Differential undercount of Mexican immigrant families in the U.S. Census

Edward Kissam
Werner-Kohnstamm Family Fund, 4810 Trinidad Ave., Oakland, CA 94602, USA
Tel.: +1 510 482 9979; E-mail: edkissam@me.com

Abstract. The U.S. Census is well known to have been compromised by differential undercounting of various populations, in particular, ethnic minorities and immigrants. This study focuses on the undercount in the Mexican immigrant population through a review of both quantitative and qualitative research. It challenges some traditional assumptions, such as that the Hispanic population is homogenous and that Census Bureau estimates of its undercount are valid for all sub-populations of Hispanics. It identifies a key reason for the undercount of the Mexican immigrant population, low-income households living in unusual and/or concealed housing units that are not in the Census Bureau’s Master Address File (MAF) and are unlikely to be enumerated or subsequently captured in the bureau’s Non-Response Follow Up (NRFU). Indeed, it seems probable that non-standard housing conditions in the local community account for one-third to one-half of the total undercount of minorities and immigrants, playing a greater role in undercount than traditional explanations such as lack of respondent motivation or inability to respond due to language or literacy constraints. Recommendations are offered to ameliorate the chronic differential undercount of this population, including employing a multi-variable analysis of census data, using community-based local address canvassing, and revising Census Bureau data processing procedures.

Keywords: Census, differential undercount, Mexican immigrant population, American Community Survey, Master Address File, low-visibility housing

1. Introduction

The constitutionally required count of the entire population of the United States by the Census for the purpose of congressional reapportionment also provides crucial guidance necessary for sound planning and equitable, population-based allocation of funding for national, state, and local efforts to improve individual, family, and community well-being. In FY 2015, for example, more than $589 million in federal funding was distributed based on statistical information derived, at least in part, from Census Bureau data [1]. Unfortunately, the accuracy and utility of these data, both in the decennial censuses and in the American Community Survey (ACS), the premier source for detailed information about the American people and the country’s workforce, are undermined by persistent problems of differential undercounting of ethnic minorities and immigrants. These persistent undercounts raise difficult questions about why it occurs and are troubling due to their practical implications for the conduct of public policy.

A key consideration is the evidence that the most societally marginal individuals and households are those most likely to be undercounted. The relationship between social marginality and census undercount has been demonstrated in many analyses using very different methodologies, including dual-system estimation, demographic analysis (DA), and ethnography [2,3]. Differential undercount jeopardizes the Census’ aspiration of generating statistical data that provides “a mirror of America”. Exploring how the interactions between day-to-day life in undercounted communities and census systems and procedures give rise to differential undercount is important as a basis for devising effective strategies to ameliorate the problem.

In an era when economic disparities continue to increase, assuring that social program resources are equitably allocated is especially important. Moreover, when many elections are decided by slim margins of
a few percentage points, having accurate information to identify racial disparities in redistricting is crucial for maintaining the integrity of the democratic process. Undercount is most problematic when it generates blurred portrayals of the contour of a sociogeographic landscape which is, in fact, pitted with deep pockets of undercounting. These problems are most serious in analyses of a relatively small sub-population which is systematically undercounted (e.g. “high risk children”) or when the specific population in a local geographic area – a state, county, or city – is more likely to be undercounted than others.

I review a range of factors which contribute to differential undercount in a particular minority – about 11.7 million Mexican immigrants [4] and their 6.7 million U.S.-born children [5], focusing on the ways that low-income immigrant families’ residing in low-visibility or “unusual” living quarters contributes to differential undercount I then go on to discuss ways in which the dynamics of other facets of census operations beyond address listing – non-response followup, data editing and imputation procedures, impact efforts to accurately estimate the magnitude of differential undercount in households headed by Mexican immigrants.

Practically speaking, decennial census data are especially important for programs targeted to children (since these data impact planning and allocation of federal funding for an entire decade). Even within the overall undercount of young children, there are racial disparities since minority children are disproportionately undercounted – e.g. Latino children aged 0–4 being 7.5% undercounted and Black children 6.3% undercounted while non-Hispanic and non-Black children in this age group are undercounted by 2.7% [6]. The undercount rate for these minority children is more than double the rate for the non-Hispanic and non-Black or other children.

The differential undercount of young minority children is, for many reasons, a particularly pressing and growing concern.

2. Current context and implications for Census 2020

Looking forward to Census 2020, there are many reasons for concern. A particularly important one is the extent to which the Census Bureau’s reduced budget and cutbacks in address canvassing operations will affect the quality of the Bureau’s Master Address File (MAF). The Census Bureau’s operational decision to save money by going from 100% address canvassing to 25% canvassing is reasonable as a response to budget constraints – particularly if the areas to be canvassed are well-targeted. However, the concomitant shift in address canvassing procedures to place more emphasis on “in office” address canvassing than “in field” address canvassing may turn out to be problematic.

Recent changes in the national sociopolitical environment and administration messaging about immigrants and immigration policy can be expected to negatively impact efforts to survey Mexican immigrants – now and in Census 2020. Whether or not administrative initiatives to deport undocumented immigrants from 2017 through 2020 are effective, the genie is now out of the bottle. By the spring of 2017 there was already evidence of reluctance by some undocumented and mixed-status families in immigrant communities to engage in any civic interactions (including sending their children to school, seeking medical care, or applying for other program benefits which are not conditioned on immigration status).

This development is particularly worrisome with respect to the differential undercount of Mexican immigrant-headed households because at least 6.1 million of the Mexican immigrant population is unauthorized [7]. The effects of anti-immigrant messaging on immigrant communities’ and individuals’ willingness to participate in Census 2020 by “providing information to the government” will vary from state to state and community to community because immigrants’ perceptions about the safety and desirability of civic participation are modulated by local attitudes and can, therefore, be offset to some extent by well-publicized and forceful state and local government commitments to protecting immigrants from deportation and immigrant social and civic integration.

Although the Census Bureau’s commitment to protect the confidentiality of individual census responses guaranteed by Title 13 is unquestionable, widespread concerns persist, and may well escalate, especially if there is renewed discussion of the addition of a question on immigration status to the decennial census or the ACS. It is likely that local government and community-based organizations with a demonstrable track record of defending immigrants’ rights and integrating them into community life can, by explaining the stakes for community well-being, persuade even apprehensive families to respond to the census. Although it is unfortunate that this issue needs to be addressed at all, it is now inevitable that, without adequate attention to pro-census messaging, differential undercount will escalate to extremely high levels.
A final, important factor in the undercount of Mexican immigrant families and their children in the 2020 decennial census will be the pace of the trend toward increased economic inequality, coupled with rising housing costs – especially in gentrifying areas of cities. If the current trend continues, cost of housing will continue to take up a greater and greater share of “low-skill” working families’ earnings and the likely result will be more immigrant families living in marginal housing conditions – in low-visibility concealed housing units and/or crowded living accommodations with more complex households consisting of “doubled-up” sub-families. These trends will affect all “low skilled” (i.e. less-educated) workers but particularly the most recently-arrived unauthorized Mexican and Central American immigrant workers and their families, who have settled in urban areas where inner-city gentrification has led to skyrocketing costs for rentals.

For all these reasons, there is a substantial risk that Census 2020 will be of lower quality than any previous decennial census since 1970. Even more important than the precise extent of national undercount are the geographical and sociological variations in undercount, driven by the system interactions in different areas which give rise to differential undercount of low-income minorities and immigrants, especially families with children. These disparities have a huge potential impact on empirically-based rational social policy, including provisions for delivery of health, education and social services – due in large measure to census-driven data being used for allocation of federal funding to states, but also as a result of state and local government use of census data to allocate funding and guide local service delivery and policies.

