UNITED STATES DEPARTMENT OF COMMERCE
Economics and Statistics Administration
U.S. Census Bureau
Washington, DC 20233-0001

MASTER FILE

December 31, 2002

DSSD A.C.E. REVISION II MEMORANDUM SERIES #PP-23

MEMORANDUM FOR: Donna L. Kostanich

From: Michael Beaghen *(signed 12/31/2002)* M. B.
Chair, A.C.E. Revision II Missing Data Subgroup

Prepared by: Michael Beaghen, Robert Sands
Decennial Statistical Studies Division

Subject: A.C.E. Revision II: Specifications for the Assignment of Probability of Enumeration Status, Census Day Residency and Match Status

This memorandum specifies the assignment of probabilities of enumeration status to unresolved E-sample people and of Census Day residency and/or match status to unresolved P-sample people in the Accuracy and Coverage Evaluation (A.C.E.) Revision II. It also specifies the assignment of probabilities of zero or one to people with resolved enumeration, Census Day residency, match and inmover status, and the calculation of sampling weights used in the A.C.E. Revision II estimation.

1 Introduction

The A.C.E. Revision II estimation (Kostanich 2002) utilizes both the original Accuracy and Coverage Evaluation (A.C.E.) coding results on the full A.C.E. sample and the Revision coding results on the smaller Revision sample (please note that the A.C.E. Revision II subsample of the A.C.E. is referred to as the Revision sample, and the new coding operation as the Revision coding). The missing data adjustments for the A.C.E. E sample and P sample are unchanged from those used to produce A.C.E. estimates.

The A.C.E. Revision II assignment of probability of correct enumeration, residency and match status used the method of defining imputation cells and donor pools. See Ikeda & McGrath (2001) for how this was done in the A.C.E.

Unresolved cases resulted when the information an interview collected was ambiguous or incomplete and did not allow us to determine whether the person was correctly counted. See Childers (2001) for details on the A.C.E. coding.

In the A.C.E. Revision II Revision coding those E-sample people and A.C.E. people (P-sample people) selected in the A.C.E. Revision sample are recoded incorporating the results of the Evaluation Followup (EFU), an evaluation of the A.C.E. The Revision coding generated one code based on the EFU, and another based on a recoding of the A.C.E. data. Selection rules evaluated the information in the two forms. One of these two codes was chosen as the 'best' code, that is, the code used in the estimation. A person was unresolved if the best code was unresolved. The codes applied in the Revision coding generally had the same definitions they had in the A.C.E. production. For details on the A.C.E. Revision II Revision coding see Adams & Krejsa (2002).

The people unresolved as a result of the Revision coding will be largely a different group of people than those unresolved in the original A.C.E. coding. Many of the cases coded unresolved in the A.C.E. production will be resolved based on the results of the EFU. On the other hand, cases previously coded resolved can now be coded unresolved because EFU information that is unresolved can be selected over resolved A.C.E. information. This happens when the unresolved EFU form has more or better information than the A.C.E.

2 Unresolved Cases

The Revision coding assigned each E-sample and P-sample person in the Revision sample a best code (Adams & Krejsa 2002). As mentioned earlier, for a given person record the Revision coding may differ from the original A.C.E. code. However, the definitions of the A.C.E. Revision II codes are essentially the same. The following definitions of unresolved codes represent slight modifications to the definitions given in Childers (2001), reflecting that now either the PFU or EFU interview can be unresolved, or both. The definitions of the resolved codes were unchanged from the A.C.E. and are found in **APPENDIX 2**.

P-sample Unresolved Codes:

NU Not enough information is collected during the A.C.E.'s person follow-up interview (PFU) and/or EFU to identify the P-sample person as a resident or non-resident in the block cluster. The residence status for the P-sample person is unresolved. This code is also used when the P-sample person is followed up to collect geographic information and that information is not collected.

MU The A.C.E.'s person followup interview (PFU) and/or EFU obtained no useful information to resolve the unresolved residence status for the matched person. The P-sample person's residence status is unresolved. If the match was to an E-sample person, the E-sample person's enumeration status is unresolved.

P There is not enough information collected in the PFU or EFU to determine if the possible match is a match or not. The match status of the P-sample and E-sample people is unresolved.

KI Match not attempted for the P-sample person because the person has insufficient information for matching and followup. The name is blank or incomplete or the name is complete but the person has only one characteristic. This is a computer assigned code and these people are suppressed from view from the clerical matchers.

KP Match not attempted for the P-sample person, because (1) the name is incomplete, such as "Mr. Jones", or (2) the name is not a valid name, such as "White Female" or "Donald Duck". This is a clerically assigned code.

E-sample Unresolved Codes:

UE Not enough information is collected during the A.C.E. person followup interview (PFU) and/or EFU to identify the census person as correctly or erroneously enumerated in the census. The enumeration status for the E-sample person is unresolved. This code is also used when the E-sample person is followed up to collect geographic information and that information is not collected.

GU The geographic work for the targeted extended search is unresolved. The field work was not done or the block number on the form was not in the surrounding blocks, in the block cluster, or on the map. It is not clear where the housing unit is located.

P See definition above for P-sample cases.

MU See definition above for P-sample cases.

In addition to these codes taken from the A.C.E., we had a new unresolved code, the conflicting code (Adams & Krejsa 2002), 'N', and another new code, 'ZZ', indicating the person was not in the Revision sample. These codes could be applied to both P-sample and E-sample people.

N The A.C.E. PFU and the EFU interviews had contradictory information and we could not determine which was correct. These people have unresolved Census Day residency/enumeration status.

ZZ Not in the Revision sample, did not get recoded.

3 Setting Final Codes, and Probabilities of Correct Enumeration, Residency, Match and Inmover Status, and Weights

In this section we describe the setting of probabilities of correct enumeration, residency, match, and inmover for cases where they are resolved. We also describe the setting of sampling weights for estimation. We say a person has a resolved correct enumeration, residency, match status or inmover status when the information gathered in the pertinent interviews has allowed us to code the status definitively. The resolved codes are defined in **APPENDIX 2**. It is worth pointing out that only status relevant to E-sample people is enumeration status. For P-sample people however, both the residency and match status are relevant. A P-sample person can have a resolved match status (match code of MU or NU) but an unresolved residency status. P-sample people with unresolved match status (P, KI or KP) always have unresolved residency status also.

3.1 Definitions of New Variables

The variables new to A.C.E. Revision II are listed and described below. They are specified in the subsequent sections of this chapter.

P-sample variables:

MPROB_FINAL	probability of match
RPROB_FINAL	probability of residence
IPROB_FINAL	probability of being an inmover
TESFINWT_FINAL	final person weight we use for nonmovers, outmovers and inmovers
TESFINWT_NEW	person weight that reflects the A.C.E. coding and non-interview adjustment but the Revision sampling
RSC_FINAL	final resident status code
MOVERPER_FINAL	final mover person status

E-sample variables:

CEPROBI_FINAL	initial probability of correct enumeration (before adjustment for duplicate counts)
CEPROBF_FINAL	probability of correct enumeration (after adjustment for duplicate counts)
TESFINWT_FINAL	final person weight we use for E-sample people
TESFINWT_NEW	person weight that reflects the A.C.E. coding but the Revision sampling

3.2 Setting Mover Status Codes

The final residence status code (RSC_FINAL), also called the A.C.E. status code, was set equal to the final A.C.E. status code as determined by the A.C.E. Revision II coding, FACESTAT (Adams & Krejsa 2002), except when FACESTAT indicated unresolved A.C.E. Revision II

mover status coding results, in which case RSC_FINAL reverted to the original A.C.E. RSC value. See **APPENDIX 1** or Childers (2001) for discussion on the A.C.E. variable RSC.

The final mover person status (MOVERPER_FINAL) was obtained from the RSC_FINAL, except when RSC_FINAL was unresolved or removed (U or R), in which case it was set to the A.C.E. value of MOVERPER. MOVERPER reflected the setting to outmover status of people with unresolved status from the whole household outmover path, and the setting to removed status of partial household inmovers 18-22 years old who were in group quarters on Census Day (Ikeda 2001b); see **APPENDIX 1** for a definition of the values MOVERPER can take.

