



# Research

## NOTES

Issue #1

www.ussc.gov/research

## Commission Collection of Individual Offender Data

By: Christine Kitchens  
M.A., Senior Research Associate

This edition of Research Notes explains how the Commission collects and analyzes sentencing information, and describes the Commission’s many datafiles.





The United States Sentencing Commission is an independent agency within the judicial branch of the federal government.<sup>1</sup> The Commission’s primary mission is to promulgate and amend the federal sentencing guidelines. The Commission has other responsibilities, including: (1) establishing a data collection, analysis, and research program to serve as a clearinghouse and information center for the collection, preparation, and dissemination of information on federal sentencing practices; (2) publishing data concerning the sentencing process; (3) collecting and disseminating information concerning sentences actually imposed and the relationship of such sentences to the factors set forth in 18 U.S.C. § 3553(a); and (4) collecting and disseminating information regarding the effectiveness of sentences imposed.<sup>2</sup>

To meet these responsibilities, the Commission receives and collects data from sentencing documents sent directly from the federal courts. Within 30 days of the entry of judgment in a criminal case, the chief judge of each sentencing court is required to submit documents about the sentencing to the Commission that includes (1) the Judgment and Commitment Order (J&C); (2) the Statement of Reasons (SOR); (3) any plea agreement; (4) the indictment or other charging document; (5) the Presentence Report (PSR); and (6) any other information the Commission finds appropriate.<sup>3</sup>

### Office of Research and Data

#### Contents

Sentencing Documents	1
Datafile Descriptions	2
Data Collection Documents	3
Analytic Processing	5
Data Resources	6

 AskORD@ussc.gov
  202-502-4500  
 @TheUSSCGov
  www.ussc.gov

**Research Notes** give background information on the technical details of the Commission’s data collection and analysis process. They are designed to help researchers use the Commission’s datafiles by providing answers to common data analysis questions.

# Commission Mission

The Commission's responsibilities include, but are not limited to, establishing a data collection, analysis, and research program.

## The Commission:



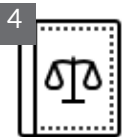
1 serves as a clearinghouse and information center for the collection, preparation, and dissemination of information on federal sentencing practices;



2 publishes data concerning the sentencing process;



3 collects and disseminates information concerning sentences imposed and their relationship to factors in 18 U.S.C. § 3553(a);



4 collects and disseminates information on the effectiveness of sentences imposed.



The Commission's largest datafile is the individual offender file. It dates back to 1991, though changes to both variables and values have occurred over time.

The Commission collects data using both optical character recognition (OCR) technology and Commission staff who interpret and collect data manually to transfer information from the submitted documents into several datafiles including the individual offender datafile, the organizational offender datafile, the appeals datafile, the re-sentencing datafile, the criminal history datafile, the enhanced drug penalty (851) datafile, and special project and report datafiles.

The individual offender datafile is the Commission's largest datafile. During fiscal year 2018, the federal courts submitted to the Commission reports of sentencing on 69,425 defendants that consisted of 321,624 individual documents.<sup>4</sup> The Commission collects information from the submitted documents pertaining to each individual offender case including: demographic information, statutes of conviction, complete guideline application information, sentencing information, criminal history points, and information about the sentence relative to the guideline range. This datafile has data back to 1991, however, changes to both the variables and the values have occurred over time, so not all analyses can be run for all time periods. The individual offender datafiles are available to researchers from the Commission's website.

The organizational datafile consists of information about organizations that have been convicted of federal violations. The Commission reviews the submitted documents and collects information including company demographic information (e.g., size, business classification), guideline application, and the details of the sentence such as fines and restitution. This datafile has data back to 1993, however, changes to both the variables and the values have occurred over time, so not all analyses can be run for all time periods. The organizational offender datafiles are also available to researchers from the Commission's website.

The appeals datafile consists of information about offenders whose sentence and/or conviction has been appealed. The appeals file contains information about the party filing the appeal, the guideline and issue appealed, and the outcome of the appeal. This datafile has data back to 1993, however, changes to both the variables and the values have occurred over time, so not all analyses can be run for all time periods.

