Variations in Fresno County ACS and 2020 Census Response Rates: Implications for Allocating Social Program Funding

By
Ed Kissam and J. Gregory Robinson
June 14, 2023

Introduction

Although policy discussion of the dynamics of decennial census undercount has tended to focus on race as the primary correlate of differential census undercount, the racial/ethnic composition of a tract, neighborhood, or community does not adequately explain variations in decennial census and American Community Survey (ACS) response and how these impact representation and allocation of funding driven by census-derived data.

Analysis of variations in sub-state and <u>sub-county</u> level of census and ACS response is important, both for Census Bureau planning and for very practical considerations. Despite expensive and extensive efforts in non-response followup (NRFU) that are meant to overcome the problem of variations in decennial census response, differential undercount assessed via dual-system estimation has always been correlated with differential undercount. Consequently, disparities in response rate are always consequential.

We recently presented analysis of small-area variations in 2020 census self-response in California in relation to socioeconomic, demographic, and sociopolitical characteristics (Robinson, Kissam, O'Hare 2023). Here we deepen our analysis of the dynamics of local variations in census response to examine response at the census tract and sub-county level. We focus on Fresno County, a large California county (about 1 million population) with a mix of rural and urban settlement. We discuss the patterns and implications of the observed relationship between census self-response and ACS response in relation to local community characteristics.

We review the correlation between patterns of decennial census non-response and ACS non-response and what this implies for equitable allocation of funding or other social investment based on both census and ACS data as well as for community-level planning. In the real world of community life, the analysis of national-level and state-level undercount for major race/ethnic groups is inadequate; equity and opportunity require fair and effective local investment.

Our Fresno County case study is particularly timely because a recent Census Bureau comprehensive analysis of the relationship between Census 2020 data, administrative record (AR) data, and 2020 Post-Enumeration Survey (PES) suggests that Hispanic-majority communities with concentrations of non-citizens such as those found in Fresno County may have

been even more seriously undercounted than indicated by the PES-based national estimate of net undercount of Hispanics.[1]

This Fresno County analysis of the dynamics of variations in census response "zooms in" to examine granular patterns of census and ACS response at the sub-county level by considering how these patterns affect a region of Fresno County referred to as "the Westside" in order to show how census and ACS response level in this particular sub-county area made up of rural, farmworker communities, 9 census tracts, a population of about 54,000, four municipalities, and 11 school districts, may impact equitable allocation of federal funding during the post-censal decade.

We conclude by discussing the implications of this granular analysis of variations in census and ACS response rate and the resulting patterns of census undercount and potential ACS sample bias both for Census 2030 planning and ad hoc efforts to account for these factors in a "curated data enterprise" similar to that proposed by experts at the University of Virginia Biocomplexity Institute or via the Census Challenge program.[2]

The Metrics of Sociopolitical Equity

Background

There is now a long history of more than half a century of intense research, policy discussion, and litigation about differential undercount in the decennial census. The discussion has revolved primarily around equitable census representation for racial/ethnic minorities. It has, at the same time, been closely linked to litigation by urban municipalities such as New York City, Los Angeles, and others with concentrations of low-income racial/ethnic minorities seeking to receive a fair share of census-driven funding. The latest is a challenge of Population Estimates data (derived in large part from decennial census data) for Detroit.[3]

The evidence of varying levels of differential undercount has mostly been drawn from findings of post-enumeration surveys (PES) and Demographic Analysis (DA), two coverage measurement programs the Census Bureau uses to evaluate the quality of the decennial census.[4] PES findings have been published that document disparities in enumeration related to race/ethnicity, type of housing unit, and tenure (homeowner vs. renter). Demographic Analysis (DA) has identified the persistent differential coverage of the Black population and the undercounting of young children.[5] A recent experimental analysis using a special dataset of administrative records suggests a serious Census 2020 undercount of non-citizens and Hispanics.[6]

A major Census Bureau initiative for Census 1990, the Alternative Enumeration Initiative, by relying on ethnographic research in diverse neighborhoods, yielded extensive high-quality findings about other dimensions of undercount—stemming from respondent attitudes about the census and patterns of response and "structural" factors such as living arrangements and conceptualizations of household.[7]

Additional insights were developed in the 1987 Los Angeles Test Census. In that study, a "triple system enumeration" (test census, test PSE, and an in-depth evaluation of both), coupled with ethnographic research, provided multi-faceted details on dynamics of differential undercount in the Los Angeles basin.[8] There has also been analysis of differential undercount of farmworkers developed using varied methodologies from 1990-2010 and extensive research in recent years on undercount of pre-school age children.[9]

An important related initiative was the Census Bureau's internal development of a Planning Database (PDB) generated by multi-factorial analysis of patterns of census response in prior censuses and insights from the ethnographic research to develop the Bureau's "Hard To Count" index that generated tract-level scores of "hardness/easiness" of enumeration.[10] Subsequently, a "Low Response Score (LRS) was developed.[11] Most recently, the California Department of Finance developed a California Hard-To-Count (HTC) Index to guide efforts in state-funded outreach efforts in 2020 to increase census response in the hardest to count tracts.[12]

Analysis of Census Response and Equity in the Current Case Study

Our analysis of the patterns of census and ACS response in Fresno County focuses more on census response in relation to multiple "structural" factors associated with tract-level response than on the ways in which racial/ethnic makeup of a neighborhood, tract, or community affects response. Despite the clear-cut evidence of "systemic racism" as a major factor in contemporary society, a sound understanding of socioeconomic, sociopolitical, and cultural factors are needed to adequately understand variations in response level within California's "minority-majority" communities. The two contrasting California counties where we have analyzed 2020 census response patterns in some depth are both ones where about half of the population is of Hispanic origin (Fresno County 53% Hispanic and Los Angeles County 48% Hispanic).

Our current case study approach does not assure the findings can be extrapolated to the entire state but, nonetheless, there are similar patterns of sociological diversity in other California counties and sub-state regions and in some other states. The two counties we have focused on, one mostly urban, one mostly rural, with a combined population of 11 million people, make up more than one-quarter of the entire state population of slightly under 40 million. Despite the Hispanic population plurality, both Los Angeles and Fresno counties are racially/ethnically diverse. As is the case in the state as a whole, the non-Hispanic White population is a minority, making up about one quarter of the population in Los Angeles and Fresno counties and one-third of the population in the state.[13]

Inevitably, in a societal context where half the population is made up of a historically-undercounted racial/ethnic group, the issue of equitable census representation must move onward beyond consideration of race/ethnicity as a "cause" of census undercount and go on beyond the implicit and dangerous assumption that each racial/ethnic group is homogeneous.

An important initial finding that highlights this point is that the difference in Census 2020 self-response between easy-to-count (ETC) vs. hard-to-count (HTC) Hispanic-majority tracts in California was 17.5 percentage points.[14] This mirrors the educational, cultural, and income

diversity within the state's large Hispanic population that results in part from differences between immigrant generations (1st generation, Generation 1.5, 2nd generation, and 3rd and higher generation).[15] A particularly important dimension of diversity related to census response among Mexican and Central American immigrants in Fresno County is legal/citizenship status.

Table 1 below shows that the "structural" analyses incorporating multiple factors affecting census response are more explanatory than ones based on race/ethnicity alone:

Table 1		
Explaining Census 2020 Self-Response at the Census Tract Level		
Independent Variables Explaining Response	*Los Angeles County Fit*	*Fresno County Fit*
Race/ethnicity variables alone-Census 2020 response	Adj. R-square=.180	Adj. R-square=.364
Multi-variable California HTC score-Census 2020 response	Adj. R-square=.371	Adj. R-square=.478
Multi-variable Census Bureau national LRS-Census 2020 response[16]	Adj. R-square=.408	Adj. R-square=.605

*Census 2020 self-response data from California SwORD tabulation October 16, 2020. Data on percent of each racial/ethnic category in a tract from Census Bureau Planning Database (ACS 2015-2019). All regressions are significant p<.001.

Several issues should be considered with respect to **Table 1**.

First of all, the ethnographic research, along with the broader sociological and anthropological literature on the diverse populations in Los Angeles and Fresno County point to the limitations of OMB race/ethnicity categories as key factors in explaining social behavior that includes census response—most notably due to ethnic diversity among Asian and among Hispanic populations.

Review of this evidence is beyond the scope of this paper. However, it is clear that the unwarranted breadth of the Asian racial category that encompasses a broad span of cultural diversity and national origin (from Japan to Afghanistan) is problematic. There is generally less public awareness of diversity among Mexican and Central American immigrants to California but as many as 10% of Mexican immigrants to Fresno County are likely to be indigenous ethnic minorities.[17] Analyses using detailed and reliable ethnic self-identification might show ethnicity to be a more powerful factor than race as defined in the OMB categories.

However, beyond the ethnic diversity of immigrants in Fresno County and most of California, immigration status also very clearly plays a major role in differentiating the socioeconomic life, sociocultural identity, and, eventually, census response within the overall Hispanic population. Mix of immigration statuses in household, extended family, and social networks must be understood to have a major role in determining community civic engagement, including census response.[18]

We focus here on the ways in which how socioeconomic and educational diversity <u>within</u> a broad race/ethnic group in OMB taxonomy (Hispanics) becomes more important in determining attitudes about the census and opportunities to respond than broad racial categorization.

Why Census Response Patterns Are Important for Understanding and In Responding To Inequitable Enumeration

The standard assessment of census equity is the national-level PES-based estimates of census coverage rates for different racial/ethnic groups. One measure of equity is the gap in coverage between various racial/ethnic groups and the non-Hispanic White population and its' relationship to census response rates. However, measuring this relationship is complex; a very thorough discussion by census expert, William O'Hare, was presented in a report presented in La Union Del Pueblo Entero v. Donald J. Trump.[19] The Census Bureau's recent experimental AR-based Census 2020 enumeration reveals further complexities as to how OMB-defined race, age, legal status, and inclusion in various administrative datasets (including the MAF) affects non-response followup (NRFU) and consequently the reliability of the final census enumeration.[20]

In **Table 2** below we compare the historical relationship between overall US population census coverage and coverage of the Hispanics to PES-detected differential undercount.

We use the term "overall occupied housing unit response rate" to refer to historical mail return rate data, i.e. excluding vacant housing units, those erroneously designated as residential, and housing units that no longer exist erroneously included in the Master Address File (MAF). We estimate the 2020 Census response rates for occupied housing units based on self-response rates overall (66.9%) and from Hispanic-occupied housing units (60.4%).