### 3. Methodology

The analysis presented here rests primarily on review of both quantitative and qualitative research on differential undercount over the past three decades. The most important underlying insights stem from the detailed analyses of quantitative data from the Causes of Undercount Survey conducted in the Los Angeles basin in 1987 in conjunction with the test of census-related operations for the 1990 Decennial Census [8]. These findings are consistent with numerous subsequent ethnographic studies of differential undercount conducted as part of the 1990 census conducted by the Census Bureau as part of its “alternative enumeration” program [9]. It has been necessary to use these older data analyses because, as a result of the Census Bureau’s budget limitations and litigation regarding its procedures for statistical adjustment of census data, the most recent analyses of differential undercount are not as useful as earlier ones. Generally speaking, coverage measurement studies in 2010 focus on more limited correlates of undercount than those examined in previous research.

### 4. Findings

#### 4.1. The Census Bureau’s reliance on single-variable analyses of net undercount is increasingly inadequate for capturing the size and profiling the characteristics of the Mexican immigrant population

The U.S. Hispanic population was estimated to be 52.7 million in 2012 of whom about 11.4 million were foreign-born Mexicans [10]. Another 6.7 million U.S.-born children live in households with a Mexican-born parent [11]. Therefore, about 18 million, or slightly more than one-third of the entire U.S. Hispanic population, live in households headed by Mexican immigrants. These Mexican immigrant households are predominantly “mixed-status” families since naturalization rates are relatively low among Mexicans, because relatively few post-1986 Mexican immigrants have had opportunities to secure legal status and since most households include U.S.-born children.

The Census Bureau’s standard coverage measurement reports consist of a series of single-variable models (e.g. race or Hispanic origin, tenure, age, gender, type of enumeration area), each of which is known to be correlated with undercount. But the analysis is inadequate both by virtue of using a single variable at a time and because it fails to include many social system variables known to be correlated with risk of undercount.

While the proportions of individuals omitted from the decennial census is generally believed to have decreased from 1970 through 2010, it is likely that the longstanding social system correlates of census undercount continue to powerfully affect likelihood of enumeration [12]. The relative importance of each specific factor can be expected to vary from census to census and from community to community. But the types of interactions between census operations and social system dynamics which affect undercount in ways de-
Table 1
Profile of U.S-born, foreign born Mexican immigrants, and Mexican immigrant farmworkers in California

<table>
<thead>
<tr>
<th>Correlate of undercount</th>
<th>US-born MX-origin (22.3 million)</th>
<th>MX immigrants (11.4 million)</th>
<th>MX immigrant MSPW’s in California</th>
</tr>
</thead>
<tbody>
<tr>
<td>Limited-English</td>
<td>11%</td>
<td>29%</td>
<td>91%</td>
</tr>
<tr>
<td>Less than high school education</td>
<td>21%</td>
<td>59%</td>
<td>90%</td>
</tr>
<tr>
<td>Home ownership</td>
<td>53%</td>
<td>45%</td>
<td>18%</td>
</tr>
<tr>
<td>Unauthorized legal status</td>
<td>0%</td>
<td>50%</td>
<td>67%</td>
</tr>
<tr>
<td>Poverty rate</td>
<td>26%</td>
<td>29%</td>
<td>31%</td>
</tr>
</tbody>
</table>

Author’s estimate based on analyses of the profile of US-born and foreign-born Mexican-origin Hispanics from the Pew Research Center (Gonzalez-Barrera and Lopez 2013), an estimate of the numbers of Mexican-born unauthorized immigrants from USCIS (Baker and Rytina 2013) and California farmworker characteristics from an earlier analysis by the author of National Agricultural Worker Survey data (Williams and Kissam 2013).

A novel factor which may have a significant impact on specific interactions between social system factors and census operations is the Census Bureau’s reliance on online response as the primary mode of census response in 2020. Specific changes in social system factors (e.g., prevalence of low-visibility irregular housing) and census operations (e.g., changes in address canvassing procedures, online modality of census response) will also affect enumerability but the types of interactions among multiple correlates of undercount described by Fein and West and others will remain in force.

4.2. Accurately calculating the undercount needs to take into account subpopulation characteristics

The Census Bureau’s official census coverage measurement report of population undercount based on dual-system estimation (DSE) reported a net undercount of 1.54% for the overall U.S. Hispanic population [13]. An immediate problem with this analysis is that it steers data users toward a de facto assumption that the Hispanic population (or individuals and households in any other OMB-defined racial group) is homogeneous and that the reported national estimate of net undercount for the overall group represents the actual undercount for all sub-populations of Hispanics.

Socioeconomic and ethnic diversity among Hispanics makes the reliability and relevance of the national-level “generic” estimate of the undercount of the Hispanic population in state or local planning which relies on census data questionable. Although the population of Mexican immigrant families is large, their socioeconomic profile differs greatly from that of the overall Hispanic population. Consequently, the Hispanic population is, in fact, heterogeneous with respect to the characteristics correlated with undercount. So one cannot use a single variable analysis as a reliable explanatory factor in understanding the dynamics of undercount across the highly diverse Hispanic population.

Also relevant here is the growing research literature on the perspectives, educational experiences and outcomes of subsequent generations of Mexican, and other, immigrants, including the ways in which educational attainment affects labor market experience [14]. This body of research emphasizes multiple dimensions of diversity within the overall Hispanic population, most of which affect enumerability.

Table 1 shows some of the most significant differences among sub-populations within the overall population of Hispanics and Mexican-origin individuals with respect to known correlates of differential undercount.

Table 1 underscores the problematic nature of using the quasi-racial variable of Hispanic origin as the single basis for estimating the undercount of families in Mexican-origin subgroups since the known social system correlates of undercount vary so much from sub-population and can be quite different in marginal groups from those affecting enumeration of the population at large.

4.3. Studies on the causes and ethnographic research yield comparable estimates of the undercount of Mexican immigrant-headed households

David Fein’s analysis of undercount in the 1986 Los Angeles Test Census in an inner-city area with a population that was 72% Hispanic (primarily of Mexican origin), 13% Asian or Pacific Islanders, and 14% non-Hispanic Whites provides the best available source of insights related to the extent of undercount in relation to each of the above characteristics of Mexican immigrant households and others [15]. Fein reports that 15% of households where the respondent was limited
in English were omitted, that 11.2% of those with no high school was omitted, and that 13.9% of those living in a rental unit were omitted.

Fein’s analysis suggests that the differential undercount of the “minority within a minority” of Mexican immigrant households is greater than generally recognized, that living in low-visibility housing is a leading factor of the undercount, that this results in systematic skewing of analyses of the socioeconomic profile of the population, and that this disproportionately affects some cities and counties – since immigrant settlement is typically clustered in well-established migrant destinations due to migration dynamics.

Fein did not analyze census omissions based on immigrants’ legal status but did report that the omission rate was 13.8% for non-citizens and that omission rate varied in relation to length of time in the U.S. Immigrants who had been in the U.S. only 5 years or less were omitted at a rate of 19.5% while those who had been in the U.S. 12 or more years were omitted at a rate of 9.9%. Fein’s analysis led him to conclude that ethnic enclosure and recency of immigration – what I refer to as “structural” (neighborhood/community) characteristics – were among the most powerful predictors of both total and partial undercount, although other social system causes of undercount continued to be important also.

A 1992 analysis prepared by Manuel De La Puente based on his review of findings from nine of the ethnographic research/alternative enumeration studies conducted by the Census Bureau’s Center for Survey Methods Research as part of the 1990 decennial census shows the extent to which interactions between ethnicity and local community housing conditions affect the undercount of diverse Hispanic sub-populations in different localities [16]. De La Puente reports that ethnographic studies showed that census omission of Hispanics varied from 5% (in New Orleans, LA) to 60% (in Long Island, NY). His summary of the results shows that 14.4% of the Mexican Hispanics were missed by the census.

A subsequent analysis from De La Puente’s analysis is one of the few that report a specific omission rate for Hispanic children aged 0–6; his tabulation shows that 17.6% were not enumerated – more than half (57%) of them being missed as a result of total household omission [17].