The re-sentencing datafile consists of offenders who have been re-sentenced during each fiscal year. The type of re-sentencing (e.g., R. Crim. P. 35(b)) is collected along with the new sentence length and any changes to guideline information that may apply as part of the re-sentencing. This datafile goes back to 2008, however, changes to both the variables and the values have occurred over time, so not all analyses can be run for all time periods.

The Commission's criminal history datafile was established in 2016. It consists of detailed information about the prior criminal history convictions of federal offenders sentenced in a fiscal year. Each offender may have multiple prior events and each prior event may have multiple prior offenses associated with it. The datafile contains information about the types of offenses in each event, when the offender was sentenced for each event, and if the probation officer assessed criminal history points for the event.<sup>5</sup>

# Data Collection Documents

Within 30 days of judgment, the chief judge is required to submit documents about the sentencing to the Commission.

## 1. INDICTMENT (OR OTHER CHARGING DOCUMENT)



An indictment is a written statement of the essential facts constituting the offenses charged by a grand jury.

## 2. PLEA AGREEMENT (PLEA)



A written agreement between the government and the defendant in which the defendant agrees to plead guilty to one or more of the charged offenses, often in return for the government's agreement to dismiss other charges or to support a specific sentence.

## 3. PRESENTENCE INVESTIGATION REPORT (PSR)



A PSR is prepared by a probation officer and contains information about the offense and offender, the statutory range of punishment, the guideline calculation, and any bases for imposing a sentence above or below the guideline range.

## 2. JUDGMENT AND COMMITMENT ORDER (J&C)



The J&C is a written record of the defendant's conviction(s) and the sentence the court imposes.

## 5. STATEMENT OF REASONS (SOR)



The SOR is a form on which judges explain the reasons for imposing a sentence outside the guidelines range.

The enhanced drug penalty (851) datafile was established in 2016.<sup>6</sup> It consists of information about drug trafficking offenders who were eligible for enhanced drug penalties due to prior convictions for serious drug felonies or serious violent felonies.<sup>7</sup> The data includes information about whether the government sought the enhanced penalty and whether the enhanced penalty remained in effect at sentencing. The enhanced drug penalty datafile is available to researchers from the Commission's website.

The Commission also creates special project and report datafiles from time to time, which generally consist of data collected by the Commission staff for a special research topic—data that was not regularly collected in the fiscal year datafiles. The information collected in these datafiles may be used to assist the Commissioners in making policy decisions or to help inform Congress about a particular topic. Selected special project datafiles are available to researchers from the Commission's website.

The Commission collects information only on federal criminal cases involving at least one felony or Class A misdemeanor conviction (*i.e.*, cases involving only petty offenses are not recorded on the Commission's datafiles).<sup>8</sup> All the offenders in the Commission's individual offender datafile were convicted as adults in the federal system.<sup>9</sup> Death penalty case information is not submitted to the Commission because these cases are considered non-guidelines cases.<sup>10</sup> Records of any federal criminal case in which the defendant was acquitted of all charges, or where all charges were dismissed are not maintained by the Commission. Cases where a diversionary sentence was imposed also are not kept in the Commission's individual offender datafile. Finally, cases are not included if all the offense conduct occurred before the federal sentencing guidelines were promulgated.<sup>11</sup>

Court personnel send the required documents to the Commission electronically. Prior to 2005, the Commission received documents only in paper copy format. In 2005, the Commission established its Electronic Submission System ("ESS") and, by June of 2006, all documents were being stored electronically in the Commission's Oracle database. Court personnel log into the ESS, enter case identifying information (*e.g.*, offender name, docket identification number), identify the status of the documents (*i.e.*, note which documents are attached, waived, or omitted), indicate if the documents involve an amendment to a prior judgment or a re-sentencing, and then attach PDF versions of the documents. In addition, districts can submit cases to USSC through the Probation and Pretrial Services Automated Case Tracking System (PACTS). The ability to submit documents through the PACTS system eliminates the need to separately submit documents into two different systems. All 94 federal judicial districts now submit electronic documents to the Commission.

The Commission collects information only on federal criminal cases that involve at least one felony or Class A misdemeanor conviction.

