Journal of Sociological Research
ISSN 1948-5468
2013, Vol. 4, No.2

Detecting "Real" Population Changes with American Community Survey Data: The Implicit Assumption of Treating Between-Year differences as "Trends"

Carlos Siordia[1*]

[1] Department of Epidemiology, Graduate School of Public Health, University of Pittsburgh

[*]Address correspondence to Carlos Siordia, PhD, 130 North Bellefield Ave, Pittsburgh, PA 15213. Phone: 1-142-383-1708. Email: cas271@pitt.edu

Doi:10.5296/jsr.v4i2.4883 URL: http://dx.doi.org/10.5296/jsr.v4i2.4883

Funding

This work was supported by the National Institute of Aging at the National Institutes of Health (grant number T32 AG000181to A. B. Newman).

Abstract

BACKGROUND: The American Community Survey (ACS) in the United States (US) collects detailed demographic information on the US population. Pressures to use year-to-year population estimates to analyze "trends" (i.e., between-year differences on the characteristics of interest) have motivated the need to explore how single- or multi-year estimates can be used to investigate changes in US population over time. **OBJECTIVE:** The specific aim of this manuscript is to provide empirical evidence that between-year differences in population characteristics have difference levels of uncertainty around point-estimates. **METHODS:** Six ACS Public Use Microdata Sample (PUMS) single year files from 2005 through 2010 are used to empirically show the heterogeneity of uncertainty in "between-year differences" on level of education, for a birth cohort born between 1960 and 1970 of non-Latino-whites and Mexican Latinos/as. **RESULTS:** The data show the precision of the education estimate decreases as the specificity of the population increases. For example, Mexican's 99% confidence intervals have wider and more time-varying bandwidths than non-Latino-whites. **CONCLUSIONS:** Inferring meaningful population change requires the challengeable assumption that between-year differences are not the product of data artifacts. Harvesting reputable ACS data demands further research before between-year differences can be treated as "real change."

KEY WORDS: ACS; PUMS; trend; Mexican; education; standard errors; estimates; sampling;

1. INTRODUCTION

Detecting how the characteristics of a population change over time is of interest to both the academic and private sectors. The American Community Survey (ACS) is a population based and yearly survey administered by the US Census Bureau and may one day be the primary data source for investigating population trends. The ACS in the United States (US) collects detailed demographic information on about three million people each year to infer the characteristics of the more than the 300 million people in the US population. The ACS rolls from month-to-month to produce "timely estimates"—making it the largest and most current source of demographic information. Data from the ACS help fulfill legal mandates and significantly influence the distribution of hundreds of billions of dollars from the United States (US) federal government to local stakeholders (Reamer, 2010). The ACS replaces the "long form" previously administered during decennial counts. The main goal of the ACS is to provide estimates on the characteristics of the various "communities" in the US. ACS data is widely used. Because of its importance, a recent publication investigated how well ACS data counted naturalized citizens (Van Hook & Bachmeier, 2013). A growing number of highly trained statisticians, demographers, survey methodologist, economist, psychologist, geographers, mathematicians, and many others make up the ranks of the admirable army tasked with fulfilling the enterprise of providing high quality data to presumably help maintain the social equilibrium on which US democracy is build.

Using samples to infer the characteristics of the population is filled with challenges. For example, the ACS uses survey responses from about 3 million people to infer the characteristics of more than the 300 million people in the US population. Note that because only one person (referred to as the "reference person") participates with completing the survey, it may be technically said that about 80% or more of the people represented in the ACS data are described through a 'proxy' (e.g., mother of the family)—a point worth noting since proxy reports may introduce hard to quantify measurement inaccuracies (Siordia, 2012). Developing the simple random sample so frequently discussed in the assumptions of inferential statistics is not possible for many logistical reasons—forcing ACS administrators to use complex stratified sampling frameworks that attempt to account for unit-non-response amongst other intricacies. At stake in the discussion being offered in this paper is the view that ACS data can represent the "actual figures that would have been obtained by interviewing the entire population" (US Census Bureau 2010). It could be said that giving every individual in the US population and "equal chance" of selection is challenging when trying to capture responses to more than 3 million of them during a 12-month period. The ACS is a large scale enterprise seriously attempting to conquer the vast challenges present when only using 1% of the population to describe 100% of them—a stunning but constitutionally mandated task.

The desire to study population trends is of high interest: federal government can project where resources will need to be distributed in the future (Bchir et al, 2013; Bishaw & Semega, 2008); local governments can help assess the needs of their constituents (e.g., Heckman & LaFontaine, 2010); academicians can paint a contemporary and moving picture of population characteristics (Cherlin, 2010; Crosnoe & Benner, 2012); political enterprises can geographically locate potential pools of potential voters to gauge their level of civic engagement (e.g., Syvertsen et al, 2011) or appeal to important issues such a economic inequality (Kochhar, Taylor, & Fry, 2011); research on health (Alkema et al, 2012; Fuller-Thomson et al, 2009); and those in the marketing sector can use large scale social data to project where the next big-box store should be located (see Morritt & Weinstein, 2012). Estimating trends with ACS data is challenging (see Martin, Schoeni, & Andreski, 2010) in part because the ACS exists only to meet legal demands set forth by the US constitution and other federal mandates. Any subsequent use of the data is secondary and not a motivating interest in the production of ACS questions or sampling techniques. However, the ACS is staffed with humans and as a consequence is subject to the push and pulls of social powers. Data scientist interested in the purity of information collection protocols and careful interpretations of data must temper their objections in the face of administrative demands. This manuscript hopes to give voice to some of their concerns.

