A Summary Review of Research Relevant to Housing Units Missing from the Census Bureau's Master Address File (MAF)

by Ed Kissam

INTRODUCTION AND OVERVIEW

This summary review of research on differential undercount in the decennial census focuses on a specific cause of undercount—the undercount that results when the place where individuals or a family live has not been identified and included in the Census Bureau's Master Address File (MAF), its list of census addresses.

Historically, the problem has been that the decennial census (and the American Community Survey) are primarily mail surveys. If a household did not receive a mailed census form they generally could not return it, although in some areas the decennial census uses an alternative mode of enumeration—leave/enumerate (where census enumerators leave a form for a household to complete) or update/enumerate (stopping by to add a housing unit to the MAF and, at the same time, enumerating the household). In Census 2020 this problem will persist even though the Bureau will encourage online responses—because if there is no address on file for the respondent they will not receive a link to allow them to respond online.

Census undercount is correlated with a range of factors which include individual respondent socioeconomic characteristics, household characteristics, and neighborhood and community characteristics. Various 'undercount' analysts focus on one set of factors or another. Historically, the most common focus has been on the well-established correlation between race/ethnicity and undercount. However, beginning in the 1970's, Census Bureau researchers began to look more deeply into the ways in which social system functioning and census operational systems interacted to give rise to differential undercount of minorities.

A practically important insight which has emerged from the research literature on causes of census undercount is that race/ethnicity does not "cause" undercount directly but that economic/social disadvantage, coupled with language and cultural factors, does; the most economically and socially marginal populations are those with the highest risk of undercount. So it is not surprising to find that the interaction of multiple factors associated with undercount give rise to a "mega-undercount" of immigrants.

This dynamic strongly shows that the most severely undercounted sub-population among immigrants are the most recently arrived, undocumented, low-income immigrants with the least robust social networks. For example, in a top-quality Los Angeles area study (Fein 1989) it is estimated that households headed by non-citizen immigrants were twice as likely to be omitted from the census than those with citizen heads of household and that recent immigrants were twice as likely to be missed than long-term settlers.[1]

Although limited literacy, language limitations, long working hours, minimal familiarity with the census process are all factors which contribute to the census undercount of immigrants, the research indicates that one of the most important factors in immigrant undercount stems from immigrants living in crowded, sub-standard housing—a factor some have referred to as a "structural" cause of undercount because in the neighborhoods where low-income immigrants and other low-income ethnic/racial minority families live, the Census Bureau's address list fails to include many of the low-visibility, "unusual" living quarters.[2]

Incomplete Census Address Lists, Total Household Omission, and a Strategy to Ameliorate It

The omission of housing units from the MAF cannot be overcome by messaging to encourage census response or by improved followup on non-response—because families with no address are operationally invisible. The Census Bureau cannot usually recognize that the families in a housing unit it does not know exists have not been enumerated. Moreover, total household omission (resulting from a housing unit not being in the MAF), by definition, means that everyone in the household, not just a single individual, is left out of the census count.

This inherent problem of deficiencies in the decennial census (and ACS) sampling frame further suggests that even current estimates of housing unit omission based on dual-system estimation using a post-enumeration survey (PES) are biased downward—because this theoretically valid method of assessing survey accuracy is extremely difficult to implement practically because it requires that the dual sampling frames (in this case, the decennial census and a post-enumeration survey to assess the accuracy of the census) are developed independently.

Nonetheless, despite its limitations, the quantitative research indicates that statewide at least 2.2% of all housing units in California are omitted from the MAF. This problem accounts for about half of aggregate differential undercount in the decennial census.[3]

The proportion of housing units omitted from the MAF varies from census tract to census tract and is almost certainly highest in low-income minority neighborhoods and, within those neighborhoods, highest for immigrants, and within the immigrant population, highest for the most marginal sub-population, recent, undocumented immigrants. The quantitative research suggests that in these neighborhoods overall, that 3% or so of housing units may be omitted from the MAF and that, in some community contexts, the problem is still more serious.[4]

The observation that MAF problems occur in particular sorts of "hot zones" where housing conditions are sub-standard and where chaotic, crowded, "irregular" housing is more prevalent provide an important insight. The nature of the problem suggests a solution, namely that a "targeted" local address canvassing initiative can make substantial contributions to improving the MAF.