Each case submission of electronic documents is reviewed during the analytical case processing to determine if it is missing critical documents. The Commission considers the PSR, J&C, and SOR to be the three critical documents for data collection. The submission is also reviewed to ensure that the documents provided match the document status information submitted by the district (e.g., if the document status states that a PSR was attached, the PSR must be one of the documents attached to the submission). If the Commission does not receive a critical document, or the document submission status does not match the actual documents attached, the case submission is rejected. If a case is rejected, an e-mail describing the problem is sent back to the court personnel that submitted the case.<sup>12</sup>



The Commission considers the PSR, J&C, and SOR to be the three critical documents for data collection.

If the case submission does include all the critical documents, then analytical case processing is completed. In general, cases are processed in order of sentencing date, with cases having the oldest sentencing date being processed first.<sup>13</sup>

During the analytic case processing phase, the OCR technology extracts information for specific data fields (e.g., gender), and then staff in the Commission’s Office of Research and Data review each of the documents and enter the remaining demographic, sentencing, and guideline application information for each offender. First, the PSR is reviewed and available demographic information is entered.<sup>14</sup> Additional identifiers are also taken from the PSR to help the Commission match its data to other federal criminal datafiles.<sup>15</sup> Next, staff reviews the J&C and enters the case disposition (guilty plea vs. trial), statutes for each count of conviction, and the statutory ranges for each count of conviction. The statutes of conviction are coded

from the J&C and the statutory ranges are coded from the PSR.<sup>16</sup> The disposition is entered based on information taken from both documents (e.g., the J&C may note that the offender was convicted of counts 1-3 at trial, but the PSR must be reviewed to determine if the type of trial was jury or bench). The J&C is also reviewed for sentence length information and financial aspects of the offender’s sentence.<sup>17</sup> Guideline application information is then extracted from the PSR, adjusting for any factual determinations (e.g., changes in the amount of drugs for which the offender is held accountable) that the court has documented on the SOR.<sup>18</sup> The SOR is then further reviewed to determine the final sentencing range and whether the sentence is inside or outside of the range.<sup>19</sup> Finally, for any case involving a sentence outside of the guideline range, information about the reasons why the sentence is outside of the range is collected from the SOR.

Generally, courts use the current standard form (AO245B) to impose judgment in a criminal case and provide reasons for that judgment. The PSR generally provides information to the court in a standardized format.<sup>20</sup> When possible, data entry screens at the Commission are generally designed to mimic the way the information is presented on the forms (i.e., the same number of checkboxes in the same order).

Whenever older forms or district-modified forms are provided to the Commission, data collection issues can arise. For instance, the current SOR form clearly distinguishes between a sentence outside of the guideline range due to a departure and a sentence outside the guideline range due to a variance.<sup>21</sup> When districts submit older or non-standard SORs, this distinction may not be as clearly indicated. Information such as who initiated the departure or sentence outside the guideline system may be missing. The Commission attempts to determine the fields that can be filled-in from these alternate forms. Matters of legal interpretation are resolved by the Commission’s Office of General Counsel.



Conversely, districts sometimes submit other documents in addition to, or in place of, standard forms. Some districts submit SORs that document a sentence is outside of the guideline range, and the SOR refers to a sentencing transcript for the reasons. The district may submit portions of or the entire sentencing transcript. In these instances, the Commission examines the transcript to attempt to determine the reasons why a sentence is outside of the range. Some SORs may also simply indicate that the reason the sentence is outside of the range is the “plea agreement.” In these instances, the plea agreement is reviewed to try to determine the reason the sentence is outside of the range so that the specific reasons can be coded.<sup>22</sup> More Commission resources are needed for cases that include one or more non-standard forms, or cases where the standard form is not fully or correctly filled-out.

Sometimes the documents received for a single case contain conflicting information. When this happens, the Commission attempts to determine which information is accurate and should be entered. For example, occasionally the offender’s name is not consistent on all the documents received for a case. Sometimes this is due to misspellings or confusion over the order of names. Sometimes the offender has given one name initially to law enforcement and then at some point his or her true name is discovered and is used on subsequent documents. The Commission tries to determine the correct spelling/order of the name and enter it that way into the datafile.

