Identifying Indigenous Mexican and Central American Immigrants in Survey Research

Susan Gabbard;[1,] Edward Kissam;[1] James Glasnapp;[2] Jorge Nakamoto;[1] Russell Saltz;[3] Daniel Carroll;[3] Annie Georges[1]

[1]Aguirre Division, JBS International, Burlingame California
[2]Palo Alto Research Center, Palo Alto California
[3]Employment and Training Administration, U.S. Department of Labor, Washington DC

Acknowledgement: The findings and conclusions in this report are those of the authors, and do not necessarily represent the views of the U.S. Department of Labor. We would like to thank the farm workers and interviewers for their participation. For further information, contact: Susan Gabbard, PhD; JBS International, Inc., Aguirre Division; 555 Airport Boulevard Suite 400; Burlingame, CA 94010; E-mail: sgabbard@jbsinternational.com; Phone: 650.373.4949; Fax: 650. 745-1179.

Key words: Indigenous, farm workers

Identifying Indigenous Mexican and Central American Immigrants in Survey Research

ABSTRACT

This paper examines alternative approaches to identifying indigenous populations that use multiple questions from the National Agricultural Workers Survey (NAWS). The results show that survey respondents do not consistently identify as indigenous across questions on race, language and place of birth. Using the combined responses on all measures results in a higher percentage of respondents being identified as indigenous compared to using individual measures. There is only a modest overlap among question domains with most respondents being identified solely on their response to a single question (either race or language). For Mexican respondents, we examined whether their municipality of birth as reported in the NAWS is identified as indigenous based on the 2010 Mexican Census of Population and Housing data.

Identifying Indigenous Mexican and Central American Immigrants in Survey Research

It has become increasingly important to accurately measure the growing indigenous populations in the United States. At the same time, the identification of individuals who are indigenous is complex. For example, the International Labor Organization and United Nation's legal framework states that identifying indigenous individuals is fluid, varying with "lineage" (being descendants of a distinct cultural/social group prior to European colonization); it also varies with social context such as language, culture and social organization characteristics, all of which also varies over time (Deruyttere, 1997).

This paper examines how Mexican and Central Americans responded to U.S. survey questions on race, language and place of birth. The research sought to explore whether variants of standard survey questions would better identify indigenous Mexican and Central American population. In addition, the analysis looked at both consistency in responses across questions as well as whether single question or multiple question measures would best identify this population.

Identifying and enumerating indigenous individuals through surveys implies defining and measuring multiple and varying components of social identity such as lineage, cultural orientation, self-identification, language, and community residence or affiliation. Each of these components is based upon sound sociological reasons. For example, community of origin plays a significant role in an individual's social and economic relationships both within the originating community and among migrants within the United States. Self-identification with indigenous culture, whether specific to the originating community or at a more pan-ethnic level, also affects various aspects of social and economic relations. Finally, language skills and use, even if context-sensitive and imperfectly measured, influence not only social relationships, but also workforce participation and service access.

Multiple approaches have been used in Mexico's decennial census to enumerate indigenous groups. In the 1980 and 1990 censuses language was used as the deciding factor to identify indigenous individuals; in their 2000 census however, both language and self-identification were used to identify indigenous status. Researchers at Mexico's El Consejo Nacional de Población (CONAPO) reported tabulations based on self-identification, language and the sociocultural identity which was measured using information about respondents' village of residence (Fernández, García & Ávila, 2002). The study showed that using self-identification or language gave rise to substantial differences in the estimated size of the indigenous population. For example, the size of Mexican indigenous populations varies by as much as 17 percent when using the single criterion of language compared to the dual criteria of language and self-identification (Fernández et al., 2002).

The challenges in estimating indigenous population size across definitional criteria are compounded in the United States by the ways in which survey questions that elicit self-identification are worded. For example, standard survey questions on race, ethnicity or language such as those from the U.S. Census could fail to elicit accurate responses necessary to enumerate indigenous groups for three primary reasons. First, indigenous groups in the United States include a diverse array of individuals who are from multiple countries with distinct histories and languages. Meta-categories used in

standard ethnicity or race question might not be intuitively understood by indigenous individuals who often recognize themselves as members of a specific community and possibly as members of a language group such as Mixteco. Such individuals might not identify with meta-categories of ethnicity such as Latino or Hispanic or with meta-categories of race such as American Indian/Alaska Native but would instead select "Other" for their race (Kissam, Herrera & Nakamoto,1993; Martin & Gerber, 2005; Huizar Murillo and Cerda 2004).