Consideration of the reasons why some sorts of living accommodations fail to be included in the MAF also provides insights about promising components of the solution. The strategy of relying on partnerships between local government and grassroots community groups to conduct local address canvassing to improve the MAF provides a way to deploy local knowledge in knowing where and how to look for hidden or low-visibility housing units. It is reasonable to expect that, trusted community groups can do better at overcoming social and linguistic barriers and building local, low-income minority families confidence that acknowledging an irregular housing unit is occupied will not result in reprisals (since the confidentiality of census data is diligently protected).

Consequently, local address canvassing conducted via collaboration between local government and grassroots community-based organizations in a special pre-census preparation process in 2017 known as LUCA (Local Update of Census Addresses) can make a major contribution to overcoming the pervasive problem of total household omission. If the LUCA process is successful in identifying even half of the housing units which would otherwise be omitted from the MAF, it can ameliorate the differential undercount of minorities and immigrants in the decennial census by at least 0.6% and generate somewhere in the range of $4-8 billion in additional census driven federal funding for California in the decade from 2021-2030.[5]

RESEARCH RELEVANT TO HOUSING UNITS NOT IN THE MAF

This research literature, which includes post-enumeration surveys in past decennial censuses, ethnographic studies of the causes of census undercount, and a unique study of undercount conducted as part of the 1986 Los Angeles Test census[6] - provides valuable insights about the importance of Census Bureau's Master Address File (MAF) quality as a foundation for a successful census. However the implications of this research for understanding the extent to which omissions of housing units from the Census MAF compromised census quality, have not received much attention and, consequently, census improvement efforts have focused on other factors related to census undercount.

The studies discussed below and their implications are crucial because they provide quantitative insights about the proportion of aggregate census undercount which can be attributed to total household omission (which is highly correlated with an address not being in the MAF), and the extent to which undercount stems from partial household omission (typically

thought of as resulting from respondent error). This, in turn, makes it possible to assess the specific contribution that MAF improvement can make to census accuracy.

Research providing insights on variations in MAF Quality

An important implication of multi-variate models of census undercount which include neighborhood characteristics is that MAF quality is correlated with community housing characteristics and that it varies from neighborhood to neighborhood.

Fein and West-Findings from the Los Angeles Test of Census-related Operations (1986)

The most methodologically thorough research on total household omission as a result of problems in the Census Bureau's master address file is a Census Bureau study led by David Fein and Kirsten West. The study can best be seen as a "triple enumeration" one because it includes a comparison of the housing units identified in the test census, in the post-enumeration survey conducted as part of the test census, and the study's special survey, the "Causes of Undercount Survey" conducted in the Los Angeles area.

There are three major documents based on this research: a 1988 paper by Fein and West (David Fein and Kirsten West, "The Sources of Census Undercount: Findings from the 1986 Los Angeles Test Census", Survey Methodology, Vol. 14, December, 1988), Fein's 1989 Ph.D. dissertation for Princeton University (David Fein, "The Social Sources of Census Omission: Racial and Ethnic Differences in Recent U.S. Censuses", Ph.D. dissertation, Department of Sociology, Princeton University, 1989) , and a subsequent 1990 paper by Fein (D.J. Fein, "Racial and Ethnic Differences in U.S. Census Omission Rates," Demography, 27(2), 285-302, 1990).

The Causes of Undercount Survey (CUS) consisted of a stratified cluster sample of 190 blocks in the Los Angeles Test of Census-related Operations (TARO) in 1986. It included nine jurisdictions, Bell Gardens, City of Commerce, East Los Angeles, Montebello, Monterey Park, Pico Rivera, south El Monte, South San Gabriel, and West Whittier-Los Nietos. Accompanying the CUS study were three ethnographic research studies, the most relevant of which were conducted by Diego Vigil (1987) and Camilo Garcia Parra (1987).