In other cases, the SOR states that the sentence is within the guideline range, but the sentence provided on the J&C is outside of the final guideline range provided on the SOR or statutory range provided in the PSR. In a case such as this, the Commission will review the documents to determine if a typographical error has been made on one of the documents (sometimes the final sentencing range on the SOR is mistyped for the given final offense level and criminal history category). If the conflict of information cannot be resolved, then the case is coded as having a sentence outside the guideline range with no reason provided. Multiple Commission staff members will review these cases to try to resolve these inconsistencies in the data.<sup>23</sup>

The Commission devotes significant resources to collecting data and trying to ensure that the data is as accurate and complete as possible.<sup>24</sup> In addition to the above reviews, the Commission reviews all cases sentenced outside of the applicable guideline range after they are analytically coded to verify critical information such as the sentence length, final guideline range, and the reasons given by the court for imposing a sentence outside of the range.

The Commission has also developed a computerized edit program that identifies out of range values, unusual values, and logical inconsistencies between variables. Each case involving one or more of these “edits” is reviewed to verify the unusual or inconsistent information. Each year, about 60 percent of

## Phases of Case Processing

- 1 Staff review the submitted documents to verify that the submission meets all document requirements and that defendant and court data is correct.
- 2 Staff review the PSR and extract offender and case identifiers as well as demographic information and statutory ranges for each count of conviction.
- 3 Staff review the J&C to extract information about the disposition, counts of conviction, type(s) of sentence imposed, sentence length, and monetary penalties.
- 4 Staff review the SOR to extract the final guideline range. Staff compare the total sentence with the final guideline range to determine whether the sentence is inside or outside the guideline range.
- 5 Staff review the PSR to extract information about the guideline application and the offender’s criminal history.



The Commission devotes significant resources to collecting data and trying to ensure that the data is as accurate and complete as possible.



the cases are reviewed in regard to one or more of the data fields.

The Commission produces an *Annual Report*, discussing the Commission's work during the past fiscal year, a summary of the cases sentenced that year, the Commission's budget, etc. The Commission also produces the *Sourcebook of Federal Sentencing Statistics*, which displays information from the Commission's fiscal year datafiles in various tables and charts. These resources and many others are available on the Commission's website at [www.ussc.gov](http://www.ussc.gov).<sup>25</sup>

The Commission releases its data on an annual basis to the public for research purposes. Persons interested in obtaining the Commission's publicly released datafiles may visit the Commission's website at <https://www.ussc.gov/research/datafiles/commission-datafiles>. The datafiles are made available for download in SAS or SPSS format along with a codebook describing the available fields and values. In addition, the Commission's data is available on a website maintained by the Federal Justice Statistics Resource Center (FJSRC),<sup>26</sup> which is sponsored by the Bureau of Justice Statistics and developed by the Urban Institute. This website makes the Commission's data and corresponding datafile documentation (i.e., codebooks describing the datafile composition) available for download. The FJSRC website also has data from other federal criminal justice agencies available for download, as well as a point-and-click feature that allows users to select variables from pull-down menus and create cross-tabs without downloading the data. The Commission also provides data and datafile documentation to the Inter-University Consortium for Political and Social Research at the University of Michigan (ICPSR).<sup>27</sup> Note that the case identifiers (e.g., offender's name, social security number) are removed from all publicly released data.

Further information about the Commission's public access policy can be found at: [https://www.ussc.gov/sites/default/files/pdf/about/policies/20180814\\_Public\\_Access\\_Documents\\_Data.pdf](https://www.ussc.gov/sites/default/files/pdf/about/policies/20180814_Public_Access_Documents_Data.pdf). For any data related questions, please contact the Commission at [askORD@ussc.gov](mailto:askORD@ussc.gov) or call 202-502-4500. To keep current with new Commission news and publications, please follow the Commission on the Twitter account [@theusscgov](https://twitter.com/theusscgov).

## Endnotes

1 The United States Sentencing Commission was established by the Sentencing Reform Act of 1984. Its principal purposes are (1) to establish sentencing policies and practices for the federal courts, including guidelines regarding the appropriate form and severity of punishment for offenders convicted of federal crimes; (2) to advise and assist Congress, the federal judiciary, and the executive branch in the development of effective and efficient crime policy; and (3) to collect, analyze, research, and distribute a broad array of information on federal crime and sentencing issues.

2 The Commission's research and data duties are described in 28 U.S.C. § 995(a)(12) through (16).

3 28 U.S.C. § 994(w)(1).

