Census 2000 Evaluation I.3
September 29, 2003

Evaluation of Nonresponse Followup - Mover Probe

FINAL REPORT

This evaluation reports the results of research and analysis undertaken by the U.S. Census Bureau. It is part of a broad program, the Census 2000 Testing, Experimentation, and Evaluation (TXE) Program, designed to assess Census 2000 and to inform 2010 Census planning. Findings from the Census 2000 TXE Program reports are integrated into topic reports that provide context and background for broader interpretation of results.

Don Keathley
Planning, Research, and Evaluation Division

TABLE OF CONTENTS

TABLE OF CONTENTS i

LIST OF TABLES ii

EXECUTIVE SUMMARY iii

1. BACKGROUND 1
   1.1 1990 Census 1
   1.2 Census 2000 2

2. METHODOLOGY 3
   2.1 Input Files 3
   2.2 Non-ID Extract File 3
   2.3 DRF2 3

3. LIMITATIONS 4

4. RESULTS 4
   4.1 Introductory Questions, 'Mover' Check Box 4
   4.2 Matching Non-ID EQs with Marked 'Mover' Check Boxes to the DRF2 5
   4.3 Continuation Forms, No MAFID 6
   4.4 Multiple EQs 6
   4.5 Potential Mover Probe-only EQs Representing NR Inmover Households that Census 2000 would not have otherwise Enumerated 7
   4.6 Some Comparisons 9

5. CONCLUSIONS AND RECOMMENDATIONS 9

REFERENCES 10

APPENDICES 12
   Appendix A 12
   Appendix B 13
   Appendix C 14
   Appendix D 17
   Appendix E 22
   Appendix F 23
   Appendix G 24

LIST OF TABLES

Table 1. Matching Results for Non-ID EQs 5

Table 2. Counts of Potential Mover Probe-only EQs, by RSTATUS 7

Table 3. Final Counts of Potential Mover Probe-only EQs, by RSTATUS 8

Table C.1. Non-ID File Extract Variables 14

Table C.2. DRF2 Variables 15

EXECUTIVE SUMMARY

THE MOVER PROBE

Census 2000 was made up of various data collection operations; some of them attempted to gather census-day information after April 1, 2000. Two of these operations were the nonresponse followup and coverage improvement followup. Nonresponse followup enumerators attempted to enumerate housing units in mailback areas from whom Census 2000 had not received a questionnaire. Coverage improvement followup enumerators attempted to do the same for mostly vacant and deleted units in mailout/mailback, update/leave, and urban update/leave areas. Enumerators used an enumerator questionnaire in both operations.

All enumerator questionnaires contained a mover probe. The mover probe allowed enumerators to identify households in housing units in the two followup operations that moved into the housing unit *after* April 1, 2000 and did not return a census questionnaire for their census-day addresses (*nonresponse inmover households*). Enumerators in both operations would then attempt to complete a separate enumerator questionnaire for every *nonresponse inmover household* for the address at which it lived on census day. These enumerator questionnaires represented the *nonresponse inmover households'* completed census questionnaires.

The purpose of this evaluation is to determine the usefulness of the mover probe.

MOVER PROBE-ENUMERATED HOUSEHOLDS THAT CENSUS 2000 WOULD NOT HAVE OTHERWISE ENUMERATED

In Census 2000, there were 105,480,101 occupied housing units in the United States (the 50 states and the District of Columbia). Of these, 22,850 would not have been enumerated by Census 2000 without the mover probe. This represents 0.02 percent of the total U.S. occupied housing unit count.

There were a total of 45,507,823 enumerator questionnaires from the two followup operations. The 22,850 enumerator questionnaires representing occupied housing units enumerated only by the mover probe represents 0.05 percent of all nonresponse followup and coverage improvement followup questionnaires.

There were 125,585 enumerator questionnaires that had information indicating that they represented *nonresponse inmover households*; only 18.19 percent (22,850) of these represented households that Census 2000 would not have otherwise enumerated without the mover probe.

RECOMMENDATIONS

We recommend a reevaluation of the mover probe in future census and test census operations. There are three reasons for our recommendation. First, even though the U.S. occupied housing

unit count increased by only 0.02 percent, it was still an increase. Using the mover probe improved the accuracy of the U.S. housing unit count (and other related statistics).

Second, the mover probe was a small part of the EQ - the cost of keeping it in future census operations might be negligible.

Finally, only 18.19 percent of all the enumerator questionnaires representing *nonresponse inmover households* were mover-probe only enumerations. One factor possibly contributing to this relatively low proportion is that enumerators may not have applied the mover probe correctly, thereby falsely identifying *nonresponse inmover households*. Enumerators in future census operations will have the benefit of using computerized questionnaires, e.g., the hand-held computer. If enumerators did apply the mover probe incorrectly, then computerized questionnaires might be able to minimize or eliminate this problem (through edit checks in the instrument).

Additionally, there was no way to link separate enumerator questionnaires for the same followup-operation housing unit - one enumerator questionnaire represented the followup-operation housing unit, the other represented the housing unit at which a *nonresponse inmover household* resided on census day. We recommend developing a way to link these questionnaires. This would allow future evaluations to compare responses between the questionnaires.

1. BACKGROUND

Both the 1990 and 2000 censuses used a mover probe. The mover probe allowed enumerators to identify households that:

- moved after census day (April 1)
- did not complete and return a census questionnaire for their census-day address.

Once enumerators identified these households, they would complete a census questionnaire for their census-day address.

1.1 1990 Census

Three of the data collection operations in the 1990 Census were the list/enumerate procedure (L/E), the nonresponse followup (NRFU), and the Vacant/Delete/Movers Check. L/E enumerators visited housing units in very rural areas; they compiled basic address lists and completed census questionnaires (personal visit interviews) for these housing units. The L/E took place on and around census day (April 1, 1990).

NRFU enumerators visited housing units that had received but not returned mailback census questionnaires. NRFU enumerators attempted to determine the 1990 census-day status and complete a census questionnaire for these housing units. The NRFU was a post-census day operation.

L/E and NRFU housing units with *vacant* and *deleted* 1990 census-day classifications became part of the Vacant/Delete/Movers Check. Vacant/Delete/Movers Check enumerators visited these housing units to verify their L/E and NRFU *vacant* and *deleted* classifications. There were some housing units whose Vacant/Delete/Movers Check classifications did not match their respective L/E or NRFU classifications. Among these were housing units whose Vacant/Delete/Movers Check classification was *occupied*.