The CUS study areas are particularly relevant to efforts for targeted MAF improvement in areas with concentrations of immigrants because they are fairly typical of the kinds of California neighborhoods/census tracts which have historically had higher levels of undercount. About three-quarters (74%) of the heads of household in the survey area were Hispanic, and 10% were Asian or Pacific Islander. In these communities with a high proportion of minority residents, 24.4% of the population spoke English less than well, 52.4% had not completed high school. About one-third (30%) were immigrants and half (15%) were recent immigrants (<5 years in the U.S.).

A particularly important element in the CUS study is that it included information on the characteristics of the kinds of housing units which were missed in the MAF and, consequently, the test census. This is because the analysis distinguished cases of total household omission from partial household omission. Typically, total household omission results from a housing unit not being included in the Census Bureau's Master Address File.

In reporting their findings from the Causes of Undercount Survey Fein and West (1988) state that 50% of all "non-match" cases were ones where the address was not on the census address list, i.e. missing from the MAF.

In his detailed analysis of survey findings in his dissertation Fein (1989) reports the MAF omission rate of two different types of "low visibility" housing. Fein found that 6.7% of the housing units in the survey area were "strange/unusual" types of living accommodations (e.g. a garage, a toolshed) and that 17.6% of the units of this type were not in the MAF. This, then, implies that 1.19% of this type of housing unit was missed in the census. Another 8.9% of these housing units were "attached" low-visibility housing units and 23.5% were missed. Thus this type of housing unit accounted for 2.09% of the missing housing units. The aggregate total consisting of both types of "low visibility" housing units missing from the MAF, therefore, was 3.28%.

Fein discusses the findings from the "triple enumeration" analysis—of data from the test census itself, the PES component of the test census, and the "Causes of Undercount Survey" (CUS) which provides particularly good evidence of the extent of high immigrant undercount and reported aggregate omission rates of 15% for households where the respondent spoke little or no English, 11% for cases where the respondent had no high school education, and 13.8% for non-citizens.

Sherman et al (1997)

This report (Jennifer Sherman et al., "Finding Invisible Farmworkers: The Parlier Survey", California Institute of Rural Studies, 1997) analyzed observations of housing patterns from its high-quality street-level address canvassing/mapping of Parlier, a farmworker community in Fresno County.

The research team determined that 20% of "back houses" were not identified in Census Bureau records. When adjusting for vacancy rates in the front houses and low-visibility back houses, it appeared that 16% of the occupied housing units in Parlier were missed in the 1990 decennial census because they were not included in the MAF.

The analysis, presented by the research team as an ancillary finding from the California Agricultural Worker Health Survey, is useful in that it analyzed housing patterns down to the

block level and confirmed, as researchers had previously determined in Parlier that subtle differences in housing patterns make a substantial difference in MAF quality.

It also deserves note that the CIRS research team found that the families and individuals residing in the hidden "back houses" were more willing to respond to interviewers' questions than the occupants of the "front houses" because the "front house" landlords often were inclined to actively conceal the existence of these possibly illegal units. It also deserves note that the systemic bias inherent in omission of occupied low-visibility housing units created a significant bias in the demographic and socioeconomic profile of Parlier as a community.

Kissam and Jacobs (2004)

Research led by Ed Kissam, at Aguirre International, and Ilene Jacobs at California Rural Assistance (Edward Kissam and Ilene J. Jacobs, "Practical Research Strategies for Mexican Indigenous Communities in California Seeking to Assert Their Own Identity" in Jonathan Fox and Gaspar Rivera-Salgado, *Indigenous Mexican Migrants in the United States*, Center for U.S. Mexican Studies and Center for Comparative Immigration Studies, University of California, San Diego, 2004) focused explicitly on the differential undercount of migrant and seasonal farmworkers in California. It is particularly relevant to the undercount of immigrants because about 91% of California's farmworkers are Mexican or Central American immigrants.

California Rural Legal Assistance's research on census undercount in California farmworker communities in Census 2000 using an adaptation of the Census Bureau's alternative enumeration/ethnographic research methodology developed by Leslie Brownrigg at the Census Bureau's Center for Survey Methods Research. Five farmworker communities were included in the study. Kissam reported aggregate omissions, as had Fein and West, and categorized them either as "partial household omissions" (some individuals living at an address were not included in census responses) or "total household omissions".[7]

The households in the study blocks in these communities were almost all immigrant farmworker households and, thus, most of the omitted population were Latinos. Total household omissions ranged from 6.4% in Arvin (Kern County) to 27.6% in Parlier (Fresno County).[8] Ultimately, total household omission in the five rural communities in the study accounted for 58% of an aggregate undercount of 13.9% of the population in these farmworker communities.

