

The Sources of Census Undercount: Findings from the 1986 Los Angeles Test Census

DAVID J. FEIN and KIRSTEN K. WEST¹

ABSTRACT

This paper presents results from a study of the causes of census undercount for a hard-to-enumerate, largely Hispanic urban area. A framework for organizing the causes of undercount is offered, and various hypotheses about these causes are tested. The approach is distinctive for its attempt to quantify the sources of undercount and isolate problems of unique importance by controlling for other problems statistically.

KEY WORDS: Census; Undercount; Coverage improvement; Post enumeration survey.

1. INTRODUCTION

In the last decade or two the need to better understand the causes of undercount in the U.S. census has become pressing. As the census has become an increasingly important tool in governing the nation, conducting business, and monitoring social change (Citro and Cohen 1985; Clogg *et al.* 1986), public concern about the quality of census data has intensified. Much of this concern has arisen because it is perceived, with good foundation, that net census undercount disproportionately affects the economically disadvantaged members of society (Citro and Cohen 1985, ch. 5; Ericksen 1983). Representatives of the disadvantaged believe that as a result their constituents are being denied a fair share of public funds and political representation (Choldin 1987).

Assuming that an acceptable method could be found, one solution to the problem would be to correct the census for the bias due to differential undercount. In the fall of 1987, however, the Department of Commerce decided not to adjust the 1990 census but instead to concentrate on achieving a more complete enumeration (Ortner 1987).

Improving census coverage implies a need to understand the causes of census undercount better than ever before. Many special coverage improvement programs were implemented in the 1980 census, and these may have contributed to the achievement of historically low levels of overall net coverage error. In spite of such efforts, wide socioeconomic coverage differentials have persisted. In response, the Census Bureau has embarked on a broad research program to identify the causes of undercount, concentrating on population subgroups that are especially difficult to enumerate.

This paper presents results from a study of the causes of census undercount in a hard-to-enumerate, largely Hispanic area in Los Angeles. The approach is distinctive for its attempt to quantify the sources of undercount and isolate problems of unique importance by controlling for other problems statistically.

Though the putative inequities mentioned above result from net census coverage error (omissions less erroneous enumerations), to keep the analysis manageable only census omissions are investigated here. Omissions in the U.S. census deserve a higher position on the research agenda

¹ David J. Fein and Kirsten K. West, Undercount Research Staff, Statistical Research Division, U.S. Bureau of the Census, Washington, D.C. 20233. The views expressed in this paper are those of the authors and do not necessarily reflect those of the Census Bureau.

because they are more numerous, vary more systematically with socioeconomic characteristics, and have been more politically controversial than erroneous inclusions.

The paper begins by describing a system for classifying the causes of undercount. Methods and results are presented next. A concluding discussion summarizes the implications for coverage improvement.

2. RESEARCH MODEL

The research model is presented in Figure 1. It represents undercount as a problem that occurs primarily at the household, rather than the individual, level. This specification is consistent with the basic sources of undercount in a census based on contacting each household rather than every individual in the population.

Three different household-level undercount problems are distinguished in the top margin of Figure 1: the omission of an entire household due to failure to enumerate a physical housing unit, the omission of an entire household in an enumerated housing unit, and the omission of only some members in a household where others are enumerated. Each of the three undercount problems can originate in census operations, in the society being enumerated, or in an interaction between operational and social system features. The following discussion is restricted to errors associated with the mailout/mailback methods used in the 1986 Los Angeles test census for a largely low income, Hispanic population.

2.1 Implementation of Census Operations

Operational difficulties during the census can cause the omission of housing units, of households in enumerated units, and of individuals in enumerated units. Occupied housing units can be missed because they are never added to the address lists or because they are on the lists but are erroneously deleted (U.S. General Accounting Office 1980). Given that a housing unit is correctly listed, all of the persons living in that unit may still be missed by the census due to misclassification of occupied units as vacant during nonresponse followup (U.S. Bureau of the Census 1987b; Ericksen 1983).

For questionnaires which households complete and mail back there are relatively few procedures for detecting missing persons. Procedures aimed at improving within household coverage include a question asking respondents if they were uncertain about including anyone and a clerical consistency check between a roster of household members requested at the beginning of the questionnaire and the number of persons for whom data are provided later on in the form (U.S. Bureau of the Census 1987b; Edson 1987). These procedures "cause" within household omission if they do not operate as intended due to errors in the administration of edit followup. Similarly, errors by enumerators during mail nonresponse followup may result in failure to add persons who should have been added.

Another important census operation is public information. Census publicity programs are designed to motivate mail response and reduce deliberate concealment by educating people about the uses of census data, the importance of complete reporting, and the confidentiality of census records. The extent to which such programs can reduce within household omission is unknown.

2.2 The Social System

At each stage of the census, data collection procedures come into contact with a social system which has many attributes that can impede enumeration. These attributes include unwillingness

Sources of Undercount for . . .

Missing Housing Units	Missing Households in Enumerated Units	Missing Persons in Enumerated Households
<ul style="list-style-type: none">• Address list omission• Erroneous deletion	<ul style="list-style-type: none">• Misclassification of occupied housing units as vacant	<ul style="list-style-type: none">• Failure of coverage edit procedures• Enumerator errors• Ineffective public information
<ul style="list-style-type: none">• Physical characteristics affecting unit visibility	<ul style="list-style-type: none">• Factors affecting household visibility• Factors leading to refusal	<ul style="list-style-type: none">• Factors causing unwillingness to report• Factors causing respondent definitional error

Census Implementation

Social System

Figure 1. Research Model

to report some or all household members, inability to report in a manner consistent with census definitions, and low "social visibility" of household members or the housing units in which they live. (Social visibility is the degree to which household members and housing units possess characteristics which make them perceptible to outsiders.)

