. However, such comparisons suggest that respondents, quite logically, are more likely to report experience of such sensitive issues in the self-completion module compared with the face-to-face interview. Of those aged 16 to 59 years in the year to March 2018 CSEW who reported being a victim of force in the last 12 months in the self-completion module, only 12% reported being a victim of domestic violence in face-to-face interviews (15% for women and 9% for men). Therefore, estimates of the number of incidents of domestic violence from the main face-to-face interview (employing any methodology) should be treated with caution.

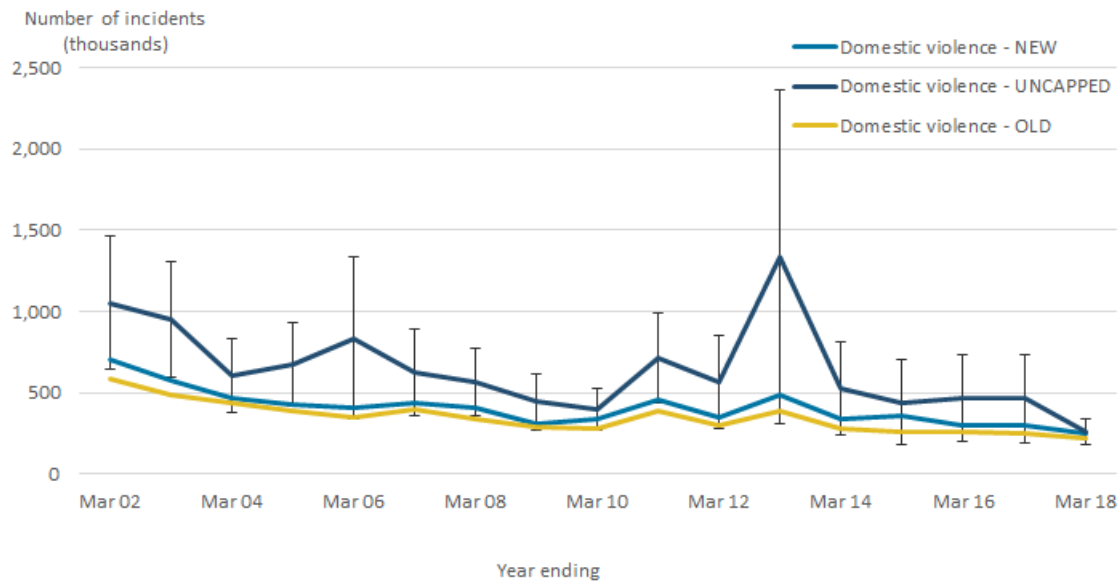
[Table 4](#) shows that the new methodology picks up a substantially higher number of domestic violence crimes. The estimated number of incidents of domestic violence has increased between 7% and 39% each year compared with those previously published.

There is a similar scale of increase in estimates of incidents of acquaintance violence (increases between 4% and 41% each year compared with previously published estimates). Additionally, as expected, there is a smaller scale of increase in estimates of incidents of stranger violence (between 2% and 20% each year compared with previously published estimates). Stranger violence is considerably less likely to be subject to high levels of repeat victimisation.

Figures 5, 6 and 7 show estimates for all breakdowns of violence by relationship of victim to perpetrator alongside uncapped estimates. For each breakdown, the new methodology enables us to include more incidents than previously published while not interrupting the ability to measure trends over time. In comparison, the uncapped estimates are much more volatile over time and even when we see sharp rises and falls between years, they are often not statistically significant. Importantly, the new estimates we are publishing typically fall inside the confidence intervals of the uncapped estimates.

**Figure 5: Trends in Crime Survey for England and Wales domestic violence by differing methodologies**

**England and Wales, year ending March 2002 to year ending March 2018**



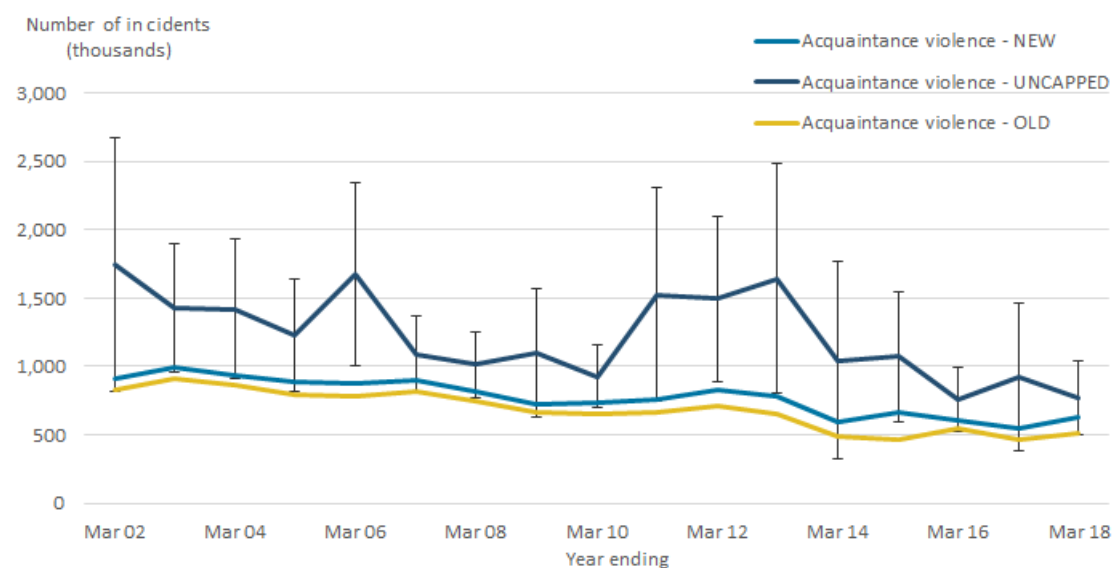
**Source: Office for National Statistics - Crime Survey for England and Wales**

**Notes:**

1. Data relate to adults aged 16 years and over.
2. The margins of error around the uncapped estimates are represented by the 95% confidence intervals.