Vacant/Delete/Movers Check enumerators asked the occupants (the household) of the *occupied* Vacant/Delete/Movers Check housing units if they had moved into the housing unit after April 1, 1990. If they had, then the enumerators asked the household if it had completed and returned a census questionnaire for the address at which they resided on census day. If they replied ‘no’, then the enumerator completed a census questionnaire for their census-day address.

See U.S. Bureau of the Census (1993) for more details on the 1990 census data collection operations.

1.2 Census 2000

Two data collection operations in Census 2000 were the NRFU and coverage improvement followup (CIFU). Enumerators in NRFU attempted to obtain completed Census 2000 forms from households in mailback areas that did not respond by mail, through the internet, or a telephone questionnaire assistance operator (U.S. Bureau of the Census, (2002a)). CIFU enumerators attempted to improve Census 2000 coverage of housing units (mostly vacants and delete units) in mailout/mailback, update/leave, and urban update/leave areas (U.S. Bureau of the Census, (2003)). Enumerators used enumerator questionnaires (EQ) to collect data in both operations.

Both operations took place after April 1, 2000 (census day). This meant that a household occupying a housing unit during NRFU or CIFU could be different from the same housing unit's census-day household. A series of introductory questions on the EQ allowed enumerators to identify these households - this series of questions was the first half of the mover probe (Appendix A).

If the household for a NRFU or CIFU housing unit was different from its census-day household (an *inmover household*), then enumerators would ask the *inmover household* if it had completed and returned a census questionnaire for its census-day address; this question was the second half of the mover probe (card K - Appendix B). If the *inmover household's* response was 'no' or 'could not remember' (a *nonresponse (NR) inmover household*), then enumerators filled out a separate EQ for the census-day address at which the *NR inmover household* resided. This EQ represented the *NR inmover household's* completed Census 2000 form. Census 2000 could then include the *NR inmover households'* data in any relevant census statistics (e.g., national occupied housing unit count).

NRFU and CIFU enumerators had a labeled EQ for every housing unit in NRFU and CIFU, respectively. A labeled EQ contained address information for a given NRFU or CIFU housing unit. Enumerators completed labeled EQs for the housing units whose addresses appeared on the labels. These are the EQs that enumerators used in applying the mover probe.

NRFU and CIFU enumerators also had unlabeled EQs. Unlabeled EQs contained no address information (they were blank forms). These are the EQs that enumerators used when they collected data for a *NR inmover household's* census-day address.

This paper evaluates the usefulness of the Census 2000 mover probe: how well did it do in picking up housing units and households that Census 2000 would not have otherwise enumerated? We used EQs from the United States only in this evaluation (we excluded EQs from Puerto Rico).

2. METHODOLOGY

This section describes how we arrived at the numbers in Tables 1 and 2 in section 4. Section 2.1 lists the input files; sections 2.2 and 2.3 describe the files and list the variables we used from each file. Appendix C describes the variables in detail; Appendix D describes how we used the variables to obtain the numbers in the tables.

2.1 Input Files

We used two data files in our analysis:

- Non-ID Extract File
- Decennial Response File - Stage 2 (DRF2)

2.2 Non-ID Extract File

The non-ID extract file is a hierarchal file - it contains address-level, household-level (within an address), and person-level (within a household) data for all unlabeled EQs, only. We used both address- and return-level data in our analysis. Decennial Systems and Contracts Management Office (DSCMO) created the non-ID file. See Table C.1. for details on the variables.

To obtain the numbers in Table 1, we used the variables:

RT	⇛ record level - subsetting variable
PROCID	⇛ processing ID - matching variable
MAFID	⇛ subsetting variable
FCUISF	⇛ item F ('mover' check box indicator) - subsetting variable

For Tables 2 and 3, we used:

RT	⇛ record level - subsetting variable
MAFID	⇛ matching variable
FLAG	⇛ match/geocode indicator for an address - subsetting variable
FCUISF	⇛ item F ('mover' check box indicator) - subsetting variable

2.3 DRF2

The DRF2 is split into three sets of files. One set of files contains housing unit-level records; the second set contains census questionnaire (return)-level records; the third set contains person-level records. There is one file for each state and one Puerto Rico file in each set. These files contain data from all returns from all Census 2000 data collection operations. DSCMO created the DRF2 files. We used return-level records, only. See Table C.2 for details on the variables.

For Table 1, we used:

RCPID	⇛	processing ID - matching variable
RUID	⇛	MAFID - subsetting variable
RISMOV	⇛	item F ('mover' check box indicator) - subsetting variable
RSOURCE	⇛	source of return - analysis variable

For Tables 2 and 3, we used:

RUID	⇛	MAFID - matching variable
RPRSTAT	⇛	return and PSA household status - analysis variable
RISMOV	⇛	item F ('mover' check box indicator) - subsetting variable
RSTATUS	⇛	status of return - analysis variable

3. LIMITATIONS

There was no way to match labeled EQs to unlabeled EQs for the same NRFU/CIFU housing units. For example, we could not match on address - the only address on a labeled EQ was for the NRFU/CIFU housing unit while the only address on an unlabeled EQ was for the *NR inmover household's* census day residence. Therefore, we could not do any analyses that required matching the labeled and unlabeled EQs, e.g., we were unable to verify if an unlabeled EQ had a matching labeled EQ that indicated that the NRFU/CIFU housing unit contained an *inmover household*.

4. RESULTS

Section 4.1 provides an introduction, describing the introductory questions and 'mover' check box on the EQ in some detail. Sections 4.2 through 4.4 show a series of steps we took to arrive at our final tallies. Section 4.5 shows the number of households that the mover probe picked up that Census 2000 would have otherwise missed. Section 4.6 shows some comparisons.

4.1 Introductory Questions, 'Mover' Check Box

Enumerators applied EQs in NRFU and CIFU. Enumerators completed labeled EQs for the households living in NRFU and CIFU housing units; they completed unlabeled EQs for households living in NRFU and CIFU housing units that both:

- lived in a different housing unit on census day and
- did not return a completed census questionnaire for their census-day address.

Households meeting both of these conditions were *non-response (NR) inmover households*. Unlabeled EQs contain census information for these households for the housing unit at which they lived on census day.