## 4.4. Mexican immigrant farmworkers living along the Pacific seaboard have been seriously undercounted in two censuses

Mexican immigrant farmworkers living along the Pacific Seaboard are a well-studied subpopulation among Mexican immigrants overall and a subject of long-standing interest to the author. There is abundant evidence that they were undercounted in the 2000 and 2010 decennial censuses [18,19]. The extent of undercount within this socially and economically disadvantaged sub-population of predominantly Mexican migrants underscores the fact that system factors, not race or ethnicity per se or the arbitrary construct of “Hispanic origin”, are the primary causes of undercount.

Virtually all migrant and seasonal farmworkers in California (95%) are Mexican immigrants [20]. They are a “small” group in terms of the national population but a large one regionally in California: 630,000 workers and an additional 820,000 non-working dependents [21]. Another 400,000 farmworkers and family members very similar to the California farmworker population live in Oregon and Washington. Thus, this Pacific Seaboard population of migrant and seasonal farmworkers is about 1.8 million. About two-thirds of this population is made up of family households with children. Therefore, estimates of farmworker undercount are a useful indicator of the extent of undercount within a very low-income sub-population of Mexican immigrant families.

The author’s study of census undercount in California farmworker communities in the 2000 decennial census used an adaptation of the Census Bureau’s alternative enumeration methodology [18]. Five farmworker communities were included in the study. Aggregate omissions were categorized, as Fein and West had done, either as “partial household omissions” (some individuals living at an address were not included in census responses) or “total household omissions”. The study found a 13.9% overall population undercount. However, as was the case in the Census Bureau’s ethnographic studies, the research team found substantial variations in undercount from community to community as well as in the extent to which total household omission vs. partial household omission contributed to undercount, although each of the study neighborhoods was made up of predominantly Mexican immigrant farmworker families.

A subsequent coverage measurement study of census undercount in rural California conducted in 2010 was based on a survey of 423 households conducted
in 33 “hard to count” rural census tracts in agricultural, predominantly Hispanic immigrant areas of ten California counties in the San Joaquin Valley, Central Coast, and South Coast [19]. The approach used in this study was adapted from the Census Bureau’s post-enumeration survey. Three-quarters of the households in the study area were Hispanic households and included many migrants and almost half (44%) were farmworker households. This study estimated there was an aggregate undercount of about 9% of the overall study area population (including Hispanic and non-Hispanic households and excluding an estimated 1% homeless who were not enumerated).

4.5. Community housing context powerfully affects the undercount of Mexican immigrant families

Two facets of housing conditions are linked to undercount. The first is living in a housing unit which is “low visibility”, “irregular”, or “unusual” in terms of Census Bureau address listing which leads to total household omission. The second is living in crowded housing. When there is a “complex” household, i.e. multiple families or social units living in housing classified as a single housing unit, where there are, functionally, multiple households, the more peripheral social units/households and individuals are most likely to be omitted.

The Los Angeles research, the alternative enumeration/ethnographic research and the farmworker research make it clear that sub-standard housing conditions – both irregular low-visibility housing units and crowded housing – are most prevalent in low-income neighborhoods with high levels of ethnic enclosure where planning/zoning and housing code regulations are least likely to be enforced.

It seems probable that housing conditions in the local community account for one-third to one-half of the total undercount of minorities and immigrants – because lower-income households are much more likely than affluent households (which are, typically, overcounted) to live in low-visibility or hidden housing units not included in the Master Address File (MAF), the Census Bureau’s address list. Consequently, even if national-level estimates of differential undercount were sound, they would not be reliable in many states, counties, communities, or neighborhoods.

This is a problem common to all low-income neighborhoods, but it is most serious in those with high concentrations of immigrants because it is common practice for more settled waves of immigrants to facilitate “irregular” or “unusual” housing accommodations for relatives for fellow migrants in their hometown migrant-sending networks [22]. Undercounting is already likely in communities and neighborhoods which are migration destinations for Mexican and Central American immigrants because the vast majority of these newly-arriving migrants face serious constraints on their labor market and housing mobility, due to lack of legal status, limited-English, and low levels of educational attainment. It is likely that the maturity of a migration network and level of current migration, as well as prevalence of immigrants, also affects undercount by facilitating very low-income immigrant-headed households’ ability to secure “unusual” housing accommodations.

4.6. Low-visibility housing results in total household omission and is a leading cause of undercount

Living in a low-visibility housing unit (e.g. for immigrants in urban areas, living space above a commercial establishment, a basement, for those in suburban areas or rural areas, a garage or backyard shed, a barn, camper, or trailer) leads to total household omission because the housing unit is not included in the MAF. There is very little chance that housing units not included in the MAF can be enumerated in the course of the Census’ Non-Response Follow Up (NRFU); almost all result in total household omissions.

A substantial amount of the overall undercount of low-income minority households, including Mexican and Central American immigrant households stems from total household omission. In presenting their findings from the Los Angeles Causes of Undercount Survey Fein and West report that 50% of all “non-match” cases (i.e. those including omissions and erroneous enumerations, were ones where the address was not on the census address list, i.e. missing from the Master Address File (MAF) [23]. The Mexican and Central American households are particularly likely to live in low-visibility housing because the breadwinners’ (typically both husband and wife’s) labor market mobility is limited due to low home-country educational attainment, limited-English, and lack of legal status, all factors which leave them stranded in low-wage, often unstable or seasonal immigrant-dominated occupations such as construction, restaurant service, housecleaning, home health care.

In his detailed analysis of the Los Angeles Causes of Undercount Survey findings Fein reports the MAF omission rate of two different sub-types of “low visi-
bility” housing [8]. 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 housing units of this type were not in the MAF. This, then, implies that 1.19% of this sub-type of housing unit was missed in the census. Another 8.9% of the housing units in the survey area were “attached” low-visibility housing units and 23.5% of these were missed. Thus, this second sub-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%. None of the individuals living in these housing units can be enumerated. Because their living quarters were not in the MAF they do not receive a census form and the NRFU addresses only the problem of census forms which are mailed to a household listed in the MAF but not returned.

Therefore, the actual undercount of both housing units and persons in these sorts of neighborhoods is likely to be at least 3% higher than the official coverage measurement estimate based on DSE (if one assumes that the missed housing units have about the same number of household members in them as those that are enumerated).

The same “structural” issues relating to housing arrangements and housing stock affect census undercount in rural areas with high concentrations of immigrants. A research team conducting the California Agricultural Worker Survey analyzed observations of housing patterns from its high-quality street-level address canvassing of Parlier, a farmworker community in Fresno County, in relation to census undercount [24]. They determined that 20% of all “back houses” (informal housing units constructed behind a main house) were not identified in Census Bureau records. When adjusting for vacancy rates in the front houses and low-visibility back houses, the team concluded that 16% of the occupied housing units in Parlier were missed in the 1990 decennial census because they were not included in the MAF. Their analysis is useful because it analyzed housing patterns down to the block level and showed how local neighborhood-level differences in housing conditions make a substantial difference in MAF quality.

My subsequent research in 2010 in 33 rural California hard-to-count tracts [19] identified a correlation between type of housing unit and total household omission similar to that reported by Fein and West in the urban Los Angeles basin. Overall, 7.2% of the surveyed residences lacked standard mail delivery. Some-
National Agricultural Worker Survey data provide an indication of the extent to which families in at least one major sub-group of recent Mexican immigrants with children may be at risk of total household undercount as a result of living in a trailer which is not included in the MAF. In 2011–2012, one-third of farmworker families with children who had lived in the U.S. for 5 years or less (32%) were living in a trailer. In contrast, only 8% of the overall U.S. population lives in a trailer. However, the proportion of the rural Mexican immigrant population (the sub-group for which we have relevant data) living in trailers varies greatly from community to community. In a 2007 report, JBS International found that nationally 15% of farmworkers living off their employers’ farms lived in trailers and 21% of those living on a farm lived in them [27].