**Figure 6: Trends in Crime Survey for England and Wales acquaintance violence by differing methodologies**

England and Wales, year ending March 2002 to year ending March 2018



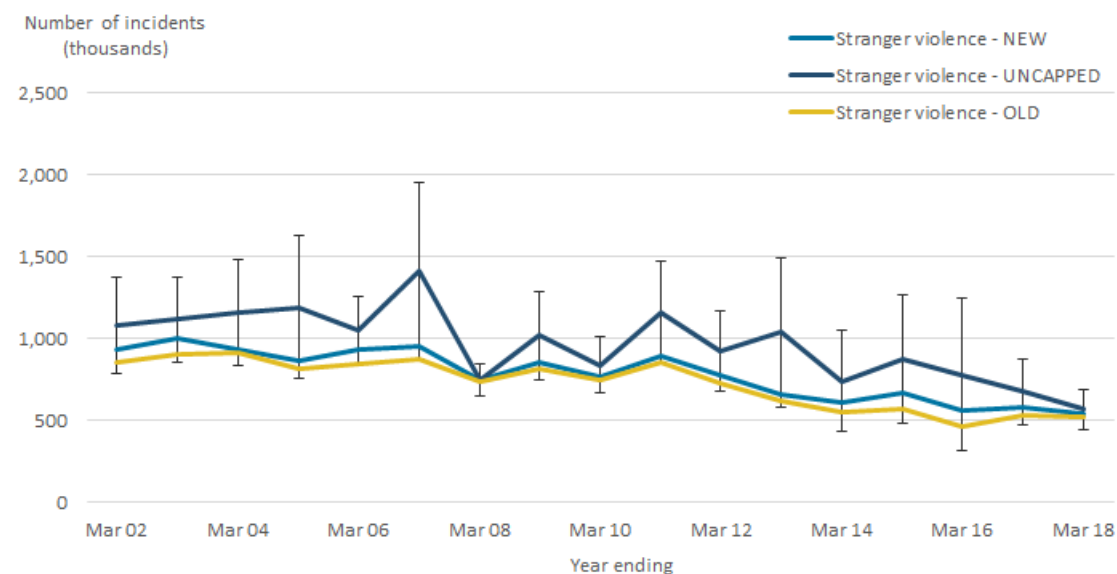
Source: Office for National Statistics - Crime Survey for England and Wales

**Notes:**

1. Data relate to adults aged 16 years and over.
2. The margins of error around the uncapped estimates are represented by the 95% confidence intervals.

**Figure 7: Trends in Crime Survey for England and Wales stranger violence by differing methodologies**

England and Wales, year ending March 2002 to year ending March 2018



Source: Office for National Statistics - Crime Survey for England and Wales

**Notes:**

1. Data relate to adults aged 16 years and over.
2. The margins of error around the uncapped estimates are represented by the 95% confidence intervals.

## Violence by sex of victim

Some criticism of the way the CSEW measures repeat victimisation looked at sex distributions. Analysis of uncapped incidents of violence was cited<sup>2</sup> showing the cap of 5 masked the true sex distribution of violence (owing to the repeat nature of crimes that are more prolific among women, such as domestic violence).

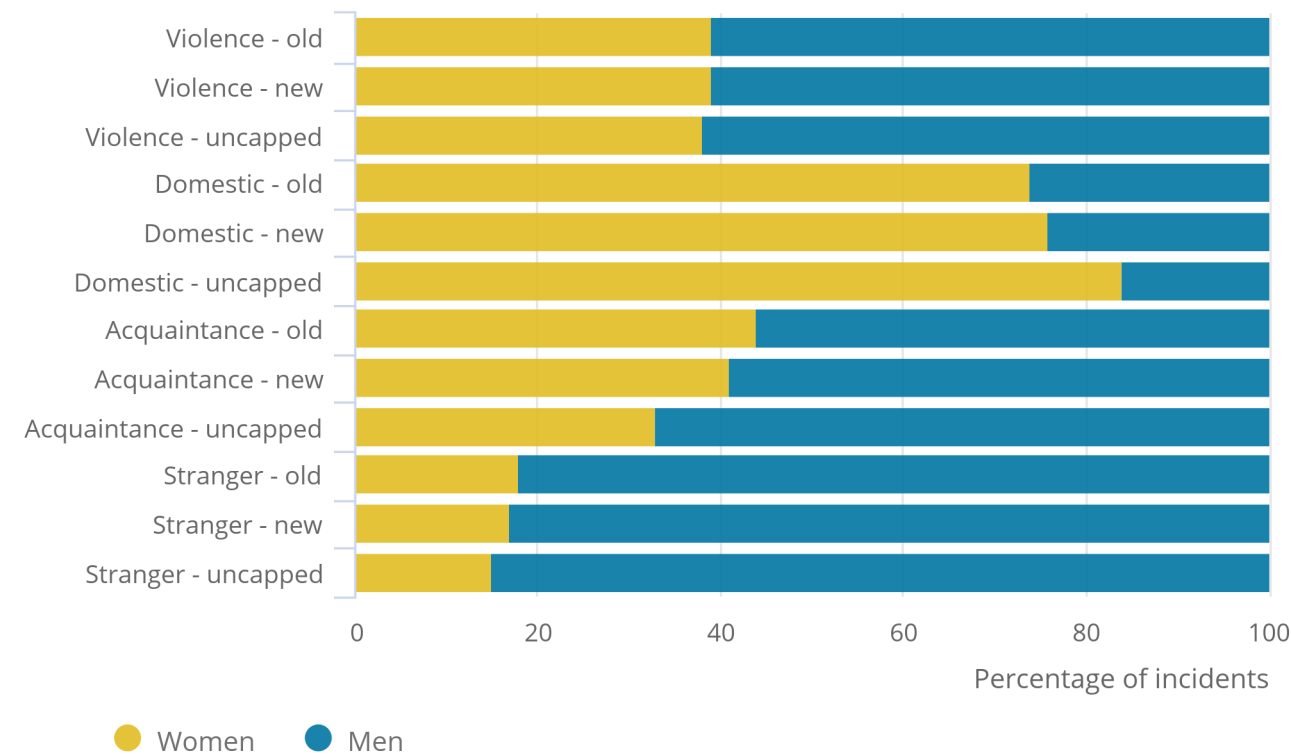
As can be seen in Figure 8, the uncapped estimates of violence show a more extreme sex distribution of uncapped domestic violence than the estimates previously published. For example, when using uncapped estimates for the year ending March 2017<sup>3</sup>, women experienced 84% of domestic violence incidents reported by respondents to the survey. This compares with 74% of the (smaller number of) domestic violence incidents we previously published and 76% of the number of incidents using our new methodology.