P-sample Coding of RSC_FINAL

If FACESTAT = I, N, O, R, U then RSC_FINAL = FACESTAT
If FACESTAT = BD, DD, DK, NA, NI then RSC_FINAL = RSC

P-sample Coding of MOVERPER_FINAL

If RSC_FINAL = I then MOVERPER_FINAL = 2
If RSC_FINAL = N then MOVERPER_FINAL = 1
If RSC_FINAL = O then MOVERPER_FINAL = 3
If RSC_FINAL = R and MOVERPER = 1 then MOVERPER_FINAL = 1
If RSC_FINAL = R and MOVERPER = 2 then MOVERPER_FINAL = 2
If RSC_FINAL = U and MOVERPER = 1 then MOVERPER_FINAL = 1
If RSC_FINAL = U and MOVERPER = 3 then MOVERPER_FINAL = 3

3.3 Assigning Match, Residence and Inmover Probabilities to Resolved P-sample

The probability of match status and residency status have the same interpretation for the A.C.E. Revision II as they did in the A.C.E. However, in the A.C.E. Revision II we introduce the probability of being an inmover, IPROB_FINAL. Note that conflicting cases receive a probability of residency of 0.5, while their match status is their MER match status. See **Section 6** for a discussion on conflicting cases. Note that EUF1 is the match code assigned in the MER and CODE_FINAL is the highest match code for input into the missing data processing (Beaghen 2002). The match codes seen below are defined in **Section 2** or in Childers (2001). '*' indicates that you should go to **Section 4** for specifications on how probabilities were set for people with unresolved codes.

MPROB_FINAL

Set to 1 if CODE_FINAL is M, MR, or MU
Set to 0 if CODE_FINAL is DP, FP, GP, NC, NL, NN, NP, NR, NU, MN, ZZ, any inmover code, RSC_FINAL = I or R
*Set to imputed probability of match if CODE_FINAL is KI, KP, P

If CODE_FINAL is N then do
Set to 1 if EFU1 is M, MR or MU
Set to 0 if EFU1 is DP, FP, GP, NC, NL, NN, NP, NR, NU, MN, ZZ or any inmover code
*Set to imputed probability of match if EFU1 is KI, KP, P

RPROB_FINAL

Set to 1 if CODE_FINAL is M, MR, NC, NP, NR
Set to 0 if CODE_FINAL is DP, FP, GP, NL, NN, MN, or RSC_FINAL= I or R
*Set to imputed probability of residence if CODE_FINAL is KI, KP, MU, NU, P
Set to 0.5 if CODE_FINAL is N

IPROB_FINAL

Set to 1 if CODE_FINAL is IK, IR, IU
Set to 1 if CODE_FINAL is a NC, NL, NN, NP, NR, NU, MR, MU or MN and RSC_FINAL = I
Set to 0 if CODE_FINAL is ID, IF, IG, IN, ZZ, N, DP, FP, GP, KI, KP
Set to 0 if RSC_FINAL = R, N, O or U

3.4 Assigning Enumeration Probabilities to Resolved E-sample

CEPROBI_FINAL and CEPROBF_FINAL are analogs to the A.C.E. variables CEPROBI and CEPROBF. CEPROBI described the probability of correct enumeration before taking into account an adjustment for duplicates found inside the A.C.E. search area. CEPROBF was the probability of correct enumeration after taking into account the adjustment for duplicates. CEPROBI_FINAL is based on the Revision coding. However, the Revision coding did not conduct a new duplicate search. Thus to calculate CEPROBF_FINAL we use the duplicate counts taken from the Measurement Error Reinterview (MER) coding (Davis & Raglin, 2001), RDUPCNT1 and RDUPCNT2. There was a duplicate search conducted in the MER and since it took account of the EFU results the duplicate counts are closer to what Revision duplicate counts would have been. '*' indicates that you should go to **Section 4** for specifications on how probabilities were set for people with unresolved codes.

CEPROBI_FINAL

Set to 1 if CODE_FINAL is M, MR, CE
Set to 0 if CODE_FINAL is DE, EE, FE, GE, KE, MN, ZZ
*Set to imputed probability of correct enumeration if CODE_FINAL is UE, MU, P, GU
Set to 0.5 if CODE_FINAL is N

CEPROBF_FINAL

$$\text{CEPROBF\_FINAL} = \text{CEPROBI\_FINAL} * \left(\frac{1 + RDUPCNT1}{1 + RDUPCNT1 + RDUPCNT2} \right)$$

3.5 Final Weights

The final weight is TESFINWT_FINAL for the P-sample, including inmovers, and the E-sample. For the E-sample this reflects the sampling, including TES sampling. For the P-sample nonmovers and outmovers it reflects sampling, including TES sampling, and also the non-interview adjustment. For the P-sample inmovers it reflects sampling and the non-interview adjustment. TESFINWT_NEW is a weight that reflects the A.C.E. coding and non-interview adjustment but the Revision sampling. It is required for the A.C.E. Revision II estimation.

The following variables were needed in creating the weights:

PWGHT	the A.C.E. sampling weight for a P-sample person (Ikeda 2001a)
EWGHT	the A.C.E. sampling weight for an E-sample person (Ikeda 2001a)
TESWGT	the TES weight for the cluster (Ikeda 2001a)
RPWGT	the MER sampling weight for a P-sample person (Davis & Raglin 2001)
REWGT	the MER sampling weight for an E-sample person (Davis & Raglin 2001)
NIWGTO_FINAL	the A.C.E. Revision II non-interview adjusted weight for Census Day (Ikeda 2002)
NIWGTI_FINAL	the A.C.E. Revision II non-interview adjusted weight for A.C.E. Interview Day (Ikeda 2002)
TES_FINAL	the flag indicating whether the person is a TES person and requires the TES weight, TESWGT (Wolfgang 2002)

The instructions for calculating the weights is as follows below.

P-sample TESFINWT_FINAL for non-inmovers (RSC_FINAL = N, O, R, U)

If CODE_FINAL = ZZ then TESFINWT_FINAL = 0
Else if TES_FINAL = 1 then TESFINWT_FINAL = NIWGTO_FINAL*TESWGT
Else TESFINWT_FINAL = NIWGTO_FINAL

P-sample TESFINWT_FINAL for inmovers (RSC_FINAL = I)

TESFINWT_FINAL = NIWGTI_FINAL

P-sample TESFINWT_NEW for all A.C.E. P-sample

TESFINWT_NEW = TESFINWT * RPWGT/PWGHT

E-sample TESFINWT_FINAL

If TES_FINAL = 1 then TESFINWT_FINAL = REWGT*TESWGT
Else TESFINWT_FINAL = REWGT

E-sample TESFINWT_NEW

TESFINWT_NEW = TESFINWT * REWGT/EWGHT.

4 People with Unresolved Residency/Enumeration Status Only

In the A.C.E. E-sample people received unresolved enumeration status only as a result of an incomplete or ambiguous PFU. Likewise, P-sample people with only unresolved residency status (that is, excluding from consideration people with insufficient information) could only result from an incomplete or ambiguous PFU. In the Revision coding, people with unresolved enumeration status and people with only unresolved residency status could arise from either an incomplete or ambiguous PFU or EFU.

The residency status for Revision P-sample people and the enumeration status for Revision E-sample people often changed from the A.C.E. to the Revision coding because the Revision coding processed both the original information from the PFU and the new information from the EFU. Thus while the EFU information resolved many cases that were unresolved in the A.C.E. on account of the PFU, EFU cases with incomplete or ambiguous information were a new source of unresolved cases in the Revision coding. There were about the same number of E-sample unresolved cases in the Revision as in the A.C.E., more than six million, with about half of these representing new unresolved cases. In contrast, the Revision coding generated substantially more P-sample unresolved cases than the A.C.E., 4.6 million versus 2.7 million. We saw this increase because the entire Revision P-sample except insufficient information people went to EFU,

including whole households of non-matched people who did not go to PFU. These people were assumed in the A.C.E. to be resolved and could have become unresolved because of the EFU.

Note that P-sample people with only unresolved Census Day residency status are coded NU and MU. E-sample people with unresolved enumeration status are coded UE and MU.

4.1 Differences from A.C.E. Methodology - After Followup Groups

In A.C.E. Revision II the imputation cells are generally based on after followup groups defined by the results of the person followup (PFU) and evaluation followup (EFU). This is in contrast to the original, unimplemented A.C.E. missing data plan, where the imputation cells were based exclusively on information obtained before any followup was conducted; that is, the imputation cells were based on before followup groups. An ad hoc fix to the A.C.E. production missing data methodology was effected by creating the after followup groups for 'potential fictitious' and 'lived elsewhere on Census Day' (see Cantwell & Childers 2001). The PFU information was deemed to be highly relevant to resident or enumeration status. The new cells also showed greater discrimination in assigning probabilities of correct enumeration and residency. For example, the probability of residency for the potential fictitious/lived elsewhere ranges from 0.12 to 0.18, whereas the lowest probability of residence in the original plan was 0.76.