Kissam (2012)

In a second collaborative coverage measurement study of census undercount in rural California conducted by California Rural Legal Assistance in collaboration with Aguirre International in connection with Census 2010 (Ed Kissam, "Census Enumeration of Immigrant Communities in

California: Dramatic Improvements but Challenges Remain", paper presented to the Conference on Hard-to-Survey Populations, American Statistical Association, December, 2012) the research team focused on 33 "hard to count" census tracts in eleven rural California counties in the San Joaquin Valley, Central Coast, and South Coast. It was estimated that total household omission accounted for about 7% of an aggregate undercount of about 9.9% of the study area population-- i.e. two-thirds of the aggregate undercount.[9] The methodology used in this study was not the ethnographic/alternative enumeration methodology but, rather, one adapted from the Census Bureau's post-enumeration survey.[10]

Although the "hard to count" census tracts in this Census 2010 coverage measurement study were more diverse and had less dense concentrations of immigrants than in our earlier research on Census 2000 undercount in farmworker communities, it included many immigrant households.   Three-quarters of the households in the study area (75%) were Latino households and 44% of all the households were farmworker households.

Nonetheless, the research team found in this rural area correlates of undercount similar to those reported by Fein and West. For example, the researchers found that 27% of the housing units categorized by field researchers as "unusual" had not been enumerated while only 3% of the households living in standard, usually single-family, homes had been missed. The study also found that 17.7% of the households where a recent immigrant was the census respondent had not been enumerated, although Census Bureau targeted mailings of Spanish-language census forms had dramatically improved response.

The Census Bureau's Alternative Enumeration/Ethnographic Research from the 1990 Decennial Census

The overall research design for this national initiative is described comprehensively by Leslie Brownrigg and Manuel de La Puente, the census bureau researchers who managed the effort (Brownrigg and De la Puente 1992). De La Puente provides a comprehensive summary of findings from the 29 Center for Survey Methods Research/Census Bureau alternative enumeration/ethnographic research community studies (Manuel de La Puente, "Using Ethnography to Explain Why People Are Missed or Erroneously Included by the Census: Evidence from Small Area Ethnographic Studies", Center for Survey Methods Research, U.S. Census Bureau, 1993).

Building on insights about differential undercount from earlier ethnographic research in inner-city African-American communities (Valentine and Valentine 1971) and ethnographic research in the 1986 Los Angeles Test Census (Vigil 1987), Brownrigg and De La Puente standardized the coverage measurement methodology used by independent researchers in the alternative approach to measuring census undercount. Several studies of these studies are particularly relevant to the challenges of successful LUCA-based MAF improvement, address canvassing in

general, and strategies to construct an improved MAF by adding low-visibility housing units which had previously not been included.

Alfredo Velasco details the dynamics through which crowded housing and irregular housing units interacted to give rise to undercount in a Latino neighborhood of San Diego (Sherman Heights) which he characterizes as a "pockmarked slum" (Velasco 1992).

Martin Dale Montoya in a study of Woodburn, OR, a farmworker town with a high concentration of undocumented indigenous immigrants, documented the consequences of the Census Bureau mis-classifying a motel as a "special place" when, in fact, it was the residence of numerous migrant farmworkers (Montoya 1992). He reports that the census enumerated only 44 of 92 farmworkers living in the motel. Montoya goes on to note that his findings are consistent with those from the other ethnographic alternative enumeration studies,

> The Ad Hoc household relationships found in Woodburn concur with the complex housing arrangements discussed in other census research: for instance, Rodriguez and Hagan (1991:10), identify 15 different "complex housing arrangements" in their study of Houston's Latino immigrants; Stepick and Stepick (1990:35) report 13 variations of "complex household composition" in their study of Haitians in Miami; Camilo Garcia Parra (1989:9), reports that "established Hispanics" in Washington State provide newcomers with housing; Victor Garcia (1992:12), in Guadalupe, (CA), also describes "irregular housing set ups" to accommodate temporary boarders.