The specific aim of this manuscript is to provide empirical evidence that between-year differences in population characteristics have different levels of uncertainty around point-estimates—where the precision of the estimate decreases as the specificity of the population increases. Before the torrent of technical language arrives, the reader should note that the main argument being made throughout is that inferring meaningful population changes from between-year differences requires the questionable assumption that differences are not the product of data artifacts. This position does not equate an argument for the high fallibility of population estimates with ACS data. In stark contrast, the mere fact that the quality of a survey's product can be challenged in a public discourse speaks well of the integrity of the enterprise and shows how rich resources in ACS microdata can be harvested to quantify an estimate's plausible precision. It is only because the US Census Bureau is so forthright about ACS products that this investigation is possible. Pointing out between-year differences may be the product of the quality of samples from year to year is intended to give pause to those who would unintentionally misuse ACS products—as the ACS discourages inferring trends with their data. As will be further argued below, harvesting data to detect "real changes" in the US population over time may require either the formation of larger multiple-year files or the ascertainment that between-year data products are minimally impacted by random (e.g., sample selection) and non-random (i.e., item non-response and their allocation algorithms) processes. Challenges with the latter issue may increase if survey participation levels continue to decrease.

The goals of the project are accomplished by estimating 'college educated' (at least one year of college education or more) in a cohort that was approximately born between 1960 and 1970 and following them with six single-year ACS Public Use Microdata Sample (PUMS) files from 2005 to 2010. In 2005, they ranged in age from 35 to 45. It is assumed that in most cases most individuals at this age range have completed most of their formal education. Under this assumption, it is expected that the level of educational attainment in the population cohort born between 1960 and 1970 will remain relatively stable—as few in their group may be expected to increase their level of education during mid-life ages. Please note no minimum or maximum thresholds are established for determining which between-year percent change can be considered notable or inconsequential—quantifying meaningful

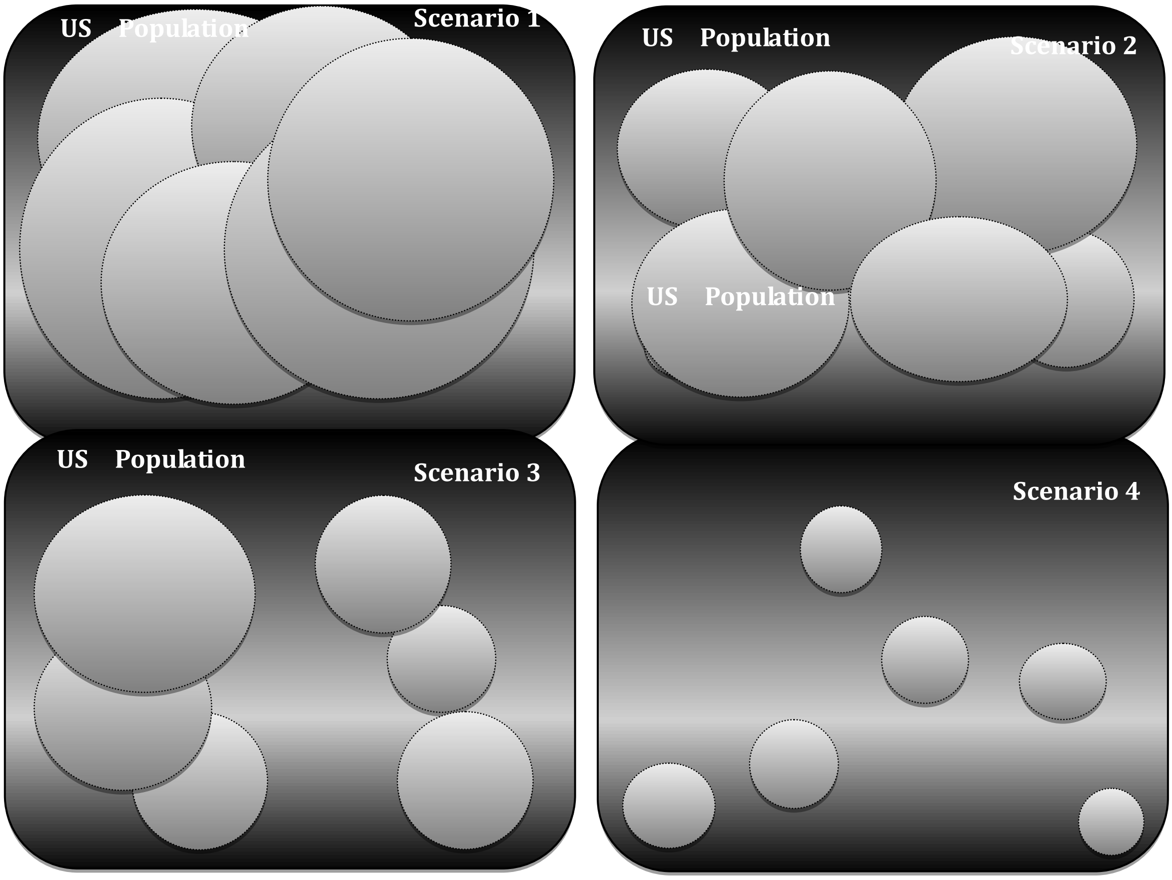
Journal of Sociological Research
ISSN 1948-5468
2013, Vol. 4, No.2

changes are not the primary interest of the current project. The analysis focuses on showing how the level of precision (e.g., 99% confidence intervals) in the estimates of college education for the 1960-1970 cohort vary over time, between a majority and minority group stratified by sex. Please note delineating "birth cohort" is not an exact science (see Siordia & Leyser-Whalen, 2013). Individuals approximately born during the 1960 and 1970 are primarily chosen for pragmatic reasons as no overwhelming factor—other than the fact that they were born in the middle of the Vietnam War era—links them as a cohesive birth group.