The most important social system factors causing housing unit omission are those affecting the social visibility of units. Some kinds of units are easier to find and more likely to appear on commercial address lists than others. Social system sources of omission for households in enumerated units include factors depressing the visibility of household members and refusal to report.

All three broad sets of social system causes are implicated in within household omission: unwillingness to report, definitional problems, and the differential social visibility of household members. Willingness to report can be approached by considering the perceived costs and benefits of reporting for respondents (Dillman 1978). There has been much discussion of the perceived costs of census reporting. People may fear that disclosure of adult males will jeopardize welfare eligibility, that persons illegally in the country will be deported, that reporting more persons than allowed by a lease will prompt landlord troubles, and that police will be informed of the whereabouts of lawbreakers (Bailar and Martin 1987). Such fears may cause noncompliance when there is disbelief in the Census Bureau's promise of confidentiality.

The sources of definitional error are quite different from those of concealment. Definitional errors arise in the complexities of household living arrangements, as conditioned by respondents' abilities to understand and apply census enumeration and residence rules (Hainer *et al.* 1988).

Having mentioned some of the major sources of undercount, we will now examine the extent to which they occurred during the 1986 Los Angeles test census.

3. METHODS

3.1 Data Sources

This study takes an intensive look at undercount in a March 1986 test census conducted in the northern half of Los Angeles County. The population was low income and largely Hispanic. Nearly two-thirds (65%) of the heads of households enumerated in the census were of Spanish origin and 13% were Asian. Residences in this part of Los Angeles were largely single family dwellings (73%) and small apartment buildings (15%). Owners lived in half (51%) of the occupied units, in contrast with nearly two thirds (65%) of all occupied units nationwide (U.S. Bureau of the Census 1987a: 106, table 18; U.S. Bureau of the Census 1987c: 712, table 1285).

The data analyzed are from the 1986 Los Angeles test census itself; the Post Enumeration Survey, or PES, conducted to measure test census coverage; and a special followup to the PES—the Causes of Undercount Survey. The census enumerated 109,900 housing units and was intended primarily as a test of planned 1990 census operations.

The Post Enumeration Survey (PES) was one of these operations. The purpose of the PES, conducted in July 1986, was to identify census omissions and erroneous enumerations (Diffendal 1988). It did this by attempting to match PES to census records. When a PES person's record was found in the census it was termed "matched"; otherwise the person was considered "nonmatched".

Three kinds of PES households are distinguished here, depending on whether all, some, or none of their members were matched to the census. "Complete match" households contain only persons in the PES who were matched to persons in the census. "Partial nonmatch"

households contain at least one person who could not be matched and at least one person who was matched to the census. "Total nonmatch" households include only persons who could not be matched to the census.

These three household types are distinguished to allow examination of problems associated with housing unit omission, omission of entire households in enumerated units, and omission of persons from households that were partially enumerated. Completely matched households are included for reference purposes, to represent households correctly enumerated in the census.

A special followup survey – the Causes of Undercount Survey – was conducted in November 1987 to obtain additional information needed to compare these household types. The survey obtained information on census characteristics for nonmatched persons, as well as some new household and housing unit data not available on the census or PES files.

The entire partial nonmatch stratum and nearly all households in the total nonmatch stratum were selected for reinterview. Eight total nonmatch households had to be omitted because several items needed to reinterview them were missing. Households in the complete match stratum were subsampled to reduce survey costs.

The distribution of the 966 completed Causes of Undercount Survey interviews by household type is shown in the right-most column of Table 1. This table also gives the unweighted numbers for all 5814 PES households and the 1420 cases in the Causes of Undercount Survey sample. The overall response rate for the survey was 68%, reflecting considerable success in locating households in a transitory urban area despite the 16 months intervening between the survey and the PES.

3.2 Analysis Plan

There are several parts to the analysis. PES total nonmatch households are examined first. Two sets of comparisons are made: 1) of missed housing units with enumerated housing units and 2) of missed households in enumerated units with enumerated households. Missed housing units were expected to contain a higher percentage of clustered housing units and unusual unit types and locations than enumerated units. Missed households in enumerated housing units were expected to be smaller, contain adults who were less frequently at home, and move more often than enumerated households. Most of the explanatory variables for housing unit and household omission were obtained either from the census Address Control File or from the PES matched file, and thus are available for all 193 total nonmatch households in the sample.

Table 1

Numbers of Households in the PES and Causes of Undercount Survey Sample,
and Numbers of Completed Interviews, by Household Type.

Household Type	Post Enumeration Survey	Causes of Undercount Survey	
		Sample	Completed Interviews
Complete Match	4,871	489	382
Partial Nonmatch	738	738	484
Total Nonmatch	205	193	100
All Types	5,814	1,420	966

The second part of the analysis compares partial nonmatch with complete match households to identify factors responsible for within-household omission. Two sets of explanatory factors are distinguished, those indicating inadvertent or “definitional” errors and those representing reasons for deliberate concealment. Indicators for definitional errors include large size and complex composition of households, poorly-spoken English and educational deficits. Concealment indicators include presence of recent immigrants, welfare reciprocity, crowded housing, and disbelief in census confidentiality. It was hypothesized that partial nonmatch households would score higher on the definitional and concealment indicators than would complete match households.

The analysis begins with bivariate relationships between each of the explanatory factors and partial omission and then considers multivariate relationships. The source for many of these indicators was the Causes of Undercount Survey; hence, only data from interviewed households are used.

In the final part of the analysis, characteristics of four types of individuals are compared: persons *matched* in complete match and partial nonmatch households, and those *nonmatched* in partial and total nonmatch households. Characteristics compared include age, sex, education, relationship to the household head, and citizenship status.