**Figure 8: Distribution of Crime Survey for England and Wales violent incidents between men and women, by differing methodologies**

England and Wales, year ending March 2017

Figure 8: Distribution of Crime Survey for England and Wales violent incidents between men and women, by differing methodologies

England and Wales, year ending March 2017



Source: Office for National Statistics - Crime Survey for England and Wales

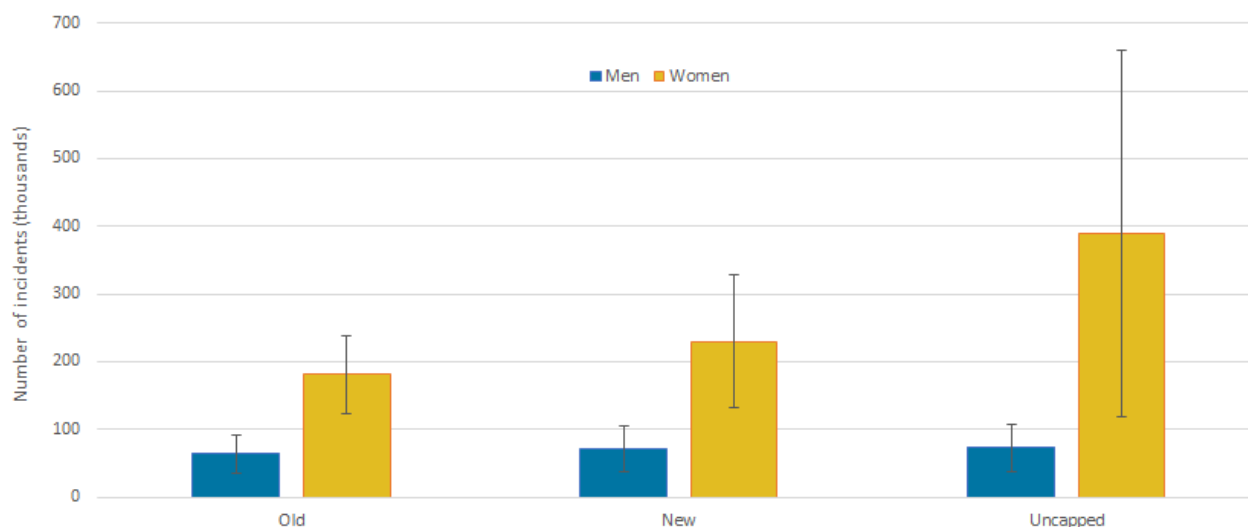
Notes:

- 1. Data relate to adults aged 16 years and over.

Figure 9 shows the associated confidence intervals around estimates of domestic violence in the year ending March 2017 by sex for each version of the methodology. For this particular year, no methodology produces confidence intervals that do not overlap for domestic violence incidents.

**Figure 9: Confidence intervals around estimates of Crime Survey for England and Wales domestic violence for men and women by differing methodologies**

England and Wales, year ending March 2017



Source: Office for National Statistics - Crime Survey for England and Wales

**Notes:**

1. Data relate to adults aged 16 years and over.
2. The margins of error around the uncapped estimates are represented by the 95% confidence intervals.

Given that the uncapped confidence intervals around these estimates are so wide, the new methodology will be equally (and possibly more) effective in detecting significant differences between the number of incidents experienced across the sexes. The increased volume of incidents being included in the data using the new methodology better reflects the distribution of incidents between men and women than our previously published ones (without the same large confidence intervals that we see in the uncapped data). This is demonstrated in the year ending March 2017 data, where the new methodology identifies the same significance differences in the number of domestic violence incidents experienced between the sexes as the uncapped estimates.

## Confidence intervals for the new methodology

[Table UG2](#) presents all the headline estimates we usually publish as part of our quarterly bulletins (as well as the associated confidence intervals). [Table U1](#) presents estimates and associated confidence intervals using the uncapped data.

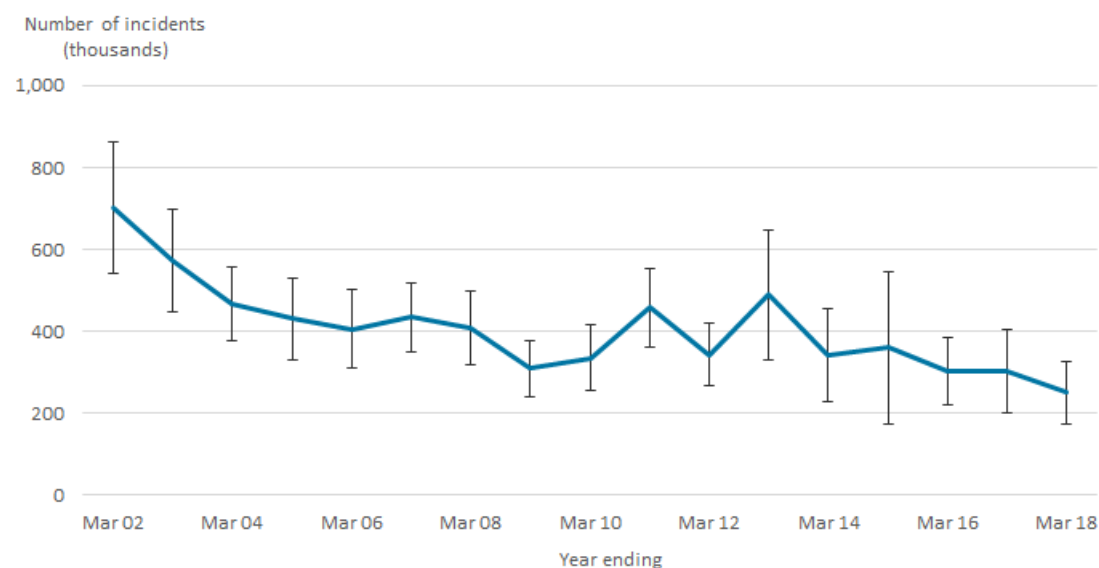
The new methodology widens the confidence intervals associated with estimates for the number of incidents of violence. This is to be expected given the greater volatility introduced into the estimates. For example, confidence intervals produced using the new methodology for the year ending March 2018, as compared with those associated with the data we have previously published, are larger for all forms of violence. The confidence interval for domestic violence is 23% wider, for stranger violence is 9% wider and for acquaintance violence is 55% wider. However, when compared with the uncapped confidence intervals ([Table U1](#)) these are still preferable in the sense that they provide greater precision around the point estimates.

