

January 22, 2009

MEMORANDUM FOR THE RECORD

From: Suzanne M. Dorinski
Statistical Support and Consulting Staff
Office of Statistical Methods and Research for Economic Programs

Subject: Public Use Documentation of the Imputation Methodology for the 2007 Census of Juveniles in Residential Placement

This memorandum presents a streamlined version of the imputation methodology for the 2007 Census of Juveniles in Residential Placement (CJRP). **There is a longer version of this memorandum which is for internal use only because it contains individual juvenile offender records, which are confidential.** The individual juvenile offender records have been removed from this version.

This document uses intentional white space to improve readability. The document is available as a PDF file, because the original formatting might not be reproduced on the reader's computer.

The imputation methodology for the 2003 CJRP and succeeding years is completely different from the methodology used for the 1997, 1999, and 2001 censuses. Since the methodology has completely changed, the references section does not refer to any imputation methodology documentation for censuses before 2003.

Imputation programs written for the 2001 CJRP and earlier censuses used the section and question number as variable names. It's very easy to make a typing mistake while using that convention. That naming convention also makes it more difficult to read the program code and debug it. For the 2007 CJRP, we assigned variable names that are more descriptive. The naming convention is shown in Table 1 on page 2. The section and question number for each item are shown in parentheses.

Table 1. Naming conventions used in 2007 CJRP Imputation System

Concept	2007 CJRP	2006 JRFC	2006 CJRP
Persons assigned to beds	total_2007 (S1Q10b)	total_2006_jrfc (S1Q5b)	total_2006 (S1Q10b)
Persons assigned to beds age 21 or older	adults_2007 (S1Q11)	adults_2006_jrfc (S1Q6)	adults_2006 (S1Q11)
Persons under age 21 assigned to beds	kids_2007 (S1Q12b)	kids_2006_jrfc (S1Q7b)	kids_2006 (S1Q12b)
Persons under age 21 assigned to beds due to offenses	kid_offenders_2007 (S1Q13b)	kid_offenders_2006_jrfc (S1Q8b)	kid_offenders_2006 (S1Q13b)
Persons assigned to beds for reasons other than offenses	kid_nonoffenders_2007 (S1Q14b)	kid_nonoffenders_2006_jrfc (S1Q9b)	kid_nonoffenders_2006 (S1Q14b)
Juvenile offender ID	kid_id (S2Q1)		
Juvenile offender's sex	kid_sex (S2Q2)		
Juvenile offender's birth date	kid_birth_month (S2Q3) kid_birth_day (S2Q3) kid_birth_year (S2Q3)		
Juvenile offender's race	kid_race (S2Q4)		
Agency that placed the juvenile offender in facility	kid_placed_by (S2Q5)		
Juvenile offender's most serious offense	kid_offense (S2Q7)		
Juvenile offender's adjudication status	kid_adjudication_status (S2Q9)		
Juvenile offender's date of admission	kid_admitted_month (S2Q10) kid_admitted_day (S2Q10) kid_admitted_year (S2Q10)		

Table of Contents

I. Introduction 4

II. Summary of the Files 6

III. Critical Item Facilities 7

IV. Questionnaire Items Eligible for Imputation 9

V. Imputation Rates 10

VI. Collapsed Facility Type Codes 12

VII. Imputation Methodology for Section I Data 15

VIII. Imputation Methodology for Item Nonresponse in Section II Data 16

IX. Imputation Methodology for Section II Data for Critical Item Facilities 18

X. Imputing Missing Data for Juvenile Offenders in Tribal Facilities 20

XI. Special Handling for Critical Item Units in the District of Columbia (DC) 22

XII. Caution When Comparing State Data over Time 23

XIII. References 24

Attachment A: Percentage of Juvenile Offenders Held in Critical Item Facilities by State of Facility 25

Attachment B: Section II Item Imputation Rates by State of Facility 27

Attachment C: Data Quality Issues in CJRP Over Time 29

I. Introduction

The CJRP is a mail canvass census that was first conducted in 1997. The CJRP asks juvenile residential custody facilities in the U.S. to describe each youth assigned to a bed in the facility on the last Wednesday in October. The census is not sent to adult facilities, or facilities exclusively for drug or mental health treatment, or abused or neglected children. The census is normally conducted in odd-numbered years; the collection scheduled for 2005 was delayed several months until early 2006. The reference date for the 2007 CJRP was Wednesday, October 24, 2007.

CJRP replaced the Census of Public and Private Juvenile Detention, Correctional, and Shelter Facilities, also known as the Children in Custody census, which had been conducted since the early 1970s. Previous censuses collected data on the facilities and the juvenile offenders that were held in the facilities.

CJRP collects an individual record on each juvenile held in the residential facility, with information on the juvenile's sex, date of birth, race, who placed the offender there, most serious offense, court adjudication status, and date of admission to the facility. These data are requested for all offenders under 21 years of age in the facility.

The National Center for Juvenile Justice, the research division of the National Council of Juvenile and Family Court Judges, maintains the CJRP databook online. The databook contains a set of pre-defined tables detailing the characteristics of juvenile offenders in residential placement facilities. Tables are currently available for 1997, 1999, 2001, 2003, and 2006.

The project sponsor is the Office of Juvenile Justice and Delinquency Prevention (OJJDP). The Office of Statistical Methods and Research for Economic Programs (OSMREP) is responsible for imputation in the CJRP, while the Governments Division (GOVS) is responsible for data collection and editing.

For the 2007 CJRP, we mailed out questionnaires to 3,116 facilities. Of the 3,116 facilities, there were 136 that had closed, 14 that were added after mail out, 21 temporarily out-of-scope (i.e., they did not hold offenders on the reference date), and 62 that were permanently out-of-scope (i.e., they never hold juvenile offenders). Thus, there were 2,911 in-scope facilities on reference day. 2,902 of the 2,911 facilities responded to the 2007 CJRP, for a 99.7 percent unit response rate. 9 facilities refused to participate in the 2007 CJRP, but we imputed records for those facilities.