Table 2: The Historical Relationship Between Level of Census Response and PES-Estimated Coverage of Hispanics				
Census Response and Coverage Indicators: 1990-2020	*Census 1990*	*Census 2000*	*Census 2010*	*Census 2020[21]*
1. Overall Occupied Housing Unit Response rate %	74.7	80.1	79.3	76.6
2. US Hispanic Occupied Housing Unit Response rate %	63.4	71.4	69.7	67.8
3. Differential Hispanic vs. Overall HU Response rate %	-11.3	-8.7	-9.3	-8.8
4. PES-estimated overall coverage %	-1.61	+0.49	+.01	-0.24
5. PES-estimated Hispanic coverage %	-4.99	-0.71	-1.54	-4.99
6. Hispanic differential undercount %	-3.88	-1.20	-1.55	-4.75
7. Ratio of Hispanic undercount to level of census response	**.34/1**	**.14/1**	**.16/1**	**.54/1**

Estimating census coverage in relation to census response rate is very important as a basis for adequately assessing census equity because the currently available PES-based and Demographic Analysis (DA)-based estimates of coverage and differential undercount at the national level as well as the synthetic PES estimates at the state level obscure the actual landscape of census equity. They present a smoothed-out surface-level picture for the nation and state political

jurisdictions which obscures the jagged actual "underwater" local landscape where we can observe deep pockets of undercount and mounds of overcount using the ratio of census response to eventual published tabulations of coverage patterns.

Moreover, mapping these patterns at the census tract level makes it possible to more adequately assess the implications for allocation of census-driven social program funding and, ultimately for finding improved approaches to promoting response in low-response areas—especially as there is now improved understanding that non-response is determined by many intertwined factors in addition to respondent motivation.

As noted previously, we do not believe that racial/ethnic equity is the only or, indeed, the most significant indicator of census equity—since there is such overwhelming evidence that race/ethnicity combined with other "structural" factors affect census response in diverse neighborhoods and communities. Moreover, in California, and in our specific case study counties Fresno, and Los Angeles, geographic patterns of housing are no longer as closely linked to racial/ethnic housing segregation as to socioeconomic status.

Figure 1 below, based on Census Bureau 2020 PES data, shows that, even at the national level, being counted in the census is monotonically related to self-response with lower response leading to higher omission and vice versa.

Figure 1
2020 National-Level Relationship of Tract-level Household Omission
In Relation to Self-Response

2020 HH Omission By Tract-level Self-Response Rate

Ratio of HH omission in lowest-response decile of tracts as compared to highest-response decile of tracts=3.24

Specific factors leading to HH omission are discussed subsequently

*** Methodological issues stemming from factors affecting prevalence of vacant HU's and quality of vacancy-delete operations in 2020 may slightly modify this analysis*

***Further issues stemming from quality of unduplication procedures may also affect the PES-based estimate of net undercount in high-response vs. low-response tracts*

2020 Self-Response Rate Decile and HH Omission

Source: U.S. Census Bureau Decennial Statistics Division, "2020 Post-Enumeration Survey" (May 2022 release)
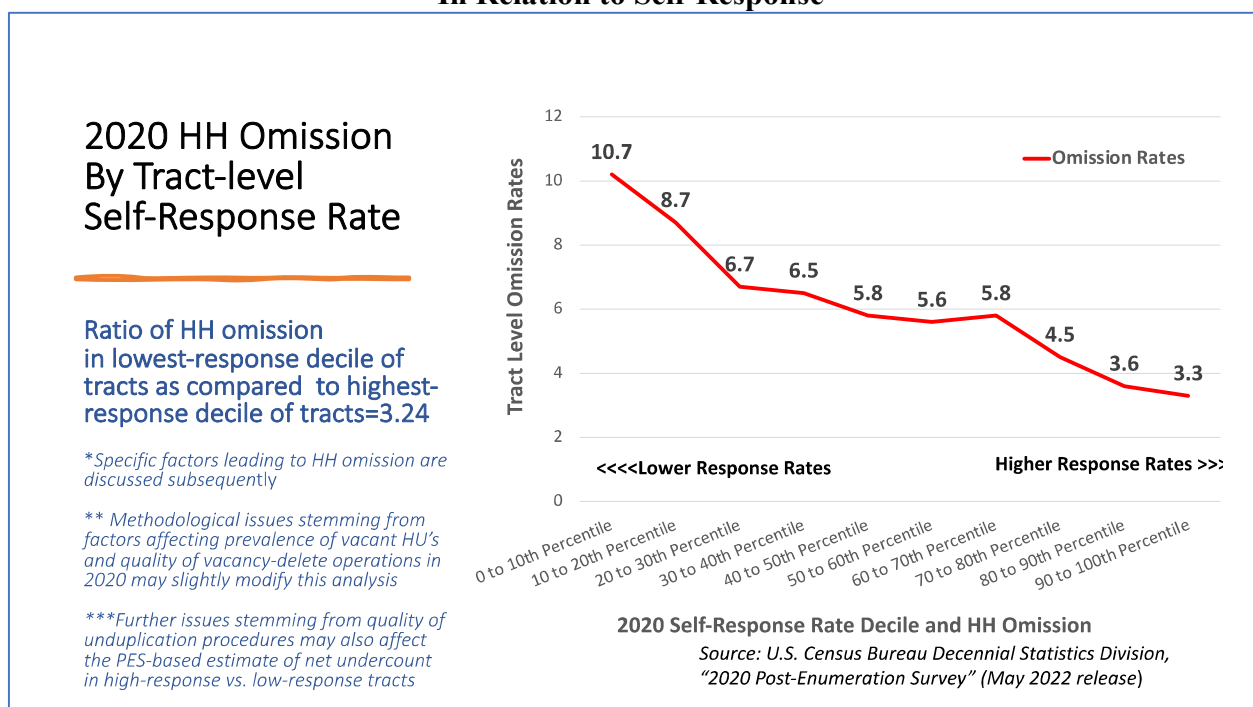
Figure 1 indicates that the differential self-response rates have not been fully compensated for in the course of NRFU. The PES-based analysis we draw on in Figure 1 (i.e. Appendix Table 6, "Census Coverage Estimates for People in the United States by State and Census Operations: 2020 Post-Enumeration Survey Report", June, 2022) shows statistically significant undercount of the households in the lowest and next lowest deciles of self-response (0-20th percentile) and the next to highest (80th-90th percentile).

These tract-level patterns of response and eventual local-level census overcount/undercount and subsequent Population Estimates Program data throughout the post-censal decade are important for understanding census equity because a good deal of census-driven social program funding ultimately flows to organizations that make a broad range of investments and provide multiple services to relatively small geographic service areas.

In Fresno County, where we "zoom in" to look at the impact Census 2020 is having on disadvantaged communities, the paradigm case is ESEA Title IA funding for K-12 schools where funding flows to school districts not directly to individual households. At the same time, there is an even wider range of indirect impact patterns on federal, state, and local investment in community development, public health, and on policy/legislative dialogue.

The Local Patterns of Census 2020 Self-Response in Fresno County

Overall Patterns of Response <u>within</u> Fresno County

Fresno County is a Hispanic-majority county and is in one of hardest-to-count sub-state regions in California although it achieved a remarkably good overall 2020 Census self-response rate (68.9%), actually exceeding the national average of 66.9% self-response. Nonetheless, there was a substantial differential in the average response rates when analyzed by quintile as measured by the California HTC score. **Table 3** below summarizes the relationship between HTC score and reported self-response rate.

Table 3 Overview of Fresno County Tract-Level 2020 Census Self-Response by HTC Quintile			
Quintile Ranked By HTC Score *(1=highest 5=lowest)*	**Average HTC Score for tracts in quintile**	**Average 2020 Tract-level Self-Response Rate (SwORD)**	**Difference from 2010 Mail Response Rate**
Quintile 1	96	59.7	-2.6
Quintile 2	77	64.7	+0.3
Quintile 3	57	66.2	+1.7
Quintile 4	30	72.7	+4.5
Quintile 5	9	79.3	+5.5
Overall County	**52**	**68.9**	**+2.1**

* Robinson's analysis here is based on Census Bureau Planning Database 2013-2017 and California SwORD data

Table 3 shows that California's Get Out The Count campaign was moderately successful in increasing self-response in the harder-to-count tracts and is consistent with the findings from California's post-census analysis of outcomes and an in-depth analysis of campaign outcomes in the southern San Joaquin Valley by the Equity Research Institute at the University of Southern California. [22] But that there was still a 19.6 percentage point differential in average self-response between the highest and lowest HTC quintiles of tracts in the county.

Moreover, our comparison of 2020 self-response rates to the 2010 mail response rates, shows a 8.1 point spread in the <u>2010-2020 changes in response rate</u>. Response in the easiest-to-count quintiles of tracts improved greatly (+4.5% in Quintile 4 and +5.5% in Quintile 5) but decreased by 2.1% in Quintile 1,and remaining basically unchanged at +.03% in Quintile 2.

Variation in tract-level rate of self-response in relation to HTC score and the resulting response gap between easier and harder to count tracts suggests that many of the earlier structural correlates of local response rate and consequent undercount identified in the PDB remained the same as they had been previously.

Contrary to popular opinion, the Census 2020 self-response data does not necessarily indicate that California's proactive initiative "improved" the quality of census data in the county but simply that it <u>did</u> have some success in mitigating the extraordinary challenges stemming from spillover of the intense national dialogue about adding the citizenship question to Census 2020, unique challenges presented by the COVID-19 pandemic (even prior to NRFU), and longstanding census response barriers of language and literacy.[23]

Digging Deeper: Local Neighborhood/Community Characteristics That Explain Census 2020 Self-Response Traces in Fresno County

As noted previously, both the California HTC score and the Census Bureau's LRS score provide good explanations of tract-level census response rate in Fresno County—even after the interventions provided by the Get Out The Count campaign that were designed to ameliorate disparities in response rate. However, we went on to explore whether streamlined alternative regressions might provide further insights into the specific dynamics of census response in this distinctive local county context.

Our streamlined regression model, relying on a mix of socioeconomic, sociopolitical, and structural community characteristics drawn from the Census Bureau's Planning Database (PDB), does very well at predicting self-response (showing an adjusted R-square=.713, p<.001) somewhat better than the "generic" HTC and LRS models.

Prominent factors explaining level of response are the following:[24]

- % of tract population in HH below poverty (standardized Beta= -.293)
- % of renter-occupied HH's (standardized Beta = -.388)
- % of non-citizen headed HH's in the tract (standardized Beta = -.153)

- % of tract population in rural area (standardized Beta = -.413)
- % of tract population in urbanized area (standardized Beta = -.129)

It deserves note that these factors explaining tract-level response rate are all "structural" ones and, like the HTC model, do not include race/ethnic composition of a tract as an independent variable.

It is, however, noteworthy that inclusion of the variable of % non-citizen headed HH's in the tract is an important consideration. We believe this consideration is important, because this PDB variable, along with educational attainment, are, at least in labor-intensive agricultural counties like Fresno County, where the predominant immigrant population consists of Mexican and/or Central American farmworkers, is a proxy indicator of the proportion of households where the head of household lacks legal status or where the household is "mixed" legal status.[25]

This is our "preferred" regression model because it is highly explanatory. However, it is important to recognize the data limitations, particularly the fact that each ACS tract-level variable in the PDB has a margin of error (consistently documented in the PDB documentation). It is also important to note that, given the differential in ACS response that is correlated with hardness-to-count of a tract, the value of the underlying variables may be systematically impacted by sampling error.

We note, also, that this model does not adjust for tract-to-tract variations in vacancy rates. Other limitations are, of course, that tract-level composition may have changed throughout the ACS 5-year dataset period or from 2019 to 2020. It is also useful to note that there is a reported margin of error for the ACS-derived variables used in the analysis.