4 See Table 2 of the Commission's 2018 SOURCEBOOK OF FEDERAL SENTENCING STATISTICS.

5 See U.S. SENTENCING COMM'N, *Guidelines Manual*, §4A1.1(a)-(e) (Nov. 2018) [hereinafter USSG] to determine how criminal history points are assessed.

6 Section 21 U.S.C. § 851 provides that the government must give the defendant notice of its intent to seek these enhanced penalties.

7 Prior to December 21, 2018, drug trafficking offenders who had been convicted of a felony drug offense were eligible for these enhanced penalties. Congress amended this provision through the First Step Act of 2018, Pub. L. No. 115-391 (2018), to both limit these enhancements to only offenders convicted of "serious drug felonies" but also to expand the eligibility for this enhancement to drug trafficking offenders who had been convicted previously of a "serious violent felony."

8 If the case also involves a petty offense of conviction, information related to the petty count of conviction (e.g., the statute(s), special assessment, etc.) is not recorded on the Commission's datafile. Only the information about the other felony or Class A misdemeanor count(s) are entered in the datafile.

9 An offender may be younger than 18 at the time of sentencing, however, the offender was convicted and sentenced as an adult. The individual offender datafile does record the offender's age at sentencing. For more information about younger offenders, see the Commission publication *YOUTHFUL OFFENDERS IN THE FEDERAL SYSTEM* (2017) at <https://www.ussc.gov/research/research-reports/youthful-offenders-federal-system>.

10 The 2018 *Guidelines Manual* specifies that the sentencing guidelines do not apply to Class B or C misdemeanors. USSG §1B1.9. Additionally, the sentencing guidelines do not specifically contain any provisions that include the death penalty. If a statute of conviction includes a provision for the death penalty and the offender is ultimately sentenced to death, then this statutory punishment "trumps" the guideline sentence.

11 The guidelines first became effective on November 1, 1987.

12 Note that the Commission will accept case submission if the court personnel specifies that a document does not exist (e.g., the PSR is waived). In addition, once the case submission end-date for inclusion in the Commission's fiscal year datafile has passed (generally this is usually late January/early February after the end of the fiscal year in September), then all rejected cases are pushed back into the analytical case processing queue and the data is entered to the extent practicable. These partial-document cases are included in the Commission's annual datafiles, although more data fields may be missing than cases with all the documents present.

13 Occasionally, the Commission is interested in examining a subset of cases sentenced within a specific period of time. Specific date filters are applied so that cases sentenced within these dates will be given preference in the processing queues.

14 See ADMINISTRATIVE OFFICE OF THE UNITED STATES COURTS, *THE PRESENTENCE INVESTIGATION REPORT: FOR DEFENDANTS SENTENCED UNDER THE SENTENCING REFORM ACT OF 1984* (MARCH 2006) [hereinafter AO PSR REPORT], at [http://cdn.ca9.uscourts.gov/datastore/library/2013/02/26/Horvath\\_presentence.pdf](http://cdn.ca9.uscourts.gov/datastore/library/2013/02/26/Horvath_presentence.pdf). That report provides, "The aim of the

presentence investigation is to provide a timely, accurate, objective, and comprehensive report to the court. The report should have enough information to assist the court in making a fair sentencing decision and to assist corrections and community corrections officials in managing offenders under their supervision.” *Id.* at I-1. It also states, “The report is designed to provide the court with a complete and concise picture of the defendant.” *Id.* at I-2. The PSR includes basic demographic information that the Commission collects, including gender, race, ethnic origin, citizenship, number of dependents, and education level. The PSR also includes a text description of custody issues (e.g., dates of arrest, bail/bond status), a description of the offense conduct, the guideline computation, the offender’s criminal history, and personal information such as financial status and family information. The presentence report is a confidential document. *Id.* at VI-3. The Freedom of Information Act (FOIA) and the Privacy Act do not apply to the federal judiciary. *Id.* Therefore, the PSR is considered a sensitive document due to the nature of some of the information and is not publicly released/available.

15 Identifiers include social security number, Federal Bureau of Investigation number, U.S. Marshals Service number, Probation and Pretrial Services Automated Case Tracking Electronic Case Management System (PACTS/ECM) identification number, Bureau of Immigration and Customs Enforcement (ICE) number, and probation office. Twice each year, the Commission matches its data to data maintained by the Administrative Office of the U.S. Courts and notifies districts of any cases that have not been received by the Commission. The Commission undertakes this effort to assess the completeness of its datafile and to make efforts to improve it. The Commission will also request any documents missing from cases that have been received. Additionally, the Commission may match its data against these and other datafiles for special projects.