The 2007 questionnaire is divided into two sections: Section I, which collects general information about the facility, and Section II, which collects individual person data for juvenile offenders held at the facility.

This document includes the response rates and describes the imputation methodology that we used to make complete data sets for analytical purposes. We summarize the numbers of facilities

and records reporting on our 2007 edited and imputed files in Chapter II. In Chapter III, we discuss facilities, referred to as critical item facilities, that were only able to respond to a critical subset of the requested data. We did not impute for every item in the questionnaire, and Chapter IV covers the items that were eligible for imputation.

The discussion of imputation rates begins in Chapter V and continues in Chapters X through XII. As described above, the unit response rate is very high and leads to very low imputation rates for Total Persons, Total Adults, Total Juveniles, Total Juvenile Offenders, and Total Juvenile Nonoffenders, as seen in Chapter V.

Special caution should be taken when using the State data over time. There is a marked increase in high imputation rates for 2007. The exact imputation rates by State for 2007 can be found in Attachment B. Items that are highlighted have imputation rates that exceed 30 percent, and caution should be taken when using these data. Attachment C shows the items within each State that are more than 30 percent imputed for the period from 1997 through 2007.

A description of the imputation methodology is provided in Chapters VI through IX. References are included in Chapter XIII.

II. Summary of the Files

Table 2. Summary of records on the 2007 edited file

628	facilities that hold offenders and reported only critical items
59	facilities that hold nonoffenders only and reported only critical items
307	facilities that hold nonoffenders only and reported more than critical items
<u>69,995</u>	juvenile offender records from 1,908 facilities that reported more than critical items
<u>70,989</u>	records on the 2007 edited file

Table 3. Summary of facilities on the 2007 edited file

628	facilities that hold offenders and reported only critical items
59	facilities that hold nonoffenders only and reported only critical items
307	facilities that hold nonoffenders only and reported more than critical items
<u>1,908</u>	facilities that hold offenders and reported more than critical items
<u>2,902</u>	facilities in the 2007 edited file

Table 4. Summary of records on the 2007 imputed file

16,932	juvenile offender records from facilities that hold offenders and reported only critical items
59	facilities that hold nonoffenders only and reported only critical items
307	facilities that hold nonoffenders only and reported more than critical items
<u>69,995</u>	records for juvenile offenders from 1,908 facilities that reported more than critical items
<u>87,293</u>	records on the 2007 imputed file

Table 5. Summary of facilities on the 2007 imputed file

637	facilities that hold offenders and reported only critical items
59	facilities that hold nonoffenders only and reported only critical items
307	facilities that hold nonoffenders only and reported more than critical items
<u>1,908</u>	facilities that hold offenders and reported more than critical items
<u>2,911</u>	facilities in the 2007 imputed file

Table 6. 2007 CJRP head counts

103,900	people in residential placement
696	adults
103,204	juveniles
86,927	juvenile offenders
16,277	juvenile nonoffenders

III. Critical Item Facilities

In follow-up interviewing, GOVS attempted to collect as much data as possible to fill in both sections of the questionnaire. The following data items were deemed critical:

Section I:

- Question 5 (type of facility)
- Question 10a and 10b (total persons assigned to beds in the facility)
- Question 11 (number of persons age 21 or over assigned to beds in the facility)
- Question 12a and 12b (number of persons under age 21 assigned to beds in the facility)
- Question 13a and 13b (number of offenders under age 21 assigned to beds in the facility)
- Question 14a and 14b (number of nonoffenders under age 21 assigned to beds in the facility)

Section II:

- Question 2 (whether the facility is all-male, all-female, or holds both sexes)
- Question 5 (placement agency)
- Question 6 (placement agency's government level)
- Question 7 (offense code)
- Question 9 (adjudication status)

In previous CJRP data collections, the critical items field was set to either 0 or 1, with 1 indicating that the facility responded only to the critical items. In the 2007 CJRP, we have expanded the possible values for the critical items field to include 4 statuses. (See Table 7 on the next page.)

Table 7. Values for the critical items field

Critical item field value	Meaning
0	Facility is neither a critical item facility nor a refusal.
1	Facility responded only to the critical items.
2	Facility responded to Section I of the questionnaire, but date of birth, offense, and date of admission are missing for all records in Section II of the questionnaire.
3	Facility is a refusal; all data on the file for that facility have been imputed.

Facilities with the critical item field set to 1 have only one record per facility on the edited file, and the information in Section II for those facilities is used to generate the juvenile offender roster for each facility.

Facilities with the critical item field set to 2 are imputed in the same manner as facilities with the critical item field set to 1. When date of birth, offense, and date of admission are missing for every juvenile in the facility, we have to impute for every juvenile in the facility, and we want to avoid using the same donor over and over within the facility.

Refusal facilities are also imputed in the same manner as facilities with the critical item field set to 1, to minimize the number of records within a facility imputed using the same donor. The critical item field value of 3 is the way to readily identify refusals on the file.

IV. Questionnaire Items Eligible for Imputation

The following items were eligible for imputation in the 2007 CJRP.

Section I:

- Question 10b (total persons assigned to beds in the facility)
- Question 11 (number of persons age 21 or over assigned to beds in the facility)
- Question 12b (number of persons under age 21 assigned to beds in the facility)
- Question 13b (number of offenders under age 21 assigned to beds in the facility)
- Question 14b (number of nonoffenders under age 21 assigned to beds in the facility)

Section II:

- Question 2 (juvenile offender's sex)
- Question 3 (juvenile offender's birth date)
- Question 4 (juvenile offender's race)
- Question 5 (placement agency)
- Question 7 (juvenile offender's most serious offense)
- Question 9 (juvenile offender's adjudication status)
- Question 10 (juvenile offender's date of admission to the facility)

Earlier versions of CJRP have included questions about the numbers of locked doors in a facility. The 2007 CJRP included a set of questions about locked doors which did not ask for numbers of locked doors. The 2007 questions asked under what circumstances doors were locked. There was no imputation performed for these questions.