Review of all 29 community ethnographic alternative enumerations (beyond the scope of this summary review) underscores the extent to which overall undercount varies from community to community. It is consistent with findings from later, more quantitative, analyses of housing unit omission in the MAF, for example, in observations on undercount in American Indian communities.

Census Bureau-Decennial Statistical Studies Division, 2010 Census Coverage Measurement Memorandum Series #2010-G-06 (2012)

This memorandum reports the Census Bureau's findings from a special analysis of its post-enumeration survey research in connection with the 2010 decennial census. It provides an excellent, up-to-date quantitative analysis of the extent of missing housing units in census address lists and tabulations of the extent of the problem in different contexts.

Another document emerging from the Census 2010 coverage measurement research has been confusing to some readers because the Bureau reports "net undercount of housing units", that is, overcount minus undercount. The memorandum discussed here is more useful because it breaks out the undercount of housing units, i.e. the proportion of housing units which were

omitted from the MAF, separately from those which were over-counted. The report shows that nationally 3.2% of all housing units (occupied and vacant) were missed and that 1.9% of all occupied housing units were missed. Moreover, it reports the proportion of housing units omitted for various strata. One stratum consists of states; the reported rate of housing unit omission in the MAF file in California overall (occupied and vacant units) is 2.2% and the proportion of <u>occupied</u> housing units missed is at least 1.3%.[11]

The report's tables include analysis of housing unit omissions by type. This analysis is useful as an element in identifying the census tracts where address canvassing, and specifically, LUCA-based local address canvassing, can have the greatest impact on MAF improvement. For example, as has been known for a long time, housing units occupied by renters are more undercounted than homeowners; it is reported that 2.5% of the renter-occupied housing units were omitted. Also useful is the analysis showing that housing units which were trailers were particularly likely to be omitted (8.9%) as were small multi-unit buildings (3.7%).

The report also breaks out omitted housing units by type of census enumeration area (mailout-mailback, update/leave, update/enumerate). Unfortunately, the update/leave and update/enumerate operations do not seemed to have worked as well as might be hoped in 2010—e.g. with 6.8% of the renter-occupied housing units in update/enumerate areas being omitted. This is, of some importance because these enumeration modalities do, in some areas, provide a means of overcoming underlying MAF problems.

Salvo and Lobo (2013)

Joseph Salvo of the New York City Planning Department (Joseph J. Salvo and Peter Arun Lobo, "Misclassifying New York's Hidden Units as Vacant in 2010: Lessons Gleaned for the 2020 Census", <u>Population Research Policy Review</u> (2013) 32:729-751) provides a very thorough and useful quick review of the literature on problems with address listing. Included below is an excerpt from the article since it summarizes the relevant research so succinctly.

> *Undercounts of the population became a major topic of discussion heading into the 1990 census, and subsequent results confirmed that fewer households than expected returned their forms (Anderson and Fienberg 1999). Questions also arose about the quality of the address list being used to conduct the enumeration (U.S. General Accounting Office 1992). Evaluation research found that a sizable number of people were not captured because housing units in small multi-unit buildings were missed (Childers 1993). Some research even found that ''whole household misses'' played a significant role in the widening gap in census coverage between white and black respondents (Brownrigg and Wobus 1993). Large ethnographic studies in a number of different localities cited ''irregular housing,'' consisting of informal conversions from single family to multi-family arrangements, as being one reason for the undercount of the population (Ellis 1995; de*

la Puente 1995).[3] Changes in the nation were producing new housing arrangements, some of which were increasingly difficult to detect and incorporate (Ericksen 2012). In an effort squarely aimed at improving the 2000 census, a report issued by the National Research Council's Panel on Alternative Census Methods challenged the Census Bureau to better incorporate small multi-unit buildings into the enumeration. The report identified ''structures that have units without unique labels or units that are not clearly distinguishable'' as a key problem in taking the census (Cohen et al. 1999).

Salvo's analysis addresses an important issue related to MAF improvement and the potential of LUCA to improve census quality which has not been addressed much elsewhere: the problems ensuing from failure to delete vacant units from the MAF. Salvo's analysis shows that failure to delete vacant housing units overloaded non-response followup (NRFU) as local census offices were confronted with an impossibly heavy workload. This is relevant because there is, in addition to Salvo's analysis, additional analysis showing that overly high NRFU workloads impaired census quality.