Confidence intervals for domestic, acquaintance and stranger violence as measured by the new methodology are presented in Figures 10, 11 and 12 respectively.



**Figure 10: Trends in Crime Survey for England and Wales domestic violence with new methodology**

England and Wales, year ending March 2002 to year ending March 2018



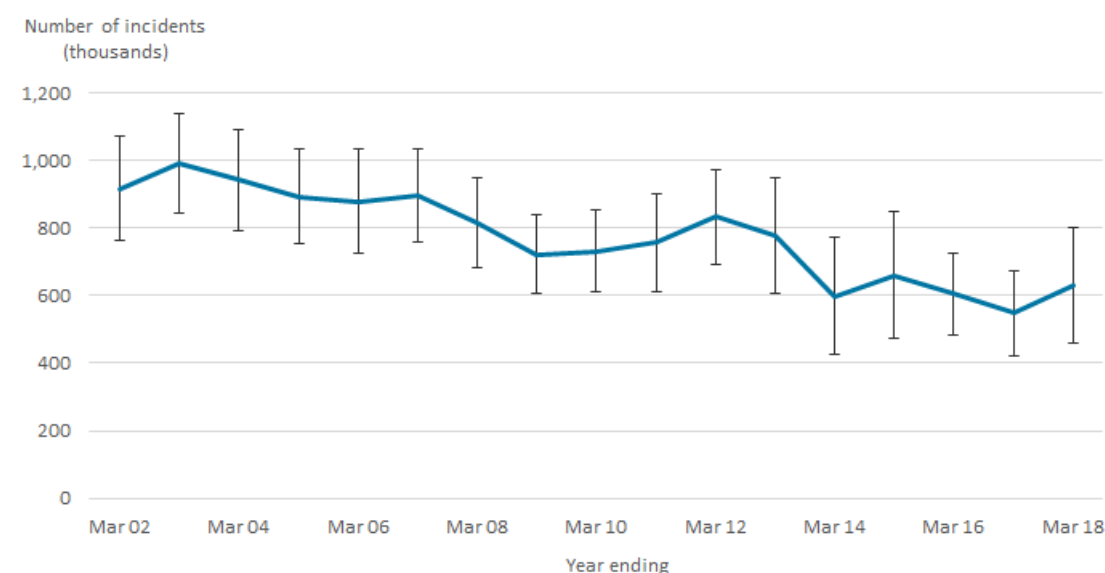
Source: Office for National Statistics - Crime Survey for England and Wales

**Notes:**

1. Data relate to adults aged 16 years and over.
2. The margins of error around the uncapped estimates are represented by the 95% confidence intervals.

**Figure 11: Trends in Crime Survey for England and Wales acquaintance violence with new methodology**

England and Wales, year ending March 2002 to year ending March 2018



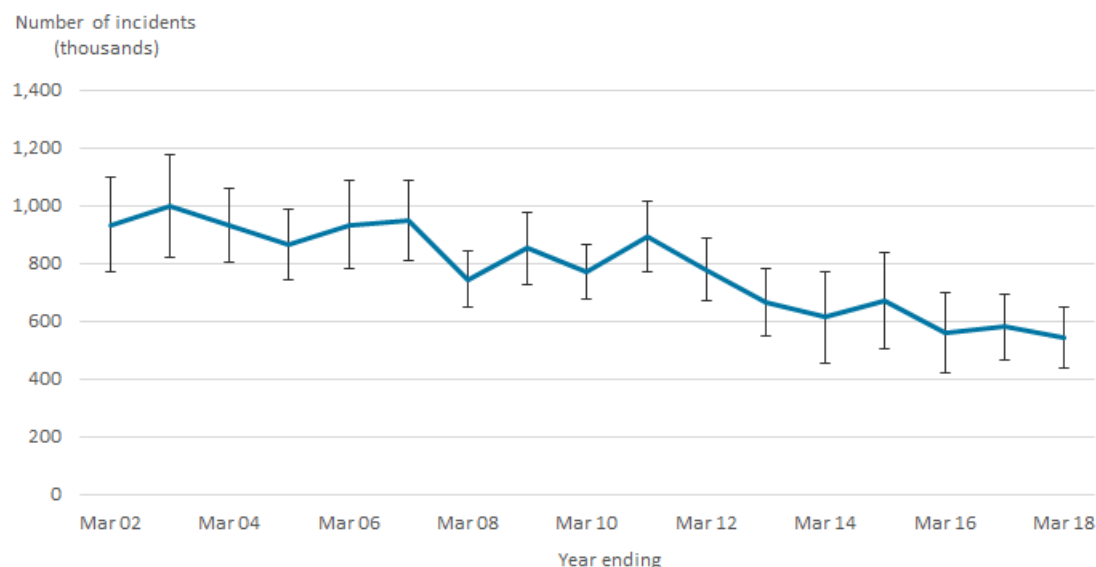
Source: Office for National Statistics - Crime Survey for England and Wales

**Notes:**

1. Data relate to adults aged 16 years and over.
2. The margins of error around the estimates are represented by the 95% confidence intervals.