V. Imputation Rates

The facility imputation rates for Section I are shown in Table 8 below. The facility imputation rate is:

$$\frac{\text{Number of facilities with imputed data for Section I item}}{\text{Number of facilities in CJRP data file}} \times 100.$$

The only missing data in Section I was for the 9 facilities that refused to participate in the 2007 CJRP, so the imputation rates for each item in Section I are $(9/2911) \times 100$.

Table 8. Section I facility imputation rates in the 2007 CJRP

Item	Percent Imputed
Total persons	0.3
Adults	0.3
Juveniles	0.3
Juvenile offenders	0.3
Juvenile nonoffenders	0.3

The item imputation rates for Section II are shown in Table 9 on the next page. The juvenile offender item imputation rate is:

$$\frac{\text{Number of juvenile offender records with imputed data for item}}{\text{Number of juvenile offender records in CJRP data file}} \times 100.$$

Table 9. Section II item imputation rates in the 2007 CJRP

Item	Percent Imputed
Sex	7.5
Birth month	20.3
Birth day	20.3
Birth year	20.3
Race	20.6
Placed by	2.8
Offense	23.2
Adjudication status	8.5
Admitted month	20.3
Admitted day	20.3
Admitted year	20.3

There are 86,927 juvenile offender records on the imputed file. 21,889 (25.2 percent) juvenile offender records have at least one item imputed.

VI. Collapsed Facility Type Codes

For imputation purposes, we need to assign a collapsed facility type code (Cat) to every facility. Cat is the variable on the 2007 file that contains the collapsed facility type code (column 621).

The 2007 Cat code is assigned with the following procedure:

1. If the agency checks only one facility type box on the 2007 form, the checked box is mapped to the appropriate collapsed facility type and Cat is set. See Table 10 below. Note that if the agency only checks “other,” we don’t assign a collapsed facility type here.

Table 10. Cat codes for collapsed facility types

Cat	Collapsed Facility Type	Check box on 2007 form
0	Detention Center	S1Q501
1	Shelter	S1Q508, S1Q509
2	Reception / Diagnostic Center	S1Q503
3	Training School	S1Q502
5	Ranch, Camp, or Farm	S1Q506, S1Q507
6	Halfway House / Group Home	S1Q504, S1Q505

Note: In the table above, the check box naming convention is S1Q5____, where S1 means section 1, Q5 means question 5, and the last two digits match the checkbox on the questionnaire.

2. If the agency checks more than one facility type box on the 2007 form, we determine the collapsed facility type based on a hierarchy. The Cat code is determined by the highest checked box on the list. If an agency checks boxes that indicate it is both a reception center and a training school, it will be coded as a training school, since the training school is higher up in the hierarchy than the reception center is. The hierarchy is shown in Table 11 below.

Table 11. Hierarchy to be used when more than one collapsed facility type indicated on 2006 CJRP

Cat	Collapsed Facility Type	Check box on 2007 form
3	Training School	S1Q502
0	Detention Center	S1Q501
2	Reception / Diagnostic Center	S1Q503
5	Ranch, Camp, or Farm	S1Q506, S1Q507
1	Shelter	S1Q508, S1Q509
6	Halfway House / Group Home	S1Q504, S1Q505

Note: In the table above, the check box naming convention is S1Q5___, where S1 means section 1, Q5 means question 5, and the last two digits match the checkbox on the questionnaire.

3. If the Cat code is still missing, we look at the information provided on the 2006 Juvenile Residential Facility Census. We use a hierarchy here, because an agency could have checked more than one box on the 2006 JRFC. The Cat code is determined by the highest checked box on the list. If an agency checks boxes that indicate it is both a detention center and a shelter, it will be coded as a detention center, since the detention center is higher up in the hierarchy than the shelter is. The hierarchy is shown in Table 12 below.

Table 12. Hierarchy to be used when more than one collapsed facility type indicated on 2006 JRFC

Cat	Collapsed Facility Type	Check box on 2006 JRFC
3	Training School	S1Q13_02
0	Detention Center	S1Q13_01
2	Reception / Diagnostic Center	S1Q13_03
5	Ranch, Camp, or Farm	S1Q13_05, S1Q13_07
1	Shelter	S1Q13_08, S1Q13_09
6	Halfway House / Group Home	S1Q13_04, S1Q13_06

Note: In the table above, the check box naming convention is S1Q13___, where S1 means section 1, Q13 means question 13, and the last two digits match the checkbox on the questionnaire.

4. If the Cat code is still missing, we hold the Cat code constant from the 2006 CJRP.
5. For the refusals in the 2007 CJRP, we hold the facility type answers constant from the 2006 CJRP. This was a change from previous approaches implemented due to data quality concerns from one facility on the 2006 JRFC.
6. If a facility is still missing a Cat code, OJJDP will assign the code. [This situation occurred for 1 facility in the 2007 CJRP. The facility is new, and marked the “other” box and specified that it’s an independent living facility. It was classified as group home / halfway house.]

2,493 Cat codes were assigned because the facility checked only one facility type box on the 2007 CJRP form, 377 Cat codes were assigned because the facility checked multiple facility type boxes on the 2007 CJRP form, 15 Cat codes were assigned based on the facility responses on the 2006 JRFC form, 25 Cat codes were assigned based on the 2006 CJRP Cat codes, and 1 Cat code was assigned based on guidance from OJJDP.