Second, to avoid discrimination, indigenous peoples in Mexico and Central America may be reluctant to self-identify as indigenous. In the U. S., demeaning terms referring to indigenous identity are commonly used pejoratively by non-indigenous Mexicans, and U.S. Spanish-language radio and television programs routinely incorporate racial jokes and pejorative stereotypes of indigenous individuals (Morales, 2008). Studies analyzing labor market disparities for Latinos showed in the U.S. indigenous populations face a hierarchical labor market structure that relegates them to less desirable jobs (Hall & Patrinos, 2006; Patrinos, 2000). Indigenous individuals may tend to identify with the dominant groups to avoid discrimination (Morales, 2008). Language itself can be the mechanism of acculturation, deculturation, and alienation. In Mexico, although there are up to 90 distinct languages spoken by members of indigenous communities (INEGI, 2004) there are increasingly more indigenous individuals whose dominant language is Spanish. Individuals from Mexican and Central American indigenous groups are hesitant to report they speak an indigenous language for fear it will similarly expose them to discrimination while in the U.S. (Morales, 2008).

Third, children of immigrants may identify as ethnically indigenous based on their parents' village/migration network without speaking the indigenous language, or being born in an indigenous community. Thus, a survey question that directly asks individuals about their language or community affiliation may still undercount those who are indigenous if it does not capture parental community of origin or indigenous language exposure during childhood.

Data on U.S. hired farm labor force provides a unique opportunity to explore how indigenous individuals identify themselves through a variety of questions. Farm work is often a gateway job for new immigrants, particularly those from rural areas. Immigrants from Mexico and Central America comprise 80 percent of the U.S. farm labor force (Carroll, Samardick, Bernard, Gabbard & Hernandez, 2005). Studies suggest that an increasing number of recent Mexican and Central American immigrants to United States are of indigenous origin (Huizar Murillo & Cerda, 2004; Runsten & Reimer, 2005). Moreover, research identified rural U.S. states as new settlement areas among Mexican and Central American immigrants (Bump, Lowell, & Pettersen, 2005; Massey & Espana, 1987; Mines, 1981). This research shows that much of the rural-to-rural migration in the southeastern U.S. is from Guatemala to Florida, and the overwhelming majority of these Guatemalan immigrants are indigenous Mayan (Bump, et al., 2005); however Mexican indigenous migrants, including Zapotecs and Mixtecs, are now making up a larger portion of the migrant stream. In the western U.S. the predominant flows of indigenous migrants are of Mixtec, Zapotec, or Triqui origin (Mines, Nichols, & Runsten 2010).

The National Agricultural Workers Survey (NAWS), the leading data source on U.S. farm workers includes several questions that can identify indigenous individuals. Beginning in 2005, the NAWS incorporated variants of language and place of birth

questions that were specifically designed to better identify indigenous respondents. These changes include supplementing the question on primary language use with questions that ask about all adult languages spoken as well as childhood language exposure. In addition, beginning in 2009, place of birth information was collected for each parent as well as the respondent. Using the NAWS data collected from these questions, this paper addresses two questions. First, how well do survey questions on race, language, and place of birth identify indigenous Mexicans and Central Americans? Second, how well does a multi-variable indicator using combinations of race, language and place of birth identify indigenous Mexicans and Central Americans?

METHODS

The primary data source is the National Agricultural Workers Survey (NAWS), a U.S. Department of Labor-sponsored survey that collects extensive data from a nationally representative sample of U.S. farm workers. Trained interviewers administer the survey in person in Spanish or English depending on the dominant language of the respondent. In cases where the respondent's dominant language is an indigenous language, interviewers administer the questionnaire with the assistance of a translator. Additional information about the NAWS including its design, sampling and weighting can be found at http://www.doleta.gov/agworker/naws.cfm. The Mexican 2010 Census of Population and Housing was collected by the Instituto Nacional de Estadística y Geografía e Informática (INEGI). INEGI's published tables included Municipio (municipality) level data on the number of individuals age 5 and older who speak an indigenous language. This information was matched to the NAWS place of birth information that was obtained from Mexican respondents in the NAWS.