**Figure 12: Trends in Crime Survey for England and Wales stranger violence with new methodology**

England and Wales, year ending March 2002 to year ending March 2018



Source: Office for National Statistics - Crime Survey for England and Wales

**Notes:**

1. Data relate to adults aged 16 years and over.
2. The margins of error are represented by the 95% confidence intervals.

Owing to very little change in estimates of non-violent crimes, confidence intervals for these crime types typically remain similar to those previously published. Confidence intervals for all CSEW crime (excluding fraud and computer misuse) have become a little wider than those previously published; for example, the year ending March 2018 confidence interval is 12% wider.

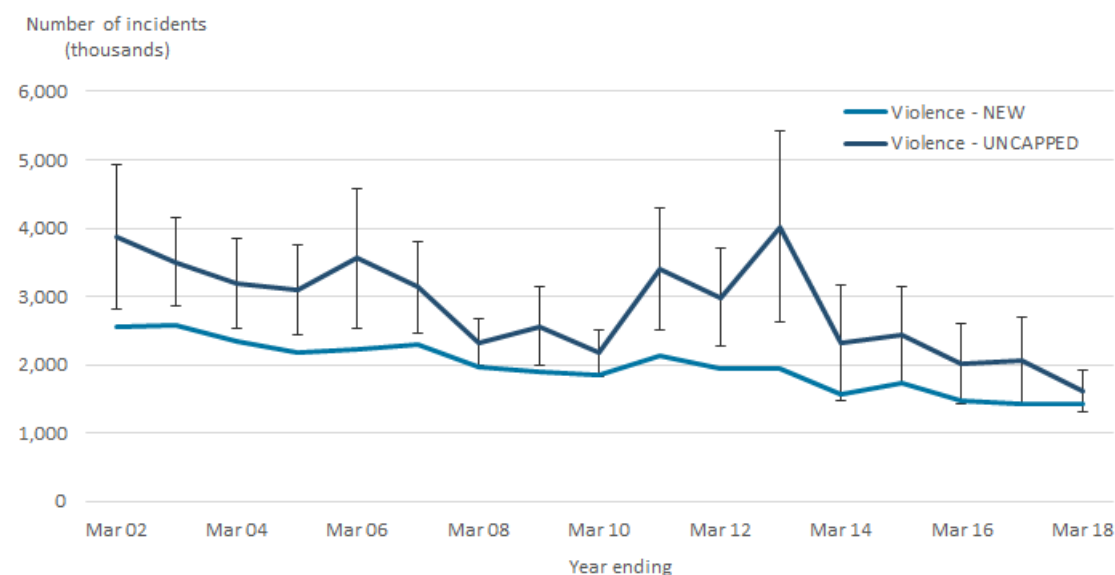
## Availability of uncapped variables for estimating incident numbers

The uncapped estimates ([Table U1](#)) have large confidence intervals, which will be reported alongside estimates with appropriate caveats to aid in interpretation of these data.

Figure 13 shows how volatile the trend in uncapped violence is, compared with the estimates we have calculated using the new methodology. The uncapped data display large increases or decreases in violence; however, despite these being large differences, they are rarely statistically significant (see the associated confidence intervals). If we had moved to uncapped estimates, it is quite possible we could be reporting increases (or decreases) between 30% and 40% in violence that were not statistically significant. The confidence intervals are also very large when estimates are broken down by types of violence ([Table U2](#)).

**Figure 13: Trends in Crime Survey for England and Wales violence by differing methodologies**

England and Wales, year ending March 2002 to year ending March 2018



Source: Office for National Statistics - Crime Survey for England and Wales

**Notes:**

1. Data relate to adults aged 16 years and over.
2. The margins of error are represented by the 95% confidence intervals.

**Notes for: Adults: Impact on Crime Survey for England and Wales data**

1. All CSEW crime excluding fraud and computer misuse has been used for this analysis, since data on fraud and computer misuse are only available in partial format from the year ending March 2016. Both fraud and computer misuse offences are unaffected by the 98th percentile methodology since high-order repeat victimisation does not appear to be common for these crime types. Consequently, only small changes as a result of our weight adjustment are expected for these offences.
2. [UNESCO, ESRC and Lancaster University research briefing: Violence and society, November 2014 \(PDF, 256KB\)](#).
3. Data have been presented for the year ending March 2017 as opposed to the year ending March 2018, as the confidence intervals around estimates of violence in the latter year are uncharacteristically small and not representative of a typical survey year.

## 5 . Children: What are the 98th percentile values for the number of incidents in a series?

There is much more variability in the 98th percentile values from the children aged 10 to 15 years survey data than the adult data. This is due to the considerably lower sample size for the children's survey (around 3,000 children aged 10 to 15 years in each year compared with currently around 35,000 adults aged 16 years and over in each year).

[Table 5](#) details the three-year rolling 98th percentiles for the number of incidents in a series by crime type. These have been calculated based on “broad” as opposed to “preferred” measures of crime<sup>1</sup>. The variability in the 98th percentile values for robbery in the three years ending March 2013 (a value of 20) and criminal damage to personal property in the year ending March 2017 (a value of 12) are particularly large. These values result from higher levels of repeat victimisation being recorded in these years. As in the adult survey, violence is the crime type most impacted by this change in methodology. Robbery, personal theft and criminal damage to personal property are only largely impacted in select years.

The changes in incident numbers have mostly affected violence, robbery and criminal damage offence categories. These categories have all seen consistent upward changes compared with those previously published. Personal theft offences have predominantly remained unchanged since the level of repeat victimisation for these headline offences is low. While for some years the 98th percentile may jump dramatically, this percentile is typically only applied to a very small number of cases in the dataset. Hence, the impact on the estimates is not as extreme as one might assume from looking at the 98th percentiles alone.