VII. Imputation Methodology for Section I Data

The only missing data in Section I was for the refusal facilities.

Handling refusals

The 9 facilities that are refusals for the 2007 CJRP did respond to the 2006 JRFC, so we have prior year data for all 9 facilities. We calculated average 1-year growth rates by imputation cell and then applied those rates to the prior year data. When we apply the growth rate to prior year data, we round the result to the nearest whole number.

The 1-year growth rate is calculated for every facility that reported data in both the 2006 JRFC and the 2007 CJRP. The 1-year growth rate is the 2007 data item divided by the 2006 data item.

The imputation cell is all the facilities within a given state and Cat (collapsed facility type) code. If there are fewer than 15 respondents or less than 70 percent response in the imputation cell, we collapse the imputation cell to the national level. Collapsing was not required for the 2007 CJRP.

For all 9 refusals, we applied the average 1-year growth rate to the 2006 JRFC number of juvenile offenders to impute the juvenile offenders for the 2007 CJRP.

We applied the average 1-year growth rate to the 2006 JRFC number of juvenile nonoffenders to impute the juvenile nonoffenders for the 2007 CJRP.

We then added the imputed number of juvenile offenders to the imputed number of juvenile nonoffenders to impute the number of juveniles for the 2007 CJRP.

We applied the average 1-year growth rate to the 2006 JRFC number of adults to impute the adults for the 2007 CJRP.

We then added the imputed number of juveniles to the imputed number of adults to impute the total number of persons for the 2007 CJRP.

VIII. Imputation Methodology for Item Nonresponse in Section II Data

Change for 2007

The basic methodology for dealing with item nonresponse is still the hierarchical hot deck, which was used in the 2003 and 2006 CJRP collections. However, for the 2007 CJRP, we impute juvenile offenders in tribal facilities separately from the juvenile offenders in all other facilities. Juvenile offenders in tribal facilities are not eligible to be donors for juvenile offenders in non-tribal facilities. For more details, please see Chapter X, “Imputing Missing Data for Juvenile Offenders in Tribal Facilities.”

Randomly imputing day of birth, month of birth, or day of admission

The first missing items that are imputed are:

- kid_admitted_day if both kid_admitted_month and kid_admitted_year are reported,
- kid_birth_month if kid_birth_day is missing but kid_birth_year is reported, and
- kid_birth_day if both kid_birth_month and kid_birth_year are not missing.

When kid_admitted_month and kid_admitted_year are both reported, but kid_admitted_day is missing, we impute kid_admitted_day by randomly selecting a day based on kid_birth_month. This prevents the imputation of days that do not exist, such as February 30th.

Some of the records will only have reported kid_birth_year, while other records will only have reported kid_birth_month and kid_birth_year for the birth date. We randomly assign a birth month if kid_birth_month is not reported, then randomly assign a day of the birth month based on kid_birth_month. This prevents the imputation of days that do not exist, such as February 30th.

Age and stay calculations

The reference date of the questionnaire is October 24, 2007. Some facilities may report based on an alternative reference date. If an alternative reference date has been used, the date is shown in the alternative reference date field of the data file (columns 612 through 619).

We calculate an age for all records where it is possible to do so. If the facility is reporting based on an alternative reference date, the age of the juvenile offender is calculated as of the alternative reference date; otherwise the age of the juvenile offender is calculated as of October 24, 2007.

End users calculate a length of stay variable, which is based on the date that the juvenile offender was admitted to the facility. If a facility is reporting based on an alternative reference date, the length of stay is calculated as of the alternative reference date; otherwise the stay is calculated as of October 24, 2007. We calculate length of stay for all records where it is possible to do so.

Hierarchical hot-deck for item nonresponse

The imputation methodology for item nonresponse in Section II data is a hierarchical hot-deck. The record requiring imputation is matched to a pool of records where all the information is not missing, and then a donor record is selected. The missing values in the record requiring imputation are copied from the donor record. The match is first tried on all available information. If a match is not found, the match is made less restrictive until a donor record is found.

The definition of records where all the information is not missing includes those records for which we only imputed `kid_birth_month`, `kid_birth_day`, or `kid_admitted_day`. These records are considered eligible donors because if `kid_birth_year` is not imputed, we have a good idea how old the offender is, and if `kid_admitted_month` and `kid_admitted_year` are not imputed, we have a good idea how long the offender has been in placement.

The available information for matching is the Cat code, the state where the facility is located, and any reported data for `kid_sex`, `age`, `kid_race`, `kid_placed_by`, `kid_offense`, `kid_adjudication_status`, and length of stay.

When imputing `kid_adjudication_status`, those records with `kid_adjudication_status = 08` (convicted in adult criminal court) are **never** part of the pool of potential donors. We confirm with the sponsor that there should not be imputed values of 08 (convicted in adult criminal court) on the final data file.

The advantage of the hierarchical hot-deck method is that imputed values should be consistent with the rest of the juvenile offender record, because the donor record is a juvenile offender record that has passed the edits.

IX. Imputation Methodology for Section II Data for Critical Item Facilities

Changes for 2007

The basic methodology for dealing with nonresponse in critical item facilities is the same as it was in the 2003 and 2006 CJRP collections. However, for the 2007 CJRP, we impute juvenile offenders in tribal facilities separately from the juvenile offenders in all other facilities. Juvenile offenders in tribal facilities are not eligible to be donors for juvenile offenders in non-tribal facilities. For more details, please see Chapter X, “Imputing Missing Data for Juvenile Offenders in Tribal Facilities.”

We introduced a new classification of critical item facility in 2007. We noticed in the 2006 CJRP that some facilities will provide a roster of juvenile offenders, but not much information about the individual offenders. If date of birth, offense, and date of admission are missing for all the juvenile offenders in the facility, we really don't have much information to work with.