Every EQ contained the mover probe. Four introductory questions made up part of the mover probe (Appendix A). Responses to these questions indicated whether a household moved into a NRFU/CIFU housing unit after April 1, 2000. Households that responded 'Yes' to question S1 and 'No' to question S2 were *inmover households*. These questions were applicable to labeled EQs, only.

Every EQ also had a 'mover' check box (item F in Appendix E). A marked box indicated that the household was a *NR inmover household*. The 'mover' check box was applicable to unlabeled EQs, only.

4.2 Matching Non-ID EQs with Marked 'Mover' Check Boxes to the DRF2

The non-ID extract file contained records for all unlabeled EQs, only. We matched all of the non-ID extract EQs with marked 'mover' check boxes (*non-ID mover EQ*) to the DRF2; we matched on processing ID. Table 1 shows the results of the matching:

Table 1: Matching Results for *Non-ID Mover EQs*

Matching Category	Description	Frequency
1	the *non-ID mover EQ* matched to a DRF2 return (EQ) with a marked 'mover' check box (*DRF2 mover EQ*)	89,779
2	the *non-ID mover EQ* did not match to a *DRF2 mover EQs* but did match to a DRF2 return with a blank 'mover' check box (*DRF2 non-mover return*)	0
3	the *non-ID mover EQ* did not match to any DRF2 return.	35,806
Total	All *non-ID mover EQs*	125,585

We obtained the counts in Table 1 in two steps. First, we matched all *non-ID mover EQs* to DRF2 returns with a marked 'mover' check box (*DRF2 mover EQ*) only; category 1 shows the counts of the *non-ID mover EQs* that matched. Then we matched all *non-ID mover EQs* that did not match to any *DRF2 mover EQ* to all other returns on the DRF2 (DRF2 returns with a blank 'mover' check box). Categories 2 and 3 show the results of this matching. See Appendix D for the operational details.

Non-ID mover EQs in categories 1 and 2 represent *NR inmover households* that Census 2000 would potentially not have otherwise enumerated without the mover probe (*potential mover probe-only EQs*). We considered category 2 EQs to be *potential mover probe-only EQs* because we assumed that the non-ID file 'mover' check box code was the correct code, i.e., their matching DRF2 returns had erroneously blank 'mover' check boxes.
Non-ID mover EQs in category 3 represent households that Census 2000 dropped from further processing (most of these EQs represented households that Census 2000 was unable to geocode, making them ineligible for further processing). Therefore, we considered category 3 EQs as not being *potential mover probe-only EQs*.

Table 1 shows 89,779 *potential mover probe-only EQs*, all from category 1. This result meant that all *non-ID mover EQs* matching to the DRF2 were consistent in having marked 'mover' check boxes on both files. The 35,806 *non-ID mover EQs* in category 3 were not *potential mover probe-only EQs*, so we dropped them from further consideration.

4.3 Continuation Forms, No MAFID

Of the 89,779 category 1 *potential mover probe-only EQs* in Table 1, 75 of them matched to a DRF2 EQ whose return was a continuation form (RSOURCE=37), i.e., an extension to a questionnaire. Since these EQs were just extensions and not the base forms themselves, we considered them to not be *potential mover probe-only EQs* - we dropped these 75 EQs from further consideration.

Additionally, 471 of the category 1 *potential mover probe-only EQs* in Table 1 had a FLAG value of 'no MAFID' (see Table C.1 in Appendix C for the definition of MAFID). These EQs represented *NR inmover households* that had no geocoding and should not have gone through any further census processing, i.e., these EQs should not have made it on to the DRF2. This indicated that these 471 *non-ID mover EQs* were probably category 3 EQs. In turn, this meant they were not *potential mover probe-only EQs* - we dropped these 471 EQs from further consideration, also.

After subtracting the continuation form and 'no MAFID' EQs from category 1, we were left with 89,233 *potential mover probe-only EQs*.

4.4 Multiple EQs

Two or more of the 89,233 *potential mover probe-only EQs* from section 4.2 could (and did) share the same MAFID, i.e., they represented the same *NR inmover household* - we wanted to count only one EQ per household. To do this, we kept the *potential mover probe-only EQ* for a *NR inmover household* that had the lowest RPRSTAT value (the RPRSTAT value indicates the household's return and PSA household status - see Appendix F for a summary of the PSA). The lowest RPRSTAT value for a household represents its final PSA status (see Table C.2. in Appendix C for variable definitions).

There were 274 *NR inmover households* with multiple *potential mover probe-only EQs* (553 total EQs). We selected the EQ with the lowest RPRSTAT to represent each of these 274 households; we dropped the remaining 279 EQs from further consideration. This left us with 88,954 *potential mover probe-only EQs*

4.5 *Potential mover probe-only EQs* Representing *NR Inmover Households* that Census 2000 would not have otherwise Enumerated

We needed to determine which of the 88,954 *potential mover probe-only EQs* from section 4.4 represented households that were indeed missed in Census 2000. To do this, we had to match these remaining *potential mover probe-only EQs* to all DRF2 returns with a blank 'mover' check box (*DRF2 non-mover return*). *Potential mover probe-only EQs* matching to one or more *DRF2 non-mover returns* represented households that Census 2000 enumerated using other data collection operations, i.e., they did not represent *NR inmover households*. Of the remaining 88,954 *potential mover probe-only EQs*, 64,872 matched to at least one *DRF2 non-mover return*. We dropped these EQs from further consideration.

This left us with 24,082 *potential mover probe-only EQs* (88,954 minus 64,872). To determine which of these EQs were mover-probe only enumerations, we needed the DRF2 variable RSTATUS; RSTATUS indicates the status of the housing unit (address) in which an EQ's household lived on census day. Table 2 shows this distribution.

Table 2: Counts of *Potential mover probe-only EQs*, by RSTATUS

Return Status (RSTATUS) (1)	Total *Potential mover probe-only EQs* (2)	*Potential mover probe-only EQs* that were valid enumerations (3)	Proportion of Total *Potential mover probe-only EQs* (4)
Occupied	22,437	22,404	0.999
Vacant	1,111	1,104	0.994
Delete	54	0	0.000
Undetermined (occupied, vacant)	469	436	0.930
Undetermined (occupied, vacant, delete)	11	10	0.909
Total	24,082	23,954	0.995

Counts in column (2) are for all of the remaining *potential mover probe-only EQs*. Counts in column (3) are for *potential mover probe-only EQs* that are valid enumerations. To be a valid

enumeration, a *NR inmover household* had to have a *potential EQ* that was either the basic or 'other' return (EQ) for the primary selection algorithm household; the RPRSTAT variable contained this information. *NR inmover households* with RPRSTAT values of 1 (basic return) or 2 ('other' return) were valid enumerations.