The NAWS data used in this analysis span from 2005 through 2010. During this period, 11,130 farm workers were interviewed. Of the 8,520 farm workers born in Mexico or Central America, 96 percent were born in Mexico. About half of the remaining 4 percent was born in Guatemala and the remainder in other Central American countries. Data on the locations of birth from the 2009- 2010 NAWS data was merged with the 2010 INEGI data on the proportion of indigenous languages spoken in Mexico's Municipios. Almost all (98%) of NAWS respondents from Mexico had a place of birth that matched with INEGI's Municipio data. The percentage of parents' Municipio that matched with the INEGI data was also high, with 92 percent of father's and 93 percent of mother's Municipio of birth reported in the NAWS data matching with the INEGI's Municipio data. The main reasons data did not match was missing NAWS data, misspelled place of birth and respondents reporting a village or region instead of a Municipio.

We used questions on self-reported race, primary language, and childhood language to create three measures of whether a respondent was indigenous. The first measure was based on the self-reported race question which is a standard version of the required U.S. Office of Management and Budget (OMB) race question. This question included the same list of racial groups and an open-ended "Other" category. Respondents who did not select one of the race categories, e.g. "American Indian/Alaska Native" had their response recorded verbatim under the "Other" category. These written-in "Other" responses were reviewed to identify indigenous respondents. The second measure was based on the primary language respondents spoke as adults. The third measure was based on the question which asked respondents about the language adults spoke to them when they were children. See Table 1 for wording of the language questions.

The INEGI data provided the proportion of individuals age 5 and older who spoke an indigenous language in each of the 2,456 Municipios across all 32 Mexican states. We relied on CONAPO's definition and categorized Municipios as indigenous if 70 percent or more of individuals age 5 and older spoke an indigenous language. If the NAWS respondent's or the respondent's parents' place of birth met this criteria, we classified that respondent as indigenous.

Given that the number of potential indigenous respondents was small, we combined all five years of data (2005-2010) for the analysis. We estimated the proportion of indigenous Mexican and Central American farm workers using responses on race, primary adult language, childhood language exposure, and place of birth separately. We then combined responses from all measures to assess how well a multi-variable indicator identified indigenous farm workers. All estimates are weighted using sampling weights and standard errors are adjusted to account for the complex sampling design of the NAWS data. The 95 percent confidence intervals are shown for each estimate.

Analysis that used responses on race and primary language included all Mexican and Central American respondents. Analysis that used place of birth was restricted to Mexico-born respondents to the Mexican census data.

RESULTS

Table 2 shows the estimated proportion of indigenous Mexican and Central American farm workers based on race, language and place of birth. The estimated proportion of indigenous farm workers ranged from 3 to 12 percent for individual question measure on race and language. Among single question measures, the race question identified the highest proportion of respondents (12 percent), followed by childhood language (9 percent) and adult language (3 percent). While the estimate based on primary adult language was lowest, there were a few respondents who identified an adult indigenous language but not a childhood indigenous language. (More respondents identified a childhood but not an adulthood indigenous language.)

We estimated proportion of indigenous farm workers using combined responses to both race and language questions. Depending on the question used, estimates of the indigenous farm worker population ranged from a low of 13 percent when using race and primary adult language to a high of 17 percent when using race and childhood language. These estimates suggest that neither race nor primary language alone provided an accurate count of indigenous farm workers. As Table 2 indicates each question identified some individuals not picked up with the other question. The overlap among questions was relatively small with about one-fifth (22 percent) of respondent identified as indigenous by both the race and language questions. Almost half (48 percent) of the indigenous farm workers were identified solely by their responses to the race question and 30 percent solely by their responses to the primary language or childhood language questions.

The examination of place of birth was limited to Mexico-born respondents from the 2009-2010 data since the NAWS began asking the extended place of birth question in 2009. Respondents were counted as indigenous if INEGI reported that more than 70 percent of the residents of their Municipio or their parents' Municipio of birth spoke an

indigenous language. Only a small percentage of respondents (N=89) were identified as indigenous by their municipios. As shown in Table 3, in 2009-2010, an estimated 6 percent of Mexico-born farm workers were either born in an indigenous Municipio or had a parent who was born in an indigenous Municipio. Among workers identified as indigenous on any question, 12 percent were identified solely based on place of birth.