### **Notes for: Children: What are the 98th percentile values for the number of incidents in a series?**

1. The “Preferred measure” takes into account factors identified as important in determining the severity of an incidence (such as level of injury, value of item stolen or damaged, relationship with the perpetrator) while the “Broad measure” counts all incidents which would be legally defined as crimes and therefore may include low-level incidents between children.

## **6 . Children: Impact on Crime Survey for England and Wales data**

### **Trends in CSEW crime**

As a note of caution, owing to the smaller sample size for the children’s survey, trends over time are typically more difficult to discern than those in the adults’ survey. Moving to the new methodology in calculating estimates from the Crime Survey for England and Wales (CSEW) has not changed this.

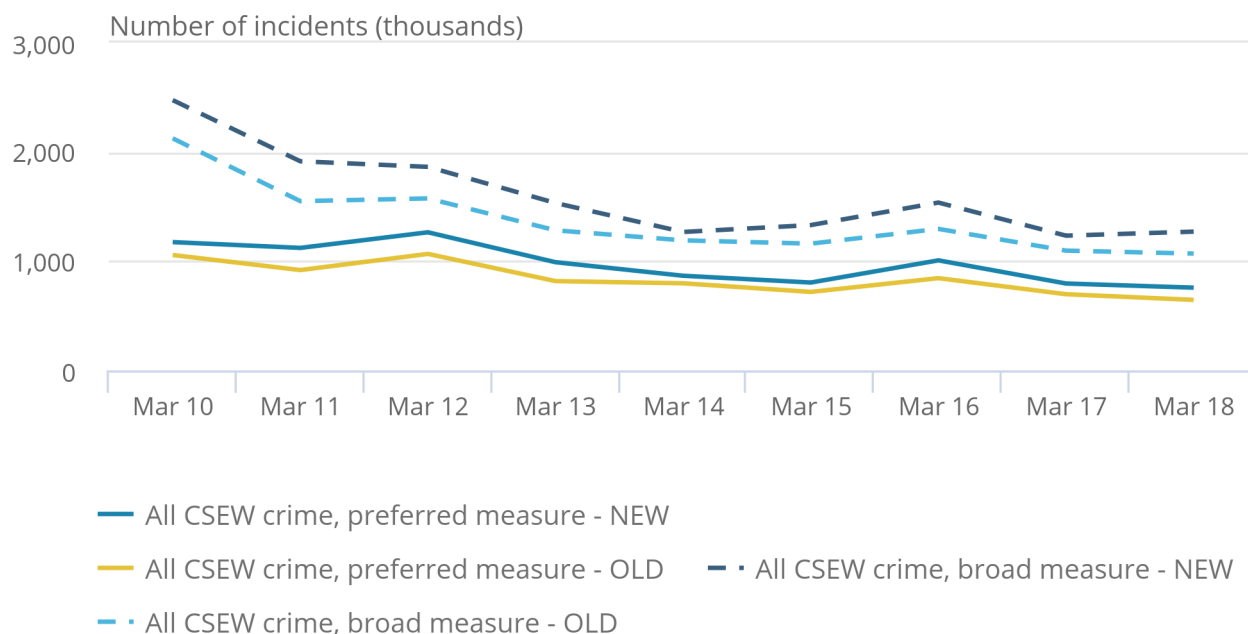
As seen in Figure 14 and [Table 6](#), trends in all CSEW crime for children aged 10 to 15 years (whether looking at the “preferred” or “broad” measure<sup>1</sup>) are relatively unaffected by the new methodology. However, the volumes we have reported are typically between 10% and 20% higher than our previously published figures. The increase in volumes between methodologies is larger than that for adults. This is owing to violence, which comprises a much larger proportion<sup>2</sup> of all CSEW crime for children aged 10 to 15 years (at around 50% to 60% in the preferred measure and around 60% to 75% in the broad measure). However, the trend in all CSEW crime remains similar because the trend in violence experienced by children aged 10 to 15 years remains largely unaffected.

**Figure 14: Trends in all Crime Survey for England and Wales crime experienced by children aged 10 to 15 years, by old and new methodology**

England and Wales, year ending March 2010 to year ending March 2018

## Figure 14: Trends in all Crime Survey for England and Wales crime experienced by children aged 10 to 15 years, by old and new methodology

England and Wales, year ending March 2010 to year ending March 2018



Source: Office for National Statistics - Crime Survey for England and Wales

### Notes:

1. Data relate to children aged 10 to 15 years.
2. Question changes during development of the children's questionnaire in the first two years should be considered when interpreting the figures; comparisons before the year ending March 2012 should be interpreted with caution.
3. The "preferred measure" takes into account factors identified as important in determining the severity of an incidence (such as level of injury, value of item stolen or damaged, relationship with the perpetrator) while the "broad measure" counts all incidents which would be legally defined as crimes and therefore may include low-level incidents between children.