Before discussing the data source, analytic approach, and results it is important to present a simple conceptual framework for why assuming between-year changes are the product of real changes in the population needs careful consideration. If the questions is: *Should investigators use ACS data to infer population changes (i.e., trends) overtime?*; then the simple answer is: no, not yet. The analysis will show that noticeable differences on a presumably stable characteristic (i.e., educational attainment in middle-age) in a birth cohort do appear over the course of six years. These "differences" could be used to draw trends but should be accompanied by a clear and upfront declaration that they may be the product of either real change in the population or data artifacts.

Figure 1 offers a simple illustration of how the between-year differences that are believed to display population trends can range from being a reflection of real change in the population to being a product of data production phenomena. The four rectangles represent four hypothetical populations and each bubble represents one of the six survey years (2005-2010) in the analysis. "Scenario 1" presents a case where each sample is widely representative of the target population—shown by the width of each circle and their overlapping. Under Scenario 1, observed between-year differences could be said to likely be the product of real change in the population. Assuming between-year differences represent real change becomes more questionable under Scenario 2 and even less probable under Scenario 3 where samples are less representative. Scenario 4 shows a case where between-year change is likely the product of data artifact—shown by the narrow and non-overlapping bubbles.

Figure 1: Source of change



Determining where single-year ACS data can be located from year to year across the spectrum of these hypothetical scenarios may be difficult. Using single-year data to infer meaningful trends would require that something like Scenario 1 be present across all the years under observation. For example, if response through mail decreases from 2005 to 2010 and the use of in-person computer-assisted interviews increases, then any possible impacts from sample selection alterations or response editing for questions from interviewer effects would have to be considered 'noise' (i.e., irrelevant heterogeneity seen as inconsequential to the quality of estimates). Because existing *public* data may not allow for this type of investigation. Internal review could be undertaken by the US Census Bureau charged with the production of ACS data. For example, algorithms could be built into data processing protocols to determine which geographical areas are producing surveys with high item missingness. Algorithms in data cleaning stages could also seek to determine within-interviewer and between-survey item non-random response patterns (i.e., fabrication of responses by field operatives). This is important becuase the bureau has disclosed that there are survey administration changes over time which includes different rates in mode of data collection (Martin, Schoeni, & Andreski, 2010).

Assuming year-to-year sample quality incomparability as highly plausible is advised. One publication by McElroy (2009) gives support to this view. McElroy (2009) utilized weighted averages to show how potential bias present when comparing trends using multi-year files can be reduced. He concludes that multi-year data are not comparable across different periods and offers advice echoed in this publication: between-year comparison must await additional ACS data. Assuming between-year sample incomparability does not necessitate the view that ACS sampling protocols are inherently flawed with systemic errors. It may be that even a superb and improving sampling methodology is limited when only using 1% of the population to infer 100% of the characteristics in the target population. We now turn to the data for evidence, not of between-year sample quality incomparability, but for actual information showing how presumably stable estimates fluctuate and vary in their level of precision over time.

2. Methods

Estimates on the amount of people with ‘at least one year of college or more’ are produced by using “single-year” Public Use Microdata Sample (PUMS) data files from the ACS. The 2005 file only contains mainland (i.e., residing within the contiguous US) Mexican-origin-Latinos (hereafter simply refer to as Mexicans) and Non-Latino-White (NLW) individuals between the ages of 35 and 45—people who were born approximately between 1960 and 1970. The “1960-1970 birth cohort” is followed for the remainder of the data extractions. Please note that the term ‘cohort’ refers to the fact that subsequent samples attempt to draw only those who are likely to have been born during the 1960 to 1970 period. The expression “likely to have been born” is used because the actual day of birth or day of participation with the survey is not provided in public microdata—their absence creates the lack of precision and need to clarify that a few in their group may have been born a few month before 1960 or after 1970. Both the Mexican and NLW groups include native- and foreign-born individuals.

The NLW racial-ethnic majority group in the US is expected to have the most stable estimates—by virtue of their size and high survey participation rates. In contrast, Mexicans are an ethnic minority group in the US and were chosen to show how the stability of estimates can be less stable in a “special population.” In addition to the sample selected from the 2005 file, the following single-year data files and age ranges are used selecting individuals from the mainland and from the same racial ethnic groups: 2006 file with people between the ages of 36 and 46; 2007 file and people aged 37 to 47; 2008 file with individuals with ages from 38 to 48; 2009 file with subjects aged 39 to 49; and the 2010 file with those between the ages of 40 and 50. There are a total of 6 samples being used in the analysis and are kept separate for year-specific estimations. Given the technical nature of this paper, it should be noted that PUMS ACS files are updated from time to time as errors are discovered. The PUMS files being used here were all downloaded in December of 2013 from the US Census website as “UNIX” zip files and have the following “last updated” dates: 2005 updated on November 9, 2010 ; 2006 updated on December 17, 2009; 2007 updated on September 23, 2008; 2008 file updated on December 9, 2010; 2009 updated on November 2, 2010; and 2010 updated on October 12, 2011.

The single-point estimate on the number of people who have at least one year of college or more is calculated for each of the six samples by using a single population weight (variable name PWGTP). Details on the production of weights are available elsewhere (US Census Bureau, 2011). Briefly, individuals are assigned their weights through a series of stages.

Assigned 'base weights' to a person's to capture initial probability of selection are changed to 'adjustment weights' to capture differences in actual and expected populations which are then filtered through a 'trimming procedure' (uses capping) which help account for unexpectedly large base weights (Keathley, Navarro, & Asiala, 2010). To account for non-interviewed persons a set of 'non-interview adjusted weights' are produced followed by a post-stratification procedure that allows state-level sums of the weights to equal population controls generated by full-count Decennial numbers (Keathley, Navarro, & Asiala, 2010). For simplicity, the average number of people represented by each actual respondent is represented by using an easy-to-understand Person Inflation Ratio (PIR) (see Siordia, 2013c), which is computed as follow: (weighted count ÷ weighted total population). The average number of people represented by each person increase as the PIR number increases—the increase in PIR may signal an increase on the potential for person associated measurement errors to increase.