Bivariate percentages are based on weighted data to compensate for the PES and Causes of Undercount Survey sampling designs, though tests for differences between these percentages used unweighted numbers. Unweighted data were used to estimate parameters of log-linear models. The effects of the PES sampling design on estimates for the final models were evaluated by adding in all two-way interactions which included the PES stratification variable. This adjustment did not greatly change the results; thus, the estimates presented here do not include the stratification variable. Because the second stage of PES sampling entailed cluster sampling of households in census blocks, the standard errors calculated are likely to underestimate the true sampling errors: they are presented only as rough guides to the significance of parameters.

4. FINDINGS

4.1 Total Nonmatch Households

Table 2 shows the final status assigned in the census to PES total nonmatch households for cases sent and not sent to nonresponse followup. Of the 193 total nonmatch cases 97, or 50%, never appeared on the census address lists. Thus, housing unit omission appears to explain why the PES could not find anyone in these households in the census.

The remaining 96 cases did appear on the census address lists. What caused these households to be missed? The explanation is probably that most of these units were census closeout interviews, where a landlord or neighbor provided only an estimate of the total number of persons in the household and not detailed information for individuals. This hunch is supported by the finding that of the 44 cases the census classified as occupied, population counts for 37 were “goldplated”. This means that the final count accepted for these households was not obtained in the usual manner by allowing the FOSDIC (Film Optical Device for Input to Computers) machines to count persons. Instead, goldplating involved accepting a total count for the household entered on the questionnaire in the field. This is likely an indication that the household was a closeout case.

Thus, the census really did not miss most of these 44 households entirely, though when it came time for PES matching, there were no individual census person records to be matched.

Table 2
 Final Status Assigned in the Census to PES Total
 Nonmatch Households By Nonresponse Followup Status:
 Numbers of Units^a

Final Status of Unit in Census	Sent to Nonresponse Followup?		
	No	Yes	Total
Omitted from the Census Address Lists	97	0	97
Included in the Census Address Lists	4	92	96
Occupied, Direct Accept ^b	1	6	7
Occupied, Gold-plated ^c	2	35	37
Vacant, Direct Accept	1	34	35
Vacant, Gold-plated	0	17	17
All Units	101	92	193

Notes: ^a N's are unweighted.

^b Direct Accept: FOSDIC person count accepted.

^c Gold-plated: Field counts accepted instead of FOSDIC.

An allowance is made for these cases in the dual system estimation method. Nevertheless, it still is true that these households were not directly enumerated.

To summarize, 50% of the PES total nonmatch households were in units which appeared to have been entirely omitted. Of the households living in units which were enumerated, 54% had been classified as vacant, possibly erroneously, and 46% had been found to be occupied. Of the total nonmatch households classified as occupied in the census, up to 84% may have been enumerated in closeout interviews.

Figure 2 compares some physical characteristics of units left off the census address lists (light bars) with units that were not left off the lists (dark bars). The top set of bars represents the basic types of housing units. Attached single family homes, such as duplexes, appear to have been a major problem in the L.A. test census. Thirty-four percent (34%) of the missed units fell into this category, in contrast to only 8% of enumerated units. Missed units were less likely than enumerated units to be detached single family homes or apartments in large buildings, suggesting that the census was more successful at finding such units.

Whether or not an interview was completed, Causes of Undercount Survey interviewers were asked to record when units they visited fit any of several "unusual unit" categories listed on the front of their questionnaires. The bottom half of Figure 2 shows that the interviewers identified a higher percentage of unusual units among units that were missing from the census address lists, 28%, than among units that were included, 7%. Unit types found to be particular problems were abandoned-looking buildings and secondary units on a lot.

Physical characteristics of units thus do appear to affect their visibility during census address list development. What might cause households to be missed in units that were enumerated?

Households may be more easily missed if they are small and mobile. Figure 3 compares characteristics of total nonmatch households in enumerated units with a combined group of complete match and partial nonmatch households – that is, households which were enumerated. Households missed in the test census (light bars) were on average considerably smaller than those where some or all members were counted (dark bars). Whereas 53% of the total nonmatch households in enumerated units had one or two members, only 35% of the enumerated households were this small.

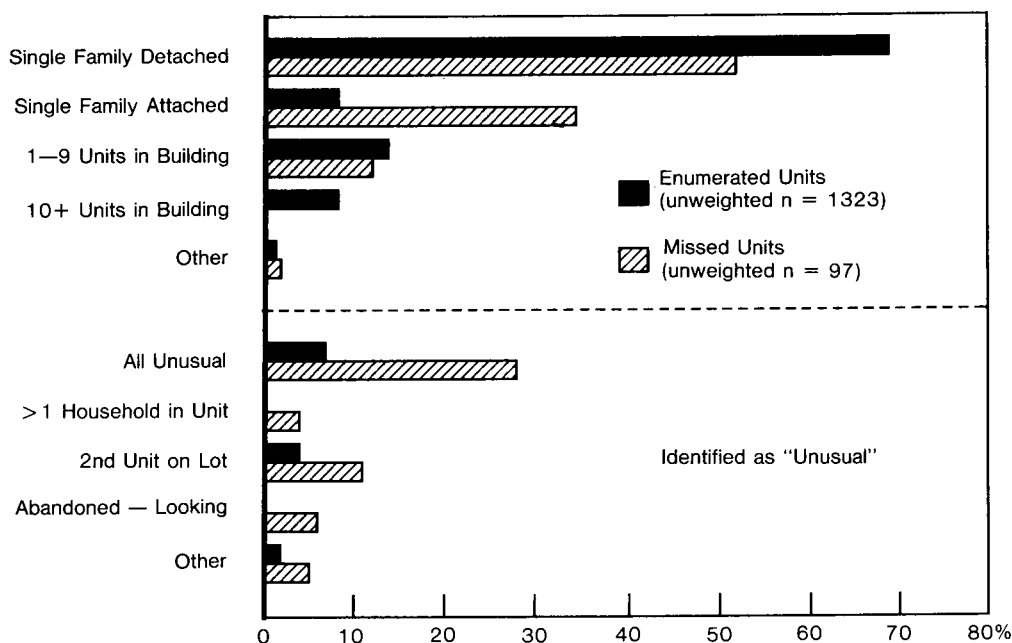


Figure 2. Physical Characteristics of Enumerated and Missed Housing Units (Weighted Percentages)