Table 2 shows non-zero counts in the 'Vacant' row. By definition, any EQ with a marked 'check' box represented a *NR inmover household* at its census-day address, therefore meaning an occupied census-day address. This implicitly meant that the census-day address was not vacant on census day. In turn, this indicated that there should have been zero counts in the 'Vacant' row for both columns (2) and (3). In light of this, we assumed that the EQs that made up the counts in the 'Vacant' row were not *potential mover probe-only EQs* - we dropped these EQs from further consideration.

Table 2 shows 54 EQs in the 'Delete' row for column (2). EQs in Table 2 represent households that Census 2000 did not enumerate in any other data collection operation and that contained the household's final PSA household status. Therefore, EQs in Table 2 represent their respective household's final census questionnaire. In turn, this meant there should have been zero counts in the 'Delete' row. We assumed that the 54 EQs in this row were not *potential mover probe-only EQs.* and dropped them from further consideration.

The previous discussion implies that all of the *potential mover probe-only EQs* in Table 2 should have had an RSTATUS of 'occupied'. By extension, this means that all of the EQs in the two 'Undetermined' rows should represent occupied census-day addresses, as well. When we combined this with the high proportion of 'Occupied,' 'Vacant,' and 'Delete' EQs in the 'Occupied' row (22,437 of 23,592 EQs in column (2) in the 'Occupied,' 'Vacant,' and 'Delete' rows in Table 2), we assumed the 'Undetermined' EQs represented occupied census-day housing units.

When we drop the 'Vacant' and 'Delete' rows and merge the two 'Undetermined' rows in Table 2, we obtain Table 3.

Table 3: Final Counts of *Potential mover probe-only EQs*, by RSTATUS

Return Status (RSTATUS) (1)	**Total *Potential mover probe-only EQs* (2)**	***Potential mover probe-only EQs* that were valid enumerations (3)**	**Proportion of Total *Potential mover probe-only EQs* (4)**
Occupied	22,437	22,404	0.999
Undetermined - all	480	446	0.929
Total	22,917	22,850	0.997

EQs making up the valid enumeration counts in Table 3 (column (3)) represent the *NR inmover households* that Census 2000 was able to enumerate using the mover probe, only. Hence, the mover probe was responsible for enumerating 22,850 households that Census 2000 would not have otherwise picked up.

4.6 Some Comparisons

The results in section 4.5 showed that there were 22,850 households (at their census-day addresses) that Census 2000 would not have enumerated without the mover probe. These households represent:

- 18.19 percent of all DRF2 EQs with a marked 'mover' check box (125,585 EQs)
- 0.05 percent of all NRFU and CIFU EQs (45,507,823 EQs)

The internet (American Fact Finder) shows 105,480,101 occupied housing units in the United States on April 1, 2000; the mover probe accounted for 22,850 of them. This means the mover probe accounted for a national occupied housing unit population increase of 0.02 percent (from 105,457,251 to 105,480,101 housing units).

5. CONCLUSIONS AND RECOMMENDATIONS

The mover probe had a limited usefulness [1]. We were able to enumerate 22,850 households (at their census-day addresses) with the mover probe that Census 2000 would not have otherwise enumerated. The impact that these households had on the total United Sates occupied housing unit count, however, was small. Their inclusion increased the U.S. occupied housing unit count by 0.02 percent (from 105,457,251 to 105,480,101 housing units). On the other hand, the mover probe was not a complete loss: Census 2000 would not have enumerated these housing units without it.

The mover probe's limited usefulness also shows in the proportion of NRFU and CIFU housing units in which *NR inmover households* resided that Census 2000 would not have otherwise enumerated; the 22,850 figure above represents only 0.05 percent of all NRFU and CIFU housing units (1 in every 2,000 NRFU/CIFU housing units).

Finally, only 18.19 percent of all *non-ID mover EQs* represented households that Census 2000 would not have otherwise enumerated without the mover probe. Most of the remaining *non-ID mover EQs* represented *mover households* that Census 2000 could not geocode or that Census 2000 enumerated in other data collection operations.

Because of the above, our initial recommendation was to drop the mover probe from future

---

[1] We confined our analysis to the national level. Analyses at smaller geographic levels might have revealed larger impacts on the corresponding occupied housing unit counts.

census operations. However, the mover probe was a small part of the EQ - the cost of keeping it in future census operations might be negligible. Additionally, the mover probe was successful in enumerating 22,850 households that Census 2000 would not have otherwise counted. Finally, part of the reason relatively few *non-ID mover EQs* represented mover probe-only households is that enumerators may not have applied the mover probe correctly [2]. Enumerators in future census operations will use computerized questionnaires, e.g., the hand-held computer. Instruments such as this could reduce the number of mover-probe application errors.

Because of these three factors, we recommend keeping the mover probe, making it a part of computerized questionnaires. This recommendation assumes that both the cost of keeping the mover probe is negligible and that enumerators in future operations will use automated questionnaires. We might want to test it in 2010 census testing. Then we can re-evaluate its effectiveness and efficiency.

If the mover probe does become part of future census operations, then we recommend developing a way of linking *NR inmover household* EQs (the unlabeled EQs) with the labeled EQs for the same NRFU / CIFU housing unit. Information on the labeled EQs might shed some light as to why only 18.19 percent of all *non-ID inmover EQs* represented households that Census 2000 would not have counted without the mover probe.