Nonetheless, our on-the-ground observations from survey work in Fresno County suggest that changes are gradual and trends are stable (e.g. more and more exurban settlement around the City of Fresno consistently increases the percentage of urbanized area in some of the tracts that were formerly classified as mostly urban cluster).

An important consideration in trying to fully understand the fine-grained dynamics of census response is that multiple factors are correlated with tract-level census response. By referring to this model as our "preferred" one we do not mean to suggest that there are not other alternative and illuminating explanations of census response. We favor this model, in part, because it does well at avoiding problems of multicollinearity that might affect the reliability of the coefficients in the model.[26] It is also consistent with the California HTC index in highlighting prevalence of poverty and renter/homeowner housing patterns as key factors determining self-response. It also supports other census planning concerns reflected in calls for improvement of enumeration in rural areas.

Digging Deeper Still: The Key Role Played by Broadband Access and Citizenship Status

Fresno County, despite being considered a rural county, actually has very few census tracts that are predominantly "rural" as defined by the Census Bureau. A limitation of our "preferred" model of Census 2020 self-response in Fresno County is that including the broadly-defined geographic variables (% urbanized area, % urban cluster area, % rural area) in the regression is somewhat problematic due to differences in conceptualization of the rural-urban spectrum in California labor-intensive agricultural counties and the national conceptualization of urban-rural.[27] Moreover, we know that there are distinct challenges in MAF-building in Fresno County exurban and rural areas, as well as in conducting vacancy-delete operations that may affect the analysis.

Consequently, we developed an alternative streamlined "non-geographic structural model" to examine factors playing a major role in explaining variations in tract level self-response without reference to Census Bureau classification of tracts on the rural-urban continuum. This model also does well, though not as well as the one incorporating rural/urban tract classification. It achieves an adjusted R-square=.563, $p < .001$, still better than the "generic" HTC model. Prominent factors explaining level of response are the following:

- % of HH's in tract with broadband access (standardized Beta=+.331
- % of renter-occupied HH's (standardized Beta = -.284)
- % of non-citizen headed HH's in the tract (standardized Beta = -. 268)
- % of foreign-born population naturalized U.S. citizens (standardized Beta=+.135)

What is perhaps most noteworthy in this model is the key role broadband access played in determining tract-level self-response in Census 2020. Fresno County presents a good opportunity to examine this particular issue because broadband access is very uneven—due both to infrastructure and household income. [28]

A closely-linked issue is that several areas in the rural "Westside" area of Fresno County were slated for Update-Leave response and that, even in some of the other areas that received standard enumeration treatment (Internet First or Choice), there are significant proportions of low-income households without broadband access.[29] Very high proportions of all households in the county have mobile phones but our observation of Census 2020 indicates they were not easy to use for census response and were, in fact, very seldom used.

It is interesting to see, also, that the proportion of the foreign-born population in a census tract or neighborhood that are naturalized is positively associated with self-response rate. This, too, is an example of a situation where local community sociopolitical context is an important consideration. Earlier research has shown that, in the Mexican-origin immigrant population in California, educational attainment is an important determinant of naturalization. Naturalized Mexican immigrants have consistently higher level of educational attainment than naturalization-eligible legal permanent residents who do not naturalize because many self-select out of the process believing that will not be able to succeed in demonstrating English-language proficiency.[30] We also learned in the course of the San Joaquin Valley Census Research Project

survey of low-income Latinos that many naturalized citizens are particularly inclined to respond to the census as an affirmation of their participation in civic life and to show that they count.[31]

In summary, the "core" national determinants of Census 2020 self-response, prevalence of poverty and/or homeowner/renter ratio in a tract play an important role in explaining local level of response in Fresno county, but there are also distinctive factors we have identified that provide a more nuanced analysis of the multiple factors and complex interactions that determined self-response rate in 2020.

To What Extent Might NRFU Have "Cured" Problems of Low-Response in Fresno County?

We looked at the issue of varying levels of self-response by "zooming in" on a sub-county region of Fresno County where we expected self-response rate to be low because of HTC scores in the tracts making up the sub-county region and because of their specific distinctive community characteristics: the "Westside" sub-county region. (See **Appendix Table 1** for details on community characteristics for this sub-county region compared to Fresno County overall, California, and the United States).

This "Westside" cluster of 9 census tracts has a population of about 54,000 that is 85% Hispanic and where only slightly more than half (55.9%) of the heads of household have even a high school education. Slightly more than one-third (38%) of the labor force is employed in agriculture[32] and the poverty rate is 31.7%. It is also an area with a high concentration of immigrants (43.5% of heads of HH are foreign-born), few of whom are naturalized (7.7%). Only 40.4% of households in the area have broadband access—although virtually all households have mobile phones.

Given these area characteristics, it is not surprising the sub-county region has an average HTC score of 84.6, putting it in the hardest-to-count quintile of tracts in the county. The average Census 2020 self-response rate in the sub-county region was 53.6%, that is 15.3% lower than the average Fresno county self-response rate of 68.9%, making for a heavy NRFU workload in this predominantly rural area.[33]

The Census Bureau report on operational metrics at the tract-level show that 15.8% of the non-responding HH's in the Westside area were resolved via proxy interview. This is consistent with our San Joaquin Valley Census Research Project (SJVCRP) data where 18.4% of low-income Latino respondents said they would be willing to participate in a proxy interview. However, more than one-third of the SJVCRP respondents who said they were willing to agree to a proxy interview said they didn't know enough about neighboring HHs to estimate HH size. An additional 5% said they might consider doing a proxy interview, depending on the enumerator's approach, available time, or which specific neighboring household it was; but only half said they knew enough to give a reliable answer about how many people lived there.

Census Bureau reports show another 7.9% of non-responding HH's in NRFU were resolved using administrative records (ARs). As might be expected, AR use in the Westside area was less than half the rate reported nationally (18.44%).[34] It is unclear how exactly this affected data quality in a sub-county area with much higher-than-average concentrations of undocumented immigrants (estimated to be in the 8%-20% range) for whom AdRec matches are infeasible or unreliable.[35]

Part of this uncertainty stems from the fact that there are not, typically, "undocumented households"; instead, many households are "mixed status" ones where some household members are US-born, legal permanent residents, or lawfully present (e.g. DACA recipients) with Social Security Numbers while others are not. The issue of Census 2020 enumeration of "mixed status" and "unauthorized households in farmworker communities is substantial because these households were less inclined to respond[36], because they are prevalent (about 19% households being mixed-status and 13% all-undocumented in one in-depth study of a San Joaquin Valley farmworker community), and because they are somewhat different in composition than the all-legal/citizen households.[37]

Another problem with quality of AdRec-based enumerations is that field research shows a higher proportion of complex households. Inevitably, in "complex households" (actually housing units) there are multiple households so reliance on an IRS return as a basis for imputing household size is flawed.[38] Moreover, in labor-intensive farmworker areas such as Fresno County with concentrations of workers lacking employment authorization, SSN-based datasets are known to be problematic because "identity loan" is prevalent.[39]

A final concern about the reliability of the post-NRFU data is that NRFU seeks to secure data for non-responding households that are in the Census Bureau's Master Address File (MAF). We estimate that the proportion of low-visibility/hidden housing units omitted from the MAF in Fresno County overall is likely to be about 3% and, based on field research by Kissam, that the proportion of households omitted from the MAF in the Westside region may be as high as 8%.[40] We also assume, given the high NRFU workload and the challenges confronting enumerators in 2020 that there were probably few non-ID responses or in-field MAF "adds" from enumerators.

The bottom line is that we do not know if the relationship between 2020 census response rate and final enumeration numbers that is observed at the national level for Hispanics provides the best estimate as to how census response is transformed into final enumeration numbers in the course of NRFU in the local context of hard-to-count sub-county areas of Fresno County or other similar areas. However we believe it is very unlikely that NRFU went better in the lower-response areas of Fresno County than in the nation overall and, as the discussion above suggests, it probably did not go as well.

Our analysis of the ratio of eventual post-NRFU census undercount to low response based on national-level PES and operational data from the Census Bureau can be considered to be something of a best-case scenario—because, as shown in the experimental AR Census 2020

analysis, these hard-to-count neighborhoods are ones where high-quality ARs are less often available and where AR matches to MAF-IDs are also problematic.

The remaining issue is that, due to application of differential privacy to Census Bureau reporting of tract-level operational metrics, it is difficult to quantify the extent to which NRFU was particularly challenging and whether data quality suffered significantly. While the published information on NRFU operations is available down to the tract-level, it must also be recognized that the operational reporting provides no information about the quality of the NRFU procedures. Where we do have some clues about the reliability of NRFU operations—due to reliance on proxy interviews and ARs—it is likely that that the hardest-to-count neighborhoods are those where the quality of NRFU procedures is also most uncertain.

Additional analysis (including analysis of the CUF to overcome small-area distortions due to application of differential privacy) would be very useful in order to improve understanding of the disparities in hard-to-count areas such as the Westside of Fresno County. Such understanding, even with the legal constraints on adjusting decennial census data would, at least, provide valuable insights to data users who are prepared to incorporate considerations about geographic patterns of differential undercount in relation to structural factors affecting census self-response and NRFU quality.

ACS Response Patterns Mirror Census 2020 Self-Response Rates But Are Lower

It is not surprising to see that ACS response patterns mirror those observed for Census 2020 self-response—but at much lower levels of response, especially for hard-to-survey populations. After all, the ACS was a replacement for the "long form" sample from the decennial census but was much more burdensome—in particular requiring high levels of literacy and numeracy skills to successfully respond. Moreover, the ACS has not had the communications budget or intensive messaging/response promotion campaign as the decennial census. Unit non-response and item non-response have both been increasing when measured at the national level.

We have not seen reports discussing the dynamics of ACS response at the local level but are very pleased that the Census Bureau has published details on patterns and that ACS response at the census tract level is now available as part of the PDB dataset.

Figure 2 below shows the response level for the ACS in Fresno County by HTC quintile and compares it to the patterns of Census 2020 self-response levels noted earlier. As expected, analysis of the correlation between ACS response level and Census 2020 self-response level shows them to be significantly correlated (Pearson correlation=.597, p<.01).