Salvo and Lobo, while focusing on MAF errors related to vacant units include very useful observations about the sorts of supplemental data sources that could be used successfully in LUCA and offer useful recommendations regarding the feasibility and utility of "continuous" MAF improvement and the training needed to make MAF improvement efforts work well.

GAO-03-60 Decennial Census: Lessons Learned for Locating and Counting Migrant and Seasonal Farmworkers (2003)

As noted above, because California farmworkers are almost all immigrants, the research on MSFW undercount is more broadly relevant to the problem of immigrant undercount. The report discussed in some depth the importance of adding low-visibility housing units to the MAF. The report is particularly useful in that it includes photos of examples of the types of housing units missed in rural communities in California.

The GAO report also includes some observations particularly relevant to the need for increased local government participation in the LUCA process. It notes, for example, that only 36% of local government entities eligible to participate in LUCA did so. Presumably, this stems in part from the fact that many did not have the resources to engage in even "office review" activities for local updates of census addresses, much less the in-field street-level address canvassing envisioned here.

The GAO report also confirms the viability of grassroots community groups' efforts in identifying low-visibility housing units which were not in the MAF. The report notes that 73% of addresses submitted to the Census Bureau's Los Angeles Region by CRLA in its 2000 census improvements were verified as valid "adds" to the Census Bureau's address list.

The Census Bureau's Comparative Ethnographic Studies of Enumeration Methods and Coverage across Race and Ethnic Groups for Census 21010 (Schwede and Terry 2013)

This Census Bureau's study of the 2010 decennial census included eight study sites, seven of which were chosen to examine the multiple factors involved in census errors (including omissions and incorrect enumerations). The study, however, has limited relevance to the issue of improving MAF quality because the methodology focused on in-person interactions (particularly in non-response followup) as a factor in erroneous enumeration.

The authors of this paper stress the insights from previous research that multiple factors interact in census undercount. Although there is no definitive report on the proportion of omitted housing units and the study relates primarily to in-person interviews (not the standard census response modality) the authors do note,

> *Of these broader categories [of erroneous enumeration], interviewer error accounted for 29 percent of the sources listed, unknowledgeable respondents for 12 percent of sources listed, and respondent concealment/refusal for 9 percent for Nonresponse Followup/Update Enumerate persons with at least one inconsistency, with other sources accounting for the remaining inconsistencies. A few persons may have been missed in "hidden housing units" on the same property.*

The authors also note the particular challenges involved in enumerating "complex" households (multiple family/social units in a single housing structure—e.g. recently arrived solo male immigrants living in crowded housing provided by *paisanos*). This issue, coupled with the problem of residents' "active concealment" of irregular, often illegal, housing arrangements which are prohibited under local zoning is related to the problem of MAF quality (since it is sometimes ambiguous under Census Bureau residence rules whether some sorts of living quarters are separate "housing units" or part of a single, larger housing unit. What is relevant to assessing MAF quality and completeness is that these sorts of addresses, most often a single address with hidden housing units at a single mail address, are actively concealed and, without special efforts, be omitted from the MAF.

SUMMARY CONCLUSIONS

The research literature indicates that omission of low-visibility housing units from the Census Bureau's MAF address list varies greatly from one community, neighborhood, census tract to another but that, statewide at least 1.3% of all <u>occupied</u> housing units in California were omitted from the Census 2010 MAF. The research also indicates that in "hot spots" where sub-standard, low-visibility housing units, and "unusual" living accommodations are prevalent, at least 3% and, in some circumstances many more, of the actual occupied housing units in a community are not included in the MAF.

This, in turn, confirms the utility and viability of a proactive LUCA-based initiative of local address canvassing so that local government entities, municipalities and counties, can, as part of their standard collaboration with the Census Bureau add these low-visibility housing units to the MAF.

Such efforts, if well-managed and adequately funded, can make significant improvements in the MAF as evidenced by the GAO review of farmworker low-visibility housing units proposed by CRLA to the Los Angeles Census Region. Almost three-quarters of the low-visibility housing units identified by CRLA community workers as additions to the MAF were subsequently determined to be valid.