Community surveys the author has conducted in the course of research on Mexican immigrant settlement throughout the rural U.S. showed that, in four of twelve different migrant-receiving communities, trailers were the predominant mode of housing for immigrants, all in the Southeastern U.S. In Adel, GA, for example, 45% of Latino (mostly Mexican) immigrant families surveyed lived in trailers and three-quarters of those families had children living with them. Of the immigrant households with children living in trailers, 83% included pre-school age children.

Although trailers are particularly common in the Southeastern U.S., it is also clear that there are other communities in the Western U.S. where clusters of unpermitted mobile home parks have become a common type of housing for Mexican immigrants. The author’s field research in rural communities over the years suggests that these informal clusters of trailers often lack mail delivery and are, therefore, at very high risk of census undercount.

4.8. Crowded housing is a factor in partial household omission and contributes to differential undercount in immigrant neighborhoods

The ethnographic and survey research indicates that the most important factor in partial household undercount is crowded housing, often resulting from families “doubling up” into complex households with multiple sub-families. This is common in low-income communities generally but particularly notable in Mexican immigrant communities.

Newcomers are, in most, though not all, circumstances provided housing by relatives, friends, or individuals from their hometown village migration network. In these circumstances, the individuals left off the census roster are the members of a distinct social/family unit who are sharing housing accommodations with a primary nuclear family. They are not included in the census respondents’ tally – usually because they are not part of the family who is considered to be living at the place.

A 2015 report from the National Research Center on Hispanic Families and Children provides valuable insights about the prevalence of these crowded complex households [28]. The analysis shows that, although Hispanic children with a foreign-born parent are not more likely to live in a household with sub-families than other Hispanic children, more of them do live in a crowded household – double the proportion of Hispanic children with U.S.-born parents (21%) and four times the proportion of White children (10%).

4.9. There are community-level variations in the ratio of total/partial household undercount

The research suggests that the ratio of total/partial household omissions of undercounted populations varies a good deal from community to community. De La Puente’s analysis of the Census Bureau’s alternative enumeration studies of Hispanic neighborhoods in the 1990 decennial census provides insights into community-to-community variations in the ratio of total household omissions to partial household omissions vary. De La Puente reports that the proportion of all individuals omitted from the census a result of total household omissions ranged from 2% in Guadalupe, California (in a neighborhood of settled Mexican immigrants living in single-family homes) to 60% among Salvadorans in Long Island, NY [29].

Total household omission was a more important factor than partial household omission in six of the nine Hispanic alternative enumeration areas analyzed by De La Puente. The ratio of total household omission/partial household omission varied but accounted for most of the aggregate population undercount in the Hispanic alternative enumeration areas: 83% in Long Island, 80% in New Orleans, LA, 79% in Sannturce, Puerto Rico, 75% in San Francisco, CA, 71% in Woodburn, OR, and 52% in Bronx, NY.

The author’s field research in a study of undercount in Census 2000 in diverse farmworker communities and neighborhoods in California provides some useful insights about the dynamics underlying these community to community variations in the ratio of total/partial household omission [18]. Total household omissions
ranged from 6.4% in Arvin (Kern County, CA) to 27.6% in Parlier (Fresno County, CA). Ultimately, total household omission accounted for 58% of an aggregate undercount of 13.9% of the predominantly Mexican population in these five rural farmworker communities. The extent and type of overall undercount in different communities was clearly correlated with distinctive local patterns of residential housing accommodations among immigrant migrant and seasonal farmworkers.

In Oxnard, CA, where there was a 34% overall undercount, about two-thirds of the aggregate undercount stemmed from partial household omission, in part because immigrant families in the study neighborhoods had built multiple additions to what were originally very small, low-cost single-family houses in order to house arriving relatives and friends from their hometown village and, in part because a popular ultra-low rent apartment housing complex was extremely overcrowded and census respondents didn’t include everyone in their housing unit on the household roster.

In contrast, in Parlier, CA, there was a similar aggregate population undercount (37.6%) but the total/partial household omission ratio was reversed – with two-thirds of the undercount stemming from total household omission due to the prevalence of actively concealed “back houses” as had been described in previous years by the research team for the California Agricultural Worker Survey.

At the same time, Arvin, CA, a community which was socioeconomically and demographically very similar to Parlier and Oxnard, had only an aggregate undercount of 13.8%, about half from total household omission, the other half from partial household omission, because most housing consists of small single-family dwellings with well-marked street addresses.

These observations are interesting in relation to O’Hare’s top-down analysis of patterns of undercount at the state and county level [30]. O’Hare reports a strong correlation (~0.76) between a state’s racial/ethnic composition and undercount of Black and Hispanic children ages 0–4. He observes that this relationship is not so powerful when looking at all counties but is more pronounced in larger, urban counties with large minority populations with a Census 2010 undercount in the 10%–13% range for children 0–4 in nine of the ten largest counties.

While substandard housing is prevalent in all the most seriously undercounted communities, the particular type of housing accommodations which prevail locally affects the specific components of census undercount. Given the fiscal challenges faced by the medium-size and smaller low-income communities that find it challenging to enforce zoning and building codes and which are also preferred destinations for Mexico-U.S. migrants, it is likely that the difficulties encountered in building an accurate MAF will become worse rather than better over time.

Also, fewer of the small and medium-size immigrant-dense municipalities with constrained finances due to local poverty and low property tax base participate in the Local Update of Census Addresses (LUCA) and, consequently, their local census address lists are worse than in the larger municipalities such as Los Angeles, Miami, Houston, and New York which invested significantly in LUCA (since they have been very aware of the fiscal consequences of differential undercount and its relationship to MAF quality for several decades). 4.10. Census residence rules are a factor in differential undercount of Mexican and Central American immigrant households

In the real world of immigrant (and other minority) communities there is a blurred line between the housing arrangements which result in census omissions discussed above that are categorized as “total household” or “partial household” omissions because of the problematic nature of Office of Management and Budget (OMB)/Census Bureau residence rules. These rules to guide census response may seem to be quite logical in defining a “household” as everyone living in a “housing unit” and seeking to have a census respondent (P1) include on the household roster everyone living there. However, even if census form instructions are clearly understood by the respondent (which is not always the case) there are cultural, cognitive, and practical barriers respondents face in successfully accomplishing this task.

Cultural concepts of household based on social/economic relationships in the social universe of immigrants clash with the Census Bureau’s definition of household which is linked to physical characteristics of housing arrangements. For example, as Mahler notes in her discussion of census omission among Salvadorans and other Hispanic immigrants on Long Island, and in my own rural community research confirms, the pervasiveness of sub-let spaces in migrant-receiving communities affects undercount [31]. In these circumstances, in potential census respondents’ minds the definition of “household” based on social relationships trumps the official definition and these periph-
eral individuals are not included on the household roster. Similar types of conflicted definitions of "household" have also been noted by anthropologists in their ethnographic study of Haitian undercount in Miami where "household" is thought as everyone who shares food [32].

There are also some particularly problematic culturally variant concepts related to categorizing a "boarding house" where there is, for example, a single main entrance but multiple partitioned living spaces. Presumably, census residence rules would suggest a single housing unit/household but the potential census respondent, often the onsite "primary family" who is actively concealing the fact that they have rented out living space, are clear that each room ("quartito") is a separate household/housing unit. These prevalent perspectives on household definition are rational with respect to the rules of the local social universe (because, as is the case with single room occupancy (SRO) occupants, the "boarders" are often unrelated and do not share economic resources). Currently, as low-income suburban neighborhoods come to be the residence for increasing numbers of doubled-up immigrant families, definitional complexities and confusions multiply—but the consequence, whether the "cause" of census omission is categorized as total or as partial household undercount, is that the peripheral socially-distinct individuals, sub-families and social units, are often not enumerated.