REFERENCES

American Fact Finder - see Appendix G

U.S. Bureau of the Census (2002a), *Nonresponse Followup for Census 2000,* Decennial Statistics Studies Division, July 25, 2002

U.S. Bureau of the Census (2002b), *Analysis of the Primary Selection Algorithm,* Decennial Statistics Studies Division, November 12, 2002

U.S. Bureau of the Census (2003), *Coverage Improvement Followup,* Decennial Statistics Studies Division, May 9, 2003

---

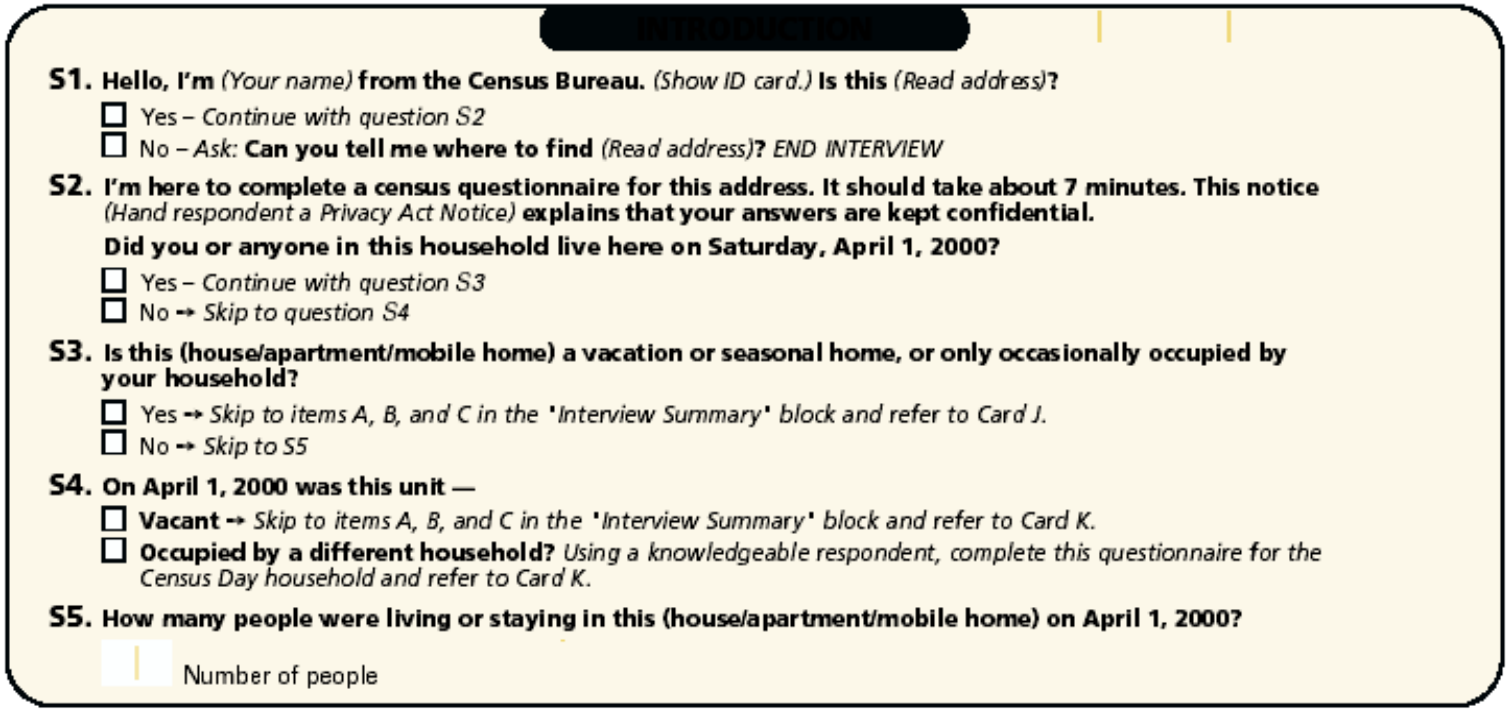
[2] Other possible contributors include:

- respondents saying they had not returned or did not remember returning a census questionnaire when, in fact, they had, resulting in multiple returns for their census-day address.

- respondents providing erroneous address information for where they lived on census day, making it impossible for the Census Bureau to geocode the address that they provided.

U.S. Bureau of the Census (1993), 1990 Census of Population and Housing, Evaluation and Research Reports: Programs to Improve Coverage in the 1990 Census , 1990 CPH-E-3.

Appendix A: Introduction Questions on the EQ (first half of the Mover Probe)



INTRODUCTION

S1. **Hello, I'm** *(Your name)* **from the Census Bureau.** *(Show ID card.)* **Is this** *(Read address)*?

- ☐ Yes – *Continue with question S2*
- ☐ No – *Ask:* **Can you tell me where to find** *(Read address)*? *END INTERVIEW*

S2. **I'm here to complete a census questionnaire for this address. It should take about 7 minutes. This notice** *(Hand respondent a Privacy Act Notice)* **explains that your answers are kept confidential.**

Did you or anyone in this household live here on Saturday, April 1, 2000?

- ☐ Yes – *Continue with question S3*
- ☐ No → *Skip to question S4*

S3. **Is this (house/apartment/mobile home) a vacation or seasonal home, or only occasionally occupied by your household?**

- ☐ Yes → *Skip to items A, B, and C in the "Interview Summary" block and refer to Card J.*
- ☐ No → *Skip to S5*

S4. **On April 1, 2000 was this unit —**

- ☐ **Vacant** → *Skip to items A, B, and C in the "Interview Summary" block and refer to Card K.*
- ☐ **Occupied by a different household?** *Using a knowledgeable respondent, complete this questionnaire for the Census Day household and refer to Card K.*

S5. **How many people were living or staying in this (house/apartment/mobile home) on April 1, 2000?**

☐☐ Number of people

Appendix B: Card K (contains second half of the Mover Probe

United States Census 2000

U.S. Department of Commerce
Bureau of the Census

CARD K

IN-MOVER

An IN-MOVER situation occurs when you learn that the household occupying the housing unit moved in *after* Census Day – April 1, 2000. For an "in-mover" situation, do the following:

STEP 1: Determine the status of the housing unit as of Census Day.

- If the unit was vacant, complete the labeled D-1(E) or D-2(E) Enumerator Questionnaire for a vacant unit.
- If the unit was occupied by a different household, locate a knowledgeable respondent, such as a neighbor, and complete as much information on the labeled Enumerator Questionnaire as possible for the Census Day residents.

STEP 2: Ask the in-mover the following question:

We want to make sure that your household is counted in the census. **Did you, or anyone else in your household, complete a census questionnaire from your previous address?**

- If the in-mover did complete a questionnaire for his/her previous address thank the person and do nothing further.
- If the in-mover did not complete a census questionnaire or doesn't know if one was completed for his/her previous address, conduct an interview for the in-mover's household using a blank, unlabeled short-form Enumerator Questionnaire.
  - ▸ Conduct the interview as if you were interviewing the person at his/her previous address.
  - ▸ Enter the in-mover's previous address in the address label area on the front of the questionnaire. If the address does not include a house number and street name, or complete rural route and box number, try to obtain as much address information as you can. **DO NOT ENTER THE NONRESPONSE FOLLOWUP ADDRESS ON THIS FORM.**
  - ▸ Enter an "X" in Item F. MOV of the Interview Summary.