The research shows that LUCA-based local address canvassing is a viable strategy for MAF improvement in both urban and rural California communities—because neighborhoods in both have neighborhoods which are similar in that low-visibility housing is prevalent. This strongly suggests that it will be very important to expand the number of municipalities and counties participating in LUCA in order to optimize the positive impact of MAF improvement on the overall count of minorities and immigrants in California. Large municipalities (e.g. Los Angeles) have participated actively in LUCA since this component of the decennial census process was first introduced but small communities have not. Initiatives to incentivize their participation and provide them with technical assistance and financial assistance in LUCA-based local address canvassing will be an important element in California's efforts toward a "California Complete Count".

Unfortunately, there is good reason to believe that the quality of the Census 2020 Master Address File may be worse than in previous decennial censuses—because funding constraints have forced the Census Bureau to cut back on its own address canvassing efforts.[12] Locally-driven efforts to collaborate with the Census Bureau in improving its address lists in areas with concentrations of low-visibility housing should give special priority to identifying and adding low-visibility housing units where immigrants live—because the process of immigration is so closely intertwined with "unusual" and low-visibility housing accommodations. This should include areas known to have concentrations of farmworkers since almost all are immigrants and about two-thirds are undocumented.

It is likely, also, that continued widening of the economic gap between less-educated and limited-English Mexican and Central American immigrants and "mainstream" residents in California communities will make Census 2020 particularly problematic. Especially when one assesses the income gap between households headed by immigrants and US-born heads of household using the supplemental poverty measure developed by the Census Bureau and refined for use in California by the Public Policy Institute of California, it appears like that there will be more crowded and "unusual" housing arrangements in 2020 than today even though the overall economy has improved greatly since 2010.

END NOTES

[1] See Chapter 4, David Fein dissertation (1989), especially Tables 4.4 and 4.7. Table 4.7, for example shows that in the Los Angeles study area that 19.5% of recently-arrived immigrants (<5 years in U.S.) were missed, 17.2% of those in the U.S. for 6-10 years, and 9.9% of those who had resided in the U.S. for 20+years.

[2] The underlying core source for the Census Bureau's address list is the U.S. Postal Service mail address list. An immediate problem vis-à-vis immigrants is that the most recent arrivals must rely on PO boxes to receive their mail.

[3] Census Bureau, "2010 Census Coverage Measurement Housing Unit Results for California". This tabulation includes all housing unit omissions, including housing units which may be vacant. Other tabulations of housing unit omission (e.g. by the type of structure the housing unit is in Table 2, Census Bureau 2012 Memo) show that 1.9% of occupied housing units are missed. Joseph Salvo argues that census quality is also negatively affected by failure to identify vacant housing units (because this results in unmanageable workload in the post-census non-response followup period).

[4] For example, occupied housing units in small multi-unit buildings with 2-9 housing units are omitted at a rate of 3.6% (for renters) and 3.7% (for owners).

[5] A key factor in optimizing the efficacy of targeted local address canvassing is identifying the neighborhoods where housing patterns most impair the Census Bureau's address listing and focus on those areas. Modelling of the likely impact of LUCA-based local address canvassing to improve MAF quality necessarily needs to include assumptions as to the possible amount of improvement. The estimate presented here of the eventual impact of LUCA-based MAF improvement acknowledges that the MAF cannot be perfected in each of California's 7,500 census tracts but assumes that the 20% (1,500) of California census tracts with the most serious problems of MAF quality can be identified and that the local government/grassroots community groups can rely on local knowledge of housing conditions to further target actual street-level canvassing.