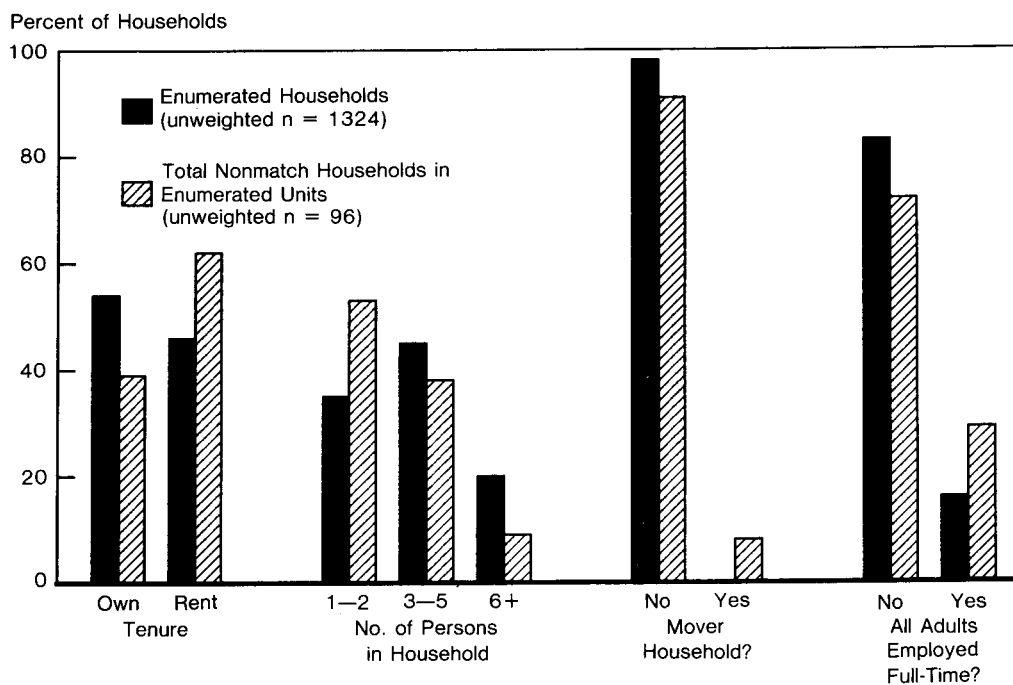


Figure 3. Characteristics of Enumerated Households and Total Nonmatch Households in Enumerated Units (Weighted Percentages)

Indicators of the propensity to move include home ownership and actual household mobility in the four months between the census and the PES. Households missed in the census were more likely to be renters and movers (61% and 8%, respectively) than were enumerated households (46% and 0%, respectively). The percentage of households in which all adults were employed full-time in March 1986 was greater by 12% for omitted households than for enumerated households, though the number of interviews for omitted households was too small for this difference to be statistically significant.

These results support the hypothesis that missed housing units and households missed in enumerated units possess attributes which reduce their visibility during a census.

4.2 Partial Nonmatch Households

From total nonmatch households, the focus shifts to the factors associated with partial household omission. In this phase of the analysis, 484 partial nonmatch households were compared with 331 complete match households. Single person households were excluded from the 382 complete match households in the Causes of Undercount Survey sample, since they were not at risk of partial omission.

Two different sets of explanatory factors were considered. The first represents household characteristics thought to be associated with definitional errors, described earlier as errors resulting from inconsistencies between household membership as understood by the Census Bureau and by census respondents. The second set of indicators represents factors thought to be associated with the deliberate concealment of household members.

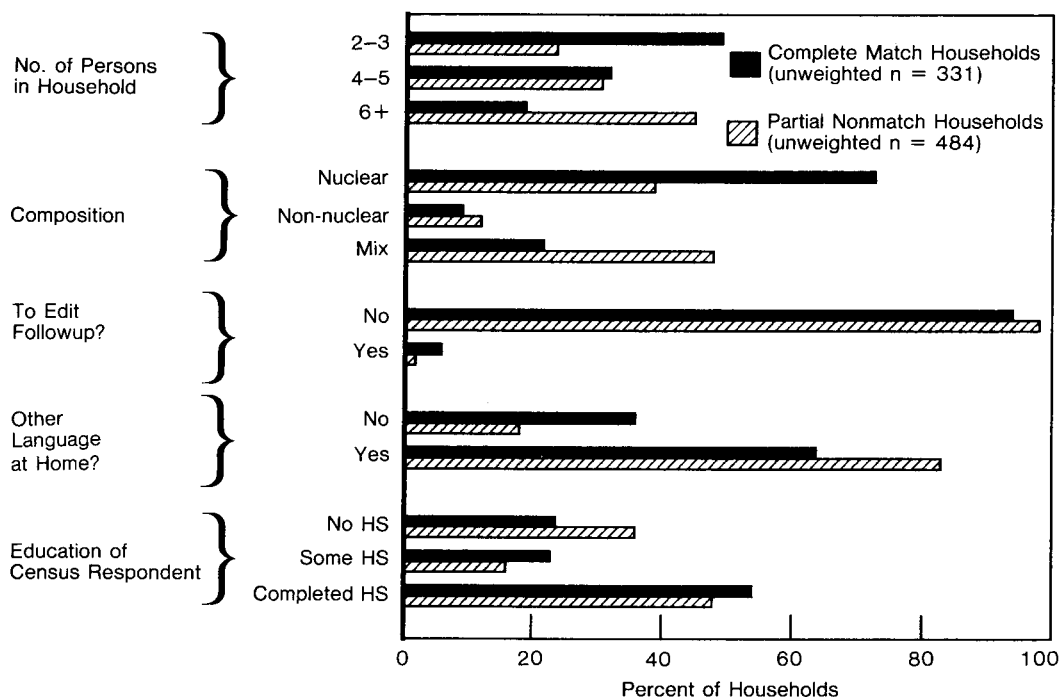


Figure 4. Definitional Error Indicators for Partial Household Omission: Households with 2+ Persons (Weighted Percentages)