Laurie Schwede points to another problem with the residence rules in an excellent Census Bureau report on complex households [33]. In these complex households, the census respondent sometimes believes that a person who is not necessarily living permanently in the household should not be included on the household roster and, in fact, the issue as to what their "usual" residence might be is confusing to all concerned. This is a common issue in cases, for example, where a single mother and her newborn child are living with her parents (with uncertainties about whether she will soon move in with the child’s father, remain at home, or live on her own) or where extended family members from a village migration network are provided a place to stay for an indeterminate period of time. In both these types of cases, it is likely that a young child will be part of the social unit that is not enumerated. Observations in the course of an ethnographic study of indigenous Mexican immigrants in Fresno County show the extent of this problem; for example, one household fluctuated in size from six to sixteen persons over the course of 18 months because of the coming and going of sub-family units [34].

In the course of research on Hispanic response to census forms and procedures in 1992, the author and his research colleagues learned about another conceptual barrier to Hispanic immigrant respondents’ compliance with the official residence rules for census reporting; some census respondents believed they should not provide information about another sub-family living in the housing unit with them because the information didn’t “belong” to the respondent [35]. In field research in farmworker communities over the past two decades, my research also found that in the larger complex households consisting of several sub-family units, the likely census respondent actually did not know much about some of the co-residents in the housing unit – even basic information such as age [36]. This was most problematic in "ad hoc" households of solo male migrants but, in some cases, there were also young couples with children living in these crowded, complex households [37]. Other Hispanic immigrants mistakenly believed that only citizens should be listed for the census count.

Moreover, as O’Hare and other researchers have pointed out, because census respondents usually provide information in a systematic way – starting with the adults and listing children subsequently in decreasing order of age – the youngest children are most likely to be omitted in enumeration of a large, complex household [38].

Continuing current efforts to instruct census respondents to follow the Census Bureau’s official residence rules are reasonable but hold out little promise—because cultural/conceptual factors are more powerful than detailed written instructions on a government form. Independent household/social units in complex households need to be given the flexibility to submit their own “household” census forms. This flexibility might be provided, for example, by modifying and refining data processing algorithms to avoid deletion of multiple census responses from a housing unit with a single postal address that appear to be duplicates but, in fact, are not.

4.11. Undercount of young children in Mexican immigrant households is related to recency of immigration and changing migration patterns

O’Hare and his colleagues recently analyzed the undercount of children in Latino households in Census 2010 and found there was a 7.1% net undercount of Latino children under 5 years of age. They also note that the undercount of these Latino children accounted
for 37% of the overall undercount of U.S. children in this age group. Their analysis points to living in a hard-to-count place, i.e. a low-visibility housing unit or a hard-to-count household, i.e. a complex household as important factors in this serious undercount [39]. The undercount of children in Mexican immigrant households is probably still higher.

Fein’s analysis of the undercount of an immigrant population of predominantly Mexican immigrants in the Los Angeles basin shows that immigrant undercount is correlated with recency of immigration. He found that undercount of the immigrant population in the U.S. for 5 years or less was double that of those who had lived in the U.S. for 12 or more years (19.5% vs. 9.9%) [40]. The author’s own research on rural farmworker undercount has shown a similar relationship with 17.7% of the population in households headed by recent immigrants in the U.S. for 5 years or less being omitted [41].

The link between recency of immigration and undercount, has indirect but important implications for the current and future undercount of the pre-school age children of Mexican immigrant parents – because Mexico-U.S. migration patterns show clearly that the bulk of Mexican migrants have, in the past, been teenagers and young adults [42]. Changes in the volume, composition, and age at which migrants leave their hometown to join the south-north migration flow inevitably impact the dynamics of undercount within this population.

Recently-immigrated Mexican families almost always include pre-school age children because the newly-arriving immigrants are in prime childbearing years. Based on analysis of data from the National Survey of Family Growth, Emilio Parrado shows that Hispanic immigrant women’s fertility peaks between 0–5 years after migration, explaining the high proportion of U.S.-born children in Mexican immigrant households and also implying that there are pre-school age children in most of these households of recent immigrants [43]. Data from the Mexican Migration Project, for example, show that male migrants on their first trip to the U.S., were on the average 27.0 years old while women were 28.8 years old [44]. While some young couples spend the money on border-crossing fees needed to bring Mexico-born children to the U.S., the fact that about four out of five children of a Mexican immigrant parent are U.S.-born suggest that most are born relatively soon after young migrant couples in their twenties have settled in the U.S.

Changing migration patterns are also impacting family formation binationally and, indirectly, the undercount of children in the households of young immigrants who have settled in the U.S. Up through 2001 Mexico-U.S. migration was dominated by young men – teenagers and young adults who came to the U.S. as solo migrants to send remittances home to wives raising children in households in Mexican hometowns [45]. Post-9/11 border control, coupled with growing reluctance of young women to stay at home while men went north has resulted in more young couples, including families with children, migrating north [46]. Although male migrants continue to outnumber female migrants, Rogelio Saenz reports that the sex ratio of Mexican migrants (number of males per 100 females) dropped from 146 in 2003–2007 to 125 in 2008–2012 [47].

There is further evidence as to how changing migration trends are affecting the composition of the settled Mexican immigrant population. An analysis of National Agricultural Worker Survey data, for example, shows that in 1996–2000 only 10% of recently-arrived (virtually all Mexican) farmworkers lived in households with children; however, by 2010–2012, almost one-quarter (24%) of recently-arrived farmworker households were ones with children and two-thirds (68%) of these were households with children aged 0–5 [48].

The evidence from migration research and observations about the changing demographics of the Mexican immigrant population, then, shows that the number of pre-school age children in recently-arrived Mexican immigrant households is increasing and that they will probably be more seriously undercounted if recent migrants increasingly reside in actively-concealed housing accommodations.

4.12. Overall census omission of Mexican immigrants and their children is higher than generally recognized

In summary, the survey and ethnographic research show that the DSE-based generic estimate of Hispanic undercount leads to a serious underestimate of the extent of undercount among disadvantaged sub-populations among Hispanics, as by the review presented here of research on census undercount of Mexican immigrants and a sub-group among them, Mexican-origin farmworkers.

Fein’s survey-based estimate that 13.9% of non-citizens (most of them of Mexican origin) in the 1986 Los Angeles study area were omitted from the census is remarkably close to Manuel de la Puente’s estimate.
based on the 1990 ethnographic research that 14.4% of households headed by a person of Mexican origin were omitted from the 1990 census [13] and my estimate of 13.9% undercount in farmworker communities in 2000 [15]. This is about nine times the “official” estimate of generic Hispanic net undercount of 1.54% based on recent coverage measurement research [49].

To be sure, much of the research providing the basis for estimating the undercount of Mexican immigrant families has taken place in hard-to-count low-income neighborhoods and census undercount is less extreme in more affluent neighborhoods. However, even if we take into account the fact that only a minority of Mexican immigrant families live in the low-income neighborhoods where the undercount has been estimated at around 14%, it would be reasonable to estimate that the rate of omissions of persons in the overall Mexican immigrant population is actually around 8%.

If this is indeed the case, the actual differential undercount of individuals in Mexican immigrant families is still about five times higher than the official estimate of net undercount of Hispanics in 2010. Innovative new operations such as targeted mailing of bilingual census forms (begun in 2010) have been very helpful in ameliorating undercount in linguistically-isolated neighborhoods – but the MAF-related problems persist.

The ethnographic research and overall sociological research make it clear that children growing up in Mexican immigrant households are even more likely than the overall population of Hispanic families to live in neighborhoods and communities where low-visibility housing is prevalent and to live in a low-visibility housing unit themselves [50].

The children of a relatively recently-arrived undocumented parent are still more likely to live in housing where the entire housing unit is missed than is a long-term middle-aged settled Mexican immigrant who has accumulated the financial resources to buy his own house and whose children have left home. There is also little doubt that partial household undercount also contributes to the undercount of young children in families with an undocumented parent – because a disproportionate number of the recently-arrived parents and some of the unmarried teenage mothers are living in a doubled-up household where their status and their children’s status as “household members” is uncertain.