Figure 2: Fresno County Census 2020 Self-Response and ACS Response by Quintile
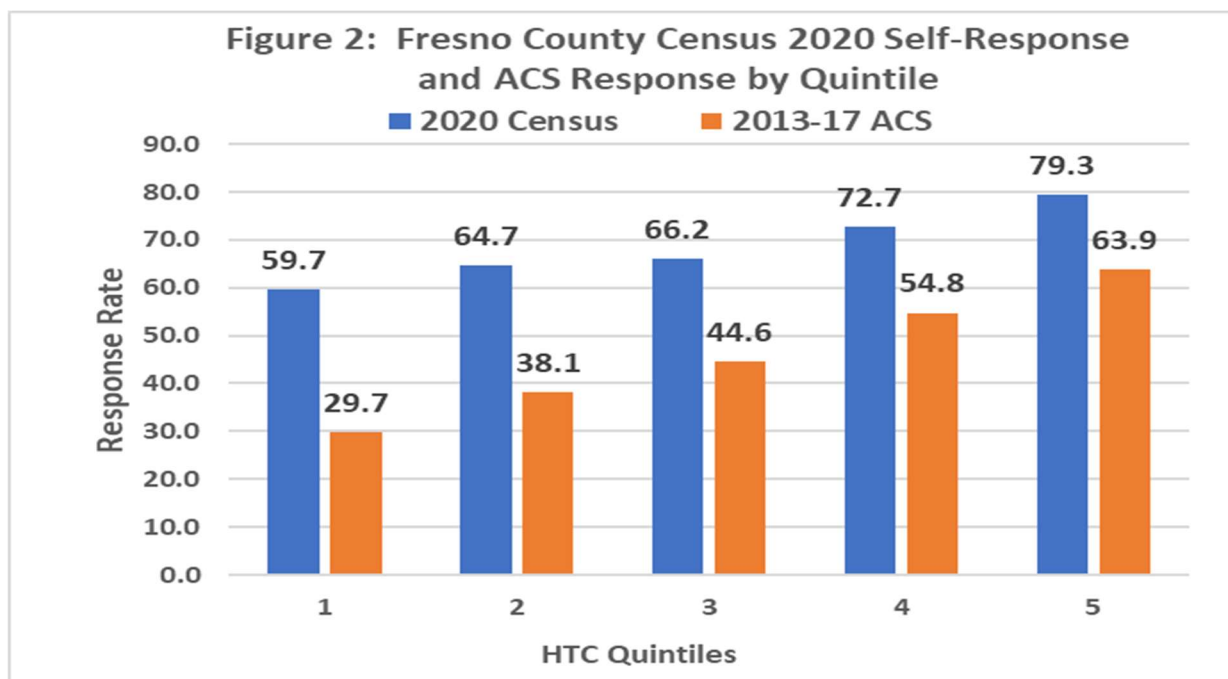
* The response rates in Figure 2 are based on the 2013-2017 ACS data in the planning database. The average ACS response rate in the 2015-2019 data is 1.0% lower (46.1% in the newer dataset vs. 47.1% in the earlier).

The correlation between Census 2020 self-response and ACS response levels in the continuum from harder-to-count to easier-to-count tracts shown in Figure 2 is cause for concern because it suggests that lowered response will synergistically undermine both the reliability of population count and the demographic/socioeconomic profile of neighborhoods and local communities. Moreover, the response gap between average response in the hardest-to-count and easiest-to-count quintile of tracts is even greater for the ACS (34.2%) than for 2020 Census response (19.6%). In hard-to-survey Quartile 1, the ACS self-response rate of 29.7% is only one-half the 2020 census response rate.

Specific Factors Affecting Local ACS Response in Fresno County

A regression model relying solely on "structural" factors similar to the one we use to explain decennial census self-response highlights key factors related to ACS response in Fresno County: Adjusted R square=.814, p<.001 Key factors significantly correlated with ACS response level are (in descending order of importance):

- % in tract without HS education: (B= -.436)
- % of HH's in poverty (B= -.301),
- % renter HH's (B= -.195)
- % limited-English Spanish-speaking HHs (B= -.116)

This analysis of factors related to ACS response provides valuable clues about potential ACS sample bias. It is not surprising to see that educational attainment is so prominent a factor in ACS response. We are not clear whether ACS responses are adequately weighted to compensate for lower response in the HTC tracts in Fresno County and other California Latino communities with concentrations of immigrants. ACS responses are weighted to conform to the census profile of each area sampled; a discussion of ACS weighting procedures (Rothbaum et. al 2023) indicates that weighting is done based on the poverty rate, rental rate, density of housing units, demographics (race, ethnicity, age, and sex) <u>but only at the county level</u>. Since the extreme variations in ACS response level we note in Fresno County are at the sub-county level, weighting at the county level would not seem to resolve the problem of varying ACS response being correlated with level of poverty and educational attainment.

Limitations of The Planning Database Data on Educational Attainment

It is important to recognize the cut-point in the PDB variable of high school education as distinguished from higher level is not ideally-suited to profiling the educational attainment in the HTC tracts with concentrations of Mexican immigrants. This makes proper weighting still more of a challenge. Although is likely not necessary to have a high school level of education to respond to some of the 44-question ACS survey, it is, at the same time, likely that response is difficult for the typical Mexican immigrant with less than 9 years of schooling.[41]

Table 4 below provides a rough idea of the situation—based on San Joaquin Valley Census Research Project (SJVCRP) survey data on Fresno County low-income Latino educational attainment and National Agricultural Worker Survey (NAWS) data on California farmworkers.[42]

Table 4 Educational Attainment of Farmworkers And Low-Income Latinos in Fresno County		
Educational Attainment	*NAWS 2015-2019*	*SJVCRP 2018*
0-6 years of school	46%	65%
7-9 years of school	23%	18%
10-12 years of school	24%	12%
> More than high school	7%	6%

Consequently, the correlation between educational attainment and ACS response may be even tighter than would be indicated by our regression using the PDB variable of % without a high school education in a tract because the average functional literacy (even in Spanish) of Mexican and Central American immigrants with less than a high-school education is much lower than that of US-born adults lacking a high school education, most of whom have at least gotten to 9th or 10th grade.[43]

Even if one looks simply at proportion of the population (adults 25+) lacking a high school education, the Mexican-origin population of Fresno county has a disproportionate level of

individuals with lower educational attainments: more than 40% in the Fresno County hardest-to-count tracts (HTC Quintiles 1 and 2) lack a high school education, as compared to 10.6% of the overall U.S. population.[44]

The excellent and illuminating analysis by American Community Survey researchers (Rothbaum et al 2021) provides valuable insights with implications for assessing the extent of sample bias resulting from concentrations of Mexican immigrants in Fresno County. The researchers note that at the national level, there is evidence of 2020 ACS sample bias vis-à-vis immigrants.

... the 2020 ACS data show a notable decrease in the noncitizen population, although this characteristic tends not to change much from year to year. The noncitizen population remained between 21.7 and 22.6 million during the years 2016 to 2019, but the 2020 ACS estimates the population at 20.1 million, declining 1.6 million from 2019. Some of this decline may be true demographic change. However, much of the observed decline is likely due to nonresponse bias, since the foreign-born—and noncitizens in particular—disproportionately respond to the ACS via in-person interview methods like CAPI that were curtailed in 2020. (Rothbaum et. al, p. 8)

The ACS researchers go on to describe innovative efforts to address weighting challenges related to apparent bias seen in the overall educational profile of 2020 ACS respondents. The other indicator of sample bias they identified was an unaccountable increase in the median household income in 2020. Rothbaum and his colleagues argue that the serious problems observed in the 2020 ACS were an anomaly. This is surely correct, to some extent, because the operational challenges in 2020 were daunting. However, there is, at the same time, reason to believe that these problems may have affected ACS inclusion of the less-educated immigrant sectors of the U.S. population in previous years.

The practical consequences of combined HTC-linked disparities in self-response rate, compromised NRFU quality, and resultant differential undercount in HTC tracts being correlated with ACS sample bias deserve further analysis beyond the scope of this paper. Nonetheless, what needs to be highlighted here is that the combination of differential census undercount and ACS sample bias stemming from differential response among lower-income households with lower educational attainment is particularly detrimental when it comes to use-cases where funding formulae require that both be used to allocate social program funding, our example being ESEA Title IA funding.

Implications of the Analyses of ACS and 2020 Census Response in Fresno County

We see our analysis of small-area variations in Census 2020 self-response and ACS response as a case study of the dynamics of census response and differential undercount with relevance to other similar counties across the U.S., most obviously the agricultural counties with concentrations of farmworkers[45], but more generally areas with high levels of socioeconomic diversity and neighborhoods with concentrations of Mexican and Central American immigrants.

We discuss some of the implications on the next page.

More Integrated Research and Planning To Better Understand and Respond To Multiple Interacting Factors That Affect Census Response

Our current analysis, showing that adoption of online response as the primary mode of census response had both positive and negative consequences for census accuracy and quality, points to the need for better integrated research and planning.

Internet response is a cost-effective mode of data collection that clearly worked well in 2020 for most U.S households since 80% of census self-response was via this modality. At the same time, our case study analysis of available data on statewide California response patterns and in Fresno County shows the extent to which this mode of data collection widened the gap between the digital "haves" and "have nots".

Operational decisions about decennial census treatment via Internet First or Internet Choice (data collection mode) were reasonably made but, apparently without much attention to how geographic patterns of broadband access intersected with digital literacy, local populations' accustomed modes of online interactions, or how concerns about online privacy as distinct from mail response or personal contact with an enumerator, might ultimately impact census accuracy. In particular, ethnographic research can provide crucial insights about how educational attainment, literacy level, and digital literacy specifically, would affect respondent ability of use this mode of response.

The bottom line here for assessing 2020 census quality and for future planning is that the major factors in low levels of self-response are not simply respondent motivation but, rather, the overall operational design of survey research. As Census Bureau research moves forward, and particularly as ACS design shifts more toward online data collection, it will be crucial to integrate currently-siloed applied research projects more thoroughly in order to stem the widening gap in response observed in different community contexts. Current research attention to ways to improve enumeration of undercounted populations and geographic areas is a very welcome development

For example, the Census Bureau's recent report on findings from an experimental initiative to conduct an AR-based census raises questions about the extent and patterns of differential undercount while, at the same time, providing important insights about the limitations and possibilities for more effectively using ARs in the enumeration process. The practical utility of this analysis stems in part from its having examined the fine-grained patterns of relationships between different facets of census operations (e.g. the availability and utility of ARs by type of address, by response mode, for sub-populations, by race/ethnicity, and other characteristics including immigrants with different legal statuses).[46]

The lack of appropriate research in preparation for 2020 is evident in that response options such as non-ID response were sensible but untested and apparently were instituted with little organization-wide awareness of the prevalence of doubled-up complex households and low-visibility/hidden households in low-income hard-to-count neighborhoods in states such as

California where housing costs had risen dramatically.[47] The Census Bureau has published metrics of the overall use of non-ID response (8%) but this has not been broken out by characteristics of respondents or by geographical analysis of how that modality of response went.[48]

Software design for Internet response apparently gave little attention to design of a user interface that would be user-friendly for the sorts of hard-to-count populations that relied primarily on mobile phones for connectivity. At the same time, the Census 2020 campaign to promote census response was impeded by Census Bureau guidance discouraging California outreach workers from assisting households in responding.

Our research shows how urgent the need is for more investment in ethnographic research as a basis for better planning to overcome the chronic problem of uneven self-response since our analysis shows it inevitably leads to eventual differential undercount, a phenomenon that has been evident for at least four decades.