If we try to impute those records as merely having item nonresponse, we run the risk of using the same donor over and over within the facility, creating what looks like duplicate records in the facility. To minimize that risk, we now handle such facilities like critical item facilities, and have assigned them a code of 2 (Facility responded to Section I of the questionnaire, but date of birth, offense, and date of admission are missing for all records in Section II of the questionnaire) in the critical item field.

Background

The edited file has one record per critical item facility if the critical item field is set to 1 (Facility responded to only the critical items) or 3 (Facility is a refusal; all data on the file for that facility have been imputed).

If the critical item facility holds juvenile offenders, the Section II data on the record refers to all the juveniles held by that facility. GOVS tried to find out as much as possible about the types of juveniles held in critical item facilities.

If `kid_sex = 1` in Section II of the critical item facility record, that means the facility only holds males, while `kid_sex = 2` means that the facility only holds females, and `kid_sex = 3` means that the facility holds both males and females.

Some critical item facilities were unable to indicate for which types of offenses they held offenders (so `kid_offense = 88` or `99` for those critical item facilities), some critical item facilities were able to indicate that they held offenders for offense codes applicable to both underage

persons and adults (so kid_offense = 97 for those critical item facilities), and some critical item facilities were able to indicate that they held offenders for those offense codes applicable to underage persons only (so kid_offense = 98 for those critical item facilities).

We generate the required number of juvenile offender records for each critical item facility and assign kid_id to each juvenile offender record for the critical item facility. Kid_id is a 15 character juvenile identifier. We number the records sequentially within each critical item facility, starting at 000000000000001. We also replicate the available reported information for each juvenile offender record within the critical item facility.

If we know that the facility only holds males or only holds females, we do not consider kid_sex to be imputed.

The edited file may have multiple records if the critical item field is set to 2 (Facility responded to Section I of the questionnaire, but date of birth, offense, and date of admission are missing for all records in Section II of the questionnaire). For example, the facility may have two sets of offenders who have been placed in the facility by two different types of authorities. If date of birth, offense, and date of admission are missing for all the offenders in that facility, we don't have much information to work with. We handle these facilities as critical item facilities rather than item nonresponse facilities to minimize the amount of duplication in the imputed data.

Hierarchical hot-deck for critical item facilities

The hierarchical hot-deck methodology used for item nonresponse was modified for critical item facilities. Instead of finding a matching donor pool for an individual juvenile offender record, we find a donor pool for the critical item facility and then randomly select donors from the pool without replacement. This modified version of the hierarchical hot-deck requires that the donor pool have at least as many juvenile offenders as the critical item facility. This requirement ensures that the imputed juvenile offender records for the critical item facility are not duplicated within the facility.

The available information for matching is the Cat code, the state where the facility is located, and any reported data for kid_sex, age, kid_race, kid_placed_by, kid_offense, kid_adjudication_status, and stay.

When imputing kid_adjudication_status, those records with kid_adjudication_status = 08 (convicted in adult criminal court) are **never** part of the pool of potential donors. OJJDP does not want any imputed values of convicted in adult criminal court on the imputed file.

The advantage of the hierarchical hot-deck method is that imputed values should be consistent with the rest of the juvenile offender record, since the donor record is a juvenile offender record that has passed the edits.

In the 2003 and 2006 CJRP files, we only used the top 2 levels of the hierarchical hot-deck for critical item facilities. For the 2007 CJRP, we used as many as 4 levels of the hierarchical hot-deck for critical item facilities.

We had to use 4 levels in states with large numbers of juvenile offenders held in critical item facilities where the facility reported a relatively uncommon value for who placed the juvenile in the facility. Specifically, this happened for:

- a critical item facility in Arizona with juveniles placed by school official, parent or guardian, or the young person himself or herself,
- a critical item facility in the District of Columbia with juveniles placed by a social services agency, and
- a critical item facility in New York with juveniles placed by a social services agency who were awaiting an adjudication hearing in juvenile court.

X. Imputing Missing Data for Juvenile Offenders in Tribal Facilities

For the 2007 CJRP, we handled the juvenile offenders in tribal facilities separately from the juvenile offenders in all other facilities. While juvenile offenders in tribal facilities were not used as donors for juvenile offenders in all other facilities, we did use juvenile offenders in all other facilities as donors for juvenile offenders in tribal facilities.

If race was missing for a juvenile offender in a tribal facility, we set race to 4 (American Indian / Alaska Native, not of Hispanic origin) and set the race imputation flag to 1.

Item nonresponse for juvenile offenders in tribal facilities was handled in the same way as item nonresponse for juvenile offenders in all other facilities. See Chapter VIII, “Imputation Methodology for Item Nonresponse in Section II Data,” earlier in this document for more information.

Missing data for juvenile offenders in tribal critical item facilities was handled in the same way as missing data for juvenile offenders in all other critical item facilities. See Chapter IX, “Imputation Methodology for Section II Data for Critical Item Facilities,” earlier in this document for more information.

There are 113 tribal juvenile offenders in the 2007 CJRP. 77 of those offenders (68 percent) were held in critical item facilities. Almost all the missing data for tribal offenders is due to the

critical item facilities.

The tribal facility item imputation rates for Section II are shown in Table 13 below.

Table 13. Tribal facility section II item imputation rates in the 2007 CJRP

Item	Percent Imputed
Sex	68.1
Birth month	68.1
Birth day	68.1
Birth year	68.1
Race	68.1
Placed by	68.1
Offense	68.1
Adjudication status	69.0
Admitted month	68.1
Admitted day	68.1
Admitted year	68.1

Due to the high rates of missing data on the juvenile roster for tribal offenders, non-tribal offenders were used as donors for 53 (68 percent) of the 78 tribal juvenile offender records requiring imputation.