CONCLUSION AND IMPLICATIONS

Indigenous identity is a complex, fluid, and nuanced phenomenon which is not easily captured by a single survey question, in part because identifying as indigenous is sensitive to social contexts. A multiple variable indicator appears to have the greatest advantage in identifying survey respondents as indigenous. The small overlap between the race and language questions indicates that single question measures miss a significant portion of the population. Supplementing the standard race question with questions on language experience captures greater numbers than either question domain individually. Lack of familiarity with the OMB standard racial classifications may partially explain why individuals with childhood and adult indigenous language use are more likely to select the "Other" than the "American Indian/Alaska Native category".

If the trend towards declining use of indigenous languages continues, the language question may become less useful. This decline is illustrated by the larger percentage of the overall indigenous respondents reporting indigenous languages in childhood versus adulthood. However, language characterization may change as there are growing shifts toward a perspective that native indigenous languages should be encouraged rather than discarded (Hornberger, 1997). In a similar fashion, the low number of individuals identified as indigenous based on place of birth may also stem from the decline in indigenous language speakers, as this was the underlying basis used here for identifying indigenous Mexican communities.

Parallel strands of questions regarding cultural/ethnic identity (i.e. racial self-identification and language use) have important practical consequences for program planning and policy development as well as for adequate evaluation of existing programs and policies. For example, improved data on language experience and current language use can provide valuable guidance for U.S. public health and education program design, and for administering social programs struggling to respond to an increasingly diverse Latino population.

As the United States becomes increasingly multi-ethnic and the proportion of our national population consisting of Latin American heritage increases, it is likely that overall concepts of ethnic identity will become increasingly nuanced. The transnational migration among indigenous peoples makes them an important emerging group within the Latino population because of their unique challenges in labor relations, service delivery, and social and civic integration. Surveys can more closely monitor the shift in the multi-ethnic composition of the Latino and Hispanic populations by including questions regarding childhood and adult languages, and household members' language preferences. The responses to these questions will provide valuable insights not only for improving service delivery and quality, but also for proactive efforts to integrate this group into an increasingly pluralistic society.

REFERENCES

Bump, M. N., Lowell, B. L., & Pettersen, S. (2005). The growth and population characteristics of immigrants and minorities in America's New Settlement States. In E. Gozdziak & S. F. Martin (Eds.), *Beyond the gateway: Immigrants in a changing America* (pp.19-56). Lanham, MD: Lexington Books.

Carroll, D., Samardick, R. M., Bernard, S., Gabbard, S., & Hernandez, T. (2005). *Findings from the National Agricultural Workers Survey (NAWS) 2001-2002: A demographic and employment profile of United States farm workers*. Washington, DC: US Department of Labor.

Deruyttere, A. (1997). Indigenous peoples and sustainable development: The role of the Inter-American Development Bank." Accessed at http://idbdocs.iadb.org/wsdocs/getdocument.aspx?docnum=363093

Fernández, P., García, J. E., & Ávila, D. E. (2002). Estimaciones de la población indígena en México. *Consejo Nacional de Población (CONAPO): Situación Demográfica de México*: 169-182.

Hall, G., & Patrinos, H. A. (Eds.). (2006). *Indigenous peoples, poverty and human development in Latin America: 1994–2004.* New York: Palgrave Macmillan.

Hornberger, N. H. (1997) *Indigenous literacies in the Americas: Language planning from the bottom up.* Hawthorne, NY: Mouton De Gruyter.

Huizar Murillo, J., & Cerda, I. (2004). Indigenous Mexican migrants in the 2000 US Census: Hispanic American Indians. In J. Fox & G. Rivera-Salgado (Eds.), *Indigenous Mexican Migrants in the United States*. Center for Comparative Immigration Studies University of California, San Diego. Accessed at: http://ccis.ucsd.edu/publications/books/indigenous-mexican-migrants/

Kissam, E., Herrera, E., & Nakamoto, J. M. (1993). *Hispanic response to Census enumeration forms and procedures: Report prepared by Aguirre International for the Census Bureau*. Burlingame, CA: JBS International, Aguirre Division. Accessed at http://www.jbsinternational.com/pdf/port_mc_3-census.pdf.