Therefore, it is likely that the overall undercount of young children in undocumented and mixed-status Mexican immigrant households is still higher than that of the overall undercount of of persons in Mexican immigrant-headed households – in the 8%–10% range. However, focused research on this issue, using the type of approach used in the Causes of Undercount Survey, would be valuable in more definitively quantifying the prevailing extent of undercount in this specific sub-population, the extent to which it varies from place to place, and determining at what pace the most disadvantaged arriving immigrants move from crowded and/or low-visibility housing into more easily-identifiable single-family homes or apartments.

Concurrently, attention should be given to the ways in which changing fertility patterns among Mexican immigrant women and changing patterns of migration (e.g. age at first migration, ratio of solo male migration to whole-family migration, changing costs for unauthorized border-crossing, overall flow of Mexico-U.S. unauthorized migration) affect the undercount of young Mexican immigrant children [51]. There is not yet enough research providing solid ongoing analysis of changing patterns of solo male, solo female, and whole-family Mexico-U.S. migration. However, the composition and volume of South-North migration flows can change rapidly and dramatically over a period of a few years (as has been the case over the past decade). Migration experts, for example, point to multiple factors which have resulted in declining Mexican migration and the increasing flows of Central American unaccompanied minors (most of them teenagers), as well as mothers and children. However, these flows have fluctuated greatly from year to year due to changes in home country conditions and immigration policy. Because these migration flows are so variable, they may significantly affect the findings from demographic analysis, as well as modulating the level of undercount of Mexican immigrants.

4.13. Erroneous imputations and flawed estimates of duplications further undermine the reliability of the official report of net undercount of Mexican immigrant families

The proportion of cases where the Census Bureau needs to impute whether or not a housing unit is occupied, or to impute the characteristics of the persons living in it if it is believed to be occupied but non-responsive, is higher in hard-to-count neighborhoods than in easier-to-enumerate ones because NRFU is less successful in these neighborhoods for various reasons, including enumerators’ lack of non-English language skills, family distrust of strangers, lack of landline telephones, and volume of workload in low mail-response tracts among other factors.
Therefore, it is important to consider the extent to which erroneous imputations in these immigrant-dense low-income neighborhoods where low-visibility housing is more prevalent may add to the underlying undercount stemming from total household and partial household omission and further affect the official estimate of net undercount of Mexican immigrants and their children.

It is important to keep in mind that reported net undercount is the sum of overcount (duplicate enumerations) and undercount (census omissions) and that, consequently, the Census Bureau’s reported tally of enumerated persons includes imputations. In its 2010 Census Coverage Measurement report, the Census Bureau estimated a 7.7% rate of omissions, 3.2% of duplications, and 2.4% rate of imputations for all Hispanics [52].

The question then arises as to the extent to which Census Bureau identification of duplications and procedures for imputations affect the estimate of size, demographic, and socioeconomic profile of the Hispanic population in general and Mexican immigrant households in particular. This is a concern because the Census Bureau’s adjustment for imputations, and duplications lowers the officially-reported estimate of net undercount rate for Hispanics from the 7.7% rate of observed omissions down to the officially-reported estimate of 1.54% undercount.

The Census Bureau’s approach to imputation and tabulations in reports assessing different aspects of census operations can be justified technically but the reliability of these estimates of the level of duplications and imputations is questionable when it comes to determining the actual undercount of low-income neighborhoods and the households where Mexican immigrant parents and their children live.

It is likely that flawed imputations decrease the total count of Mexican immigrants because occupied low-visibility “back houses” or other unusual and/or concealed housing units, if included in the MAF at all, may be incorrectly classified as vacant if there is no returned questionnaire from that housing unit and if NRFU subsequently fails to secure a response. In many cases, even if the property owner or neighbor is contacted in the course of NRFU they will not be inclined to report that an unpermitted and actively concealed housing structure on their own or a neighbor’s property is occupied.

In cases where a low-visibility “back house” is correctly imputed as being occupied, but household information is missing, household size and characteristics are likely to be imputed based on the nearest housing unit which was enumerated, often the main house. When this happens, the demographic and socioeconomic characteristics of the nearest enumerated household, typically, an older, settled, more financially-secure immigrant household, will be attributed to the actual occupants of the “back house”, typically a younger, more recently-arrived couple with children, a single mother with children, or teenage migrant newcomers, usually young men.

While the recent coverage measurement research does not allow a definite estimate as to how the initial census omissions or subsequent flawed estimates of the numbers of duplications and imputations might affect the official estimate of the net undercount of the Mexican immigrant population specifically (because the short form has no detail on ethnic origin and nativity), it is very unlikely that the Census Bureau’s reported level of duplications among the overall Hispanic population (3.2%) accurately represents the actual number of duplications in enumerating the Mexican immigrant population, especially since the 2010 analysis of housing unit coverage very surprisingly reports that housing units occupied by Hispanics, Blacks, Asians, and households of some other race are all overcounted [25].

If we were to assume, for example, that the actual rate of duplications in enumerating Mexican immigrant households, was only half the overall rate of presumed duplications for Hispanics (i.e. 1.6%) and that about half of the imputations of vacant housing units (i.e. 1.2%) are erroneous, the adjusted estimate of net undercount would be higher by 2.8% than the reported one. This implies that the post-NRFU procedures for imputation and adjustment for presumed duplicate responses (which may, in fact, be valid responses from sub-families in a complex household) probably contribute to a coverage measurement estimate which fails to recognize the magnitude of undercount among Mexican immigrant households.

Here too, the inherent problem stemming from the dynamics which Fein and West describe as a conflict between the census system and social systems in undercounted communities continues to be that the most socioeconomically marginal households are those most likely to be non-respondents who cannot be successfully contacted in the course of non-response followup. In these circumstances imputation inevitably magnifies the underlying systematic differential undercount.

The problems stemming from flawed imputation in communities and neighborhoods with concentrations of low-income Mexican immigrants (or other immi-
In many states and communities with the most concentrated immigrant populations, the undercount is even higher for states and communities with the most concentrated immigrant populations. Nationally, about 13% of the U.S. population is foreign-born but in California, for example, 27% of state residents are, in New York state 22%. Immigrants are even more concentrated in some major immigrant cities such as Los Angeles where 34% of the population is foreign-born and Houston where 29% are.

5. Recommendations

The evidence that structural causes of undercount, housing and neighborhood conditions, may play a greater role in undercount than respondent motivation or inability to respond due to language or literacy constraints has important practical implications. I review a promising new strategy to improve the MAF and then, other practical, feasible steps that will serve to ameliorate the chronic differential undercount of low-income minority and immigrant-headed households.

My discussion looks both at the over-arching issues related to census enumeration of all individuals who live in low-visibility housing and the specific implications for enumeration of Mexican immigrants in Census 2020. The challenges we now face in assuring an accurate census are greater than they have been for decades – but there are some strategies which can, at least, decrease the risk of a mega-undercount of Mexican immigrants, and others, in the upcoming decennial census.

The strategies I discuss are also applicable to ongoing Census Bureau operations and, arguably, may be even more crucial for successful implementation of the American Community Survey than for Census 2020 since systematic differential undercount of immigrants so seriously skew the demographic and socioeconomic profile of minorities and immigrants. The practical stakes for fair and equitable allocation of federal funding driven both by numbers and ACS-derived population characteristics used to determine program eligibility (e.g. living below the poverty level, having less than a high school education) are, of course, even higher for states and communities with the most concentrated immigrant populations. Nationally, about 13% of the U.S. population is foreign-born but in California, for example, 27% of state residents are, in New York state 22%. Immigrants are even more concentrated in some major immigrant cities such as Los Angeles where 34% of the population is foreign-born and Houston where 29% are.

5.1. Employ a multi-variable model to assess differential census undercount and craft effective responses

A multi-variable analysis of census undercount has great promise as an alternative to Census Bureau coverage measurement studies which examine undercount in relation to only a single variable at a time – e.g. race/Hispanic origin, tenure, type of housing unit. These “conceptual snapshots” are relevant but not as theoretically powerful or practically useful in efforts to ameliorate differential undercount as more thorough multi-variate models and ethnographic “in field” observations of the dynamics of undercount. Moreover, they are often interpreted (incorrectly) to imply that race/ethnic identity is a “cause”, not simply a correlate, of undercount.