## Trends in violent crime

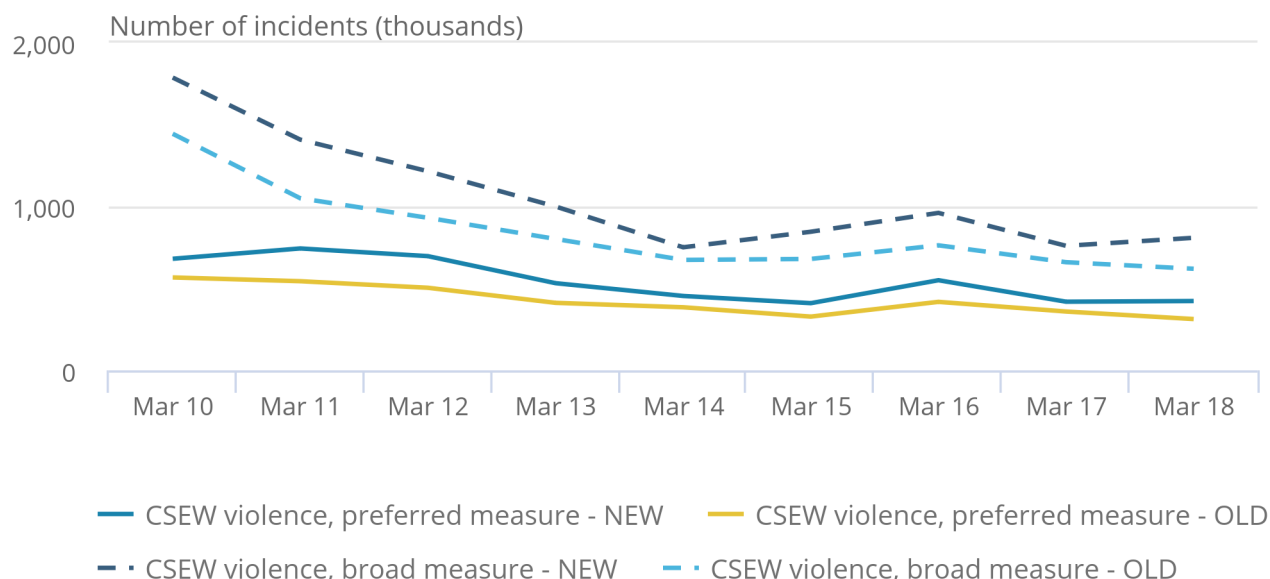
As with adults, repeat victimisation is more pronounced in violent crime than any other crime type for children aged 10 to 15 years. Trends for total violence are displayed in Figure 15 and [Table 7](#). The time series for both the preferred and broad measure are mainly unaffected, though the volume increase resulting from applying our new methodology ranges from around 10% to 40%.

**Figure 15: Trends in Crime Survey for England and Wales violence experienced by children aged 10 to 15, by old and new methodology**

England and Wales, year ending March 2010 to year ending March 2018

## Figure 15: Trends in Crime Survey for England and Wales violence experienced by children aged 10 to 15, by old and new methodology

England and Wales, year ending March 2010 to year ending March 2018



Source: Office for National Statistics - Crime Survey for England and Wales

### Notes:

1. Data relate to children aged 10 to 15 years.
2. Question changes during development of the children's questionnaire in the first two years should be considered when interpreting the figures; comparisons before the year ending March 2012 should be interpreted with caution.
3. The "preferred measure" takes into account factors identified as important in determining the severity of an incidence (such as level of injury, value of item stolen or damaged, relationship with the perpetrator) while the "broad measure" counts all incidents which would be legally defined as crimes and therefore may include low-level incidents between children.

[Appendix Table 9](#) within the [Crime in England and Wales, year ending September 2018](#) release presents the new time series for the different breakdowns of violent crime. The volumes increase for every year in comparison with previously published estimates regardless of the breakdown (apart from "Violence without injury" for the preferred measure in the survey year ending March 2018, which remained the same). "Violence with injury" estimates increase by between 12% and 46% and "Violence without injury" estimates increase (other than the survey year ending March 2018) by between 11% and 44%.

### Notes for: Children: Impact on Crime Survey for England and Wales data

1. The “Preferred measure” takes into account factors identified as important in determining the severity of an incidence (such as level of injury, value of item stolen or damaged, relationship with the perpetrator) while the “Broad measure” counts all incidents which would be legally defined as crimes and therefore may include low-level incidents between children.
2. This should not be interpreted as children experiencing more violence than adults; children aged 10 to 15 years are not asked about “household” offences within the survey, such as burglary or vehicle theft, so violent offences will naturally comprise a larger proportion of crimes committed against children than against adults.

## 7 . Implementation of new methodology for measuring repeat victimisation in published material

Our new methodology for measuring repeat victimisation within the Crime Survey for England and Wales (CSEW) has been implemented for the first time in the [Crime in England and Wales: year ending September 2018](#) release, published on 24 January 2019. All future crime releases will use this new methodology. Estimates based upon the previous methodology (incident numbers capped at 5) will no longer be produced and previous publications have not been updated.

The following suite of data related to the changes to our methodology for adults aged 16 years and over implemented back to the year ending December 1981 (unless specified otherwise) were published on 24 January 2019:

- [Appendix tables \(A\)](#) – numbers of incidents and incidence rates per 1,000 population; associated tables (for example, bulletin tables and quarterly data tables) have also been updated accordingly.
- [Annual trend and demographic tables \(D\)](#) – numbers of times victims were victimised (for the year ending March 2018), proportions of incidents experienced by repeat victims and percentages of incidents reported to the police; these were held over from the [Crime in England and Wales: year ending March 2018](#) release, published on 19 July 2018.
- [User guide tables \(UG\)](#) – confidence intervals around CSEW estimates; these were held over from the [Crime in England and Wales: year ending March 2018 release](#), published on 19 July 2018.
- [Uncapped CSEW tables \(U\)](#) – estimates, including confidence intervals, based on removing the caps on numbers of incidents entirely.

The following suite of data related to the changes to our methodology for children aged 10 to 15 years implemented back to the survey year ending March 2010 were also published on 24 January 2019:

- [Appendix tables \(A\)](#) – numbers of incidents and incidence rates per 1,000 population; associated tables (for example, bulletin tables) have also been updated accordingly.
- [User guide tables \(UG\)](#) – confidence intervals around CSEW estimates; these were held over from the [Crime in England and Wales: year ending March 2018](#) release, published on 19 July 2018.
- [Uncapped CSEW tables \(U\)](#) – estimates, including confidence intervals, based on removing the caps on numbers of incidents entirely.

In due course, we will supply the [UK Data Service](#) and our own [Secure Research Service](#) (SRS) with updated datasets containing the new crime category variables based on the 98th percentile caps. They will also contain crime category variables with the removal of the caps altogether for specialist users to access to conduct their own analyses. However, owing to the large number of years that datasets will need to be supplied for, this process will take some time to fully complete. We intend to supply all updated datasets at the same time, to avoid users accessing non-comparable data based on different incident-capping methodologies. We will keep users updated on our progress within our quarterly Crime in England and Wales releases.

## 8 . Appendix: Refinement to the weighting methodology

Our previously published [methodological note](#) in October 2017 covered all but the Background to trimming component weights sub-section in this appendix (as a final decision had not then been made on how we would be adjusting the survey weights). This information has been included again here for completeness.