Fundamentally ethnocentric misconceptions of the distinction between households (as defined by many in hard-to-count communities and the IRS) and housing units being considered households (as defined by OMB/Census Bureau residence rules)—clearly disadvantaged some of the hardest to count sub-populations in agricultural counties such as Fresno. The SJVCRP research, for example, showed that 12% of the Latino respondents only received mail at a mailbox/postal address they shared with other households and that 3% had no mail address at all. It will be crucial to field research to determine the prevalence and variation in prevalence of complex households in hard-to-count tracts and consider changes in both questionnaire design and census operations (e.g. sending 2 or 3 questionnaires to each address and encouraging each household in a doubled-up housing unit to respond).

An important practical outcome from heightened Census Bureau attention to the need for more nuanced understanding of the multiple factors entering into differential levels of survey response in diverse U.S. neighborhoods and communities, would be a shift in investment toward partnerships to develop long-term, continuous community collaboration with multiple stakeholders working toward building digital literacy and broadband access

More Nuanced Understanding of Race, Ethnicity, and Sociocultural Identity Is Needed In Order to Improve Census Response in Hard-To-Count Communities

Our analysis shows that tract-level self-response to the decennial census and ACS is, indeed, correlated with local community race/ethnic profile as defined in the OMB/Census Bureau taxonomy but less powerfully than by "structural" factors. This has a number of practical implications.

Within the local sociocultural context of day-to-day life in Fresno County, a large Hispanic-majority area, there are many nuanced gradations of individual, household, neighborhood, and community identity. We do not see our research as indicating that personal ethnic identity is

unimportant, but that the "textures", the fine-grained patterns of self-identification and group identity are very poorly captured in the OMB/Census Bureau crudely-articulated and arbitrary taxonomy of race/ethnicity, and, specifically, the umbrella quasi-racial designation of "Hispanic".

Current efforts to revise tabulation of racial/ethnic identity to include identification of multi-racial populations is welcome but inadequate without more underlying research on conceptualization of ethnic and personal identity. Our analysis of the situation in Hispanic-majority communities and counties such as Fresno and Los Angeles is simply one strand in a broader call for more sensitivity to the reality that personal, household, and community identity emerge from an interaction between conceptualizations of ethnicity framed in terms of language and cultural heritage, national origin, and legal status.[49]

Census Bureau planning, both for the decennial census and the ACS, and in Latino communities in California and throughout the U.S., will require greater attention to the fact that immigration/citizenship status is an important overlay over "Hispanic" identity and response in communities with concentrations of immigrants and to more explicitly address the idea that everyone residing in the U.S. is to be counted in the census.

The San Joaquin Valley Census Research Project, for example, gained many insights into the complex dynamics about how status-linked identity powerfully and variably affected interactions between young DACA recipients and their undocumented parents in discussions about the pros and cons of a household responding or not responding to the census. What became evident in examining factors affecting household response in this hard-to-count population was that a decision to respond or not to respond was only made after extensive internal discussion within the household. As in other sectors of U.S. life, trust in "the government" was fragile and concerns about privacy were unaffected by Census Bureau messaging about its processes for safeguarding respondent confidentiality. The Census Bureau's conceptualization of messaging to promote census response as being directed toward individuals rather than as triggers for household, network, and community discussion was, consequently, sub-optimal.

Here too, more investment in ethnographic research is called for. Research by the Census Bureau's Center for Behavioral Science Methods is making valuable contributions and the market segmentation analysis emerging from CBAMS was relevant to Census 2020 messaging but not very well integrated into the overall communications campaign. More research and better integration of research findings into efforts to promote survey response should be a priority.

Breaking The Vicious Cycle Where Social Disadvantage Leads to Low Response, Differential Undercount, and Subsequent Inequitable Allocation of Social Program Funding

Our analysis shows that multiple factors interact to determine local levels of census response. The consistent prominence of community levels of poverty (% of tract population in poverty) as explanatory factors for tract-level survey response, and the somewhat correlated mix of homeowner and renter households as important factors, highlight the need to give careful

attention to the ongoing challenges of effectively addressing the ways that socioeconomic advantage or disadvantage determine community-level census and ACS response.

Since enumeration is inextricably linked to local level of response, as shown by our review of the historical and current relationship between census response and differential undercount, more priority will need to be given to assuring equitable opportunity to respond. That, in turn, implies, at the very least, vigorous and ongoing steps toward overcoming the digital divide (not just as part of pre-census messaging) since Internet response is, in fact, a cost-effective 21st century tool for data collection, coupled with provisions to demolish the language barrier that impedes census response from many immigrant families.[50] A more robust language access program is also crucial and feasible.

We understand that the political pressures experienced by the Census Bureau are skewed toward efforts to decrease survey cost rather than toward quality improvement. However, from the perspective of survey data users, a census that reliably "mirrors America" is much more valuable than a mere count of the number of persons in the U.S.

The Use-Case of Decennial Census and ACS-Derived Data for Allocating Compensatory Education Funding

An example of an important use-case for census-derived data and the need for improved accuracy is allocation of federal funding for ESEA Title IA compensatory education. By statute, funding is allocated based on number of children living in poverty and concentration of poverty in a school district. This program, designed to overcome the educational disadvantage experienced by children in low-income families, allocates about $16 billion in funding annually.[51] Allocation is guided by number of children in poverty in each school district.

Although the funding source is the federal government, actual funding cascades downward from the Department of Education to state educational agencies and then to local educational agencies, K-12 school districts.[52] In FY21-22 California allocated slightly less than $1.8 billion in ESEA Title IA funding. Fresno County received slightly over $124 million in funding. Based on Small Area Income and Poverty Estimates (SAIPE) derived from decennial census and ACS data as well as other data sources,[53] the school districts in the hard-to-count Westside sub-county region in our study 10 LEAs received about $8.2 million.[54] National and state level PES-based analyses of differential undercount for the broadly-defined OMB-defined race/ethnic groups is illuminating but not adequate to improve educational equity in disadvantaged communities.

Clearly, the financial stakes are high. Recent efforts to rely on a blended base for Population Estimates used in this funding are an important step forward.[55]

Equitable educational opportunity for children in these marginalized neighborhoods and communities where the majority of parents are, themselves, educationally disadvantaged, rests squarely on census and ACS data accuracy; census undercount of Hispanic children and likely underrepresentation of households in poverty based on ACS data fuels the ongoing cycle of

social disadvantage.[56] Although the impacts are much less direct, similar issues affect allocation of workforce skills development program funding (WIOA) that boost household earnings and children's educational opportunities. A recent report from The Census Project provides a timely and insightful review of other crucial uses of ACS data in market research and program planning.[57]

Use Of Census-Derived Data In Public Health Planning and Allocation of Funding from Multiple Sources

Andrew Reamer has written extensively and definitively on the multitude of federal funding programs that are driven by a combination of decennial census and ACS-derived data; his research and analysis provides compelling evidence of the important of data accuracy.[58] A very recent report from the Census Bureau, using Reamer's analytic framework, now updates the information on federal program funding based on census-driven data.[59] Here we would only add that census equity also has very direct impacts on planning of funding for public spending and philanthropic investment that is not statutorily census-driven.

An example of the stakes here is the use-case where census and ACS data, combined with other data sources, are used to characterize communities as socially vulnerable. The CDC's Social Vulnerability Index (SVI) and the Healthy Places Index (HPI) developed by the Public Health Alliance of Southern California were widely used in public health response to the COVID-19 pandemic. The HPI was used to assure equitable allocation of state funding (some, but not all of it, from federal sources) of COVID-19 vaccination efforts by requiring county public health departments to achieve adequate levels of vaccination in the census tracts with the lowest level of vaccination. HPI values, indicating healthier neighborhoods, are very closely correlated with ACS response rate and with Census 2020 self-response.[60] Population underestimates in farmworker areas throughout California almost certainly skewed vaccination campaign funding since vaccination rates were computed based on public health records of vaccinations as the numerator and Population Estimates program data on tract population as the denominator.[61]

Summary Conclusions

The discussion and evidence presented here based on our case study of Fresno County adds to the compelling national-level analysis presented by O'Hare and Lee (2021) that variations in Census 2020 self-response rate observed at the national level and linked to different demographic groups are likely indications of systematic differential undercount.

We have built on the recent findings by Robinson, Kissam, and O'Hare (2023) showing that Census 2020 self-response is consistently lower in hard-to-count tracts than in easier-to-count tracts and that many of those tracts are Hispanic-majority ones.

Here we have drilled deeper to explore the patterns of Census 2020 self-response and ACS response at the tract level and in sub-county regions. We have also examined how patterns of self-response vary in Fresno County PUMAs and believe that differing levels of self-response

between PUMAs affects the quality of PUMA data.[62] It would be good to explore the extent to which these within-county variations in self-response vary in other areas of the U.S. and to analyze local variations not simply in relation to LRS but to explore distinctive factors accounting for local variation.

Our current analysis uses Fresno County and the sub-county Westside region of municipalities and hamlets with concentrations of farmworkers and immigrants as a case study in order to delve into details. We cannot argue for the generalizability of the insights emerging from this case study but we can argue that the findings are relevant to many similar areas across the U.S. We would be very interested, for example, to learn if similar response patterns can be observed in other areas of the U.S. with concentrations of Hispanic households, for example, in southern Arizona, border communities in Texas, and in New York City.

From a practical planning and policy perspective it is important to move beyond the valuable evidence of disparities in census enumeration of racial/ethnic groups yielded by the PES to understand that the multiple factors giving rise to varying levels of self-response can be analyzed to map out fine-grained geographic patterns of census response that give rise to eventual underrepresentation in census data and in ACS-based community profiles.

Our analysis shows that the issue of census equity cannot be addressed simply as a matter of state-level census accuracy, i.e. the fairness of census-based Congressional apportionment, or for underrepresented populations but that it must also be understood as an issue of equity for local communities and institutions.[63] The organizational stakeholders in accurate and fair census and ACS representation are not simply federal programs and state government but, also, institutions and service programs in small communities such as Huron and Mendota in Fresno County as well as large cities such as Detroit, Los Angeles, and New York City.

Why is local community equity important? Because, practically speaking, as we note in our discussion of allocation of ESEA Title IA funding that is driven by census and ACS data, a desirable real-world outcome would not simply be to have appropriate levels of compensatory education funding nationally and at the state level, but for local schools to be allocated a fair share of available funding to respond to the distinctive educational needs of local children disadvantaged by poverty. Where the rubber meets the road with respect to social program impact is local communities' ability to have the funding they need to provide effective services to their residents.

As recognized in discussions of the vision of a "curated data enterprise", the suitability of census data for a specific use case related to allocation of particular funding or planning needs to be assessed in light of intended use. Assessment of census quality and consideration of strategies to incorporate insights about reliability of data elements needs to be linked both to programmatic context (e.g. business investment and community development, education, workforce skills development, or public health planning) but, also, in the context of life in a particular community.

We agree with other analysts that Census 2020 was an anomaly—in large measure due to the operational challenges posed by the COVID-19 pandemic. However, variations in response were also occasioned by introduction of the new modality of Internet census response which worked well overall but provided greater advantages for households, neighborhoods, and communities with more affluent households with higher levels of broadband connectivity. This surely affects allocation of funding based on decennial census and ACS data and may well have affected the accuracy of the PL 94-171 file as well.