XI. Special Handling for Critical Item Units in the District of Columbia (DC)

During our first attempt at imputing the 2007 CJRP, we calculated the percentage of juveniles held in critical item facilities by the state of facility. The percentages are shown in Attachment A. We noticed that 95.7 of all juvenile offenders in DC are held in critical item facilities. This means that an unusually large percentage of the data in DC are imputed.

Further investigation revealed that there are two large facilities in DC that account for roughly 80 percent of the juvenile offenders held in DC, both of which are critical item facilities for the 2007 CJRP. GOVS called all the critical item facilities in DC to ask if they could provide the racial distribution of the offenders. 7 of the critical item facilities in DC were able to give us the racial distribution of their offenders.

Because there were so few fully-reported juvenile offender records in DC facilities, the imputation system defaulted to using juvenile offender records from across the country. We wrote additional SAS programs to reimpute the DC facilities after all other imputation processing was completed. The special handling is last so that no other imputed values on the file would change.

For the facilities that were able to provide the racial distribution, the `kid_race_flag` is set to 0 during the special handling. We then imputed the facilities as we would any other critical item facility where race had been reported.

After we were finished with the special handling for the 7 facilities in DC, we calculated Section II item imputation rates by state. Those rates are shown in Attachment B.

XII. Caution When Comparing State Data over Time

In the 2003 CJRP documentation, we noted that 84.6 percent of all juvenile offenders in DC were held in critical item facilities, which meant that an unusually large percentage of the data in DC was imputed in 2003. For 2007, 95.7 percent of all juvenile offenders in DC are held in critical item facilities. Comparing juvenile offenders held in DC facilities across the 2003, 2006, and 2007 data collections is not recommended, due to the extreme levels of missing data for the juvenile offenders.

Attachments A and B show some high levels of juvenile roster item imputation for 2007. Illinois, Mississippi, and Wyoming had more than half of their juvenile offenders in facilities that only reported critical items. States with 30 percent or more imputation by item included Arizona, Arkansas, Colorado, Florida (kid_race), Illinois, Iowa (kid_offense), Maine, Montana, Nevada, New Mexico, New York, South Carolina, Utah, and West Virginia.

In the 2006 CJRP, both Illinois and Rhode Island had more than half of their juvenile offenders in facilities that only reported critical items. States with 30 percent or more imputation by item included Colorado (kid_offense), Illinois (kid_birth_date, kid_race, kid_offense, kid_admitted_month and kid_admitted_day), Rhode Island (kid_sex, kid_birth_date, kid_race, kid_offense, kid_adjudication_status, and kid_admitted_date), and Wyoming (kid_sex, kid_birth_date, kid_race, kid_placed_by, kid_offense, and kid_admitted_date).

Users should be aware the differences in DC data from 2003 to 2006 may be due in part to the high levels of imputation for DC in 2003, and from 2006 to 2007 due to high levels of imputation for DC in 2007. Similarly, the differences in Colorado, Illinois, Rhode Island, and Wyoming data from 2003 to 2006 may be due in part to the high levels of imputation for those states in 2006. The differences in Arizona, Arkansas, Colorado, Florida, Illinois, Iowa, Maine, Montana, Nevada, New Mexico, New York, South Carolina, Utah, and West Virginia from 2006 to 2007 may be due in part to the high levels of imputation for those states in 2007.

Attachment C shows items by state for the 1997 through 2007 CJRP data collections. If the item imputation rate was 30 percent or more for a given year, the year is shown in the cell of the table.

XIII. References

Census of Juveniles in Residential Placement Databook, currently online at <http://ojjdp.ncjrs.org/ojstatbb/cjrp/>

Dorinski, Suzanne M. "How to collapse 2003 CJRP facility types for use in imputation," memo dated October 21, 2003. (Most of the memo has been reproduced in this document.)

Dorinski, Suzanne M. "How to impute Section I data," memo dated January 8, 2004.

Dorinski, Suzanne M. "How to calculate age on individual juvenile records in 2003 CJRP," memo dated February 25, 2004. (Most of the memo has been reproduced in this document.)

Dorinski, Suzanne M. "How to impute for item nonresponse on juvenile offender records in 2003 CJRP," memo dated April 1, 2004.

Dorinski, Suzanne M. "Documentation of the Imputation Methodology for the 2003 Census of Juveniles in Residential Placement FOR INTERNAL USE ONLY," memo dated December 1, 2004.

Dorinski, Suzanne M. "Documentation of the Imputation Methodology for the 2006 Census of Juveniles in Residential Placement FOR INTERNAL USE ONLY," memo dated March 22, 2007.

Dorinski, Suzanne M. "Documentation of the Imputation Methodology for the 2007 Census of Juveniles in Residential Placement FOR INTERNAL USE ONLY," memo dated January 15, 2009.

**Attachment A: Percentage of Juvenile Offenders Held in Critical Item
Facilities by State of Facility**
(Percentages of 30 or more have been highlighted)

Facility State	Juvenile Offenders	Percentage in critical item facilities
Alabama	1,667	11.5
Alaska	321	19.6
Arizona	1,580	39.1
Arkansas	819	31.7
California	14,033	7.5
Colorado	1,810	32.7
Connecticut	359	7.2
Delaware	301	0.0
D.C.	277	95.7
Florida	5,747	26.3
Georgia	2,733	9.0
Hawaii	130	2.3
Idaho	525	8.6
Illinois	2,552	73.1
Indiana	2,802	8.5
Iowa	1,257	22.1
Kansas	1,120	15.8
Kentucky	1,117	0.4
Louisiana	1,348	17.3
Maine	203	36.9
Maryland	862	4.4
Massachusetts	1,069	22.1
Michigan	2,616	8.4
Minnesota	1,439	21.5
Mississippi	449	52.1
Missouri	1,243	14.2
Montana	208	34.1
Nebraska	734	4.8
Nevada	1,057	43.0
New Hampshire	153	14.4
New Jersey	1,664	26.0
New Mexico	443	35.7
New York	3,471	46.6