All estimates from the Crime Survey for England and Wales (CSEW) presented in the figures and tables in our crime statistics publications are based on weighted data. That is, results obtained from surveying a sample of the population of England and Wales are scaled-up to represent the entire population.

Two types of weighting are used in the CSEW sample. First, the raw data are weighted to compensate for unequal probabilities of selection involved in the sample design. These include: the over-sampling of less populous police force areas; the selection of multi-household addresses; and the individual's chance of participation being inversely proportional to the number of adults living in the household. Second, calibration weighting is used to adjust for different levels of non-response.

The new methodology for improving the way we estimate repeat victimisation, introduces volatility into the estimates between years. To ensure a usable time series, we have made some minor changes to the weights used to compensate for unequal probabilities of selection. This reduces volatility in estimates between years. Calibration weighting will remain unchanged.

### Design weights

The main units of analysis used on the CSEW are households, individuals, and incidents of victimisation. Different weights are used depending upon the unit of analysis. Some crimes are considered household crimes (for example: burglary, criminal damage to household property, theft of and from a vehicle) and therefore the main unit of analysis is the household. Other crimes are considered personal crimes (for example: assault, robbery, theft from the person) and the main unit of analysis is the individual. These design weights are calculated using several component weights.

### Component weights

The weights are based on a number of components as follows:



- $w_1$ : "Police force area" weight, which compensates for unequal address selection probabilities between police force areas.
- $w_2$ : "Address non-response" weight, which compensates for the observed variation in response rates between different types of neighbourhood.
- $w_3$ : "Dwelling unit" weight, which is simply the number of dwelling units identified at the address – in the vast majority of cases, the dwelling unit weight is 1; historically, weight  $w_3$  has been capped at 10 to limit the variance of core household and individual weights.
- $w_4$ : "Household size" weight, which compensates for the fact that the probability of any one individual being selected is inversely proportional to the number of adults in the household; the individual weight is therefore simply the number of adults in the household.

The two design weights are constructed as follows:

- Core household weight equals  $w_1$  multiplied by  $w_2$  multiplied by  $w_3$
- Core individual weight equals  $w_1$  multiplied by  $w_2$  multiplied by  $w_3$  multiplied by  $w_4$

When we explored the effects of removing the cap of 5 from our measure of the number of incidents in a series, there were some instances in which high levels of repeat victimisation (97) coincided with very high weights. In one instance, final weights of more than 6,000 per individual coincided with a series that included 97 incidents of violence. The combined effect of this meant that by uncapping the estimates, one individual was contributing over 582,000 incidents to our annual violence estimates (as compared with the individuals' contribution of just over 30,000 incidents with the cap of 5 in place).

The component weight that contributed directly to this issue was the dwelling unit weight ( $w_3$ ). However, analysis of the data indicated the same issue may arise in the future as a result of the individual component weight ( $w_4$ ), which has similar variability.

A decision was made to trim the component dwelling unit weight at 4 for the calculation of household weights in the adults' datasets. This aligns with the weighting procedures used for the children's element of the CSEW, where this approach has been applied since the year ending March 2016. A slightly different method of trimming the dwelling unit element of the household weight at the 99th percentile had been implemented on the children aged 10 to 15 years datasets prior to the year ending March 2016. The effect on the published estimates in moving to the new trimming method was deemed to be negligible.

In calculating the core individual weight, the product of the multiplication of the dwelling unit weight and individual component weights has been trimmed at 5. Although trimming of extreme weights may introduce a small amount of bias, this is more than compensated for by the resulting improvement in precision.

## Background to trimming component weights

To assess the level at which to trim the product of the dwelling unit and individual weight, we created new calibrated weights for the years ending March 2012 and March 2013 using different design weight options. We compared the results of using these weights to each other as well as to our original approach. It was important to assess at least two years of data. We included the year ending March 2013 as extreme weights compounding with count data causing extreme sample variability when existing weights were used for incident analysis. And for the second year we included the year ending March 2012 as we knew this same issue was not present.

We computed the ratio of the difference between standard errors produced using three different weighting options. These were the product of component weights  $w_3$  and  $w_4$  trimmed at 4, the product of component weights  $w_3$  and  $w_4$  trimmed at 5 and the original weights. This is called the bias ratio. If the bias ratio is between negative 0.5 and 0.5 then the coverage probability of the confidence interval under the new method is not too different from the nominal confidence level. For example, for a 95% confidence level, when the value of the bias ratio is 0.5, the coverage probability is 0.92, which is close to 0.95. When the bias ratio is equal to 1, the coverage probability falls to 0.83, which is quite far from 0.95, the stated confidence level. Neither of the new proposed weighting options produced bias ratios that indicated the confidence intervals would be substantially different to that which we already publish.

Further analysis was completed comparing the root square measurement error (RMSE) to the size of the confidence intervals produced from CSEW data with each of these different weighting options applied. We were aware that a balance needed to be achieved between these two factors. As it transpired, the version of the weight (the product of  $w_3$  and  $w_4$ ) trimmed at 5 reacted differently for each dataset. For the year ending March 2013 (which was the problematic dataset when using the original weights), there was more of an effect than for the year ending March 2012 where we did not experience any issues. Given that neither proposed weight option produced concerning bias ratios, we used the version that included more known design weight data and the preferred option appeared to be to trim the product of the dwelling unit and individual weight ( $w_3$  and  $w_4$ ) at 5.

Following on from this, all design weights have been calibrated to the full population and take into account further elements on non-response (related to age, sex and region).

## 9 . Acknowledgements

We would like to thank Joel Williams and all at Kantar Public for their advice and support during the course of this project. In particular, we would like to thank those members of the team (Neil Twist, David Xu) who worked with us on revising and reprocessing the long back series of Crime Survey data. We would also like to thank members of the Crime Statistics Advisory Committee (CSAC) and Government Statistical Service Methodology Advisory Committee (GSS MAC) for their additional support and advice. Finally, we would like to thank everyone who gave their time and shared their knowledge when responding to the consultation.