NOTE: In this situation, you will have completed two questionnaires at the followup unit.

D-1 (F) (5-11-99)

Appendix C: Variable Descriptions

Table C.1: Non-ID Extract File Variables

Variable	Description	Length	Valid Codes
FCUISF	Interview Summary Item F (return-level variable)	1	non-blank = unlabeled EQ for a mover's census-day address blank = other
FLAG	Matching/Geocoding Flag (address-level variable)	1	blank = matched to an existing MAFID 1 = no MAFID 2 = matched to an existing MAFID, late add but not in field verification (FV) 3 = MAFID is for a group quarters 4 = in FV universe
MAFID	Census ID (address-level variable)	12	characters 1-2 = FIPS state code characters 3-5 = FIPS county code characters 6-12 = sequence ID
PROCID	Processing ID (address-level record (a) return-level record (r))	a - 12 r - 14	same as for RPCID on the DRF2, except that PROCID for address-level records contain only the first 12 characters of the ID
RT	Record Type (address- and return-level records)	1	0 = address-level record 1 = return-level record 2 = person-level record (not used)

Table C.2: DRF2 Variables

Variable	Description	Length	Valid Codes
RCPID	Capture Processing ID	14	blank = no special capture ID 60000000000000 - 98509999999996 = capture ID number
RISMOV	Interview Summary Item F	1	non-blank = unlabeled EQ for a mover's census-day address blank = other
RPRSTAT	Return and PSA Household Status	2	-1 = not computed 1 = basic return for primary PSA household 2 = other return for primary PSA household 3 = basic return for non-primary PSA household 4 = other return for non-primary PSA household 5 = redundant 6 = ineligible
RSOURCE	Source of Return (recode)	2	-1 = not computed 1-12 = paper mailback questionnaire 13-16 = paper enumerator questionnaire 17-21 = NRFU paper enumerator questionnaire 22-24 = CIFU paper enumerator questionnaire 25 = paper enumerator questionnaire from T-night 26-29 = paper questionnaire for UHE 30-36 = electronic form 37 = paper enumerator continuation form
RSTATUS	Status of Return	2	-1 = not computed 1 = occupied 2 = vacant 3 = delete 4 = undetermined (vacant, occupied) 5-6 = not used 7 = undetermined (vacant, occupied, delete)

Variable	Description	Length	Valid Codes
RUID	Unit ID Number (MAFID)	12	characters 1-2 = FIPS state code characters 3-5 = FIPS county code characters 6-12 = sequence ID

Appendix D: Methods for Creating Tables 1, 2, and 3

1. Table 1

This section describes how we obtained the values in Table 1, section 4.2 of the text.

1.1 Non-ID Extract File Subsetting

1.1.1 Record Levels

The non-ID extract file was a hierarchal file, containing address-level, return (EQ)-level, and person-level records for unlabeled EQs. We distinguished the various records using the variable RT, where

RT = 0 ⇛ address-level record
RT = 1 ⇛ EQ-level record
RT = 2 ⇛ person-level record

We needed just address-level and EQ-level records, so we kept non-ID file records with RT $\in$ (0,1) only. The result was a flat EQ-level file. We matched the address- and EQ-level records on processing ID (PROCID).

Note that PROCID is 14 characters in length for EQ-level records but only 12 characters in length for address-level records; we used the first twelve positions of the EQ-level PROCID in the matching.

1.1.2 Geography

We subsetted the file we created in 1.1.1 on FIPS state code, deleting all non-ID EQs with

FIPS state code = 72 (Puerto Rico)

FIPS state code is the first two characters of the MAFID variable (see Table C.1).

1.1.3 Inmovers

We subsetted the file we created in 1.1.2 on FCUISF, keeping only those non-ID EQs with

FCUISF ≠ blank (the value for FCUISF was non-blank)

EQs with FCUISF ≠ blank are inmover EQs. This was the *NON-ID INMOVER FILE*.

1.2 DRF2 Subsetting

1.2.1 Record Levels

There were three sets of DRF2 files - housing-unit level, return level, and person level. We needed just the return-level files. We concatenated all of the return-level files, creating one DRF2 return-level file. The only variables we kept are the ones shown in Table C.2.. This was the *DRF2 RETURN FILE*.

1.2.2 Geography

We subsetted the file we created in 1.2.1 on FIPS state code, deleting all non-ID EQs with

FIPS state code = 72 (Puerto Rico)

FIPS state code is the first two characters of the RUID variable (see Table C.2).

1.2.3 Inmovers

We subsetted the *DRF2 RETURN FILE* on RISMOV, keeping only those returns with:

RISMOV ≠ blank (the value for RISMOV was non-blank)

Returns with RISMOV ≠ blank are inmover returns. This was the *DRF2 INMOVER FILE*.

1.3 Merging the Inmover Files

We merged the *NON-ID INMOVER FILE* with the *DRF2 INMOVER FILE* - we matched the two files on processing ID, using

PROCID (from *NON-ID INMOVER FILE*) = RCPID (from *DRF2 INMOVER FILE*)

The resulting file contained <u>all</u> records from both input files (matches and non-matches). This was the *MERGED INMOVER FILE*.

1.4 Category 1 Frequency

The value in the 'Frequency' column for category 1 in Table 1 comes from the *MERGED INMOVER FILE* - it's a count of all the non-ID EQs that matched to one or more DRF2 records. We counted non-ID records matching to two or more DRF2 records only once. We output the EQs in category 1 to the *CATEGORY 1 FILE*.

1.5. Subsetting the *MERGED INMOVER FILE*

We subsetted the *MERGED INMOVER FILE*, keeping only those non-ID file records that did not match to a DRF2 record. This new file was the *SUBSETTED MERGED INMOVER FILE*.