While the two new groups corrected the most obvious limitations of the first A.C.E. imputation plan, due to time exigencies the methodology of after followup groups was not fully exploited by the A.C.E. The estimation for missing data for the A.C.E. Revision II more fully develops the use of after followup groups, abandoning entirely the before followup groups. We expect this to be a significant improvement over A.C.E. missing data methodology.

The two A.C.E. after followup groups were defined on the keyed results of the PFU. In the A.C.E. Revision II to define the after followup groups we employed the keyed responses to the PFU and EFU questionnaire check boxes and the 'why' codes, which were derived from the followup results (Adams & Krejsa 2002). Using the keyed results and the why codes we identified the following:

- unresolved cases with the same history, i.e., the recipient or imputation cells;
- the resolved followup cases that share that history up to the point of being unresolved, i.e., in the donor pool.

In other words, the after followup groups were defined by the place in the followup interview where the information is ambiguous or incomplete. The ability to discriminate correct enumeration versus erroneous enumeration status, or resident versus non-resident status, was a criterion by which we judged the efficacy our imputation cells. The sample sizes of the potential donor pools posed a practical constraint on our ability to refine imputation cells. The Revision sample is just a fraction of the A.C.E. production sample.

The imputation cells mostly were defined the same way for the P-sample and E-sample unresolved because the P-sample questions about Census Day residency were the same as the E-sample questions about enumeration status on both the PFU and EFU forms. In particular, the after followup groups are defined the same way for the P-sample and E-sample, though some of these groups were subdivided in different ways. Thus the subsequent discussion applies generally both to P-sample and E-sample estimation of unresolved.

4.2 Calculating Probabilities from Donor Pools

Unresolved cases were assigned to groups called imputation cells which were associated with donor pools. To calculate the probability of correct enumeration for a given cell:

$$\text{Imputed probability of CE} = \frac{\text{Weighted CE's in Donor Pool}}{\text{Weighted Enumerations in Donor Pool}}.$$

For the P-sample, probabilities of residency and match status are calculated analogously.

4.3 Using Why Codes and Keying Results

The responses to both the PFU and EFU interviews were keyed from the interview questionnaires and constitute datasets. The responses to both the check box questions and the handwritten notes were keyed. However, we only used the answers to the check box questions to define AFU groups. We did not use the handwritten notes on the forms as they didn't allow for an algorithm to assign them into groups.

Why codes were based on the results of the PFU and EFU forms. The definitions of the why codes taken from Adams & Krejsa (2002) appear in **APPENDIX 3**. Why codes indicate the reason why an enumeration was coded correct or erroneous, or resident or non-resident, or unresolved. In the automated A.C.E. Revision II coding they are derived directly from the keyed PFU and EFU forms. In the clerical coding they generally followed the results of the forms. However, since the clerical coders have access to notes on the forms clerically assigned why codes may be superior to what one obtains based strictly on keyed results.

4.4 The After Followup Groups

It is useful to make a distinction between what we call uninformative and informative unresolved:

- uninformative unresolved; the followup was a non-interview or a partial non-interview, though there was no evidence of an erroneous enumeration or non-resident.
- informative unresolved; a followup interview was conducted and there was evidence of an erroneous enumeration or non-resident.

An example of uninformative unresolved is a person for whom the PFU and EFU were non-interviews.

An example of people with informative unresolved are those where the EFU interview indicated they moved in since Census Day, or moved out before Census Day, but did not provide the address they moved to or from. Such people define one of our AFU groups. The donor pool consists of those resolved people who indicated in the followup that they moved in after Census Day, or moved out before Census Day; these were generally people who provided the mover address in the followup. We had an analogous after followup group for people unresolved because of they indicated they were movers on the PFU interview.

Note that when one interview was uninformative unresolved but the other interview was resolved, the Revision coding chose, or coded based on, the resolved interview. On the other hand, when the unresolved interview was informative, the Revision coding could choose the unresolved interview over a resolved one. See Adams & Krejsa (2002) for details of the Revision coding.

Table 1 shows the nine after followup groups defined for cases unresolved in the EFU, and **Table 2** shows the nine after followup groups defined for cases unresolved in the PFU. In **Table 1** and **Table 2** the three letter codes in the left hand columns identify the imputation cells. They are usually the same for the P-sample and E-sample, with the exception of some of the informative EFU cells which were divided based on operational characteristics such as whether the person went to PFU and the housing unit match status. There are three housing unit match statuses: nonmatch, housing unit match and conflicting household.

People who moved in after Census Day or moved out before Census Day were the largest informative after followup group group. Another important informative after followup group was formed by people who, according to the followup, had another residence such as a vacation home, though the followup form did not indicate whether the other residence or the sample address was the Census Day residency. The larger uninformative groups are the non-interview groups and "didn't answer other residence questions" groups.

Some of the larger EFU groups were subdivided by A.C.E. operational variables such as whether or not the household went to PFU, or whether the household was conflicting. The uninformative after followup group groups tended to have imputed probabilities of correct enumeration or residency close to one, typically in the range of 0.92 to 0.99, whereas the informative after followup group groups had lower probabilities, often less than 0.25. See Beaghen and Sands (2002) for the probabilities and the numbers of recipients and donors.

See **APPENDIX 4** for specifications on how to define each individual cell, both the recipients and the donors, and for commentary discussing the choice of recipient and donor groups. In **APPENDIX 4** the cells are listed in the same order as in **Table 1** and **Table 2**, with EFU cells (**Table 1**) coming before PFU cells (**Table 2**).

A few comments on the specifications in **APPENDIX 4** are appropriate.

- Recipients and donors for EFU cells must have been sent to EFU.
- Recipients and donors for PFU cells must have been sent to PFU.
- Recipients have unresolved CODE_FINAL.
- Donors have resolved CODE_FINAL.

See the first page of **APPENDIX 4** for more details on who qualifies as recipients and as donors.

Table 1: EFU After Followup Groups

E-Sample Cells	P-Sample Cells	Informative Groups
EAD	EAD	The followed up person 'Lived elsewhere' or at an 'other residence', but the address was not given
EMO	EMO	Followed up person moved in after Census Day or out before Census Day, but Census Day address not given
ENL	ENL	Respondent indicated the followed up person 'Never lived here' at the sample address, but did not provide the Census Day address
EOR	EOR	The followed up person had an 'Other residence', but did not indicate whether sample address or the other residence was the Census Day residency
EMP	EMP	Followed up person moved in or moved out, but no move dates given
		Uninformative Groups
EKB, EKC,	EKF, EKU	The respondent indicated the followed up person 'Lived here' at the sample residence, but did not answer the other residence question
EIB, EIC	EIB, EIC, EID	Non-interview (1); the respondent answered the current residence question, but did not answer the group quarters and other residence question
E2B, E2C	E2B, E2C, E2D	Non-interview (2); the respondent did not answer the usual residence question, nor the group quarters and other residence questions
EKR	EKR	Potentially fictitious person, no respondents knew of the followed up person

Table 2: PFU After Followup Groups

E-Sample and P-Sample Cells	Informative Groups
PAD	The followed up person 'Lived elsewhere' or at an 'other residence', but the address was not given
PMO	Followed up person moved in after Census Day or out before Census Day, but Census Day address was not given
PN2	Non-interview (2); the respondent indicated the followup person 'did not live here' at the sample address, but did not indicate the other address and did not answer the group quarters and other residence questions
POR	The followed up person had an 'Other residence', but did not indicate where the usual residence was
	Uninformative Groups
POK	The respondent indicated the followed up person 'Lived here' at the sample residence, but did not answer the other residence question
PNI	Non-interview (1); the respondent answered the usual residence question, but did not answer the group quarters and other residence questions
PN3	Non-interview (3); the 'lived here' question is DK/refused, and the group quarters and other residence questions were not answered
PN4	Non-interview (4); blank questionnaire
PKR	Potentially fictitious person, no respondents knew of the followed up person

For a few groups why codes can be used directly to form donor pools. This is true for the EFU movers who provide no address, 'Movers, no address', because both the resolved mover cases and the unresolved mover cases have the same why codes, either MI for inmovers or MO for those who moved out.