In addition to showing how education-related estimates fluctuate and even simpler fluctuation will be shown by highlighting how the count (i.e., population size) of individuals in the cohort vary from year to year. The measure is referred to as Fluctuations (F) and is calculated as follows: [(new estimate – old estimate) ÷ old estimate]*100. For example, a positive F number for Mexican males between 2005 and 2006 would indicate their sample had decrease. Fluctuation in population size may occur through mortality, emigration from the US, or differences in the selection of samples. As the F number grows in distance from zero, it may indicate that sample size difference are the product of more than mortality or emigration—under the assumption that the latter two have a small impact in the birth cohort population size. Larger F positive or negative numbers may be seen as possible evidence of between year sample incomparability. Note no effort is made to determine what a 'large' number is an the reader is encourage to choose on their own.

Because the ACS uses a sample to describe the population, inferential statistics must be used. Sampling error is said to arise in data because probability sampling is used to try and ensure the representativeness of the sample. As sampling error increases it diffuses precision in the estimate. Because PUMS files only represent a subject of the full ACS sample, the imprecision of the estimate may be further aggravated. The standard errors being estimated in this analysis could be larger than those provided by using the complete and not-top-coded internal ACS data. To estimate the margin of error (MOE) around the point-estimate, the single person weight and an additional 80 person-specific-weights (variables PWGTP1-PWGTP80) are used in a SAS 9.3® algorithm (Siordia & Young, 2013). The "replicate weights method" estimates standard errors (SE) from the 80 replicate weights (US Census Bureau, 2009a) as follows:

$$SE(x) = \sqrt{\frac{4}{80}\sum_{r=1}^{80}(x_r - x)^2}$$

where x is the education estimate based on PWGTP and x_r refers to the replicate weights ranging from x_1 (i.e., PWGTP1) to x_{80} (i.e., PWGTP80). Automated procedures for using the replicate weights with a "jackknife" approach in Stata and SAS are available[1] as is SAS code

[1] https://usa.ipums.org/usa/repwt.shtml

for approximate SEs using the generalized variance function[2]. This study used the computations displayed publicly by an entity of the University of Michigan using SAS code[3]. The SAS code used in this analysis was created from scratch by the author. An investigation on the comparability between these procedures and the SAS code developed by the author is not available and it is presumed that minimal differences should exist.

It is important to note that replicate weights in PUMS files are not the replicate weights used by internal US Census Bureau data users producing population estimates. Instead, the replicate weights in PUMS file are based on the internal ACS replicate weights *and* additional adjustment for PUMS subsampling—in other words, a series of additional steps seeking to equalize final tabulations across all data sources introduce managed weights that simultaneously account for both person- and PUMS-sampling procedures. No publication has ever shown how 'PUMS replicate weights' differ from those used internally by the bureau. More technical details on the formula for estimating SEs and discussion on replicate weights are available elsewhere (US Census Bureau, 2010). Please note that the replicate weight approach being used here is more accurate than simply using the design factor to estimate SEs (see Starsinic, 2011) and that other approaches for estimating variance are available (e.g., Keathley, Navarro, & Asiala, 2010).

The algorithm developed by the author uses the SEs to compute a "99% MOE" as follows: (SE × 2.575). The MOE is said to capture deviations between the sample estimate and the "true" estimate in the population (Siordia & Le, 2013). The 99% MOE can be used to create the 99% confidence intervals by simply subtracting and adding the MOE to the point-estimate. The statistical procedures can be used to determine within what range of numbers the 'real' population estimate *may* be found. In different words, we can be 99% certain that the true population education estimate may be found somewhere between the confidence intervals. The ACS does not produce a simple random sample, which increases the amount of assumptions being made to claim this level of confidence. For example, confidence on the quality of 99% MOE requires users to assume that replicate weights correctly capture the complex selection process by which individuals are selected into survey participation. The estimation of the 99% MOE only quantifies variability in estimate cause by sampling methods and "unit" non-response (i.e., person missingness). The 99% MOE makes various other assumptions, such as: the address list used to create the sampling frame is complete and unit-non-responders re-enter the sampling frame correctly. The 99% confidence intervals represent the ranges very likely to contain the average value of the estimated characteristic (i.e., one year of college or more) that would result over all possible (i.e., hypothetical) samples with a known probability—this view rooted on frequentist statistical traditions may be challenged if the resurgence of Bayesian modeling techniques becomes more readily available (see Alkema et al, 2012; Hawala & Lahiri, 2010).

The frequentist theory argues that if all possible samples (note 'infinity' is implied) that could be produce with the PUMS sample design were independently selected under *identical* conditions each time, then they would produce an estimate of the population

[2] https://usa.ipums.org/usa/resources/repwt/Use_of_the_Public_Use_Replicate_Weight_File_final_PR.doc

[3] https://ctools.umich.edu/access/content/group/34a72eab-daa4-4d14-80e0-9150727aed6c/Technical%20-%20Statistical/sas_code_example.txt

characteristics within the confidence interval. In this hypothetical world, samples are selected in 100% identical conditions—a seemingly radical proposition in the real world filled with human produced changes. Non specialist on the topic should note this is the standard protocol in survey research. As alluded to earlier, SEs do not account for variability due to non-sampling error—like potential correlated errors produced by survey interviewers, data coders, or other process that require human decision making. For these reasons the US Census Bureau has advised that SEs be understood as only capturing the *lower bound* of the total error. Because of the various issues raised here, the 99% MOE may not actually meet the stated level of confidence and should be interpreted with great caution (see Mather et al, 2005).