[6] This research is most thoroughly detailed in two major research documents. The first is a short paper by David J. Fein and Kirsten West, "The Sources of Census Undercount: Findings from the 1986 Los Angeles Test Census", *Survey Methodology*, December, 1988. The second much more detailed discussion is in David Fein, "The social sources of census omission: Racial

and ethnic differences in omission rates in recent censuses", Ph.D., dissertation, Princeton University, 1989. Fein's dissertation includes, in addition to findings from the Los Angeles TARO a comprehensive review of previous research on differential undercount. Findings from the ethnographic/alternative enumeration studies are summarized in Manuel de La Puente, "Using Ethnography to Explain Why People Are Missed or Erroneously Included by the Census: Evidence from Small Area Ethnographic Studies", Center for Survey Methods Research, U.S. Census Bureau, 1993. De La Puente's summary includes a bibliography listing all 29 studies. Studies which are particularly relevant to MAF improvement include those by: Pamela Bunte and Rebecca Joseph (on Cambodians in Long Beach), Boanerges Dominguez and Sarah Mahler (on Mexicans in the South Bronx), Dale Martin Montoya (on farmworkers in Woodburn, OR), Nestor Rodriguez and Jacqueline Hagan (on Guatemalans in Houston), and Alex Stepick and Carole Stepick (on Haitians in Miami). Philippe Bourgois's research on inner-city attitudes about census response are particularly relevant to understanding the challenges of MAF improvement in neighborhoods with actively concealed irregular housing arrangements.

[7] Ed Kissam and Ilene Jacobs, "Practical Research Strategies for Mexican Indigenous Communities Seeking to Assert Their Own Identity", in Jonathan Fox and Gaspar Rivera-Salgado (Eds.) *Indigenous Mexican Migrants in the United States*, Center for U.S.-Mexican Studies, 2006. The research strategy in this study was adapted from the methodology used in the Census Bureau's Center for Survey Methods Research ethnographic alternative enumeration studies developed by Leslie Brownrigg and utilized in the 29 ethnographic studies in the 1990 decennial census.

[8] A resident ethnographer (Aline Doignon) and consulting anthropologist (Anna Garcia) from Ed Kissam's New Pluralism research team subsequently mapped Arvin block-by-block in 2003 in preparation for a USDA-funded study of immigrant settlement in rural areas of the U.S.--the New Pluralism Study. Their observations in the course of mapping Arvin in 2003 using the same methods we had used to map Parlier in 1989 explain the relatively low rate of total household omissions observed in Arvin in the earlier study of Census 2000 undercount. The town has remarkably orderly housing and represents a better-than-average scenario for rural MAF development. Ed Kissam, "Migration networks and processes of community transformation: Arvin, California and Woodburn, Oregon", *Journal of Latino and Latin American Studies*, 2007.

[9] Ed Kissam, "Census Enumeration of Immigrant Communities in California: Dramatic Improvements but Challenges Remain", Conference on Hard-to-Survey Populations, American Statistical Association, December, 2012. The methodology used in this study was an adaptation of the Census Bureau's post-enumeration survey approach.

[10] The CRLA study team, unlike the Census Bureau, did not have access to definitive evidence as to whether a household had been enumerated in the decennial census or not in every case. However, the field researchers definitively resolved the question as to whether the household

which had not received a form or who had received a form and not returned it had been enumerated or not for 95% of the cases by re-interviewing those who said they had been missed late after the non-response followup process was complete.   The team then imputed the enumeration status of the balance where the CRLA team's own followup was unsuccessful. These analytic considerations in this process are discussed in the paper.

[11] The tabulation of the national rate of omissions for housing units by tenure (owner-occupied, renter-occupied, and vacant) from the 2010 housing unit coverage measurement paper indicates that about two-thirds of the housing unit omissions are vacant units. In its discussion of tabulations for different strata the paper does not provide direct tabulations of the rate of omission for occupied and unoccupied housing units but this ratio can be inferred from the tabulations of sub-types. For example, overall 8.9% of housing units which were trailers were missed, while 4.9% of owner-occupied trailers were missed and 5.8% of renter-occupied trailers were missed.

[12] The Census Bureau's Federal Register notice of its address canvassing operations is optimistic that satellite imaging and administrative datasets will provide improved ways of correcting deficiencies of the MAF. Unfortunately, these sorts of high-tech solutions do very little to address the challenge of identifying low-visibility housing accommodations since some (e.g. recent immigrants housed in a garage) are obviously invisible in satellite imagery and some of the sorts of administrative records envisioned as supplemental data sources for MAF improvement (e.g. TANF records) are obviously incomplete since undocumented immigrants are not authorized to receive services and, thus, are not included in such records.