4.5 When both the PFU and EFU were Unresolved

It often happened that both the PFU interview and the EFU interview were unresolved. In that case in order to assign a cell for imputation we chose the unresolved interview that was more informative. When both interviews had the same level of information we usually chose the EFU over the PFU because we believed the EFU questionnaire questions were more sharply defined.

When both the PFU and EFU were unresolved (PFUF and EFUF both have unresolved codes), we choose between the PFU or the EFU. The Measurement group's coding specialists generally were able to pick the EFU or the PFU as the best form, and the corresponding code as the best code. However, when both EFU and PFU were unresolved they did not choose between the two (i.e., they chose both). Choosing between unresolved PFU and EFU codes is a task of the missing data processing.

The logic we applied in deciding whether to select the unresolved EFU or unresolved PFU was similar to that applied in best coding; an unresolved form that reveals more information is selected over one with less information. There were three basic rules that guided the choice:

- Choose an informative unresolved EFU or PFU over an uninformative one.
- Among the noninformative unresolved cases, sometimes one case has more information than another. For example, a PFU 'non-interview' case (why code of NI) is missing more information than an EFU case where the other residence questions were not answered (why code of ORDK). Hence we select the EFU with ORDK over the PFU with NI.
- Similarly, with competing 'informative' unresolved EFU and PFU interviews, we chose the interview that reflected stronger evidence of an erroneous enumeration or nonresidency. For example, we picked a form with a why code of MI, MO or AD over one with OR.
- EFU and PFU interviews resulting in 'Potential fictitious' (why code KR) were a special case. If there was any information about the person on the another indicating that someone was knowledgeable about that person then they were not potential fictitious. Thus when one form is potential fictitious we picked the non-potentially fictitious unresolved form.

Note that often when the PFU and EFU were both unresolved they had the same or similar followup group. When this happened we generally chose the EFU over the PFU. While this choice is not clear and may not make much difference, there were reasons why it might be preferable to choose the EFU over the PFU.

- First, the why codes for the EFU were sometimes more precisely defined. For example, NL/DL is defined in the EFU as answering question 2.a in section 4 as 'Never lived at *A.C.E. sample address*'. On the other hand, there was no place to answer 'Never lived at *A.C.E. sample address*' on the PFU, though it was not unusual for both EFU and PFU to be have a why code of NL/DL. Similarly, MP was defined in terms of not answering certain specific questions for the EFU, though not for the PFU.
- Second, the unresolved rate for the PFU was lower. According to the PFU/EFU Review (Adams & Krejsa 2001), there were about six million unresolved PFU cases as compared to about 14 million unresolved EFU cases. Since for some unresolved cases donors consist largely of cases resolved by the other form, choosing the EFU means the donor pool will be more complete. This point was particularly pertinent for the estimation for potential fictitious (KR) cases.

Table 3 Which Form to Select when both the PFU and EFU are Unresolved

EFU Groups / PFU Groups	KR	GO	N NF, Other	NI, MP, TE	ORDK, MX, AL, GQ, RC	OR	NL/DL, MI	MO, AD
BFUFLAG = blank or 0	EFU	EFU	EFU	EFU	EFU	EFU	EFU	EFU
KR	EFU	EFU	EFU	EFU	EFU	EFU	EFU	EFU
GO	PFU	EFU	EFU	EFU	EFU	EFU	EFU	EFU
NF, other	PFU	PFU	EFU	EFU	EFU	EFU	EFU	EFU
NI, TE	PFU	PFU	PFU	EFU	EFU	EFU	EFU	EFU
ORDK, MX, AL, GQ, RC	PFU	PFU	PFU	PFU	EFU	EFU	EFU	EFU
OR	PFU	PFU	PFU	PFU	PFU	PFU	EFU	EFU
NL/DL	PFU	PFU	PFU	PFU	PFU	PFU	EFU	EFU
MI, MO, MP, AD, NI-didn’t live here	PFU	PFU	PFU	PFU	PFU	PFU	PFU	EFU

Table 3 illustrates the decision rule for choosing between the EFU or PFU when both were unresolved. Observe there were seven groupings, and that they were listed in order of least preferred to most preferred. Thus the first or least preferred group is KR or potential fictitious, and the most preferred group was MI, MO, AD.

Some details are not represented in **Table 3**. These are as follows:

- BFUFLAG = blank or 0 indicates this was a case that was not supposed to go to PFU. For these cases ignore any PFU why codes (PFUFY).
- On the PFU - some NI's (non-interviews) have question 4.a filled with did not live here on Census Day. These are separated from those NI's that have question 4.a (ALIVECD) blank, DK or refused, or answered 'lived here on Census Day'.
- In the PFU/EFU Review cases were given a why code of OR that would have been given a why code of ORDK in the Revision coding. Thus PFUFY = OR cases with the answer 'No' to question 6 (AOTHRES = 2), the other residence question, are treated like ORDK.
- In the PFU/EFU Review and in the automated Revision coding based on the PFU keying, some cases with a why code of OR would have been given a why code AD in the Revision's clerical coding. These cases are identified by not having the other address written in the form (OTHADD = 1) and by the bestcode coming from the PFU/EFU Review or the PFU keying (SOURCE = 1 or 2).
- JBPb is treated like ORDK; it indicates, did not fill out or DK/refused question 6.b, 'a place where *followed up person* stayed regularly while away on a job'
- DO and DU are treated like ORDK. Person answered staying in dorm but turns out really was not a dorm.
- LH is treated like ORDK. It means lived here. It is not clear why they are unresolved.
- DE, GE, NT, X, FE, GC, NH, 4B, JBP3, JC3, VCb, DF, DP, FP, and DB are put in with the NF's. These are anomalous codes or codes unexpectedly applied to unresolved cases, and there are only small numbers of each. They comprise the 'other' group.
- OR includes other residence codes; i.e., OH, OV, OC, OJ, OM, OP, OV, OW and any of these with suffixes such as 1 or 2. It also includes a small number of OS or out of sample cases.
- HO or HO2 is treated like OR. These lived elsewhere, but gave neither other address or where they lived on Census Day.
- MICD is treated like MI.

5 Resolving P-sample Cases with Unresolved Match Status and Residency Status

In the A.C.E., P-sample people with unresolved match status and residency status came about in one of two ways. First, the A.C.E. person interview (PI) may not have provided sufficient information for match and followup. Such people have match codes of KI and KP. Second, the A.C.E. person followup (PFU) may not have collected adequate information to allow us to determine a person's match status. These people were possible matches and have a match code of P.

The Revision P-sample people with insufficient information for match and followup tended to be the same people with insufficient information for match and followup in the A.C.E., except for some rare cases with coding changes. Note that people who were insufficient information in the A.C.E. were not sent to EFU. There were about three million weighted people with insufficient information for match and followup in both the A.C.E. and the Revision samples.

In the A.C.E., P-sample people with insufficient information for match and followup were assigned a probability of Census Day residency equal to the residency rate of P-sample people who went to PFU. An improvement over production is that the cells are broken down by conflicting household status (not to be confused with conflicting interviews, a conflicting household is one where the P-sample household and E-sample household have no people in common), see Ikeda (2001a). Research after the A.C.E. production indicated this breakdown would improve the assignment of match probabilities. For the non-conflicting households we also break down the imputation cells by whether the housing unit was matched or non-matched; see **APPENDIX 5** for the specifications for these cells. Possibly matched people, who are indicated with a match code of P, were assigned the overall residency of those before followup possible matches who were resolved (that is, BFUGP=2; see **APPENDIX 1** for definition of BFUGP).

The probability of match was assigned based on the overall match rate, divided into five groups based on mover status and housing unit match status, as was done in the A.C.E., and additionally on conflicting household status; see **APPENDIX 6** for the specification for these cells.

6 Revision E Sample and P Sample Conflicting Coding

When the A.C.E. person followup (PFU) and the evaluation followup (EFU) interviews had contradictory information and it could not be determined which was correct, the Revision coding assigned the person a code of conflicting (the conflicting coding is not to be confused with conflicting households, which is where the P-sample household has no people in common with the E sample household). All cases found to be conflicting in the Revision automated recoding were sent to analysts for clerical review. By examining the handwritten notes of interviewers, the analysts could often determine which of the interviews was the better and appropriately assign a code. There were some cases where the interviews appeared to be of equal quality, such as when

both respondents were household members or both respondents were of equal caliber proxy. For these conflicting cases, the interviews seemed equally likely to be correct based on the expertise of the analysts. Therefore, probabilities of 0.5 were assigned both for correct enumeration status of Revision E-sample conflicting cases and for Census Day residency status of Revision P-sample conflicting cases. It should be noted that the recoding of the Revision samples resulted in considerably less conflicting cases than the PFU/EFU Review sample. The PFU/EFU Review sample had about 2.6 million weighted people (Adams & Krejsa, 2001) in contrast to only about 100,000 weighted people in the Revision samples.