The estimation of 'uncertainty' around the estimates does not account for how "allocations" may affect bias in the production of the estimates. Allocations reassign values to missing, ambiguous, or illogical responses in order to create complete ACS datasets. Few, even amongst specialist, realize that precision of population estimates from ACS data do not reflect the potentially non-random error produced by the imputation of missing or illogical responses. If item missingness does not occur completely at random, then the patterns of missingness have the potential to alter the quality of the estimate and MOE (Siordia, 2013a). If in addition to this complexity, the use of non-probability driven allocation algorithms (e.g., logic driven fixes and not geographically based method using hot decks) may further aggravate the stability of quantifying MOE around population estimates. When attempting to create small area estimates (Siordia, 2013b), even more complex issues dealing with geographical uncertainty (Spielman, Folch, & Nagle, 2013) and polygon fragmentation (Siordia & Fox, 2013; Siordia & Wunneburger, 2013) need to be considered (see Sun & Wong, 2010). In short, the confidence interval may serve only as a tentative guide for where the true population characteristic may be found.

The esoteric and complex issues raised here should not be used to argue that ACS data is high fallible. It would be difficult to make the argument that comparable data in terms of size, quality, transparency, and complexity exist. To be sure, the ACS data can be regarded amongst the most complete, transparent, and non-ambiguous information sources on the US population. Readers should note the US Census Bureau tries to reduce the influence of bias on estimates (produced by systematic errors) by undertaking valuable research on questionnaire design, sampling techniques, data collection protocols, and information processing procedures. The ACS aims to reduce bias in estimates by following up on 'mail person non-respondents' with phone and in-person interviews. If upon reading the various complexities challenging this enterprise you are swayed to discredit the value of ACS data, you would be advised to compare it to the vast majority of studies which only make use of a few thousand people (typically less than 0.0001% of the population) to infer characteristics of the population without providing SEs around estimates. The ACS makes use of state of the art procedures and continually seeks to develop better approaches. This paper is not a critique on their procedures but a warning to non-technicians that great care should be used when using population estimates from the ACS and that even more caution should be employed when using other data sources.

Journal of Sociological Research
ISSN 1948-5468
2013, Vol. 4, No.2

3. Results

Table 1 shows the weighted and unweighted (actual) count of sample subjects per year by race-ethnic groups and sex. Please note that from weighted numbers, there is about one Mexican male for every 7 NLW males and about 1 Mexican female for every 8 NLW females. There are notable differences in how population weights affect the count between the actual and weighted count. For example, the ACS 2005 had 16,776 actual Mexican males who when weighted represent 2,217,076 of their counterparts. The difference

Table 1: Basic sample estimates for by race-ethnicity and stratified by sex

	Mexican Males				Mexican Females			
	Weighted[1]	UW[2]	PIR[3]	F[4]	Weighted	UW	PIR	F
2005	2,217,076	16,776	132		1,937,758	16,224	119	
2006	2,309,085	17,799	130	4.2%	1,939,815	16,564	117	0.1%
2007	2,296,395	17,896	128	-0.5%	1,939,458	16,572	117	-0.0%
2008	2,334,308	17,986	130	1.7%	1,989,933	17,171	116	2.6%
2009	2,306,643	18,301	126	-1.2%	1,979,013	17,465	113	-0.5%
2010	2,252,212	18,726	120	-2.4%	2,070,821	17,790	116	4.6%

	Non-Latino-White Males				Non-Latino-White Females			
	Weighted	UW	PIR	F	Weighted	UW	PIR	F
2005	15,821,959	161,551	98		16,030,898	169,591	95	
2006	16,086,308	164,426	98	1.7%	16,135,414	170,089	95	0.7%
2007	16,058,538	164,222	98	-0.2%	16,073,807	169,434	95	-0.4%
2008	16,070,314	163,647	98	0.1%	16,058,525	168,588	95	-0.1%
2009	15,964,822	163,021	98	-0.7%	16,088,121	168,925	95	0.2%
2010	15,802,755	163,588	97	-1.0%	15,880,964	167,891	95	-1.3%

[1] Weighted number of people (using single population weight); [2] Unweighted counts, i.e., actual number of individuals in sample; [3] Average number of people represented by each actual respondent, i.e., "Person Inflation Ratio" = (weighted count ÷ weighted total population); [4] Difference from previous year, i.e., "fluctuation" = [(new estimate – old estimate) ÷ old estimate]*100

between the actual and weighted number is capture by the PIR of 132—signaling that on average every Mexican male represents 132 of his counterparts. Mexican males' have the highest PIR scores although PIR number have been decreasing since 2005—potentially indicating a change in sample selection. NLW females have the lowest PIR values at 95 for all the ACS years under investigation.

The F score for the 2005 to 2006 period reveals a very notable value of 4.2% in Mexican males. It is said to be "notable" because the Mexican male sample is said to have increased by almost 5% in 12 months, which would equate to about 92,000 people—if mortality and emigration seem like improvable factor to account for this notable 12-month population, then sample selection differences should be considered as an explanatory factor. When compared to the other F scores for Mexican males, the 4.2% stands out as does the 4.6% F score for Mexican females during the 2009 to 2010 period. Please note the relative stability of F scores in NLWs when compared to Mexicans.

Table 2 shows the estimated number of individuals with at least one year of college education or more (numbers under the "college" column). By comparing the weighted count presented in Table 1 with the weighted estimated of college educated the percent of individuals with said level of education is presented (numbers under the "%" column). The MOE number can be used to estimate the 99% confidence interval—simple add and subtract the number in the MOE column from the number in the College column. To facilitate the interpretation of the many and large numbers, the size of the confidence interval is captured with the "bandwidth" (BW) measure which simply computes as follows: [(MOE× 2) ÷ College) × 100]. BW is a relative measure in that it represents how the estimate could range as a function of the confidence interval. More simply, BW helps standardize the confidence intervals in order to compare between the various groups.