1.6 Merging to the *DRF2 FILE*

We merged the *SUBSETTED MERGED INMOVER FILE* to the *DRF2 FILE* on processing ID, using

PROCID (from *SUBSETTED MERGED INMOVER FILE*) = RCPID (from *DRF2 FILE*)

The resulting file was the *FINAL MERGED FILE* - it contains <u>all</u> records from the *SUBSETTED MERGED INMOVER FILE* only, i.e., it did not contain any *DRF2 FILE* records that did not match to the *SUBSETTED MERGED INMOVER FILE*.

1.7 Frequencies for Categories 2 and 3

The value in the 'Frequency' column for category 2 in Table 1 comes from the *FINAL MERGED FILE* - it's a count of all the non-ID records that did not match a DRF2 inmover record (from the matching in section 1.3) but did match to one or more DRF2 non-inmover records (from the matching in section 1.6)

The value in the 'Frequency' column for category 3 in Table 1 comes from the *FINAL MERGED FILE* also - it's a count of all the non-ID records that did not match to any DRF2 record.

2. Table 2

This section describes how we arrived at the values in sections 4.3, 4.4, and 4.5 of the text, including Table 2 in section 4.5.

2.1 Subsetting Category 1 EQs

Before creating Table 2, we needed to subset the *CATEGORY 1 FILE*.

2.1.1 Delete Continuation Forms (in section 4.3 of text)

We checked the distribution of RSOURCE (questionnaire source - see Table C.2) for all *CATEGORY 1 FILE* EQs. The values for RSOURCE that we found on the file were:

RSOURCE $\in$ {17, 18, 19, 20, 21}	-	NRFU EQ
RSOURCE $\in$ {22, 23, 24}	-	CIFU EQ
RSOURCE = 37	-	continuation form

We deleted the continuation form EQs from the *CATEGORY 1 FILE*; the file reflecting this deletion was *CATEGORY 1 FILE, REVISION 1*.

2.1.2 Delete EQs with no MAFID (in section 4.3 of text)

Another distribution we checked was FLAG (matching, geocoding flag - see Table C.1) for the EQs on the *CATEGORY 1 FILE, REVISION 1*. The FLAG values we found for the EQs on the file were:

blank	matched to existing MAFID
1	No MAFID
2	Matched, late, not in FV
4	In FV universe

EQs with FLAG=1 should not have matched to the DRF2, i.e., they were not category 1 EQs. We deleted these EQs from the *CATEGORY 1 FILE, REVISION 1*; the new file reflecting this deletion was *CATEGORY 1 FILE, REVISION 2*.

2.1.3 Delete Multiple Potential mover probe-only EQs (in section 4.4 of text)

Two or more EQs on the *CATEGORY 1 FILE, REVISION 2* would sometimes represent the same household.(*multiple potential mover probe-only EQs*). We wanted to keep only one EQ per household. To determine which *multiple potential EQ* to retain, we needed the RPRSTAT variable (RPRSTAT indicates return and PSA household status for an EQ - see Table C.2). We kept the *multiple potential mover probe-only EQs* with the lowest RPRSTAT value for each household with more than one EQ - we deleted all of the remaining *multiple potential mover probe-only EQs* from the *CATEGORY 1 FILE, REVISION 2*. The new file reflecting this deletion was the *FINAL CATEGORY 1 FILE*.

2.2 Subsetting the *DRF2 RETURN FILE*

We subsetted the *DRF2 RETURN FILE* on RISMOV, keeping only those returns with:

RISMOV = blank (the value for RISMOV was blank)

Returns with RISMOV = blank are mover returns. This was the *DRF2 NON-INMOVER FILE*.

2.3 Merging the *FINAL CATEGORY 1 FILE* with the *DRF2 NON-INMOVER FILE*

We merged the *FINAL CATEGORY 1 FILE* to the *DRF2 NON-INMOVER FILE* on processing ID, using

PROCID (from *FINAL CATEGORY 1 FILE*) = RCPID (from *DRF2 NON-INMOVER FILE*)

The resulting file was the *MOVER–IN-MOVER FILE* - it contained all records from the *FINAL CATEGORY 1 FILE* only, i.e., it did not contain any *DRF2 NON-INMOVER FILE* records that did not match to the *FINAL CATEGORY 1 FILE*.

2.4 Subsetting the *MOVER–NON-INMOVER FILE* - Delete Matches (section 4.5)

We subsetted the *MOVER–IN-MOVER FILE*, keeping only those *FINAL CATEGORY 1 FILE* EQs that did not match to any *DRF2 NON-INMOVER FILE* record. This subsetted file was the *UNIQUE NON-ID INMOVER EQ FILE*.

2.5 Column (2) Counts (Table 2, section 4.5)

The counts in column (2) of Table 2 show the distribution of all EQs on the *UNIQUE NON-ID INMOVER EQ FILE* by the variable RSTATUS (return status) The RSTATUS categories are uncollapsed - see Table C.2 for the values for RSTATUS.