The analytic model developed three decades back based on the Causes of Undercount Survey research in the 1986 Los Angeles Test Census provides a robust theoretical framework for understanding the actual dynamics of differential undercount [8,23]. The particular value of this framework is that it identifies three key over-arching “domains” of interacting factors that are particularly relevant in understanding patterns of differential undercount where there are interactions among 1) census/survey system and procedures, 2) social system causes of undercount, and 3) respondent/population motivation and behavior [3]. An important aspect of the research leading to development of this model is that the 1986 Los Angeles Census Test incorporated what was, in essence, a “triple-system” estimation of undercount – from standard census enumeration procedures, standard post-enumeration procedures, and a specially-designed Causes of Undercount Survey (CUS) which provided an independent basis for analyzing the adequacy of the standard dual-system estimate (DSE). Unfortunately, this strategy for high-quality census coverage measurement was, to my knowledge, almost never used again although there is one report of findings from a version of the triple enumeration methodology [53]. The decision to replace the long-form sub-sample of the decennial census with the ACS and conduct a short-form-only census in 2010 inevitably decreased the number of variables which could be used in the 2010 Coverage Measurement Studies.

Analyses led by J. Gregory Robinson resulted in the Census Bureau’s creation of the Hard-to-Count Planning Database which was used in 2000 to target census improvement efforts in the most problematic
tracts, those with the lowest mail-response rates [54]. A modified version of the database, renamed the Low-Response Planning Database, includes variables from the ACS and was used in Census 2010 for similar efforts. In both version of the database, multiple “structural” (neighborhood characteristics, housing characteristics) and household variables are used to predict low response. Although low-response rate is imperfectly correlated with ultimate census undercount, the planning database shows the utility of multi-variate analyses in generating a higher-resolution mapping of patterns related to undercount.

In an ideal world, the Causes of Undercount Survey approach might be refined to include some of additional variables used to characterize census operations in the Hard-to-Count planning database used by the Census Bureau in 2010 and the Low-Response planning database currently available in preparation for the 2020 decennial. In terms of assessing Hispanic immigrant undercount, a particularly valuable variable to include is whether bilingual forms were mailed out – since, predictably, this operational factor appears to have made an important positive contribution to improved mail response rates, decrease in non-response followup workload, and ultimately, enumeration.

Given the “deep structure” dynamics among multiple identifiable factors and interacting systems that contribute to undercount, deployment of Fein and West’s conceptual model also provides an improved basis for data users to contextualize and interpret imperfect population estimates and/or population profiles based on ACS or decennial census data. Such an analysis can also provide a foundation for synthetic adjustment of skewed demographic and socioeconomic profiles of service populations targeted by education, health, and social programs – based on deeper understanding of the systemic biases which affect production of the decennial census dataset (including sampling, data collection, imputation, and data editing).

An advantage of Fein’s regression analysis is that it distinguished but also links the major co-variant social system and housing/neighborhood factors associated with undercount with the operational consequences—likelihood of partial or total household omission. This is useful because the strategies to address these distinct limitations in census procedures are quite different and imply the need for changes in different sub-systems of the census process.

Within this broad over-arching framework of interactions between social system and census procedures, additional consideration of the distinctive factors implied in the undercount of the distinctive sub-populations – e.g. children within the overall population of Mexican immigrants, still smaller sub-groups within this group such as the children of Mexican immigrant migrant/seasonal farmworkers, or the children of undocumented single working mothers – support the view that there are related but variant ways in which interactions between social system and census operational system factors play out in each set of social circumstances to contribute to the undercount of each sub-population.

Replacing the Census Bureau’s single variable analysis of undercount based on race/Hispanic origin will also be extremely helpful for policymakers and program planners. A more finely-textured mapping of the social geography of the U.S. and the extent to which undercount affect diverse communities – both socially-defined “virtual communities” and geographically-defined communities at the state, county, municipality level has practical policy utility. This sort of high-resolution analytic mapping then would provide a better three-dimensional portrayal of a landscape with pockets of higher-than-expected undercount and, also, islands of overcount.

5.2. Improve the Census Bureau’s Master Address File (MAF) by adding low-visibility housing units identified via community-based local address canvassing

As U.S. society becomes more diverse, as economic disparities increase, and relations between different social groups and classes become more strained, it will be increasingly urgent to incorporate into the census process “social technologies” which draw, at least in part, on local cultural capital to ease the clash between standardized census operational procedures and neighborhood social system dynamics. Community-based local address canvassing is an example of an important way in which the decennial census (and possibly other survey endeavors) could, by becoming more participatory, contribute to achieving technical objectives in survey research – in this case, an improved sampling frame, a better MAF [55].

Important technological advances have been made which improve the utility of “in office” address canvassing (e.g. use of satellite imagery to add or delete housing units from the MAF) but they cannot be expected to very effectively address the challenge of adding low-visibility housing units occupied by immigrants, especially in urban areas since satellite imagery
does not successfully identify occupants of basements or in-law units.

Similarly, administrative records are useful for some phases of census operations but have serious limitations in others because many households in the typically hard-to-count populations, the most socially and economically marginal families, are under-represented in many administrative data sets [56]. This problem is still more serious with respect to the differential undercount of immigrant populations where a significant proportion of households and individuals within households lack legal immigration status – since undocumented immigrants are ineligible for many of the programs which are the source of administrative data. Even in instances where undocumented immigrants are represented in local government administrative datasets, for example, state drivers’ licenses, their inclusion varies from state to state.

An additional implication of total household omission accounting for such a substantial portion of aggregate undercount is that the potential impact of some standard census improvement initiatives will be limited. Initiatives toward census form improvement and growing availability of bilingual forms are well-justified as a practical investment in improved census quality and are particularly valuable as a contribution to decreasing the barriers faced by low-literate, limited-English respondents who are motivated to participate. However, giving priority attention to improving the MAF may be even more important than standard procedures meant to improve mail response such as advance notification letters, reminder letters, etc.

Proactive and innovative efforts to improve the MAF by identifying and adding low-visibility housing units deserve high priority in the Census Bureau strategy for census improvement.

One particularly promising approach will be to re-engineer the LUCA process to facilitate local government units partnering with local community organizations to review and improve the MAF via community-based address canvassing. This approach, one of those recommended by the Urban Institute November, 2016 convening of experts, is particularly attractive because local government implementation of LUCA has most commonly included only “in office” address canvassing, that is, comparison of the Census Bureau’s address list to locally-available administrative data. Therefore, actual “in field” address canvassing is a valuable complement to standard in-office review.

Partnering with local community-based organizations, particularly grassroots community organizations which have rapport with undercounted groups (e.g. Mexican immigrants), as well as the cultural capital, the language, communication skills, and trust needed to successfully identify actively-concealed low-visibility and irregular housing, provides a way to engage in cost-effective “in field” address listing. This type of initiative was piloted by California Rural Legal Assistance in Census 2000, not as part of LUCA but as part of an innovative experiment by the Los Angeles Census Region to rely on community outreach workers to improve enumeration of migrant and seasonal farmworkers in the months before Census Day. More than 4,000 low-visibility housing units were identified by the CRLA address canvassing teams and submitted to the Census Bureau. It was subsequently determined that 73% were valid “adds” [57]. In contrast, only about 21% of additions to the MAF proposed by local government participants in LUCA in 2010 were valid “adds” [58].