Table 2: Single-year estimates for Mexican and Non-Latino-Whites stratified by sex

	Mexican Males				**Mexican Females**			
	College[1]	%[2]	MOE[3]	BW[4]	College	%	MOE	BW
2005	454,891	20.5%	24,613	10.8%	465,766	24.0%	22,995	9.9%
2006	474,482	20.6%	27,742	11.7%	482,605	24.9%	22,208	9.2%
2007	491,008	21.4%	23,560	9.6%	501,086	25.8%	20,913	8.3%
2008	548,081	23.5%	37,785	13.8%	556,272	28.0%	28,636	10.3%
2009	537,393	23.3%	28,444	10.6%	538,248	27.2%	20,252	7.5%
2010	537,071	23.9%	22,110	8.2%	587,231	28.4%	20,290	6.9%

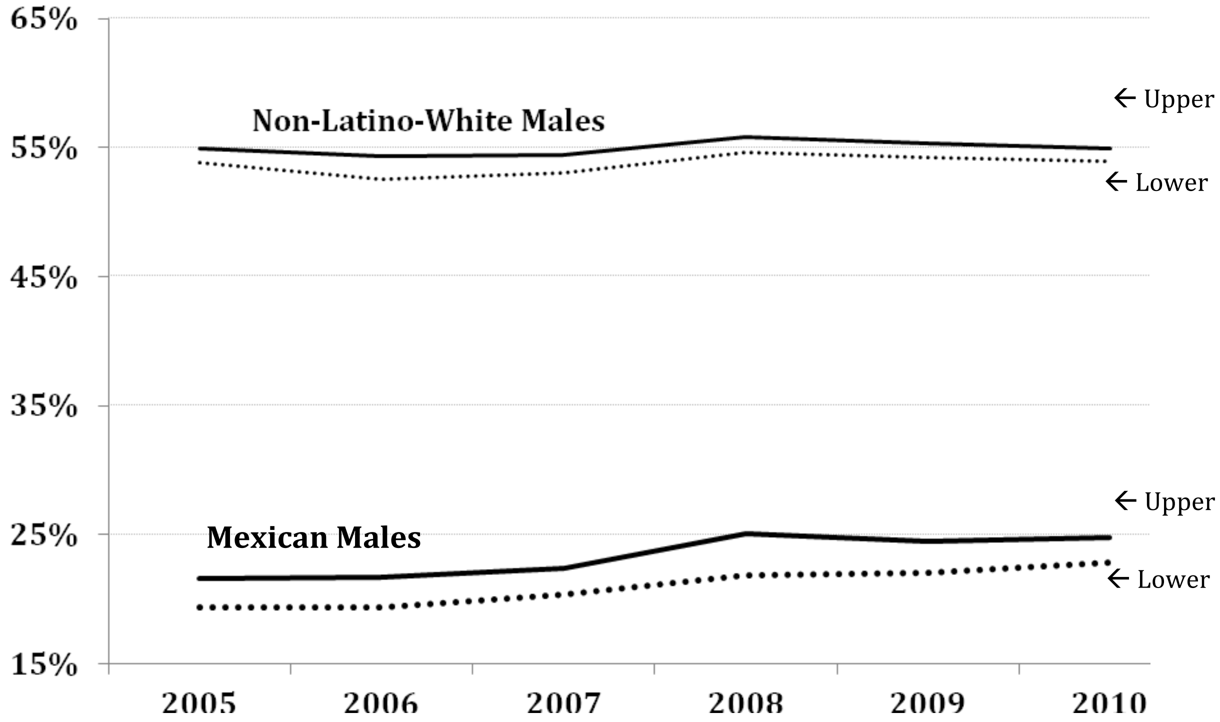
	Non-Latino-White Males				**Non-Latino-White Females**			
	College	%	MOE	BW	College	%	MOE	BW
2005	8,597,557	54.3%	86,063	2.0%	9,435,721	58.9%	69,593	1.5%
2006	8,599,516	53.5%	144,786	3.4%	9,524,071	59.0%	86,936	1.8%
2007	8,631,322	53.8%	109,513	2.5%	9,515,114	59.2%	81,693	1.7%
2008	8,878,106	55.3%	100,489	2.3%	9,739,620	60.7%	71,314	1.5%
2009	8,741,332	54.8%	85,123	1.9%	9,715,915	60.4%	149,013	3.1%
2010	8,601,088	54.4%	84,066	2.0%	9,541,912	60.1%	69,808	1.5%