Definitional Errors

Indicators for definitional errors include household size and composition, English language ability, census respondent's education, and edit followup status. Larger households, those containing more distant relatives and persons unrelated to the household head, those speaking a language other than English at home, those where the census respondent's education was low, and households not sent to edit followup were all expected to be at greater risk of definitional errors.

Figure 4 supports these hypotheses. It shows that partial nonmatch households (light bars) were considerably larger than complete match households (dark bars): 45% of the partial nonmatch households but only 19% of the complete match households contained six or more members. Whereas 40% of the partial nonmatch households contained only nuclear relatives of the household head, fully 72% of the complete match households were nuclear. Partial nonmatch households were less likely to have been sent to edit follow-up by a slight, but statistically significant, amount. Partial nonmatch households were more likely to speak a language other than English at home (83%) than were complete match households (64%). Finally, census respondents from partial nonmatch households had less formal education than those from complete match households: 36% of the census respondents from partial nonmatch households had not attended high school, in contrast with 24% of the respondents from complete match households.

Log-linear models were fitted to see whether these differences persisted at the multivariate level. The dependent variable in these models was partial household omission, with complete match households coded as 0 and partial nonmatch households coded as 1. Interactions between partial omission and each of the independent variables in Figure 4 were tested in a series of nested models. All two-way interactions among independent variables were included in each model as controls.

In the multivariate analysis, significant interactions with partial omission were found for all definitional error indicators except census respondent's education. Table 3 presents the chi square (Wald) statistics associated with the final definitional model, which excludes census respondent's education. Significant interactions of household size with composition and language other than English were also detected. Parameter estimates in Table 4 show the effects to be in the directions expected. Estimates for standardized parameters, obtained by dividing

Table 3
Chi Square Statistics For Testing Two-Way Interactions
in the Final Definitional Error Model^a

Variables	Interactions with . . .			
	Size	Composition	Edit Followup	Language at Home
Partial Omission	38.1**	42.3**	6.3*	5.2*
Size	—	112.0**	.9	50.0**
Composition	—	—	1.6	1.3
Edit Followup	—	—	—	1.0

** : $p < .01$

* : $p < .05$

^a Log Likelihood $X^2 = 42.2$, $df = 45$, $p = .5922$.

Table 4
Parameter Estimates for Interactions Between Definitional Error Indicators and
Partial Household Omission in the Final Model

Marginals with Partial Nonmatch Household and . . .	Parameter Estimate	Standard Error	Standardized Parameter Estimate
Household Size:			
2-3 Persons	-.34	.06	-5.7
4-5 Persons	-.02	.05	-.4
Composition:			
All nuclear	-.36	.06	-6.0
All non-nuclear	.22	.09	2.4
Edit Followup Status			
Not sent	.25	.10	2.5
Other Language at Home?			
Yes	.10	.05	2.0

parameter estimates by their standard errors, indicate that the effects of size and composition are about the same in magnitude and that both are larger than the effects of edit followup and language spoken at home.

Concealment Indicators

Factors hypothesized to cause concealment of household members by census respondents include: fear that persons illegally in the country would be deported, fear that disclosure of adult males would jeopardize welfare aid, and concern that reporting more persons than allowed by a lease would bring landlord troubles. Indicators for these factors were, respectively, whether the household contained recent immigrants, defined as persons entering the country in or after 1980; whether anyone in the household was receiving welfare during the census month; and the average number of persons per room in the household. Nonresponse to the census mailout was also included as a general indicator of failure to perceive positive benefits from responding to the census. Finally, belief in census confidentiality was included to see whether it helped to reduce fears resulting in concealment.

Figure 5 shows that all of these indicators were related to partial omission at the bivariate level. For example, recent immigrants were present in 26% of the partial nonmatch households (light bars), but only 12% of the complete match households (dark bars). Whereas 24% of the partial nonmatch households reported receiving welfare, only 15% of the complete match households did so. Partial nonmatch households were considerably more likely to exhibit crowding: 63% contained more than one person per room, in contrast to only 34% of the complete match households. Partial nonmatch households were also somewhat less likely than complete match households to have returned their census questionnaires by mail or to believe in census confidentiality.

Again, loglinear models were fitted, with partial omission as the dependent variable and the concealment indicators as independent variables. All two-way interactions with household size were included as controls, since other things being equal, larger households would be more likely to exhibit crowding and contain recent immigrants than small ones.

This time, two variables did not survive preliminary testing: mail nonresponse and belief in census confidentiality. Before completely dropping the confidentiality variable, tests were performed to see if interactions of partial omission with presence of immigrants, welfare reciprocity, and crowding depended on belief or disbelief in confidentiality. Belief in confidentiality was not found to affect these relationships.