**Attachment A: Percentage of Juvenile Offenders Held in Critical Item
Facilities by State of Facility**

(Percentages of 30 or more have been highlighted)

Facility State	Juvenile Offenders	Percentage in critical item facilities
North Carolina	1,039	9.8
North Dakota	203	16.7
Ohio	4,316	7.5
Oklahoma	897	17.4
Oregon	1,291	16.2
Pennsylvania	5,511	21.4
Rhode Island	317	0.0
South Carolina	1,201	46.8
South Dakota	509	29.7
Tennessee	1,320	17.0
Texas	7,034	10.8
Utah	1,050	38.7
Vermont	44	2.3
Virginia	2,163	6.0
Washington	1,526	22.9
West Virginia	575	31.3
Wisconsin	1,493	15.7
Wyoming	329	55.0
TOTAL	86,927	19.5

Attachment B: Section II Item Imputation Rates by State of Facility
(Imputation rates of 30 percent or more have been highlighted)

Facility State	Offenders	Kid_sex	Kid_birth_			Kid_race	Kid_placed_by	Kid_offense	Kid_adjudication_status	Kid_admitted_		
			month	day	year					month	day	year
Alabama	1,667	2.0	11.5	11.5	11.5	11.8	0.8	11.8	3.6	11.5	11.5	11.5
Alaska	321	18.4	19.6	19.6	19.6	19.6	4.0	24.6	21.2	19.6	19.6	19.6
Arizona	1,580	32.5	39.1	39.1	39.1	39.2	0.8	39.1	33.1	39.1	39.1	39.1
Arkansas	819	28.8	33.2	33.2	33.2	33.1	4.2	32.7	1.7	32.4	32.4	32.4
California	14,033	2.6	7.7	7.7	7.7	8.3	0.6	14.0	6.3	8.1	8.1	8.1
Colorado	1,810	3.0	32.7	32.7	32.7	32.9	4.4	47.5	3.1	32.7	32.7	32.7
Connecticut	359	0.3	7.2	7.2	7.2	7.2	3.9	35.1	0.3	7.2	7.2	7.2
Delaware	301	0.0	0.0	0.0	0.0	0.0	0.0	0.0	7.6	0.0	0.0	0.0
D.C.	277	40.1	95.7	95.7	95.7	2.9	0.0	95.7	43.7	95.7	95.7	95.7
Florida	5,747	0.1	27.3	27.3	27.3	30.3	0.0	29.4	2.5	26.3	26.3	26.3
Georgia	2,733	6.3	9.0	9.0	9.0	9.1	3.7	12.0	6.1	9.0	9.0	9.0
Hawaii	130	0.0	2.3	2.3	2.3	2.3	0.0	3.1	2.3	2.3	2.3	2.3
Idaho	525	2.7	8.6	8.6	8.6	9.1	0.0	9.1	6.1	8.6	8.6	8.6
Illinois	2,552	15.9	73.1	73.1	73.1	73.1	1.7	73.1	17.3	73.1	73.1	73.1
Indiana	2,802	7.2	8.6	8.6	8.6	12.2	1.8	8.6	8.5	8.5	8.5	8.5
Iowa	1,257	3.1	22.4	22.4	22.4	22.1	1.4	30.6	1.8	22.2	22.2	22.2
Kansas	1,120	7.4	15.8	15.8	15.8	16.9	0.0	21.5	9.9	15.8	15.8	15.8
Kentucky	1,117	0.4	0.4	0.4	0.4	0.4	0.4	3.9	1.5	0.4	0.4	0.4
Louisiana	1,348	6.8	25.5	25.5	25.5	25.5	0.0	25.6	11.6	25.5	25.5	25.5
Maine	203	33.5	36.9	36.9	36.9	36.9	33.5	36.9	33.5	36.9	36.9	36.9
Maryland	862	0.0	4.8	4.8	4.5	4.4	0.0	12.2	1.2	4.4	4.4	4.4
Massachusetts	1,069	4.1	22.2	22.2	22.2	22.1	4.4	22.5	12.3	22.8	22.8	22.8
Michigan	2,616	7.0	8.4	8.4	8.4	9.2	5.2	9.8	8.4	8.4	8.4	8.4
Minnesota	1,439	14.4	21.5	21.5	21.5	21.8	12.2	23.7	16.1	22.8	22.8	22.8
Mississippi	449	9.6	52.1	52.1	52.1	52.1	0.0	52.3	14.7	52.1	52.1	52.1
Missouri	1,243	13.1	14.4	14.4	14.4	14.2	0.1	15.4	11.7	14.3	14.3	14.3
Montana	208	20.2	34.1	34.1	34.1	34.1	9.1	34.1	1.9	34.1	34.1	34.1
Nebraska	734	1.5	4.8	4.8	4.8	4.8	4.5	5.3	0.1	4.8	4.8	4.8
Nevada	1,057	7.6	43.0	43.0	43.0	43.0	0.0	43.0	7.7	43.0	43.0	43.0
New Hampshire	153	0.0	22.9	22.9	22.9	14.4	0.0	14.4	0.0	14.4	14.4	14.4
New Jersey	1,664	17.4	26.0	26.0	26.0	26.0	15.9	26.0	19.2	26.0	26.0	26.0
New Mexico	443	35.2	35.7	35.7	35.7	35.9	0.0	42.0	35.4	35.7	35.7	35.7
New York	3,471	4.9	46.7	46.7	46.7	47.9	1.7	47.9	9.8	46.8	46.8	46.8
North Carolina	1,039	5.8	9.8	9.8	9.8	9.8	4.0	22.7	3.3	9.8	9.8	9.8
North Dakota	203	12.3	16.7	16.7	16.7	17.2	16.7	19.2	0.0	16.7	16.7	16.7