Our analyses showing that tract-level, neighborhood, and community profile with respect to immigration/citizenship status suggest that the acrimonious national debate about adding a citizenship question to the census (and in some proposals, a still more problematic query about authorized vs. unauthorized immigration status) may have played a role in varying response levels too.[64] However, we believe that it can also be argued that the challenges in Census 2020 were not so new but, rather, that longstanding chronic "systemic" problems were exacerbated by a perfect storm of unanticipated developments.

The primary analysis here is quantitative but it has been informed by prior ethnographic and survey research, including what is now a rich research literature on the experience of immigrants now settled in the U.S. We have focused on counties and communities with concentrations of Hispanics because our resources are limited but we believe, at the same time, that similar analysis and research is called for to better understand equity and inequity for other communities, most obviously, Alaska Native and American Indian (AIAN) ones, given the evidence of serious disparities they face in terms of equitable census representation.

Many commenters responding to the Census Bureau's call for planning input have already urged that planning for Census 2030 give priority to improving enumeration of hard-to-count groups and we are delighted to see that Census Bureau internal research and planning includes broad collaboration focused on reducing differential undercount.

Our analysis here supports those recommendations and points to several specific issue areas. A priority is the need to concurrently increase broadband access and digital literacy especially in communities with lower levels of functional literacy such as those we discuss here where there are concentrations of Mexican immigrants.

Another facet of data collection improvement will need to be improved language access for minority language households. Operational priorities also include ongoing efforts to enhance Master Address File building so as to include more low-visibility housing units and facilitate census enumeration, and revisions to the definition of household and/or questionnaire design and operational improvements to improve enumeration of "secondary" households in complex doubled-up housing units.

Our analysis of the linkage between the dynamics of decennial census and ACS response support the recommendations put forward by The Census Project and others on using the ACS as a testbed for research to improve Census 2030.

APPENDIX

**Appendix Table 1: Comparison of Selected Community Characteristics
for the United States, California, Fresno County, and the Westside sub-county region**

Community Characteristics	Westside sub-county region (9 census tracts)	Fresno County	California	U.S.
Hispanic	85.1%	52.3%	39.4%	18.9%
Asian-alone	1.1%	9.7%	15.9%	6.1%
Black-alone	0.9%	4.7%	6.5%	13.6%
% Foreign-born	43.5%	20.6%	26.5%	13.6%
% Non-citizen	35.81%	12.3%	12.5%	6.7%
% naturalized citizen	7.7%	8.4%	14.1%	6.0%
Estimated % undocumented	8%-20%	7.7%	6.0%	3.2%
% in poverty	31.7%	23.1%	12.3%	11.6%
% 25+ with less than a HS Education	55.9%	25.1%	15.8%	11.1%
% renters	58.5%	48.1%	44.5%	35.4%
% with broadband connectivity	40.4%	60.8%	90.4%	87.0%
% vacant units	6.6%	6.5%	7.0%	10.6%

*Community Characteristics—Average tract level for Fresno County and Westside, from PDB ACS15-19 tabulations, rounded, tract characteristics (not weighted by population). California and US data from Census Quick Facts. Weighting tract characteristics by the estimated 2015-2019 population in each tract makes only minimal changes in the analysis.

** Estimated Fresno County undocumented, Robert Warren, Center for Migration Studies, estimated Westside undocumented Ed Kissam, undocumented in California and U.S. from Pew Research Center
https://www.pewresearch.org/hispanic/interactives/u-s-unauthorized-immigrants-by-state/
https://www.pewresearch.org/short-reads/2020/08/20/key-findings-about-u-s-immigrants/

***U.S. and California non-citizens from Kaiser Family Foundation (2019) based on ACS 1-year estimates https://www.kff.org/other/state-indicator/distribution-by-citizenship-status/?currentTimeframe=1&sortModel=%7B%22colId%22:%22Location%22,%22sort%22:%22asc%22%7D

**** Housing vacancy rate for U.S. and California from American Housing Survey, Table 5, Gross Vacancy Rates by State 2005-2022.

[1] J. David Brown et.al. "Real-Time 2020 Administrative Record Census Simulation: A New Design for the 21st Century", May 5, 2023, U.S. Census Bureau

[2] Sallie Keller et. al, "A 21st Century Curated Data Enterprise: A Bold New Approach to Create Official Statistics", University of Virginia Biocomplexity Institute, 2022 https://biocomplexity.virginia.edu/sites/default/files/census/reports/UVA_Census_CDE_report_spring2022.pdf

[3] https://detroitmi.gov/news/city-files-federal-lawsuit-challenging-2021-us-census-population-estimate

[4] https://www.census.gov/programs-surveys/decennial-census/about/coverage-measurement/pes.2000.html#list-tab-400924250 These have been referred to variously over the past half-century: PEP in 1980, PES in 1990, ACE in 2000, ACS in 2010, and, once again, PES in 2020.

[5] For an early discussion of undercount of Black males, see J. Gregory Robinson, Bashir Ahmed, Prithwis Das Gupta & Karen A. Woodrow (1993) Estimation of Population Coverage in the 1990 United States Census Based on Demographic Analysis, Journal of the American Statistical Association, 88:423, 1061-1071, DOI: 10.1080/01621459.1993.10476375. There is also an excellent discussion of U.S. undercount of young children in the international context, William O'Hare "An international perspective on the undercount of young children in the U.S. Census" Statistical Journal of the IAOS 33 (2017) 289–304 289 DOI 10.3233/SJI-161008 and a recent Census Bureau report on the current situation from Eric Jensen, "Despite Efforts, Census Undercount of Young Children Persists", U.S. Census Bureau, March 10, 2022.

[6] Hansi Lo Wang, "The 2020 census may have missed a big share of noncitizens, the bureau estimates", National Public Radio, May 8, 2023 https://www.npr.org/2023/05/08/1174824891/2020-census-noncitizens

[7] 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. A presentation by Joanne Pascale, "A Meta-Analysis of Ethnographic Studies of Undercounted Populations in the U.S.", AAPOR presentation, May, 2022 provides a broad overview of the research and continuing issues.

[8] Fein DJ, West K. The Sources of Census Undercount: Findings from the 1986 Los Angeles Test Census, Survey Methodology, December, 1988.

[9] A comprehensive and up-to-date review of the research on census undercount is William O'Hare's book, **Differential Undercounts in the U.S. Census**, Springer Link/Open Access, 2019. https://link.springer.com/book/10.1007/978-3-030-10973-8 The analysis of 1990 decennial census farmworker undercount was based on a meta-analysis of diverse survey data on farmworkers. See Gabbard, S., Kissam, E., and Martin, P. (1993). The impact of migrant travel patterns on the undercount of Hispanic farm workers. In *Proceedings of the Bureau of the Census Research Conference on Undercounted Ethnic Populations.* Richmond, VA: Bureau of the Census. See also Kissam E. Census Enumeration of Immigrant Communities in California: "Dramatic Improvements but Challenges Remain. Report to California Rural Legal Assistance and The California Endowment", JBS International, September, 2010. In reviewing the 2020 Census data, the Economics and Statistics Administration determined that the decennial census data was not the "best" data for allocating federal funding for farmworker program funding (now WIOA 167). The U.S. Department of Labor and, subsequently, the Legal Services Corporation, adopted alternative approaches to guide migrant program funding.

[10] Bruce, Antonio, J. Gregory Robinson, Jason E. Devine. "A Planning Database to Identify Areas That Are Hard-to-Enumerate and Hard-To-Survey in the United States." Population Division, U.S. Bureau of the Census, 2012.

Available at http://www.asasrms.org/Proceedings/H2R2012 See, also, Bates, Nancy and Mary H. Mulry. 2011. "Using a Geographic Segmentation to Understand, Predict, and Plan for Census and Survey Mail Nonresponse," Journal of Official Statistics, Vol. 27, No. 4, 2011, pp. 601–618.

[11] Chandra and Nancy Bates (2017), "The Low Response Score (LRS): A Metric to Locate, Predict and Manage Hard-to-Survey Populations", Public Opinion Quarterly, Vol. 81, No. 1, Spring 2017, pp.144–156 https://www.census.gov/content/dam/Census/topics/research/erdman_bates_2017.pdf

[12] California Census 2020, "Identifying California's Hard To Count In Census 2020", https://census.ca.gov/california-htc/ California has, for several decades, invested in outreach efforts to promote census response—but the outreach was primarily media-based in past decades. What was new in 2020 and facilitated by the development of the California HTC index was "tailored" outreach by a wider range of local community-based organizations identified on the basis of their experience in serving and ability to provide response encouragement and assistance to diverse hard-to-count sub-populations in each region of the state. The outreach campaign conducted "on the ground" by hundreds of community-based organizations across the state was provided technical support from the state and the HTC played a central role in managing outreach. https://census.ca.gov/wp-content/uploads/sites/4/2021/04/California-Census-2020-Outreach-and-Communication-Campaign-Final-Report.pdf

[13] 25% in Los Angeles County https://www.census.gov/quickfacts/fact/table/losangelescountycalifornia/RHI425221 and 27% in Fresno County https://www.census.gov/quickfacts/fresnocountycalifornia

[14] The source of the 17.5 percentage points estimate is the difference of the response rate for HTC Quintile 1 and HTC Quintile 5 based on the tabulation of Census 2010 response data from the 2019 tract-level Planning Database and the 2020 self-response rates provided by SwORD (California's Statewide Outreach and Rapid Deployment) data. The SwORD database had daily tract-level updates of Census 2020 response from the Census Bureau. For details see https://census.ca.gov/wp-content/uploads/sites/4/2021/04/California-Census-2020-Outreach-and-Communication-Campaign-Final-Report.pdf The California HTC index included a variable for broadband access quite similar to that in the Census Bureau Planning Database but derived from California Public Utilities Commission data that were believed to be more accurate. The poverty metric used as an indicator of poverty in the California HTC also diverged slightly from the similar component of the Planning Database using an indicator of 150% poverty—a useful local adaptation, given the fact that California rental costs that make up such a major component of low-income household budgets are higher than in the rest of the U.S.

[15] The Pew Research Center did an early summary in 2004 emphasizing generational differences in identity and perspectives on social issues. https://www.pewresearch.org/hispanic/2004/03/19/generational-differences/ There is now a robust research literature on dimensions of Hispanic diversity with the work of Carola Suarez-Orozco being particularly insightful and detailed in describing multiple dimensions of diversity and linkages to civic participation. https://www.umb.edu/editor_uploads/images/college_education/6_It_s_Sort_of_My_Calling_The_Civic_Engagement_and_Social_Responsibility_of_Latino_Immigrant_Origin_Young_Adults.pdf More recently, a leading sociological researcher, Roberto Gonzalez, in his research on Dreamers and DACA recipients (Lives in Limbo: **Undocumented and Coming of Age in America** University of Californioa Press, 2015) has documented for an important sub-group of Generation 1.5 immigrants, many of them of Hispanic origin, the complex overlay of legal status. The SJVCRP research in Fresno County provides insights about their perspectives on Census 2020 response and their role in household decisions in mixed-status families about responding or not.