We assigned conflicting P-sample people the match status they had been assigned in the Measurement Error Reinterview, or MER (Krejsa & Raglin, 2001). The Revision coding determined residency status and not match status, except for some cases whose match status was corrected based on the results of the Matching Error Study. Thus there was little change in match status between the A.C.E. and the MER. However, since the MER took into account any new information from the EFU we assigned its match status.

7 Missing Data Processing Variables

This section describes missing data processing variables that are on the missing data output files. These variables serve to document the missing data processing. They may provide researchers insight into the processing and may facilitate alternative missing data imputations.

Each Revision P-sample and E-sample person has the variables CELL, PCELL and ECELL.

PCELL — Every person with an unresolved PFU code (PFUF) was assigned to a PFU cell. This assignment is indicated in PCELL. PCELL is three characters long and is in the form of the three character name of an imputation cell. For example, PCELL = 'PMO' indicates that this person with unresolved PFUF was assigned to the PFU cell PMO.

ECELL — Every person with an unresolved EFU code (EFUF) was assigned to a EFU cell. This assignment is indicated in ECELL. ECELL is three characters long and is in the form of the three character name of an imputation cell. For example, ECELL = 'EMO' indicates that this person with unresolved EFUF was assigned to the EFU cell EMO.

CELL — Every person who had either an unresolved PFUF or an unresolved EFUF had a non-missing value of CELL. If only PFUF unresolved, CELL was set to PCELL. If only EFUF was unresolved, CELL was set to ECELL. When both EFUF and PFUF were unresolved, CELL was set to either ECELL or PCELL according to the rules given in **Section 4.5**.

The probability of correct enumeration or residency for a case was set equal to the value of CELL only if the CODE_FINAL was unresolved. Recall that when only one of either the PFUF or the EFUF was unresolved, then the best code is not necessarily unresolved. Thus there are people with positive values of CELL who are resolved.

MCELL Every Revision P-sample person with unresolved match status has a positive value of MCELL. For these people MPROB_FINAL = MCELL.

8 References

Adams, T. & Krejsa, E., 2001: *ESCAP II: Results of the Person Followup and Evaluation Followup Forms Review,* Executive Steering Committee for A.C.E. Policy II (ESCAP II) Report 24.

Adams, T. & Krejsa, E., 2002: *A.C.E. Revision II Measurement Subgroup Documentation,* DSSD A.C.E. REVISION II MEMORANDUM SERIES #PP-6.

Beaghen, M., 2002: *A.C.E. Revision II - Specifications for Creating A.C.E. Revision II Edited Measurement files,* DSSD A.C.E. REVISION II MEMORANDUM SERIES #PP-18.

Beaghen, M., & Sands, R., 2002: *Results from the Imputation of Unresolved and Conflicting Cases,* DSSD A.C.E. REVISION II MEMORANDUM SERIES #PP-57.

Cantwell, P., & Childers, M., 2001: *Accuracy and Coverage Evaluation Survey: A Change to the Imputation Cells to Address Unresolved Resident and Enumeration Status,* DSSD Census 2000 Procedures and Operations Memorandum Series Chapter Q-44.

Cantwell, P., & Ikeda, M., 2001: *Accuracy and Coverage Evaluation Survey: Specifications for the Missing Data Procedures; Revision of Q-25,* DSSD Census 2000 Procedures and Operations Memorandum Series Chapter Q-62.

Childers, D., 2001: *The Design Document*, DSSD Census 2000 Procedures and Operations Memorandum Series Chapter S-DT-1.

Davis, M., & Raglin, D., 2001: *Creation of Master Data Variance Files for Coverage Evaluations, Planning, Research and Evaluation Division,* TXE/2010 Memorandum Series: CM-GES-S-01-R2.

Ikeda, M., 2001a: *Accuracy and Coverage Evaluation Survey: Specifications for Data Requirements for Missing Data Input and Output Files,* DSSD Census 2000 Procedures and Operations Memorandum Series Chapter Q-40.

Ikeda, M., 2001b: *Accuracy and Coverage Evaluation Survey: Some Notes Related to Accuracy and Coverage Evaluation Missing Data Procedures*, DSSD Census 2000 Procedures and Operations Memorandum Series Chapter Q-77.

Ikeda, M. 2002: *A.C.E. Revision II - Specifications for the Noninterview Adjustment*, DSSD A.C.E. REVISION II MEMORANDUM SERIES #PP-21.

Ikeda, M., & McGrath, D., 2001: *Accuracy and Coverage Evaluation Survey: Specifications for the Missing Data Procedures; Revision of Q-25*, DSSD Census 2000 Procedures and Operations Memorandum Q-62.

Kostanich, D., 2002: *A.C.E. Revision II: Design and Methodology*, DSSD A.C.E. REVISION II MEMORANDUM SERIES #PP-30.

Krejsa, E., & Raglin, D., 2001: *ESCAP II: Evaluation Results for Changes in A.C.E. Enumeration Status*, Executive Steering Committee for A.C.E. Policy II (ESCAP II) Report 24.

Sands, R., 2002: *Accuracy and Coverage Evaluation Revision II Survey: Specifications for Data Requirements for Estimation Input Files*, DSSD A.C.E. REVISION II MEMORANDUM SERIES #PP-23.

Wolfgang, G., 2002: *A.C.E. Revision II - Specifications for Computer Preprocessing for Missing Data Input Files for the A.C.E. Revision II*, DSSD A.C.E. REVISION II MEMORANDUM SERIES #PP-19.

APPENDIX 1 A.C.E. Variables Set in Missing Data Processing

BFUGP P-Sample Before Follow-up Match Code Group
0=RSC of I or R
1=Matches needing follow-up
2=Possible matches
3=Nonmatches needing follow-up from partial household nonmatches
4=Nonmatches needing followup from whole-household nonmatches
5=Nonmatches from conflicting households
6=People resolved before follow-up
7=People with insufficient information for matching
8=People who are potentially fictitious or potentially lived elsewhere on Census Day
9=RSC of N or U but FINMAT is blank

MOVERPER Person Mover Flag - usually set during A.C.E. interview as:
1=nonmover
2=inmover
3=outmover
When not set from interview, it was coded as:
1=residence status code of N
2=residence status code of I
3=residence status code of O
If residence status code equaled U or R then MOVERPER = 1 unless the person came from the whole household outmover path in which case MOVERPER = 3.

RSC Computer Residence Status Code (A.C.E. Status Code)
Possible codes on file are:
N(nonmover)
I(inmover)
O(outmover)
U(unresolved)
R(remove)
There is an edit that changes RSC of U to I if MOVERPER=2 (inmover).
There is another edit that changes RSC from I to R if a partial household person inmover with AGE=18-22 was in a GQ on Census Day.

APPENDIX 2 A.C.E. Match Codes

P-Sample Person Match Codes

Matched Resident

M The P-Sample and the census people were matched.

MR The P-Sample follow-up interview determined that the matched person with unresolved resident status is a resident as of Census Day.

Not Matched Resident

NP The P-Sample person is not matched to an E-Sample person. There was no follow-up for the whole household nonmatches from person interviews with household members and the whole household nonmatches were not conflicting household nonmatches. The P-Sample person is considered to be a resident on Census Day.

NC The P-Sample nonmatch was found on the census roster. This person in a partial nonmatch household was not matched to the census because only name was collected in the census for this person in a large household and the census person was not data defined. No follow-up interview is necessary. The P-Sample person is considered to be a resident on Census Day.

NR The P-Sample person is identified as a resident in the block cluster on Census Day during the A.C.E. person follow-up interview.

Non-resident

FP The P-Sample person is fictitious in this block cluster. The person is included in the independent roster in error during the CAPI interview. This person is not included in the final list of P-Sample people.

NL The P-Sample person did not live at the sample address or in the block cluster on Census Day and was listed as a non-mover or out-mover in error. This person is removed from the list of P-Sample people since the person was collected during the person interview in error.

NN The P-Sample person is identified as a non-resident in the block cluster on Census Day during the A.C.E. person follow-up interview, because the person lived in group quarters or had another residence where the person should have been counted on Census Day according to census residence rules. This person is removed from the list of P-Sample people, since he or she was collected during the person interview in error.