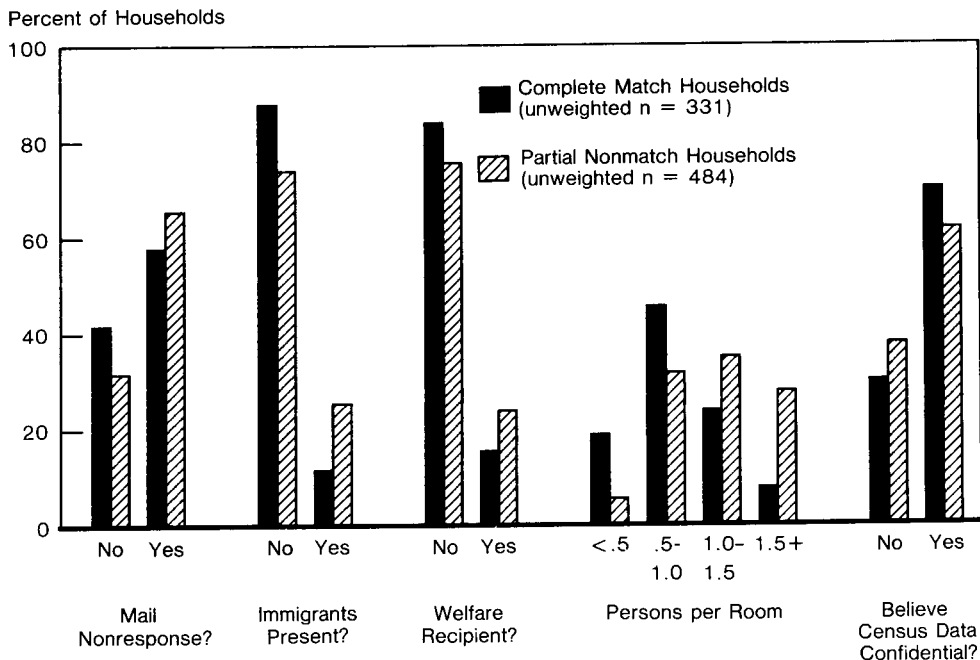


Figure 5. Concealment Indicators for Partial Household Omission: Households with 2+ Persons (Weighted Percentages)

Table 5
Chi Square Statistics For Testing Two-Way Interactions
in the Final Concealment Model^a

Variables	Interactions with . . .			
	Size	Immigrants	Welfare Assistance	Crowding
Partial Omission	2.9	11.3**	10.1**	16.7**
Size	—	.2	7.5*	221.7**
Recent Immigrants	—	—	1.6	30.0**
Welfare Assistance	—	—	—	5.4

** : $p < .01$

* : $p < .05$

^a Log Likelihood $X^2 = 103.8$, $df = 150$, $p = .9985$.

Table 5 shows that three of the remaining concealment variables immigrants, welfare, and crowding interacted significantly with partial household omission in a model which included all two-way interactions with size and all two-way interactions among independent variables. Standardized parameter estimates (see Table 6) suggest effects of roughly equal magnitude for the three indicators.

It is noteworthy that the relationship between partial omission and size vanished when crowding was included (see Table 5), suggesting that the effects of size were due to its association with crowding rather than scale alone. Crowding was also strongly associated with the presence of recent immigrants.

4.2 Person Characteristics

For the final part of the analysis of individual-level characteristics associated with undercount, four kinds of persons were compared: persons the census counted in complete match and partial nonmatch households, and persons the census missed in partial and total nonmatch households.

Figure 6 shows differences between the percentages in 10 year age groups for persons in complete match households and each of the three other groups. It shows an excess in the 20-29 year old group for persons missed in partial and total nonmatch households relative to persons in complete match households. There is also evidence of an excess in the 20-29 year age groups for persons who were enumerated in partial nonmatch households.

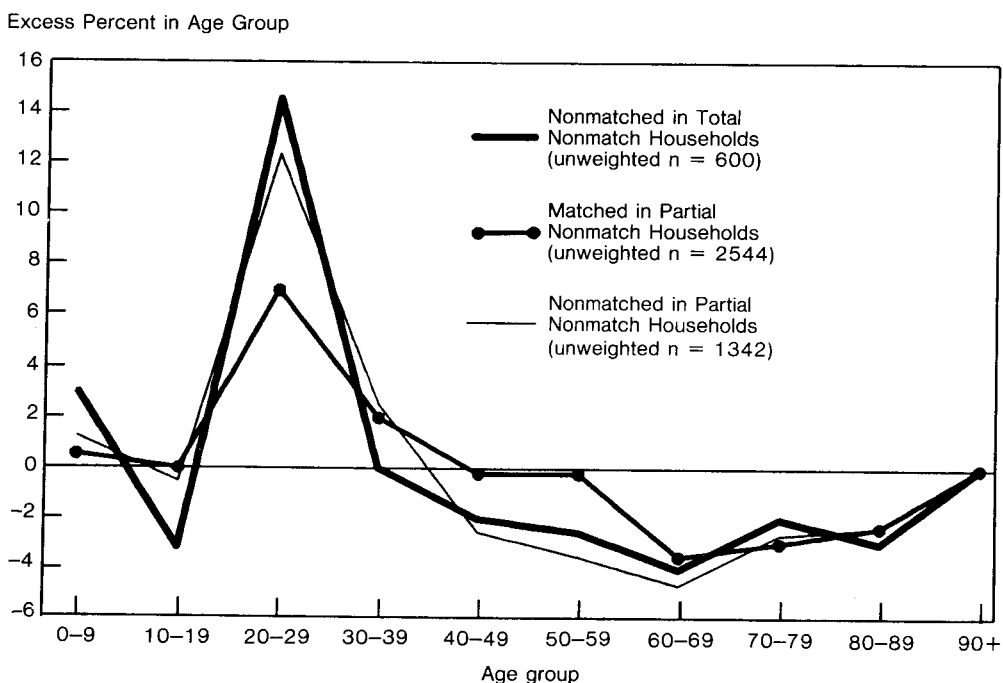


Figure 6. Excess Weighted Percentage in Age Group Relative to Persons in Complete Match Households