A first step in the direction of making the overall LUCA process more effective in improving the MAF, given the budget pressures confronting the Census Bureau in implementing Census 2020, is for those state governments committed to a fair and accurate decennial census count to provide state funding to local government entities eligible to participate in LUCA. This step is necessary because LUCA participation is voluntary. Although larger municipalities and counties with substantial populations of minorities and immigrants have participated vigorously in LUCA since the process was first introduced only about one-third of LUCA-eligible entities have gotten involved (36% registered and returned address list corrections for Census 2000, 27% in Census 2010). Many of the smaller local government units, presumably those with less financial resources, did not register and some who did register, eventually failed to ultimately submit any proposed “adds” or “deletes” or corrections to the Census Bureau [58,59].

An important second step will be to build on the important analytic work that went into developing the Census Bureau Planning database used to identify “hard to count” census tracts in 2000, now “low response rate” tracts in the current terminology. The most recent planning database (2015) now includes 550 variables derived from the 2010 decennial census and tabulations of recent American Community Survey data. It is possible to re-purpose and refine the ranking of “hard to count” tracts in order to target address canvassing efforts wisely – by revising the scoring indicator to generate a “customized” predictive model
for identifying “high-density of low-visibility housing” tracts by drawing on the rich planning database which is available.

Well-targeted locally-conducted address canvassing makes it possible for the contributions of community-based organizations to be cost effective. Review of the broader patterns of differential undercount in communities with the highest concentrations of low-income minority and immigrant households suggest that a targeting strategy which focuses on 15–20% of all census tracts would provide the optimum balance between targeting a sub-set of priority tracts and address canvassing of 100% of tracts.

The variables for such a customized planning tool designed specifically to identify census tracts with a high density of low-visibility housing include several of the relevant variables used in the Census Bureau’s analysis for reporting coverage measurement of housing units in 2010 such as trailers, multi-unit buildings with 2–9 apartments, renter-occupied housing units. However, the availability of the ACS-derived variables makes it possible to construct an even better model by including additional variables which appear to be correlated with prevalence of low-visibility housing based on the 1990 ethnographic research, the Causes of Undercount Survey, and other research on housing unit omissions. Promising variables from this research include:

- % of units with attached housing unit,
- degree of ethnic enclosure,
- % non-family households,
- % of households in poverty,
- % foreign-born head of household in U.S. < 5 years,
- % HH heads with < 9 years educational attainment,

Analytically, it is challenging to determine the economic pressures and housing conditions which lead specifically to crowded housing as distinguished from those which lead to living in low-visibility and/or concealed housing units. Nonetheless, common sense, as well as the survey research and ethnographic studies, point toward indicators of economic stress in securing housing as proxies for prevalence of low-visibility housing units.

Testing alternative models to rank areas based on predicted density of low-visibility housing in communities with diverse and distinctive patterns of housing would be desirable. The 2018 End-to-End Test of address canvassing in Pierce County, WA would be a useful place to examine a prototype model as well as to assess general adequacy of current address canvassing procedures by adding community-based address canvassing to the test design because the area has a substantial population of Mexican immigrants.

The third step is to encourage local governments and/or local Complete Count Committees to advocate that a portion of state-funded LUCA grants be used to support “in field” address canvassing by locally knowledgeable community organizations – in part because it has such promise as a tool for MAF improvement and in part because it provides a good opportunity to build solid partnerships and begin messaging about the importance of the census for local communities early on.

A robust strategy for efforts by Census Bureau stakeholders and potential partners, including local government entities, but also philanthropic organizations (which have huge stakes in securing accurate census and ACS data), and local businesses will, also need to explore the possibility of direct collaboration between community-based organizations, state government, and the Census Bureau in adding low-visibility housing units to the MAF.

For example, although California has allocated $7 million in grants to local government entities to assist them in LUCA-based efforts to improve the MAF in local jurisdictions, there are inevitably going to be some areas where local government fails to participate. In those areas, where, even with the incentive of state assistance, some small municipalities, and rural county governments fail to participate, there will be gaps in statewide efforts toward address list improvement via LUCA. In those areas, which will almost certainly include some of the rural counties with concentrations of seriously-undercounted populations such as migrant and seasonal farmworkers, the option of collaborative partnerships between the Census Bureau and community-based organizations (as was the case in the 1999-2000 CRLA-Los Angeles Region collaboration) will be important so as to “plug the holes” in the statewide LUCA effort.

5.3. Revise Census Bureau data processing procedures

A potentially promising development is the Census Bureau’s adoption of new procedures for “non-ID processing” for online responses so that households that are willing and able to submit their responses online
will be able to do so successfully even if they live in a low-visibility housing unit which lacks a US mail address and which has been omitted from the MAF.

Revised data-editing procedures can avoid the problem of online responses from households living in low-visibility housing left out of the MAF being erroneously deleted when they use the postal address of the “main house” adjacent to them (e.g. if they live in a garage, backyard trailer, or basement). It will be important to assure that online census responses which are initially “accepted” are not subsequently deleted at the data-cleaning stage. The good news is that it appears that the Census Bureau is moving toward use of more sophisticated algorithms to identify duplicate response while allowing multiple valid responses from what seems to be the same address.

The Census Bureau’s commitment to counting everyone at the right location is commendable but the possibility that the inevitably-imperfect MAF will be used as a tool to delete online census responses from the most marginal families living in housing units not included in the MAF continues to be worrisome.

Many immigrants, especially the most recently-arrived, have only a post office box for communications but it will be important, especially since online census response will be the preferred modality for enumeration in 2020, to make sure the substantial promise of online response is fulfilled. Interestingly, but not entirely surprising, is the observation that in the current climate of anti-immigrant hostility, online interactions are, in many cases, favored as being “safer” than in-person interactions.

Although there are still constraints to online access among hard-to-count populations, Internet penetration is increasingly rapidly even among low-income group – in part due to ubiquitous use of smartphones and in part due to increasing availability of public-access hotspots for Wi-Fi [62]. Smartphone use is also very high among the “working poor” in Mexican and Central American immigrant populations – because they are routinely used to secure employment, especially for day laborers and intermittently employed workers such as farmworkers, maids, and handymen, or to communicate with relatives in their hometown.

Software that is optimized for smartphone and tablet as well as traditional platforms and designed to accommodate respondents whose primary language is Spanish and who may be limited in literacy has some promise for increasing the response rate within this population. Sound software design will also need to be coupled with targeted messaging emphasizing the safety and ease of online response.

6. Conclusion

It is hoped that the foregoing analysis of the actual dynamics of the undercount within hard-to-count groups like Hispanic immigrants provides valuable insights regarding strategies which the Census Bureau could use to ameliorate or even overcome this and other group’s differential undercounts. Use of community-based address canvassing as a tool for improving the MAF is an a strategy which is both immediately useful as a rapid response to shortfalls in Census Bureau budget and as an example of how more expansively conceptualized modes of community-government partnership can improve census accuracy – even in the face of major budget challenges.

Acknowledgments

I would like to acknowledge the generosity of several researchers and Census Bureau professionals in sharing their insights and expertise. I am particularly grateful to J. Gregory Robinson for his suggestions about how to adapt the Census Bureau’s planning database so it could be used to identify and prioritize census tracts expected to have much higher than average levels of low-visibility housing. At the Census Bureau, James Christy, Brian Timko, Andrea Johnson, and Michael Ratcliffe have been helpful and patient in sharing insights about address listing and explaining details about LUCA and MAF construction.

My colleague Shannon Williams also provided crucial support in data processing for several of my analyses of National Agricultural Worker Survey data. I am also grateful to Ilene Jacobs at California Rural Legal Assistance for her tireless and insightful collaboration during the three decades we have worked together in addressing issues of undercount of migrant and seasonal farmworkers.

References


[2] De La Puente M. 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; Dominguez,


[15] Fein DJ. The Social Sources of Census Omission. p. 120.


[27] Aguirre International. Identifying High Concentrations of Mi-


[29] De la Puente M. An Analysis of the Underenumeration of Hispanics.


[38] O’Hare WP. The undercount of young children, p. 94.


[40] Fein DJ. The social sources of census omission p. 152.

[41] Kissam E. Census Enumeration of Immigrant Communities in California.


[50] O’Hare WP. The undercount of young children. Table 7.1.