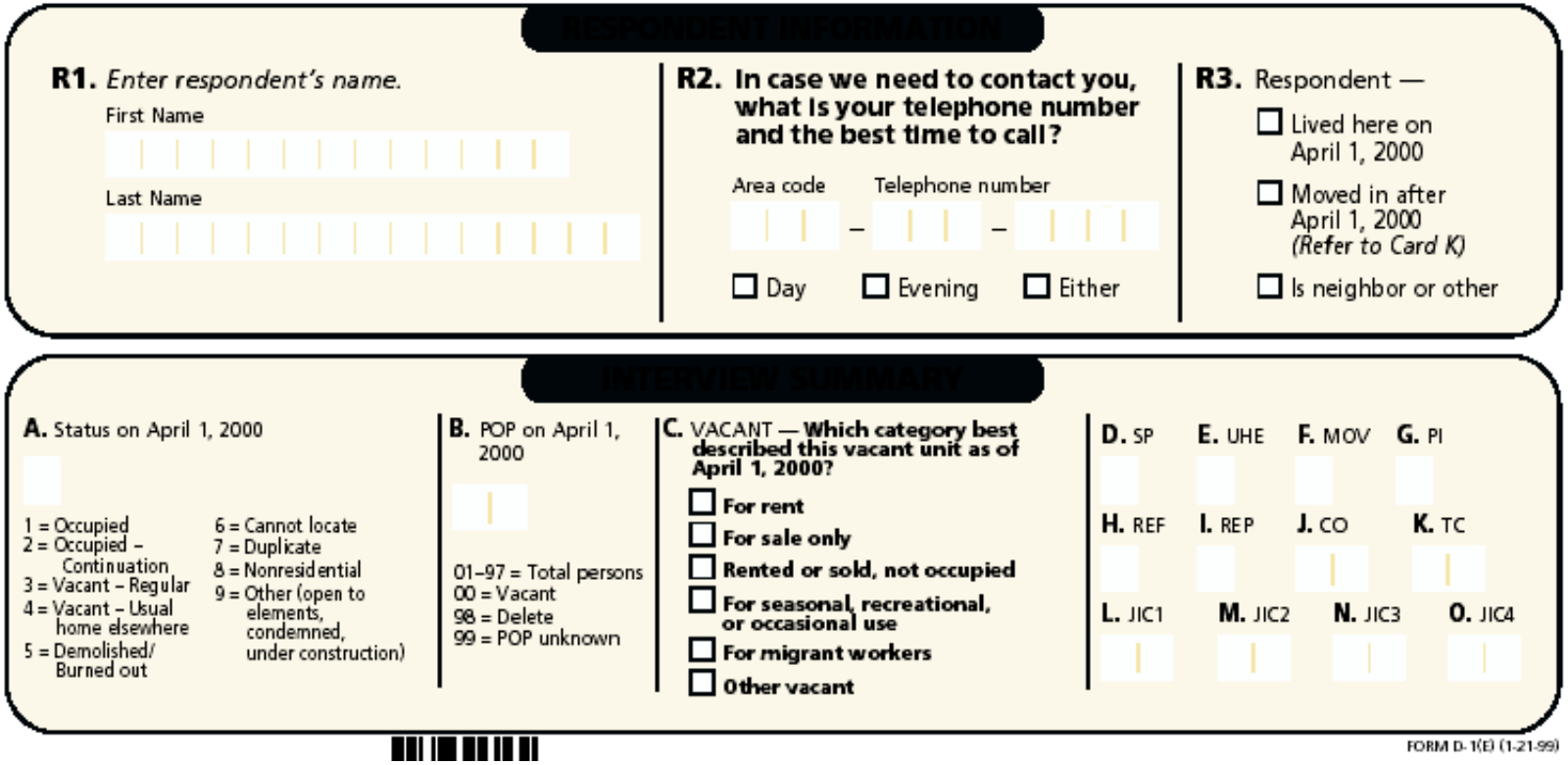
2.8 Column (3) Counts (Table 2, section 4.5)

The counts in column (3) of Table 2 show the distribution of EQs on the *UNIQUE NON-ID INMOVER EQ FILE* with RPRSTAT $\in$ (1,2), by RSTATUS. EQs with RPRSTAT $\in \{1,2\}$ are EQs that are either the basic or 'other' returns for the primary PSA household.

2. Table 3

Table 3 in section 4.5 is a subset of Table 2. It has the same columns as in Table 2. The rows are different, however - we deleted the 'Vacant' and 'Delete' rows and merged the two 'Undetermined' rows from Table 2.

Appendix E: 'Mover' Check Box (Box with the 'F. MOV' label)



RESPONDENT INFORMATION

R1. *Enter respondent's name.*

First Name

Last Name

R2. **In case we need to contact you, what is your telephone number and the best time to call?**

Area code – Telephone number – –

☐ Day ☐ Evening ☐ Either

R3. Respondent —

☐ Lived here on April 1, 2000

☐ Moved in after April 1, 2000 *(Refer to Card K)*

☐ Is neighbor or other

INTERVIEW SUMMARY

A. Status on April 1, 2000

1 = Occupied
2 = Occupied – Continuation
3 = Vacant – Regular
4 = Vacant – Usual home elsewhere
5 = Demolished/ Burned out
6 = Cannot locate
7 = Duplicate
8 = Nonresidential
9 = Other (open to elements, condemned, under construction)

B. POP on April 1, 2000

01–97 = Total persons
00 = Vacant
98 = Delete
99 = POP unknown

C. VACANT — **Which category best described this vacant unit as of April 1, 2000?**

☐ **For rent**
☐ **For sale only**
☐ **Rented or sold, not occupied**
☐ **For seasonal, recreational, or occasional use**
☐ **For migrant workers**
☐ **Other vacant**

D. SP **E.** UHE **F.** MOV **G.** PI

H. REF **I.** REP **J.** CO **K.** TC

L. JIC1 **M.** JIC2 **N.** JIC3 **O.** JIC4

FORM D-1(E) (1-21-99)

Appendix F: Purpose of the Primary Selection Algorithm (PSA)

The Primary Selection Algorithm (PSA) was applied to the response records associated with housing unit (HU) IDs. The purpose of the PSA was to select the housing unit return and person records that would be included on census files defined by subsequent processes.

More than one response to the census may be received for a given housing unit address. This occurs because there are several ways to respond to the census. A person may mail back the census form delivered to his home; he may be interviewed by a census enumerator; he may fill in a Be Counted Form and mail it in; he may fill out a form online and return it via the Internet; he may be enumerated at a group quarters (GQ) (e.g., a military base) but elect to be counted at a different address (i.e., GQ Usual Home Elsewhere (UHE)) that is a housing unit address. Each of these types of responses that arrive for the same housing unit address will create a return coded to the same Census ID. It is the job of the PSA to analyze these responses and select from among them the records that it deems most likely to represent the actual census household.

The preceding is from U.S. Bureau of the Census (2002b).

Appendix G: American Fact Finder

The occupied housing unit count shown in the report (105,480,101) comes from an American Fact Finder web site. The url for this web site is a temporary link, i.e., it will not send the reader to the page with the source table. To get to this table, go to http://www.census.gov/. From there:

1. Click on *United States Census 2000*
2. Click on *American Fact Finder*
3. Click on *2000 Summary File 1*
4. Click on *Detailed Tables*
5. Under "Choose a Selection Method", click on *List*

6. Highlight 'Nation' in the first window
7. Highlight 'United States' in the second window
8. Click on *Add* under the second window - this adds United States to the third window
9. Click on *Next*
10. Under "Choose table selection method", click on *show all tables*

11. Highlight table 'H3' in the first window
12. Click on *Add* under the first window - this adds table H3 to the second window
13. Click on *Show Result*

The resulting table shows the housing unit counts for the 50 states and District of Columbia on April 1, 2000

	United States
Total	115,904,641
Occupied	105,480,101
Vacant	10,424,540