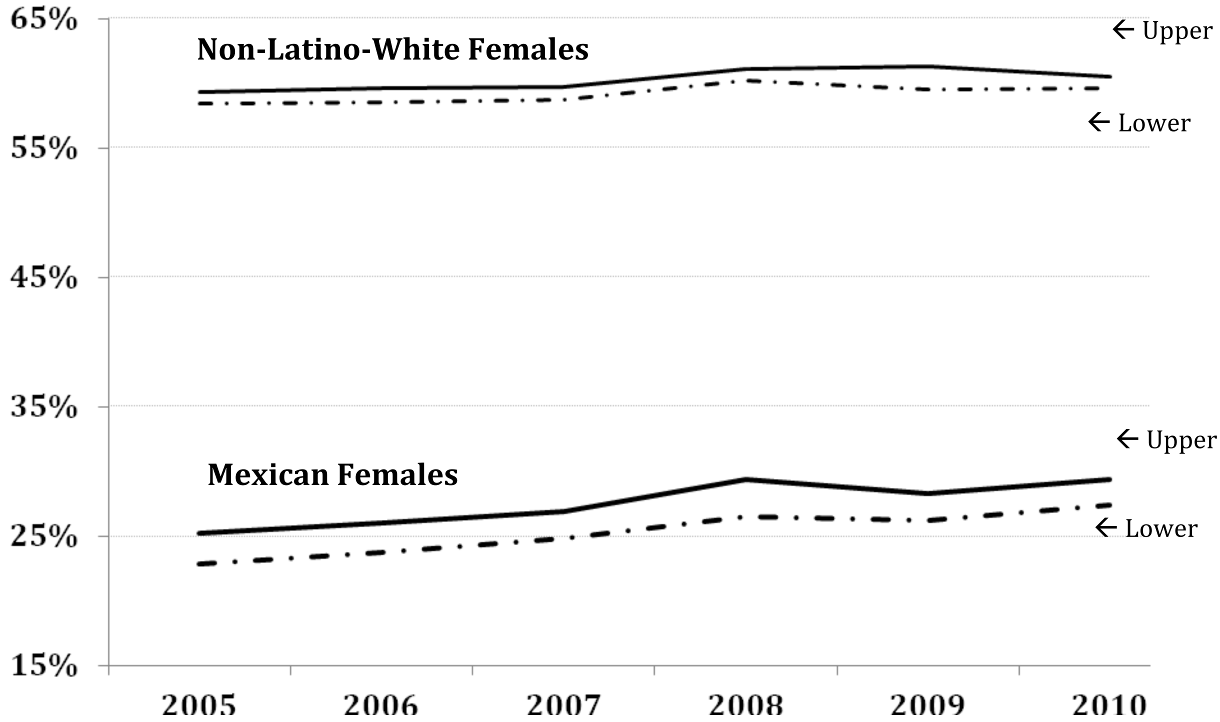
[1] Weighted number of people with at least one full year of college or more; [2] Percent with at least one full year of college or more = [(weighted disable count ÷ weighted total population from Table 1) × 100]; [3] Margin of error at 99%; [4] Relative range size where true estimate may be found, i.e., the "bandwidth" = [(MOE× 2) ÷ College) × 100]

The first thing to note from Table 2 is that Mexicans have much lower levels of individuals with at least one year of college education or more. Mexican females seem to have more education than Mexican males from the same 1960-1970 birth cohort—the same is true for NLW females relative to their NLW male counterparts. With regards to BW scores, the range of uncertainty in the estimate of educated individuals within this particular birth cohort, Mexican males have the highest BW values indicating their estimates are more volatile relative to the other groups. NLW females have the lowest BW numbers indicating that their estimates are relatively stable over time and much more precise than for the other groups.

Figure 2 graphs the computed the 99% MOE confidence intervals to displays the "2005 to 2010 trends" in what should be a relatively stable estimate of educational attainment in the 1960-1970 birth cohort of Mexican and NLWs. In general, the graph shows that all groups within the cohort have increased their level of education between the 2005 and 2010. It may be difficult to determine if this is the product of a real change in the groups or an artifact of the data—it may be that more highly educated people have higher rates of participation with the survey. The main thing to note from the graphs in Figure 2 is the space between the confidence interval lines within each group. As the between-line width grows, the location of the true population characteristic becomes less ascertainable. By way of example, note how the between line space for NLW males is much narrower than for Mexican males. The same is true when you compare the between-line width in NLW females and Mexican females. Any study of trends would need to include the variability of intervals, i.e., the heterogeneity in between-line widths for different sub-populations and their fluctuations over time. For example, note how the between-line width grows for both Mexican males and females in the 2008 single-year file. These graphs show a key challenge in trying to infer real population change with sample estimates: how do you decide where, within these liberal confidence intervals, the point-estimates should be located to draw a trend?

Figure 2: Upper and lower 99% confidence limits by race-ethnicity and sex





4. Conclusions

Detecting how the characteristics of a population change over time is important as is finding ways to make use of ACS data for this goal. The 2005 to 2010 single year ACS PUMS used in this analysis provides evidence on the heterogeneity of uncertainty in between-year differences on the level of education for a cohort approximately born between 1960 and 1970. It was argued that it is difficult to determine if 2005-2010 trends are meaningful population changes or the product of data artifacts. The investigation is limited in that it does not explore a multitude of population attributes over a wide range of grouping schemes, nor does it offer a simple solution for when to infer meaning change when evaluating trends with ACS data. Current work has begun to show that multi-year ACS files provide greater stability than single-year files (see Huang & Bell, 2012). Future efforts should first focus on developing 'quality of data measures' for single-year files and then proceed to the statistical quantification of variances. The use of ACS data to investigate population trends is not advised.

References

Alexander, C. (2001). Still rolling: Leslie Kish's "rolling samples" and the American Commu nity Survey. In Proceedings of Statistics Canada Symposium 2001, Achieving Data Quality in a Statistical Agency: A Methodological Perspective, Statistics Canada: Ottawa, Ontario, Canada.

Alkema, L., Raftery, A., Gerland., P, Clark, S. J., & Pelletier, F. (2012). Estimating trends in total fertility rate with uncertainty using imperfect data: examples from west Africa. Demographic Research, 26(15), 331-362.

Bishaw, A., & Semega, J. (2008). *Income, earnings, and poverty data from the 2007 American Community Survey*. US Department of Commerce, Economics and Statistics Administration, US Census Bureau.

Block, J. P., Subramanian, S. V., Christakis, N. A., & O'Malley, A. J. (2013). Population trends and variation in body mass index from 1971 to 2008 in the framingham heart study offspring cohort. *Plos one*, *8*(5), e63217.

Bchir, M., Vallin, J., Soomro, G. Y., Hakim, A., Karim, M. S., Scornet, C., ... & Sun, T. H. (2013). Demographic trends foreshadow major economic and social challenges. *Population: an English Selection*, *13*(2), 257-76.

Block, J. P., Subramanian, S. V., Christakis, N. A., & O'Malley, A. J. (2013). Population trends and variation in body mass index from 1971 to 2008 in the framingham heart study offspring cohort. *Plos one*, *8*(5), e63217.