Martin, E., & Gerber, E. (2005). Results of recent methodological research on the Hispanic origin and race questions. *Survey Methodology Report #2005-04.* Statistical Research Division, Washington, DC: U.S. Bureau of the Census.

Massey, D. S., & Espana, F. G. (1987). The social process of international migration. *Science*, 237, 733-738.

Mines, R. (1981). *Developing a community tradition of migration to the United States: A field study in rural Zacatecas, Mexico, and California settlement areas*. Center for US Mexican Studies University of California San Diego. Accessed at http://www.escholarship.org/uc/item/72n33714.

Mines, R., Nichols, S., & Runsten, D. (2010). *California's indigenous farmworkers: Final report of the Indigenous Farmworker Study (IFS) to the California Endowment*. Accessed at http://www.ncfh.org/pdfs/2k9/8913.pdf.

Morales, H. (2008). *Presentation to California First 5 Indigenous Families Learning Exchange*. Fresno, CA: Indigenous Learning Exchange.

Patrinos, H. A. (2000). The cost of discrimination in Latin America. *Studies in Comparative International Development*, 35, 3-17.

Runsten, D., & L. Reimer. (2005). Origins and characteristics of Mexican immigrants in San Diego: The evidence from the Matriculas Consulares. In R. Kiy & C. Woodruff, *The ties that bind us: Mexican migrants in San Diego County*. La Jolla, CA: Center for U.S.-Mexican Studies, University of California, San Diego.

Table 1
Wording of Survey Questions on Race and Language Experience

Wording of survey question	Years the question was asked
Question on primary language as adult	
In which language do you believe you are the most dominant (comfortable) conversing?	2005-2010
1. English	
2. Spanish	
3. Creole	
4. Mixtec	
5. Kanjobal	
6. Zapotec	
7. Other (Specify)	
Question on childhood language exposure	
When you were a child, in what languages did adults speak to you at home? (Check all that apply)	2005-2010
1. English	
2. Spanish	
3. Creole	
4. Mixtec	
5. Kanjobal	
6. Zapotec	
7. Other (Specify)	

Source. The National Agricultural Workers Survey.

Table 2. Percent Mexican and Central American Farm Workers Identified as Indigenous

	Number of Mexican and Central American Respondents[1]	Number of Respondents Identified as Indigenous[1]	Percent	SE	95% Lower Confidence Limit	95% Upper Confidence Limit
Indigenous identification based on:						
Race	8,501	865	12%	1%	9%	14%
Primary Adult Language	8,512	230	3%	0%	2%	4%
Childhood language exposure	8,520	627	9%	1%	7%	10%
Race and Adult Language	8,502	981	13%	1%	10%	15%
Race and Childhood Language	8,504	1,252	17%	1%	14%	19%
Interaction of Race and Language among Indigenous Mexican and Central American Farm Workers						
Race only		625	47%	3%	41%	54%
Any language		389	30%	3%	24%	37%
Race and language		240	22%	3%	17%	28%
Identification based on Municipio of birth[2]						
2009-2010	2,649	89	6%	2%	2%	9%

Source. The National Agricultural Workers Survey, 2005-2010; and INEGI, 2010.
Note. The percent and CI columns are weighted and standard errors are adjusted to account for sampling design and for clustering.
1. The numbers shown in this column are unweighted.
2. These estimates are based on Mexico-born farm workers only.

Table 3
Identifying Indigenous Farm Workers based on Combined Responses on Race, Language, and Place of Birth

	Number of Mexican Respondents[1]	Percent	SE	95% Lower Confidence Limit	95% Upper Confidence Limit
Of the three questions:					
Farm worker is indigenous based on race or language but not Municipio	235	9%	1%	7%	11%
Farm worker is indigenous based on race, language and Municipio	50	4%	2%	1%	7%
Farm worker is indigenous based on Municipio only	39	2%	0%	1%	3%
Respondent is not indigenous	2,331	85%	2%	81%	89%
Total	2,655				

Source. The National Agricultural Workers Survey, 2009-2010; and INEGI 2010.
Note. The percent and CI columns are weighted and standard errors are adjusted to account for sampling design and for clustering.
1. The numbers shown in this column are unweighted.