Attachment B: Section II Item Imputation Rates by State of Facility
(Imputation rates of 30 percent or more have been highlighted)

Facility State	Offenders	Kid_sex	Kid_birth_			Kid_race	Kid_placed_by	Kid_offense	Kid_adjudication_status	Kid_admitted_		
			month	day	year					month	day	year
Ohio	4,316	5.1	7.5	7.5	7.5	7.8	0.4	7.9	5.6	7.5	7.5	7.5
Oklahoma	897	7.0	17.4	17.4	17.4	17.4	2.7	20.3	13.9	17.4	17.4	17.4
Oregon	1,291	9.3	16.4	16.4	16.4	16.2	0.0	17.0	11.2	16.2	16.2	16.2
Pennsylvania	5,511	5.6	21.9	21.9	21.9	22.0	12.1	24.7	4.4	22.2	22.2	22.2
Rhode Island	317	0.0	0.6	0.6	0.6	0.0	0.0	0.0	0.3	0.0	0.0	0.0
South Carolina	1,201	4.2	78.3	78.3	78.3	78.0	0.7	78.5	12.5	78.3	78.3	78.3
South Dakota	509	20.8	29.7	29.7	29.7	29.9	4.9	29.9	24.2	29.7	29.7	29.7
Tennessee	1,320	3.9	17.0	17.0	17.0	17.0	1.7	23.5	11.1	17.0	17.0	17.0
Texas	7,034	8.8	10.8	10.8	10.8	11.1	0.2	11.7	8.9	10.8	10.8	10.8
Utah	1,050	29.7	38.7	38.7	38.7	40.6	14.6	38.9	15.7	38.7	38.7	38.7
Vermont	44	0.0	2.3	2.3	2.3	2.3	0.0	2.3	0.0	9.1	9.1	9.1
Virginia	2,163	3.2	8.2	8.2	8.2	6.0	2.4	9.1	2.7	6.0	6.0	6.0
Washington	1,526	20.6	22.9	22.9	22.9	26.1	2.3	23.6	12.2	22.9	22.9	22.9
West Virginia	575	17.7	31.3	31.3	31.3	31.3	11.0	31.8	12.0	31.3	31.3	31.3
Wisconsin	1,493	7.8	18.6	18.6	18.6	18.6	1.7	20.6	12.7	18.6	18.6	18.6
Wyoming	329	36.2	55.0	55.0	55.0	55.0	0.9	55.0	10.0	55.0	55.0	55.0
TOTAL	86,927	7.5	20.3	20.3	20.3	20.6	2.8	23.2	8.5	20.3	20.3	20.3

Attachment C: Data Quality Issues in CJRP Over Time

1997 through 2007 data collections

(Year is shown in table cell when imputation rate is 30 percent or more)

State	Sex	Birth			Race	Placed by	Offense	Adjudication status	Admitted		
		month	day	year					month	day	year
Alabama									2001	2001	2001
Alaska											
Arizona	2007	2007	2007	2007	2007		1999 2007	2007	2007	2007	2007
Arkansas		2007	2007	2007	2007		2007		2007	2007	2007
California											
Colorado							1999 2001 2006 2007				
Connecticut		2001	2001	2001	2001		2001 2007		2001	2001	2001
Delaware							1999				
D.C.	2003 2007	2003 2007	2003 2007	2003 2007			1999 2003 2007	2003 2007	2003 2007	2003 2007	2003 2007
Florida					2007						
Georgia							1999		1999	1999	1999
Hawaii											
Idaho											
Illinois		2006 2007	2006 2007	2006 2007	2006 2007		2006 2007		2006 2007	2006 2007	2006 2007
Indiana											
Iowa							2007				
Kansas											
Kentucky											
Louisiana											
Maine	2007	2007	2007	2007	2007	2007	2007	2007	2007	2007	2007
Maryland					2001			2003			
Massachusetts		2001	2001	2001	2001		2001		2001	2001	2001
Michigan											
Minnesota											
Mississippi	1997 1999	1997 1999 2007	1997 1999 2007	1997 1999 2007	1997 1999 2007		1997 1999 2007		1997 1999 2007	1997 1999 2007	1997 1999 2007
Missouri											
Montana		2007	2007	2007	2007		2007		2007	2007	2007
Nebraska											
Nevada					1997 2007		1997 2007	1997		2007	2007
New Hampshire											
New Jersey							1997 1999				
New Mexico	2007	1999 2007	1999 2007	1999 2007	1999 2007		1999 2007	2007	1999 2007	1999 2007	1999 2007
New York							1999 2007			2007	2007

Attachment C: Data Quality Issues in CJRP Over Time

1997 through 2007 data collections

(Year is shown in table cell when imputation rate is 30 percent or more)

State	Sex	Birth			Race	Placed by	Offense	Adjudication status	Admitted		
		month	day	year					month	day	year
North Carolina											
North Dakota							1999				
Ohio											
Oklahoma											
Oregon											
Pennsylvania							1999				
Rhode Island	2006	2006	2006	2006	2006		2006	2006	2006	2006	2006
South Carolina		2007	2007	2007	2007		2007		2007	2007	2007
South Dakota		1997	1997	1997	1997		1997	1997	1997	1997	1997
	1999	1999	1999	1999	1999		1999	1999	1999	1999	1999
Tennessee							1999 2001				
Texas											
Utah		2007	2007	2007	2007		1999 2007		2007	2007	2007
Vermont											
Virginia											
Washington											
West Virginia		2007	2007	2007	2007		2007		2007	2007	2007
Wisconsin							1999				
Wyoming	2006 2007	2006 2007	2006 2007	2006 2007	2006 2007	2006	2006 2007		2006 2007	2006 2007	2006 2007