Table 6
Parameter Estimates for Interactions Between Concealment Indicators and
Partial Household Omission in the Final Concealment Model

Marginals with Partial Nonmatch Household and . . .	Parameter Estimate	Standard Error	Standardized Parameter Estimate
Recent Immigrants:			
Immigrants Present	.19	.06	3.2
Welfare Reciprocity:			
Receiving Aid	.17	.05	3.4
Crowding:			
< .5 Persons/Room	-.49	.13	-3.8
.5-1.0 Persons/Room	-.01	.08	-.1
1.0-1.5 Persons/Room	.08	.08	1.0

Table 7
Percentage Distributions for Characteristics of Individuals by
PES Match Status and Household Type

Characteristic	PES Match Status			
	Matched in		Nonmatched in	
	Complete Match HHs	Partial Nonmatch HHs	Partial Nonmatch HHs	Total Nonmatch HHs
Sex				
Male	46.2%	50.6%	54.2%	48.2%
Female	53.8	49.4	45.9	51.8
Unweighted <i>n</i>	1667	2564	1324	582
Education				
No Formal Education	10.2	10.9	17.0	14.3
Less than High School	30.7	34.4	27.2	37.5
Some High School	20.5	20.6	19.5	19.5
High School Graduate	38.6	34.1	36.4	28.8
Unweighted <i>n</i>	1197	1560	599	315
Relationship to Head				
Nuclear Relative	86.1	83.2	63.6	85.9
Non-nuclear Relative	11.3	12.6	25.4	7.9
Non-relative	2.6	4.2	11.0	7.0
Unweighted <i>n</i>	1659	2560	1359	590
Citizenship				
Citizen Since Birth	66.2	53.5	52.6	50.4
Naturalized Citizen	9.2	9.5	6.4	6.4
Noncitizen	24.6	37.0	41.0	43.2
Unweighted <i>n</i>	1223	1567	612	316

Persons missed by the census in partial nonmatch households were slightly more likely than persons in complete match households to be males and have no formal schooling, and less likely to be citizens or close relatives of the household head (Table 7). Persons missed by the census in total nonmatch households were also slightly more likely to be noncitizens and lower in education than persons in complete match households, but displayed no differences in sex and relationship to household head. Thus, on the whole, persons missed in partial nonmatch households differed from those in complete match households in more ways than did persons missed in total nonmatch households.

In addition to biasing more census characteristics, partial household omission caused the omission of many more persons than did total household omission. Two thirds (67%) of all PES nonmatch cases were in partial nonmatch households and only one third were in total nonmatch households. Fully 82% of all PES omissions were found in housing units the census enumerated and only 18% were in missed units.

5. DISCUSSION

The findings reported here support evidence from more qualitative studies that partial household omission is the most serious undercount problem in hard-to-enumerate urban areas of the United States today. As compared with total household omission, partial omission in the Los Angeles test census accounted for twice as many missing persons, reflected more intractable sources of error, and biased more individual-level census characteristics.

The chief problems identified for total household omission were failure to include certain types of housing units in the census address lists and misclassifying occupied units as vacant. Housing units especially at risk of misclassification as vacant were those with households which were small and mobile and those in which all adults were working full-time. Experience with coverage improvement programs at the Census Bureau suggests that further reductions in housing unit omission may be possible. Such programs were responsible for adding about 10% of the units enumerated in Los Angeles. The Bureau adopted special prec canvassing procedures in the test census to find units in large multi-unit structures. Considerable success in reducing this source of error in the test census is evident in Figure 2: none of the apartment units missed were in large buildings.

The misclassification of occupied units as vacant will be more difficult to remedy. Allowing nonresponse enumerators more time per unit and improved training for certain kinds of problem households may help somewhat. Coupling these efforts with special callback procedures for smaller and more transient households and those whose members are rarely at home would also help.

It is clear that improvements at the margin of what is already a largely successful census operation will be expensive. Keyfitz (1979) and others have observed that the incremental costs from adding persons to the count soar as coverage approaches 100%. Programmatic innovations to reduce the errors observed in the 1986 test census would add to the \$2.6 billion cost projected for the 1990 census, since the methodology to be used in urban areas will be very similar to the L.A. test census.

Within-household errors will be even more difficult to address than total household omissions. The Bureau must redouble its efforts to understand the complex living arrangements and cognitive and/or cultural factors that condition how people perceive household membership. The findings reported here suggest that further efforts targeted to respondents for whom English is not a native tongue, and households containing persons only distantly related to each other may help to reduce definitional errors.

However, in light of the considerable research already performed to improve the design of the census questionnaire and the complex enumeration and residence rules to which the Bureau is bound by statute and tradition, further reductions in definitional error will require extraordinary efforts. Definitional errors are deeply embedded in cultural differences and educational deficits among hard-to-enumerate groups.

Within-household omission also was found to be strongly related to the presence of immigrants, welfare reciprocity, and crowding. That a PES-based study could detect such effects suggests that the PES succeeded in counting many persons whose presence had been concealed in the census. Some of the effects of the so-called concealment variables may be due to uncontrolled factors other than concealment, but the persistence of relationships even after household composition was added in a final log-linear model (not shown) suggests that the PES really did detect some persons who were concealed in the census. Thus, there appears to be a continuum from households that are highly resistant to enumeration to those which are less resistant, and for the latter more intensive methods like those used in the PES may be effective.

The social conditions underlying the most resistant forms of concealment present the most difficult problems for the Census Bureau. Public information programs attempting to convince people that the census is important and that census data will be kept confidential were not very effective for the hard-to-enumerate population in the Los Angeles test census, as reported by Moore and McDonald (1987), though these programs may work better under real decennial census conditions. The minimal role found for belief in census confidentiality, either in its own right or in mediating between household circumstances and concealment, suggests that the relationship between attitudes and census response behavior is not a simple one.