GP The P-Sample person is removed because the person interview was conducted at a housing unit that exists outside the sample block cluster. The person follow-up identified this housing unit as a P-Sample geocoding error.

DP The P-Sample person is a duplicate of another P-Sample person.

APPENDIX 2 A.C.E. Match Codes

MN The A.C.E. person follow-up interview determined that the matched person with unresolved resident status is not a resident in this housing unit or in this block cluster. The person is no longer in the list of P-Sample people.

E-sample Resolved Match Codes

Correctly Enumerated

M The P-Sample and E-Sample people were matched. The E-Sample person is correctly enumerated.

CE The E-Sample nonmatch is identified as correctly enumerated during the A.C.E. person follow-up interview.

MR The A.C.E. person follow-up interview determined that the matched person with unresolved resident status is a resident. The E-Sample person is a correct enumeration.

Erroneously Enumerated

GE The E-Sample person is erroneously enumerated in this block cluster, because the census housing unit is a geocoding error (i.e., counted in the block cluster in error). The E-Sample person should have been enumerated elsewhere in the census.

EE The E-Sample nonmatch is identified as erroneously enumerated from the A.C.E. person follow-up interview.

FE The E-Sample nonmatch is determined to be fictitious in this block cluster during the follow-up interview. The person may have existed, but should not have been enumerated in the census within this block cluster. The E-Sample person is erroneously enumerated in the census in this block cluster.

DE The E-Sample person is a duplicate of another E-sample person or a duplicate of a census person in a surrounding block (DE is also assigned to non E-Sample people duplicated with an E-Sample person in the same cluster).

MN The A.C.E. person follow-up interview determined that the matched person with unresolved resident status is not a resident in this housing unit or in this block cluster. The E-Sample person is an erroneous enumeration.

KE Match not attempted for the E-Sample person. The name is blank or incomplete or the name is complete but there are one or no person characteristics (computer assigned). The census name is blank or incomplete or not a valid name (clerically assigned).

APPENDIX 3 Why Codes

Table – Why Code Categories

Category	Code	Definition
No followup/ Noninterview	NF	No Followup (for PFU cases that were not sent to followup)
	NI	Noninterview (includes DK/Ref on Moving Date–EFU or Did you live here?–PFU)
	KR	Not enough knowledgeable respondents
	FE	Fictitious
	KE	Insufficient Information for Matching and Followup
	DE	Duplicate - based on the information from the form only
Easy person	LH	Lived Here on Census Day, no other residences, no staying in GQ, no moving, no special type of address
Died Before/ Born After Census Day	DB	Died Before Census Day
	BA	Born after Census Day
	DC	Died On Census Day
	BO	Born On Census Day
Movers	MO	Moved Out Before Census Day
	MI	Moved In After Census Day
	NL	Never Lived Here
	MP	Moved, no date given and cannot determine if moved in or out
Other Residences	OV	Other Residence–Visiting Friends/Family
	OW	Other Residence, Work
	OS	Other Residence, School, non-dorm
	OH	Other Residence-Vacation Home
	OM	Multiple Other Residences

APPENDIX 3 Why Codes

Table – Why Code Categories

Category	Code	Definition
	OR	Other Residence–Unresolved; Has another residence, but address not given; Has another residence, but cannot determine cycle; Has another residence, knows cycle, but cannot determine where stayed most of the time
	ORDK	Other Residence–Unresolved; Don't Know if has "other residence"
	OP	Other Residence–PFU, no notes
Group Quarters-type Situations	NH	Nursing Home
	AL	Assisted Living
	RC	Retirement Community
	DO	Dorm/sorority house/frat house
	MS	Military/Shipboard–no UHE
	UH	UHE for Military/Shipboard
	GQ	Other Group Quarters/ PFU GQ , no notes
Geocoding Issues	GO	Geocoding Section–Outside Cluster
	TE	TES–Outside Cluster/Surrounding Ring
	OC	“Outside Cluster” note for non-TES housing units

APPENDIX 4 Imputation Cells - Recipients and Donors

For all E-sample EFU Groups

For all recipients and donors	REACE = 1 (went to EFU)
For all Recipients	EFUF_FINAL and CODE_FINAL are: UE, MU or P
For all Donors	CODE_FINAL is either: N, M, MR, FE, CE, GE, KE, DE, EE, MN

For all P-sample EFU Groups

For all recipients and donors	REACE = 1 (went to EFU)
For all Recipients	EFUF_EDIT and CODE_FINAL are: NU or MU
For all Donors	CODE_FINAL is either: N, M, MR, NR, FP, GP, DP, MN, NC, NN, NL, NP; and BACESTAT ne I

For all E-sample PFU Groups

For all recipients and donors	BFUFLAG = 1 or 2 (went to PFU)
For all Recipients	EFUF_FINAL and CODE_FINAL are: UE, MU or P
For all Donors	CODE_FINAL is either: N, M, MR, FE, CE, GE, KE, DE, EE, MN

For all P-sample PFU Groups

For all recipients and donors	BFUFLAG = 1 or 2 (went to PFU)
For all Recipients	EFUF_EDIT and CODE_FINAL are: NU or MU
For all Donors	CODE_FINAL is either: N, M, MR, NR, FP, GP, DP, MN, NC, NN, NL, NP; and BACESTAT ne I

Variable descriptions are found in Sands (2002).

APPENDIX 4 Imputation Cells - Recipients and Donors

Cell	EAD (P sample and E sample)
Description	The followed up person 'Lived elsewhere' or at an 'other residence', but the address is not given
Recipients	EFUFY = AD
Donors	ORWEEK = 2 or ORMONTH =2 or ORYEAR = 2 or ORNOCYC = 2
Comments	Both recipients and donors are people who indicate they had another residence and that they lived at that other residence on Census Day. They are distinguished by whether or not the CODE_FINAL is unresolved or resolved.

Cell	EMO (P sample and E sample)
Description	Moved in after Census Day or Moved out before Census Day, but no mover address provided
Recipients	EFUFY = MI, MO or MICD
Donors	EFUFY = MI or MO
Comments	When a person indicated they moved, they received a mover why code (MO, MI) whether they were resolved, gave a good mover address, or were unresolved, did not give a good mover address. A small number of MICD are also put into this recipient pool. These codes are assigned to people who moved in on census day and the person is unresolved if the mover address is unresolved. Persons in CID = 050470006331 are excluded from the donor pool. These six persons, each with a weight of 56,617, had EFUFY = MI although the date they moved in (MOVEINDS) was indicated to be in 1996. We concluded this was a misapplication of the why code MI. We also noted that of 501 people coded MI with move in dates, only three others had move in dates before 4/1/00.

Variable descriptions are found in Sands (2002).

APPENDIX 4 Imputation Cells - Recipients and Donors

Cell	ENL (P sample and E sample)
Description	Respondent indicated the followed up person 'Never lived here' at the sample address, but did not provide the census day address
Recipients	EFUFY = NL or DL
Donors	MOVEOUT = 2
Comments	The NL and DL why codes were applied only to unresolved cases, not to resolved cases. The coding instructions indicate to code NL or DL when MOVEOUT = 2, and most cases with EFUFY = NL or DL have MOVEOUT = 2.

Variable descriptions are found in Sands (2002).

APPENDIX 4 Imputation Cells - Recipients and Donors

Cell	EOR (P sample and E sample)
Description	The followed up person had an 'Other residence', but did not indicate whether sample address or the other residence was the census day residency
Recipients	EFUFY = OC, OH, OJ, OM, OP, OV, OR, OW, OS, HO, or any of these with a suffix of 1 or 2
Donors	ORWEEK = 1 or 2 or ORMONTH = 1 or 2 or ORYEAR = 1 or 2 or ORNOCYC = 1 or 2
Comments	The donors are followed up people who indicated they had another residence and specified whether the other residence or the sample residence was the census day residency.

Cell	EMP (P sample and E sample)
Description	Followed up person moved in or moved out, but no move dates given
Recipients	EFUFY = MP
Donors	MICD in (1, 2, 3)
Comments	The MP why code is applied to unresolved cases. The coding instructions indicate to code MP when MICD is in (1, 2, 3) and indeed more than half have MICD in (1, 2, 3).

Variable descriptions are found in Sands (2002).