[16] Unlike the original HTC scores which were developed based on the basis of ethnographic research and demographic, the LRS scores are not independent of the response rates (the scores were developed via regressions using 26 ACS variables as the independent variables and the 2010 response rates as the dependent variable). The California HTC was developed using an approach similar to the original HTC one.

[17] Susan Gabbard, et al (2012), "Identifying Indigenous Mexicans and Central Americans in Surveys" http://www.asasrms.org/Proceedings/H2R2012/Identifying_Indigenous_Mexicans_and_Central_Americans_in_Surveys.pdf   See also Richard Mines, Sandra Nichols, and David Runsten, "California's Indigenous Farmworkers" http://www.indigenousfarmworkers.org

[18] In addition to the standard distinction between immigrants who are legal permanent residents, those who are naturalized U.S. citizens, and those who are unauthorized immigrants, there is a significant proportion of immigrants in what we refer to as "fragile" legal status; this includes TPS beneficiaries, DACA recipients, asylum-seekers who are lawfully present but whose status may be affected by court decisions.

[19] Declaration of Dr. William P. O'Hare Case 8:19-cv-02710-PX-PAH-ELH Document 112-4 Filed 09/01/20

[20] J. David Brown et.al. "Real-Time 2020 Administrative Record Census Simulation: A New Design for the 21st Century", May 5, 2023, U.S. Census Bureau. The meticulous discussion of issues related to availability of administrative records (ARs) and difficulties of matching them and in linking them to MAF IDs has many implications for assessing NRFU quality.

[21] The only currently available response rate data for the 2020 Census at the time of this analysis were the self-response rates based on all housing units, which include vacant units. To estimate the Census 2020 occupied housing unit response rate for all occupied units (the 76.6% in Table 2), we added the 2020 vacancy rate (9.7%) to the current self-response rate (66.9%). Although census vacancy rates are not available for demographic characteristics, we derived an estimate for Hispanics based on the vacancy rates of tracts where Hispanics represented 50 percent or more of the total tract population. In 2010, the vacancy rate of Hispanic majority tracts nationally was 9.1 percent, compared to 11.4 percent for all housing units. The 2020 vacancy rate of 9.7 for all housing units represented a decline of 1.7 percentage points from 2010, a change we assume applies to Hispanic-majority tracts. Subtracting 1.7 points from the 2010 Census Hispanic-majority vacancy rate of 9.1 percent yields a 2020 vacancy rate estimate of 7.4 percent.   The 2020 Census Hispanic-majority self-response rate is estimated to be 60.4% (O'Hare and Lee, 2021)—adding the vacancy rate of 7.4 percent gives an estimated Hispanic-majority occupied unit response rate of 67.8% (Table 2). The estimated Census 2020 occupied housing unit response rate can be adjusted when the Census Bureau released the detailed 2020 data.

[22] The California state report is at https://census.ca.gov/wp-content/uploads/sites/4/2021/04/California-Census-2020-Outreach-and-Communication-Campaign-Final-Report.pdf See, also Joanna Lee, Jennifer Ito, and Manuel Pastor, "Census 2020 in the San Joaquin Valley: An Empirical Assessment of Strategies to Activate Hard-To-Count Populations", Equity Research Institute, University of Southern California, April, 2023. https://dornsife.usc.edu/eri/census-2020-in-the-san-joaquin-valley/ California designated 10 sub-state regions for its "Get Out The Count" initiative, one of which consisted of the southern San Joaquin Valley and one small Eastern Sierras county. Fresno County's average HTC score is very close to that of the San Joaquin Valley as a whole,

[23] Kissam's research on Census 2010 participation showed that mailing of the bilingual English-Spanish questionnaire to areas with high concentrations of Hispanic households was very helpful to many respondents who would, otherwise, have had difficulty in responding. However, there are other language minorities in Fresno County (Mexican immigrants whose primary language is an indigenous one, Hmong, Cambodians, and Punjabi Sikhs), as well as Spanish-speaking households in tracts with lower concentrations of linguistically-isolated households that the bilingual questionnaire and online response option did not help. A Census Bureau paper (Siegal, Martin, and Bruno 2007) presents an excellent analysis of the distribution of linguistically-isolated households which Kissam relied on in a working paper for discussion of Census Bureau planning related to the option of mobile questionnaire assistance centers (which could not be implemented due to COVID-19 shelter in place constraints). For details, see Ed Kissam, "The Need for Questionnaire Assistance for Linguistically-Isolated Spanish-speaking Households That Do Not Receive a Bilingual Invitation or Census Questionnaire", WKF Fund, October, 2019.

[24] Our regression analysis here is based on the most recent Census Bureau Planning Database data. Planning Database with 2010 Census and 2015 – 2019 American Community Survey Data (2021). It should be remembered that these analyses use Census 2010 geography, not the updated 2020 geography because data from the pre-census ACS and Census 2010 are reported in that geographic framework. An excellent resource for understanding the relationship between the earlier and the current geography is the Census Bureau's TigerWeb Decennial website: https://tigerweb.geo.census.gov/tigerweb2020/

[25] In late 2018, the San Joaquin Valley Census Research Project conducted a split-panel survey of low-income Latinos in the region with about one-quarter of survey respondents from Fresno County. Even before prompting respondents about willingness to respond to the census if a citizenship question were included, willingness to respond was strongly correlated with household legal status. Analysis of the quantitative data can be found in Kissam, E., Mines, R., Quezada, C., Intili, J., and Wadsworth, G. "San Joaquin Valley Immigrants: Implications of Survey Findings for Census 2020". Sierra Health Foundation Center/California Institute of Rural Studies, January, 2019.

[26] The preferred model explaining 2020 census self-response does well with respect to multicollinearity diagnostics. It has modest VIF values for coefficients (all under 3.1) as well as modest values for condition index (all under 10).

[27] See https://www.ers.usda.gov/webdocs/DataFiles/53180/25559_CA.pdf?v=0

[28] Researchers at the Public Policy Institute of California (Hayes et al 2023, Hayes, Gao, and Starr 2022) have done extensive and excellent analysis of patterns of broadband access in California and point out that federal and state mapping of broadband access almost certainly overstates low-income households' access. They call for more granular mapping and highlight the fact that many factors (e.g. current awareness of and enrollment in subsidized IP subscriptions and digital literacy affect real-world level of connectivity.

[29] (Cindy Quezada, personal communication, 2020). Dr. Quezada managed the state-funded Census 2020 outreach in the Southern San Joaquin Valley (Kern, Tulare, Kings, Fresno counties) for the Sierra Health Foundation, California's regional intermediary for ACBO sub-state region VI.

[30] For a definitive quantitative analysis, also, Hans Johnson, Belinda Reyes, Laura Mameesh, and Elisa Barbour, "Taking The Oath: An Analysis of Naturalization in California and in the United States", Public Policy Institute of California, 1999. https://www.ppic.org/publication/taking-the-oath-an-analysis-of-naturalization-in-california-and-the-united-states/ Their logit model demonstrates strong correlation between educational attainment and English-language competency in explaining county-level patterns of naturalization. Kissam further analyzed this issue as part of a multi-year impact evaluation of the James Irvine Foundation's Central Valley Partnership for Citizenship. See Ed Kissam, "Impact evaluation of the Central Valley Partnership Initiative", report to the James Irvine Foundation, The Aguirre Group, 1999. The San Joaquin Valley Census Research Project also observed that naturalized Mexican immigrants were much more motivated to respond to the census than other Hispanic sub-populations (including US-born Hispanic respondents). Salvadoran immigrants, the largest local population of non-Mexican Hispanic immigrants also have lower educational attainment and barriers to naturalization. The relationship between educational attainment and naturalization is less pronounced among the Hmong, the largest Asian ethnic group in Fresno County, due to legislation waiving the English-language requirements and demonstration of civic knowledge—the Hmong Veterans' Naturalization Act of 2000 (which also benefited veterans' spouses) https://www.congress.gov/bill/106th-congress/house-bill/371

[31] Analysis of qualitative data generated in the course of the survey can be found in Kissam, E., Wadsworth, G, and Intili, J. "Troubled Reflections: Summary of Themes and Implications for Census 2020", San Joaquin Valley Census Research Project, San Joaquin Valley Health Fund, 2019.

[32] Estimate based on average employment in agriculture as reported in American Community Survey for zipcodes associated with the Westside sub-county region.

[33] We considered the extent to which the difference in reported self-response might stem from the Westside sub-county area having a vacancy rate different than that for the overall county. The PDB data shows there to be a vacancy rate in the Westside area of 6.6% (PDB15-19). The overall county vacancy rates is reported to be 6.5% (PDB15-19). Therefore, differences in vacancy rate are likely to make only a negligible contribution to the differential in response rate.

[34] Michael Bentley, "Examining Operational Quality Metrics", Decennial Statistical Studies Division, Census Bureau, April 26, 2021. https://www.census.gov/newsroom/blogs/random-samplings/2021/04/examining-operational-metrics.html

[35] The estimate of at least 7.7% undocumented population in the Westside assumes the concentration in the sub-county region is similar to that of Fresno County overall is based on analysis based on ACS data by Robert Warren of the Center for Migration Studies of New York provided to the San Joaquin Valley Census Research Project. Warren, as well as other leading immigration research experts such as Jeffrey Passel and Manuel Pastor, use an analytic approach referred to as "logical imputation" to derive estimates of the number and characteristics of undocumented immigrants from ACS data. In fact, with about one-third of households in the Westside headed by a farmworker and about half of California farmworkers are believed to be undocumented (National Agricultural Worker Survey-California Report, 2022), the actual proportion of undocumented adults in the sub-county area may well be in the 15%-20% range. The estimate of overall proportion of the U.S. population that is undocumented and the proportion in California is from Robert Warren, "US Undocumented Population Continued to Fall from 2016 to 2017 and Visa Overstays Significantly Exceeded Illegal Crossings for the Seventh Consecutive Year", Journal on Migration and Human Security 2019, Vol. 7(1) 19-22 DOI: 10.1177/2331502419830339 journals.sagepub.com/home/mhsb

[36] Patricia Leigh Brown, "The Supreme Court Should take a Field Trip to California", New York Times, June 15, 2019 https://www.nytimes.com/2019/06/15/opinion/sunday/census-citizenship-immigrants.html?rref=collection/byline/patricia-leigh-brown&action=click&contentCollection=undefined&region=stream&module=stream_unit&version=latest&contentPlacement=1&pgtype=collection

[37] Edward Kissam, "How Immigration Status Constrains Immigrant Agricultural Worker Households' Access to Health Care: Evidence from the New Pluralism Study" paper presented to the American Public Health Association annual conference, August, 2008.

[38] For a description of the 2019 Fresno County community-based address canvassing research that identified low-visibility housing units, see Ari Berman, "Hidden Figure: How Donald Trump is Rigging the Census", June, 2019 https://www.motherjones.com/politics/2018/03/donald-trump-rigging-2020-census-undercounting-people-of-color/

[39] Sarah Horton, "Identity Loan: The Moral Economy of Migrant Document Exchange in California's Central Valley", The American Ethnologist, Vol. 42, #1, February, 2015.