The findings reported here should not be generalized uncritically to the sources of undercount expected to affect urban areas in the 1990 Census. Because the data are based on a test census, errors may reflect inexperience with experimental procedures or failure to convince respondents (and census workers) that the project was as serious as the decennial census. Further, to the degree that Los Angeles is unlike other major urban areas, it may experience unique census-taking problems. For example, Los Angeles is thought to be home to more illegal aliens than any other major city (Heer and Passel 1987).

On the other hand, the net undercount rate for Los Angeles in 1980 was quite similar to the rates for other major cities, as measured in the 1980 Post Enumeration Program (Fay *et al.* 1988). Thus, what they lack in illegal aliens, these cities may make up in other hard-to-enumerate groups. Further research is needed to assess the degree to which causes of undercount differ by race, ethnicity, and other social characteristics.

It is encouraging that the causes of undercount identified in this Post Enumeration Survey-based study were reasonably consistent with more qualitative reports by ethnographers and focus groups. Also, the PES estimates for undercount from the Los Angeles test census are believed to be of high quality (Hogan and Wolter 1988). For these reasons, extension of the PES-based methodology developed in this paper to other urban (and nonurban) areas is recommended.

On the social system side, further research on how rationally people weigh the costs and benefits of responding to censuses and surveys would help to weigh the potential for improving census coverage through the Census Bureau's public information and community action programs. Better indicators for household-level reasons for concealment are also needed. Examining specific assistance programs would help to confirm the effects of welfare participation on census coverage, since not all aid would be imperiled by revealing true household composition.

Improved measurement of the sources of undercount arising in census operations is also needed. If data from census quality control programs were combined with PES matching results, error sources could be identified with greater precision.

ACKNOWLEDGEMENTS

We would like to thank the following persons for their assistance in this research: Irwin Anolik, Miriam Balutis, Gregg Diffendal, Chris Dyke, Sue Finnegan, Howard Hogan, Jan Jaworski, Pete Long, and Lynn Weidman. Betsy Martin, Jim O'Brien, and two anonymous reviewers provided valuable comments on earlier versions of this paper.

REFERENCES

- BAILAR, B., and MARTIN, E. (1987). Report on Meetings in Los Angeles, Chicago and Denver. Unpublished Census Bureau memorandum.
- CHOLDIN, H. (1987). Science and Scientists in the 1980 Census Lawsuits. Paper presented at the May 1987 meeting of the Population Association of America, Chicago.
- CLOGG, C.C., MASSAGLI, M.P., and ELIASON, S.R. (1986). Population undercount as an issue in social research, *Proceedings of the Second Annual Research Conference*. United States Bureau of the Census, Washington, D.C., 335-343.
- CITRO, C.F., and COHEN, M.L. (eds.) (1985). *The Bicentennial Census: New Directions for Methodology in 1990*, Panel on Decennial Census Methodology, National Research Council, Washington, D.C.: National Academy Press.
- DIFFENDAL, G. (1988). The 1986 test of adjustment related operations in Central Los Angeles County. *Survey Methodology* 14, 71-86.
- DILLMAN, D.A. (1978). *Mail and Telephone Surveys: The Total Design Method*. New York: John Wiley and Sons.
- EDSON, R.G. (1987). Preliminary coverage improvement results from tests for the 1990 Census. Paper presented at the August 1987 meeting of the American Statistical Association, San Francisco.
- ERICKSEN, E.P. (1983). Affidavit, Mario Cuomo, *et al.* vs. Malcolm Baldrige *et al.*, U.S. District Court, Southern District of New York, 80 Civ. 4550 (JES).
- FAY, R.E., PASSEL, J.S., and ROBINSON, J.G. (1988). The coverage of population in the 1980 Census. *Evaluation and Research Reports. 1980 Census of Population and Housing PHC80E4*, Washington, D.C.
- HAINER, P., HINES, C., MARTIN, E., and SHAPIRO G.M. (1988). Research on improving coverage in household surveys. *Proceedings of the Fourth Annual Research Conference*. United States Bureau of the Census, Washington, D.C., 513-539.
- HEER, D.M., and PASSEL, J.S. (1987). Comparison of two methods for estimating the number of undocumented Mexican adults in Los Angeles County. *International Migration Review*, 21(4), 1446-1473.
- HOGAN, H., and WOLTER, K. (1988). Measuring accuracy in a Post-Enumeration Survey. *Survey Methodology* 14, 99-116.
- KEYFITZ, N. (1979). Information and allocation: Two uses of the 1980 Census. *The American Statistician*, 33(2), 45-56.
- MOORE, J.C., and McDONALD, S.-K. (1987). The Census community awareness program: an evaluation of the potential and actual effectiveness of CCAP based on evidence from the 1986 Los Angeles Census Test. Unpublished Census Bureau report.

- ORTNER, R. (1987). Statement. *United States Department of Commerce News*, October 30, 1987.
- U.S. BUREAU OF THE CENSUS (1960). *The Post-Enumeration Survey: 1950*. Technical Paper No. 4, Washington, D.C.
- U.S. BUREAU OF THE CENSUS (1987a). 1986 Test Census, Central Los Angeles County, California. *General Population and Housing Statistics*, TC86-1, Washington, D.C.
- U.S. BUREAU OF THE CENSUS (1987b). Programs to improve coverage in the 1980 Census. *Evaluation and Research Reports. 1980 Census of Population and Housing*, PHC80-E3, Washington, D.C.
- U.S. BUREAU OF THE CENSUS (1987c). *Statistical Abstract of the United States: 1987*, (106th edition), Washington, D.C.
- U.S. GENERAL ACCOUNTING OFFICE (1980). *Problems in Developing the 1980 Census Mail List*. Washington, D.C.: General Accounting Office.