Crosnoe, R., & Benner, A. D. (2012). Families, schools, and major demographic trends in the United States. *New directions for youth development*, *2012*(135), 87-95.

Cherlin, A. J. (2010). Demographic trends in the United States: A review of research in the 2000s. *Journal of Marriage and Family*, *72*(3), 403-419.

Fuller-Thomson E, Yu B, Nuru-Jeter A, Guralnik JM, Minkler M. Basic ADL disability and functional limitation rates among older Americans from 2000-2005: the end of the decline? *J Gerontol A Biol Sci Med Sci*. 2009;64A:1333–1336.

Hawala, S., & Lahiri, P. (2010). Variance Modeling in the US Small Area Income and Poverty Estimates Program for the American Community Survey. *Proceedings, Section on*

Bayesian Statistical Science, Section on Survey Research Methods, AS [CD-ROM], Alexandria, VA: American Statistical Association.

Heckman, J. J., & LaFontaine, P. A. (2010). The American high school graduation rate: Trends and levels. *The review of economics and statistics*, *92*(2), 244-262.

Heckman, J. J., & LaFontaine, P. A. (2010). The American high school graduation rate: Trends and levels. *The review of economics and statistics*, *92*(2), 244-262.

Huang, E. T., & Bell, W. R. (2012). An Empirical Study on Using Previous American Community Survey Data Versus Census 2000 Data in SAIPE Models for Poverty Estimates. *Statistics*, 04.

Keathley, D. H., Navarro, A., & Asiala, M. E. (2010). An Analysis of Alternate Variance Estimation Methods for the American Community Survey Group Quarters Sample. Joint Statistical Meetings, Vancouver. Washington, DC: US Census Bureau

Kish, L. (1981). Using Cumulated Rolling Samples to Integrate Census and Survey Operations of the Census Bureau. U.S. Government Printing Office, Washington, DC.

Kish, L. (1998). Space/time variations and rolling samples. Journal of Official Statistics, 14, 31-46.

Kochhar, R., Taylor, P., & Fry, R. (2011). *Wealth gaps rise to record highs between Whites, Blacks and Hispanics*. Washington, DC: Pew Research Center.

Martin, L. G., Schoeni, R. F., & Andreski, P. M. (2010). Challenges in estimating trends in late-life disability from the American Community Survey. *The Journals of Gerontology Series A: Biological Sciences and Medical Sciences*, *65*(5), 517-518.

Mather, M., Rivers, K. L., & Jacobsen, L. A. (2005). *The American community survey* (Vol. 60, No. 3). Washington, DC: Population Reference Bureau.

McElroy, T. (2009). Incompatibility of trends in multi-year estimates from the American Community Survey. The Annals of Applied Statistics, 3(4), 1493-1504.

Morritt, R., & Weinstein, A. (2012). *Segmentation strategies for hospitality managers: Target marketing for competitive advantage*. Routledge.

Siordia, C. (2012). Reliability of Self-report in Very Old Mexican Americans on Basic Health Conditions. *Maturitas*, 74; 95-98.

Siordia C. (2013a). Methodological note: allocations with health insurance items in the american community survey. Journal of Frailty and Aging, 6(2):149-153.

Siordia, C. (2013b). Benefits of small area measurements: a spatial clustering analysis on medicare beneficiaries. Human Geographies, 7(1), 53-59.

Siordia, C. (2013c). Number of People in the Unites States Experiencing Ambulatory and Independent Living Difficulties. *Journal of Social Work and Disability and Rehabilitation*, in-press.

Siordia C, Fox A. (2013). Public Use Microdata Area Fragmentation: Research and Policy Implications of Polygon Discontiguity. *Spatial Demography*, 1(1), 42-56.

Siordia, C., & Le, V. D. 2013. Precision of Disability Estimates for Southeast Asians in American Community Survey 2008-2010 Microdata. *Central Asian Journal of Global Health*, 1(2): DOI:10.5195/cajgh.2012.2166-7403.

Siordia, C., & Leyser-Whalen, O. (2013). The Lucky Few and Baby Boom Generation: Relative Cohort Size, Mexican Ethnicity, and Gender on the Likelihood of Out-of-Poverty Status. *Sociological Focus*, in-press.

Siordia, C., & Wunneburger, D. F. (2013). On the Quantity, Degree, and Location of Public Use Microdata Area (PUMA) Fragmentation. *Human Geographies*, 7(2): 5-13.

Siordia C, Young RL. (2013). Methodological note: allocation of disability items in the american community survey. *Disability and Health Journal*, 6(2), 149-153.

Spielman, S. E., Folch, D., & Nagle, N. (2014). Patterns and causes of uncertainty in the American Community Survey. *Applied Geography*, *46*, 147-157.

Starsinic, M. (2011). Incorporating a finite population correction factor into American Community Survey variance estimates. Joint Statistical Meetings, Section on Survey Research Methods, presented on August 2, 2011.

Sun, M., & Wong, D. W. (2010). Incorporating data quality information in mapping American Community Survey data. *Cartography and Geographic Information Science, 37*(4), 285-299.

Syvertsen, A. K., Wray-Lake, L., Flanagan, C. A., Wayne Osgood, D., & Briddell, L. (2011). Thirty-Year Trends in US Adolescents' Civic Engagement: A Story of Changing Participation and Educational Differences. *Journal of Research on Adolescence*, *21*(3), 586-594.

US Census Bureau. (2011). ACS Accuracy of the Data (2011) on http://www.census.gov/acs/www/data_documentation/documentation_main/

Van Hook, J., & Bachmeier, J. D. (2013). How well does the American Community Survey count naturalized citizens?. *Demographic research*, *29, 1-32*.