APPENDIX 4 Imputation Cells - Recipients and Donors

Cell	EKB E-sample
Description	The respondent indicated the followed up person 'Lived here' at the sample residence, but did not answer the other residence question; and this was not a conflicting household.
Recipients	EFUFY = ORDK , JBPb[1], AL[2], GQ, RC, DO, DU, MX[3] or LH and BADINIT ≠ 4
Donors	(MI2000 = 2 or MICD = 1) and COLCYN2 =9 and GQCST = 6 and BADINIT ≠ 4
Comments	On Recipients: [1] JBPb is a why code that indicates the respondent did not answer the other residence questions about living away for a job. It represents a subset of the cases for which we had expected to have a why code of ORDK. [2] AL is a why code indicating the people were really at an assisted living home. They were coded unresolved because not all questions were asked, such as other residence questions. Thus they are like ORDK's. For similar reasons GQ, RC, DO, DU are also put with ORDK. [3] MX refers to an apartment mixup They were coded unresolved because not all questions were asked, such as other residence questions. On Donors: The recipients are people who are unresolved because they didn't answer the other residence questions, though there was no evidence of erroneous enumeration or non-resident up to that point of the other residence questions. Thus the donors are people who are resolved and who had no evidence of erroneous enumeration or non-resident up to the point of the other residence questions. This after followup group is subdivided by certain operational variables that allowed for better discrimination.

Variable descriptions are found in Sands (2002).

APPENDIX 4 Imputation Cells - Recipients and Donors

Cell	EKC E-sample
Description	The respondent indicated the followed up person ‘Lived here’ at the sample residence, but did not answer the other residence question; and this was a conflicting household.
Recipients	Same as for EKB except BADINIT = 4
Donors	Same as for EKB except BADINIT = 4
Comments	Same as for EKB

Cell	EKF P-sample
Description	The respondent indicated the followed up person ‘Lived here’ at the sample residence, but did not answer the other residence question; and went to PFU
Recipients	Same as for EKB except BFUFLAG in (1, 2)
Donors	Same as for EKB except BFUFLAG in (1, 2)
Comments	Same as for EKB

Variable descriptions are found in Sands (2002).

APPENDIX 4 Imputation Cells - Recipients and Donors

Cell	EKU P-sample
Description	The respondent indicated the followed up person 'Lived here' at the sample residence, but did not answer the other residence question; and did not go to PFU
Recipients	Same as for EKB except BFUFLAG in (1, 2)
Donors	Same as for EKB except BFUFLAG not in (1, 2)
Comments	Same as for EKB

Cell	EIB E-sample
Description	Non-interview (1); the respondent answered the current residence question, but did not answer the group quarters and other residence question. And not a conflicting household.
Recipients	EFUFY = NI and USURNOW not blank, or EFUFY = GO, NF or TE or other The rubric 'other' consists of EFUFY of DE, GE, NT, X, FE, GC, NH, 4B, JBP3, JC3, VCb, DF, DP, FP, and DB. And BADINIT ≠ 4.
Donors	GQCST in ('1', '2', '3', '4', '5', '6') and BADINIT ≠ 4
Comments	The first group of non-interviews had an answer to the EFU question about current residence (USURNOW). However, other key parts of the interview were not completed; to receive a why code of NI, at minimum the group quarters and other residence questions needed to be unanswered. Thus we choose as our recipients those who are resolved and answered the group quarters questions.

Cell	EIC E-sample
Description	Non-interview (1); the respondent answered the current residence question, but did not answer the group quarters and other residence question; and a conflicting household.
Recipients	Same as EIC except BADINIT = 4
Donors	Same as EIC except BADINIT = 4
Comments	Same as for EIC

Variable descriptions are found in Sands (2002).

APPENDIX 4 Imputation Cells - Recipients and Donors

Cell	EIB P-sample
Description	Non-interview (1); the respondent answered the current residence question, but did not answer the group quarters and other residence question; and a conflicting household.
Recipients	Same as EIC except BADINIT = 4
Donors	Same as EIB except BADINIT = 4
Comments	Same as for EIB

Cell	EIC P-sample
Description	Non-interview (1); the respondent answered the current residence question, but did not answer the group quarters and other residence question; and a conflicting household and went to PFU.
Recipients	Same as EIB except BADINIT ≠ 4 and BFUFLAG in (1, 2)
Donors	Same as EIB except BADINIT ≠ 4 and BFUFLAG in (1, 2)
Comments	Same as for EIB

Cell	EID P-sample
Description	Non-interview (1); the respondent answered the current residence question, but did not answer the group quarters and other residence question; and a conflicting household and did not go to PFU.
Recipients	Same as EIB except BADINIT ≠ 4 and BFUFLAG not in (1, 2)
Donors	Same as EIB except BADINIT ≠ 4 and BFUFLAG not in (1, 2)
Comments	Same as for EIB

Variable descriptions are found in Sands (2002).

APPENDIX 4 Imputation Cells - Recipients and Donors

Cell	E2B E-sample
Description	Non-interview (2); the respondent did not answer the usual residence question, nor the group quarters and other residence questions; and not a conflicting household.
Recipients	EFUFY = NI and USURNOW = blank and BADINIT ≠ 4
Donors	resolved EFU cases and BADINIT ≠ 4
Comments	Since the recipients are people for whom the followup obtained little or no information, all resolved cases serve as the donor pool (except that we break down this group into conflicting household, non-conflicting household).

Cell	E2C E-sample
Description	Non-interview (2); the respondent did not answer the usual residence question, nor the group quarters and other residence questions; and a conflicting household.
Recipients	EFUFY = NI and USURNOW = blank and BADINIT = 4
Donors	resolved EFU cases and BADINIT = 4
Comments	Since the recipients are people for whom the followup obtained little or no information, all resolved cases serve as the donor pool (except that we break down this group into conflicting household, non-conflicting household).

Variable descriptions are found in Sands (2002).

APPENDIX 4 Imputation Cells - Recipients and Donors

Cell	E2B P-sample
Description	Non-interview (2); the respondent did not answer the usual residence question, nor the group quarters and other residence questions; and a conflicting household.
Recipients	EFUFY = NI and USURNOW = blank and BADINIT = 4
Donors	resolved EFU cases and BADINIT = 4
Comments	Since the recipients are people for whom the followup obtained little or no information, all resolved cases serve as the donor pool (except that we break down this group into conflicting household, non-conflicting household with and without PFU).

Variable descriptions are found in Sands (2002).

APPENDIX 4 Imputation Cells - Recipients and Donors

Cell	E2C P-sample
Description	Non-interview (2); the respondent did not answer the usual residence question, nor the group quarters and other residence questions; and nonconflicting household that went to PFU.
Recipients	EFUFY = NI and USURNOW = blank and BADINIT ≠ 4 and BFUFLAG in (1, 2)
Donors	resolved EFU cases and BADINIT ≠ 4 and BFUFLAG in (1, 2)
Comments	Since the recipients are people for whom the followup obtained little or no information, all resolved cases serve as the donor pool (except that we break down this group into conflicting household, and non-conflicting household with and without PFU).

Cell	E2D P-sample
Description	Non-interview (2); the respondent did not answer the usual residence question, nor the group quarters and other residence questions; and nonconflicting household that did not go to PFU.
Recipients	EFUFY = NI and USURNOW = blank and BADINIT ≠ 4 and BFUFLAG not in (1, 2)
Donors	resolved EFU cases and BADINIT ≠ 4 and BFUFLAG not in (1, 2)
Comments	Since the recipients are people for whom the followup obtained little or no information, all resolved cases serve as the donor pool (except that we break down this group into conflicting household, and non-conflicting household with and without PFU).

Variable descriptions are found in Sands (2002).

APPENDIX 4 Imputation Cells - Recipients and Donors

Cell	EKR (P sample and E sample)
Description	Potential Fictitious
Recipients	EFUFY = KR
Donors	EFUFY = KR
Comments	Potential fictitious occur almost exclusively only when both the PFU and EFU indicate potential fictitious. These cases are resolved by the PFU (EFUFY = KR but the PFU is not KR).

Cell	PAD (P sample and E sample)
Description	The followed up person 'Lived elsewhere' or at an 'other residence', but the address is not given
Recipients	PFUFY = AD or (PFUFY = OR and AOTHRES =1 and OTHADD =1 and SOURCE = 1, 2)
Donors	AOTHRES = 1 and OTHADD = 2 and (AMOSTIME = '2' or AMOREDY ='2' or AMOREWK = '2' or AMOREMO = '2' or ASPENDCD = '2')
Comments	Both recipients and donors are people who indicate they had another residence and that they lived at that other residence on Census Day. They are distinguished by whether or not the CODE_FINAL is unresolved or resolved.