[40] Estimated by Ed Kissam based on a pilot for a California LUCA-linked initiative using local community members to conduct address canvassing to identify low-visibility/hidden housing units. The pilot included canvassing in several Fresno County tracts believed to be likely to have concentrations of hidden housing units. For details of the address canvassing approach to identifying hidden housing units see Kissam, E., Quezada, C. and Intili, J. "Community-Based Canvassing to Improve the Census Bureau's Master Address File: California's Experience in LUCA 2018," Statistical Journal of the International Association of Official Statistics, Vol. 35, December 2018. https://content.iospress.com/download/statistical-journal-of-the-iaos/sji180480?id=statistical-journal-of-the-iaos%2Fsji180480 The approach we used, "community-based address canvassing" relied on locally-

knowledgeable trained observers looking for indicators of low-visibility housing but it is possible that even this approach may miss some hidden housing units. Kissam piloted the approach in 20 blocks of a farmworker community in Riverside County and estimated that 3.65% of the housing units were not included in the MAF (based on 2019 Census Bureau LUCA data and his own canvassing). His block-level count of housing units agreed with the MAF in about half of the blocks.

[41] Functional literacy is, of course, only modestly correlated with educational attainment. Nonetheless, Kissam's census research draws on findings from the National Adult Literacy Survey that suggest Hispanic immigrants with lower levels of educational attainment have particular difficulty with highly-formatted questionnaire material. The requisite literacy tasks become still more daunting in responding to the ACS where functional competencies include the ability to reference multiple financial records and/or recall detailed financial data. Kissam, E., Herrera, E. and Nakamoto, J.M. (1993). Hispanic Response to Census Enumeration: Forms and Procedures. Report submitted to the U. S. Census Bureau under Contract No. 50-YABC-2-66027, Task Order No. 46-YABC-2-0001. For a review of the national picture of differences in literacy between immigrants and US-born/raised individuals, see Kirsch et al (2004) https://www.ets.org/research/policy_research_reports/publications/report/2004/csfc.html

[42] California NAWS data is likely to be highly representative of Fresno County farmworkers because the sampling methodology is designed to conform to county levels of seasonal agricultural worker employment. (https://www.dol.gov/sites/dolgov/files/ETA/naws/pdfs/NAWS%20Research%20Report%2015.pdf) Fresno County, with about 100,000 farmworkers is one of the top three farmworker counties in California. https://www.researchgate.net/publication/363350077_How_Many_California_Farm_Workers

[43] There are also issues about the relationship between educational attainment and functional literacy that particularly affect interpretation of Mexican immigrant literacy in Fresno County where most immigration has been from rural Mexican migrant-sending villages to U.S. farmwork; this is because of serious deficiencies in quality of Mexican rural education. A definitive discussion can be found in: Bryan Jensen, and Adam Sawyer. **Regarding Educacion: Mexican-American Schooling, Immigration, and Binational Improvement.** New York: Teachers College Press, 2013.

[44] 2019 ACS Table S0201: Selected Population Profile--Hispanic Population by Type of Origin. It is important to recognize that educational attainment is an imperfect measure of functional literacy. There have now been several decades of research on metrics of functional literacy required for specific reading tasks (including numeracy ones—"quantitative" literacy) by researchers from the Educational Testing Service (ETS) and Programme for the International Assessment of Adult Competencies (PIAAC) https://link.springer.com/chapter/10.1007/978-3-319-58689-2_9 It would be wise for the Census Bureau to conduct more research on difficulty of questionnaire response stemming not only from lexicon but, also, syntax, and formatting among lower-literacy sub-populations—especially for questionnaire navigation instructions and especially for the ACS.

[45] Aguirre International provided the Census Bureau with estimates of the number of migrant/seasonal farmworkers by county, state, and Census Bureau region for use in farmworker outreach in Census 2010.

[46] J. David Brown et.al. "Real-Time 2020 Administrative Record Census Simulation: A New Design for the 21st Century", May 5, 2023, U.S. Census Bureau

[47] Kissam, E., Quezada, C. and Intili, J. "Community-Based Canvassing to Improve the Census Bureau's Master Address File: California's Experience in LUCA 2018," *Statistical Journal of the International Association of Official Statistics,* Vol. 35, December 2018.

[48] Michael Bentley, "Census Operational Metrics: Release 2", May 28, 2021.

[49] Gabbard, S. Kissam, E., Glasnapp et al, "Identifying Indigenous Mexican and Central American Immigrants in Survey Research". http://www.asasrms.org/Proceedings/H2R2012/Identifying_Indigenous_Mexicans_and_Central_Americans_in_Surveys.pdf   See also, Ed Kissam, "The Context and Dynamics of Civic and Political Participation Among Latino Immigrants in Fresno County" https://www.wilsoncenter.org/sites/default/files/media/documents/publication/Supp%20Report%20-%20The%20Context%20and%20Dynamics%20of%20Civic%20and%20Pol%20Part.pdf

[50] Households facing language barriers to census and ACS response are predominantly immigrant-headed ones but include some US-born ones. Facilitating language access is not technically challenging in an era where Google Translate and other online tools are widely available and where rapid progress is being made in AI. Equitable language access should be understood as including attention to literacy level, and for the ACS, numeracy, since document format/screen layout and syntax, as well as lexicon, affect survey response. For an excellent discussion of this issue, even in oral communication between enumerators and census respondents, see Christina Isabelli, Yuling Pan, and Stephen Lubkemann, "Observing Census Enumeration of Non-English Speaking Households in the 2010 Census: Spanish Report", Center for Survey Research and Measurement Directorate, Census Bureau, 2012 https://www.census.gov/content/dam/Census/library/working-papers/2012/adrm/rsm2012-06.pdf

[51] U.S. Department of Education, "Final Appropriations FY21" https://www.nea.org/sites/default/files/2021-01/FY21%20Final%20Appropriations%20for%20Education%20Related%20Discretionary%20Programs%20with%20State%20Tables.pdf

[52] See Andrew Reamer, "Federal Funds from 5 Key Census-Guided Programs", Project on Government Oversight, George Washington University, 2019. https://www.pogo.org/census-project

[53] See Census Bureau documentation of SAIPE modeling at https://www.census.gov/programs-surveys/saipe/guidance/model-input-data.html Other data sources in the SAIPE modeling include SNAP administrative records data and IRS data. It is possible that reliance on SNAP data further skews estimates of the prevalence of poverty in areas with concentrations of immigrants (among them mixed-status families) since many may not apply even though US-born children are citizens and, therefore, eligible. The reliability of IRS records is also questionable in such areas because income tax reporting compliance is uneven in communities with higher-than-average levels of employment in the informal (underground) economy.

[54] Analysis by Ed Kissam based on California Department of Education data on funding allocation for FY21-22.

[55] Wiliam P. O'Hare, Implication of a Blended Base for Post-2020 Census Estimates for Young Children, The Census Project, February, 2021. https://thecensusproject.org/2021/02/16/implication-of-a-blended-base-for-post-2020-census-estimates-for-young-children/

[56] Although we do not address the issue of differential undercount of young children here, there is a robust literature on undercount of Hispanic (and other) children that shows the details of the dynamics of undercount. Research by William O'Hare has provided extensive insights into the extent of the problem. See O'Hare WP, et. al. "The Invisible Ones: How Latino Children are Left Out of our Nation's Census. Child Trends", Hispanic Institute, 2016. ftp://ftp.census.gov/cac/nac/meetings/2016   Additional intriguing details can be found in Turner K, Guzman L, Wildsmith E, Scott M. "The Complex and Varied Households of Low-Income Hispanic Children" National Center for Research on Hispanic Families and Children, January, 2015. https://www.childtrends.org/publications/the-complex-and-varied-households-of-low-income-hispanic-children

[57] The Census Project, "America's Essential Economic and Social Data At Risk: A Vision to Preserve and Enhance the American Community Survey", April, 2023 https://censusproject.files.wordpress.com/2023/05/census-acs-report-2023-v4.pdf

[58] An excellent summary can be found in Andrew Reamer, "The American Community Survey's Role in Federal Regulation and Spending", The Census Project, March 11, 2022. https://censusproject.files.wordpress.com/2022/03/reamer-acs-cp-03-11-22-rev2.pdf

[59] Ceci Villa Ross, "Uses of Census Decennial Programs Data in Federal Funds Distribution", Census Bureau, June, 2023. The report provides detailed information on allocation of funding for 353 federal programs and a useful explanation of ways in which census-derived data impact funding allocation.

[60] The high correlation between HPI and ACS response and Census Response (which is derived in large measure from ACS data) is, of course, expected but it underscores the concern that ACS sample bias may negatively impact the reliability of the HPI data for the hardest to count tracts and neighborhoods. Correlation was assessed by Ed Kissam using HPI data from the Public Health Alliance of Southern California https://www.healthyplacesindex.org for Fresno County as of April 20, 2023.

[61] It is not possible to adequately quantify the impact the skewed HPI had on vaccination in farmworker communities in California without a tract by tract estimate of population undercount. However, review of computed COVID-19 vaccination rates at the zipcode level shows a number of zipcodes where vaccination rates exceeded 100%. This use case for ACS/census-derived Population Estimates data is an important one because differentials in the range of 5%-10% significantly impacted funding for vaccination outreach. A working paper developed for outreach planning provides detailed on variation in computed vaccination levels in 121 California farmworker communities. For details see Heather Riden, Edward Kissam and Rick Mines "The Need to Micro-Target Vaccination Outreach in California Farmworker Areas". University of California Davis, Western Center for Agricultural Health and Safety, August 3, 2021. (available from edkissam@me.com)

[62] Fresno County is divided into 7 PUMAs. The Census 2020 response rate in the Fresno County PUMA with the lowest HTC score (Clovis City) is 20 percentage points higher than in the PUMA with the highest HTC Score (Fresno County West Kerman-Coalinga) that includes the Westside sub-county region we have focused on.

[63] The issue of place-based social investment has now been discussed intensely for several decades in the world of philanthropy. Similarly, in public health, there has been a tsunami of attention to social determinants of health (SDOHs). For both philanthropy and public health, accurate, fine-grained tabulations of data relevant to social disadvantage and/or health risk are crucial tools for effective response to chronic problems. For a recent discussion see Frank Ridzi, "Place-Based Philanthropy and Measuring Community Well Being in the Age of COVID-19" https://www.ncbi.nlm.nih.gov/pmc/articles/PMC8097670/

[64] An important 1988 paper by Phillipe Bourgois that was part of the Census Bureau's ethnographic research on census response documented underlying distrust of "the government" among hard-to-count populations that included immigrants in New York City in the 1980s. https://www.census.gov/content/dam/Census/library/working-papers/1990/adrm/ex90-06.pdf Testimony from a leading Latino researcher, Matt Barreto, has also suggested that the impact of the debate about the citizenship question might be long lasting. It would be valuable for the Census Bureau to conduct research designed not simply to determine level of respondent trust in the Census Bureau's privacy protection but the broader ways in which "trust in the government" affects respondent motivation. Ongoing debate about immigrants and immigration issues will almost certainly affect immigrant households' response to the ACS and to Census 2030.