16 If the statutes of conviction are completely different on the J&C from what is in the PSR, then the plea agreement is reviewed to see if the statutes listed in the plea agreement match those on the J&C – if so, then the statutory ranges are coded from the plea agreement. If those statutes are not a match, or if the PSR and plea agreement are not submitted, then the statutory ranges are left blank.

17 The J&C details all aspects of the offender’s sentence. The Commission collects many pieces of information about the sentence, including the length of the sentence, whether all or part of the sentence is time served, if any credit is given under USSG §5G1.3 (Imposition of a Sentence on a Defendant Subject to an Undischarged Term of Imprisonment), months of alternative confinement, months of probation or supervised release, hours of community service, amount of special assessment, fine amount, cost of supervision, and restitution amount.

18 This includes the amendment year of the *Guidelines Manual* used to determine the base offense level (BOL), the specific offense characteristics (SOCs), the Chapter Three adjustments, criminal history points, multiple count levels, drug/chemical types and weights, and loss amounts.

19 The SOR is part of the same form as the J&C (AO245B), see <https://www.uscourts.gov/services-forms/forms>. The SOR details information about final court findings with respect to any changes made at sentencing to information documented in the PSR as well as information about whether the sentence falls inside or outside of the guideline range and why. The Commission collects much of this information, including the final offense level, criminal history category, un-trumped (i.e., statutory constraints not taken into account) guideline range, changes to guideline application from the PSR, court findings on the mandatory minimum status, whether the sentence is within/outside the guideline range, the reasons why the sentence is outside the range, and the attribution of the origin of the departure/sentence outside the guideline system.

20 See AO PSR REPORT, *supra* note 14. This publication outlines the sections that should be included in a PSR.

21 Although the 2005 Supreme Court decision in *United States v. Booker*, 543 U.S. 220 (2005) clarified the authority of courts to impose sentences outside of the guideline range, the criteria for departing from the guideline range remained unchanged. Therefore, that variance subcategory of sentences outside of the range are captured separately in the Commission’s datafile to enable more specific analysis of outside of the range sentences.

22 Some plea agreements stipulate to a sentence outside of the range because the offender provided substantial assistance to the government or qualified for a departure under an Early Disposition Program. Other plea agreements may cite specific departure reasons such as criminal history points overrepresenting the seriousness of the criminal history or diminished mental capacity. Some plea agreements do not provide specific reasons for the sentence, but the agreements do stipulate to a specific sentence or a guideline range. The Commission records “stipulations” as the reason why the sentence is outside of the range. In cases in which no specific reason can be gleaned from the plea, the “plea agreement” indication from the SOR is entered as the only reason for why the sentence is outside of the range.

23 The data analyst initially enters the information. After all cases have been completed through analytical processing, the Commission runs the cases through an edit program. The edit program selects cases with logical inconsistencies and out-of-range values for a secondary review by members of the research staff. Any unusual case issues that cannot be resolved are referred to the Office of General Counsel for final decision.

24 See note 12, *supra*, for a description of the Commission’s data matching efforts addressing the completeness of the data.


25 The Commission’s website ([www.uscc.gov](http://www.uscc.gov)) contains other statistical reports as well as reports to Congress and *Guidelines Manuals*. For a topical index of the Commission’s publications, see <https://www.uscc.gov/research/topical-index-publications>.

26 The website for the Federal Justice Statistics Resource Center can be found at <https://www.bjs.gov/fjsrc/>.

27 The website for the Inter-University Consortium for Political and Social Research can be found at <http://www.ICPSR.umich.edu/>.

## U.S. Sentencing Commission

 AskORD@ussc.gov

 202-502-4500

 @TheUSSCGov

 [www.uscc.gov](http://www.uscc.gov)

## About the Office of Research and Data

ORD provides statistical and other social science research and analyses on specific sentencing issues and federal crime.

The office receives documents from the federal courts concerning the sentences imposed on individual offenders, analyzes and enters information from those documents into the Commission’s comprehensive computer database, and creates annual datafiles of sentencing information.