Variable descriptions are found in Sands (2002).

APPENDIX 4 Imputation Cells - Recipients and Donors

Cell	PMO (P sample and E sample)
Description	Moved in after Census Day or Moved out before Census Day, but no mover address provided
Recipients	PFUFY = MO, MI, MICD or MP
Donors	PFUFY = MO, MI
Comments	When a person indicated they moved, they received a mover why code (MO, MI) whether they were resolved, gave a good mover address, or were unresolved, did not give a good mover address. A small number of MICD are also put into this recipient pool. These codes are assigned to people who moved in on census day and the person is unresolved if the mover address is unresolved. In the PFU MP indicates the followup person was not a census day resident, either moving out before census day, or moving in after census day, and that the mover address was in unresolved. Thus the PMO group is an appropriate donor pool. Note that MP means something different when applied to the EFU.

Cell	PN2 (P sample and E sample)
Description	Respondent indicated the followed up person 'Never lived here' at the sample address, but did not provide the census day address
Recipients	(PFUFY = NI and ALIVECD = 2) or PFUFY = NL, DL
Donors	ALIVECD = 2
Comments	The NL and DL why codes were applied only to unresolved cases, not to resolved cases. Non-interviews, NI, are the Why code assigned when the key questions are not answered. However, they were also assigned when the PFU said they followup person did not live here. These NI cases have some information and thus are separated out to form their own group.

Variable descriptions are found in Sands (2002).

APPENDIX 4 Imputation Cells - Recipients and Donors

Cell	POR (P sample and E sample)
Description	The followed up person had an 'Other residence', but did not indicate whether sample address or the other residence was the census day residency
Recipients	PFUFY = OC, OH, OJ, OM, OP, OV, OR, OW, OS, HO, or any with a suffix of 1 or 2 or ((PFUFY = OR and OTHADD = 2) or (PFUFY = OR and AOTHRES = 1 and OTHADD = 1 and Source ne 1, 2)))
Donors	AOTHRES = 1
Comments	The donors are followed up people who indicated they had another residence and specified whether the other residence or the sample residence was the census day residency. During the PFU/EFU Review, PFU cases were assigned the Why code OR for cases that would be assigned ORDK in the Revision coding. The second line removes these cases from the OR group.

Variable descriptions are found in Sands (2002).

APPENDIX 4 Imputation Cells - Recipients and Donors

Cell	POK E-sample
Description	The respondent indicated the followed up person 'Lived here' at the sample residence, but did not answer the other residence question.
Recipients	(PFUFY = OR and AOTHRES ne 1 and OTHADD ne 2) or PFUFY = ORDK , JBPb, AL, GQ, RC, DO, DU, MX or LH
Donors	ALIVECD = 1 and AGROUP = 2
Comments	On Recipients: [1] JBPb is a why code that indicates the respondent did not answer the other residence questions about living away for a job. It represents a subset of the cases for which we had expected to have a why code of ORDK. [2] AL is a Why code indicating the people were really at an assisted living home. They were coded unresolved because not all questions were asked, such as other residence questions. Thus they are like ORDK's. For similar reasons GQ, RC, DO, DU are also put with ORDK. [3] MX refers to an apartment mixup They were coded unresolved because not all questions were asked, such as other residence questions. On Donors: The recipients are people who are unresolved because they didn't answer the other residence questions, though there was no evidence of erroneous enumeration or non-resident up to that point of the other residence questions. Thus the donors are people who are resolved and who had no evidence of erroneous enumeration or non-resident up to the point of the other residence questions.

Variable descriptions are found in Sands (2002).

APPENDIX 4 Imputation Cells - Recipients and Donors

Cell	PNI E-sample
Description	Non-interview (1); the respondent answered the census day residency question on the PFU (4.a), but did not answer the group quarters and other residence question.
Recipients	(PFUFY = NI and ALIVECD =1) or PFUFY = GO, TE or NF or other
Donors	ALIVECD = 1
Comments	The recipients are people for whom the census day residency question was answered 'yes', but for whom the group quarters and other residency questions were not answered. Thus we pick as donors all resolved people who answered the usual residence question. NF, GO and TE are put in with the non-interviews since they convey no interview information pertaining to census day enumeration status or residency. EFUFY = NI and USURNOW not blank, or EFUFY = GO, NF or TE or other The rubric 'other' consists of EFUFY of DE, GE, NT, X, FE, GC, NH, 4B, JBP3, JC3, VCb, DF, DP, FP, and DB. And BADINIT ≠ 4.

Cell	PN3 E-sample and P-sample
Description	Non-interview (1); the respondent answered don't know or refused to the census day residency question on the PFU (4.a), and did not answer the group quarters and other residence question.
Recipients	PFUFY = NI and ALIVECD = 3, 4
Donors	ALIVECD = 3, 4
Comments	The recipients are people for whom the census day residency question was answered 'yes, lived here on census day', but for whom the group quarters and other residency questions were not answered. Thus we pick as donors all resolved people who answered the usual residence question.

Variable descriptions are found in Sands (2002).

APPENDIX 4 Imputation Cells - Recipients and Donors

Cell	PN4 E-sample and P-sample
Description	Non-resident (1); the respondent answered don't know or refused to the census day residency question on the PFU (4.a), and did not answer the group quarters and other residence question.
Recipients	PFUFY = NI and ALIVECD = 0 or ' '
Donors	ALIVECD = 0 or ' '
Comments	The recipients are people for whom the census day residency question was answered, 'don't know' or 'refused', but for whom the group quarters and other residency questions were not answered. Thus we pick as donors all resolved people who answered the usual residence question don't know or refused.

Cell	PKR (P sample and E sample)
Description	Potential Fictitious
Recipients	PFUFY = KR
Donors	PFUFY = KR
Comments	Potential fictitious occur almost exclusively when both the PFU and EFU indicate potential fictitious, in which case we assign the recipient to the EFU cell, EKR.

Variable descriptions are found in Sands (2002).

APPENDIX 5 Residence Probability for Insufficient Information and Possible Match

	Group Description	Description of Unresolved Group or Cell (Recipients)	Description of Donor Pool
	For all recipient and donor cases - in the A.C.E. Revision II sample: REACE = 1, 2[1]	BESTF2 is: KI, KP or P	BESTF2 is either: N, M, MR, NR, FP, GP, DP, MN, NC, NN, NL, NP
	Insufficient information for match and followup	BESTF2 = KI, KP	BFUFLAG = 1 or 2
PKC	Conflicting Household	BADINIT[2]= 4	BADINIT = 4
PKH	Matched Housing Unit	BADINIT = 1	BADINIT = 1
PKN	Non-Matched Housing Unit	BADINIT = 2	BADINIT = 2
	Possible matches		
PPM	Possible matches	BESTF2 = P	BBFUGP = 2

[1]KI and KP's will generally have REACE = 2, as any cases with production codes of KI or KP did not go to EFU. However, there were a small number of new KP cases coded based on the results of EFU. P's did go to PFU, thus those in sample will have REACE = 1.

[2]Variable descriptions are found in Sands (2002).

APPENDIX 6 Match Probability for Insufficient Information and Possible Match

	Group Description	Description of Unresolved Group or Cell (Receipients)	Description of Donor Pool
	For all recipient and donor cases - in the A.C.E. Revision II sample: REACE = 1, 2[1]	CODE_FINAL is: KI, KP or P	CODE_FINAL is either: M, MR, MU, NR, NC, NP, NU
MCF	In a conflicting household	BADINIT[2] = 4	BADINIT = 4
MHN	HU Match, Non-mover	BADINIT = 1 and RSC_FINAL = N	BADINIT = 1 and RSC_FINAL = N
MHU	HU Match, Mover	BADINIT = 1 and RSC_FINAL = O, U	BADINIT = 1 and RSC_FINAL = O, U
MNN	HU Non-match, Non-mover	BADINIT = 2 and RSC_FINAL = N	BADINIT = 2 and RSC_FINAL = N
MMN	HU Non-match, Mover	BADINIT = 2 and RSC_FINAL = O, U	BADINIT = 2 and RSC_FINAL = O, U

[1]KI and KP's will generally have REACE = 2, as any cases with production codes of KI or KP did not go to EFU. However, there were a small number of new KP cases coded based on the results of EFU. P's did go to PFU, thus those in sample will have REACE = 1.

[2]Variable descriptions are found in Sands (2002).