

CARRA Working Paper Series
Working Paper #2018-08

**Individual Changes in Identification with Hispanic Ethnic Origins:
Evidence from Linked 2000 and 2010 Census Data**

Mark A. Leach
U.S. Census Bureau

Tomás Jiménez
Stanford University

Center for Administrative Records Research and Applications
U.S. Census Bureau
Washington, D.C. 20233

Paper Issued: August, 2018

Disclaimer: This paper is released to inform interested parties of ongoing research and to encourage discussion. Any views expressed are those of the authors and not necessarily those of the U.S. Census Bureau.

Individual Changes in Identification Hispanic Ethnic Origins: Evidence from Linked 2000 and 2010 Census Data

Mark A. Leach, U.S. Census Bureau
Tomás Jiménez, Stanford University

ABSTRACT

Population estimates and demographic profiles are central to both academic and public debates about immigration, immigrant assimilation, and minority mobility. Analysts' conclusions are shaped by the choices that survey respondents make about how to identify themselves on surveys, but such choices change over time. Using linked responses to the 2000 and 2010 Censuses, our paper examines the extent to which individuals change between specific Hispanic categories such as Mexican origin. We first examine how changes in identification affect population change for national and regional origin groups. We then examine patterns of entry and exit to understand which groups more often switch between a non-Hispanic, another specific origin, or a general Hispanic identification. Finally, we profile who is most likely to change identification. Our findings affirm the fluidity of ethnic identification, especially between categories of Hispanic origin, which in turn carries important implications for population and compositional changes.

Key words: Hispanic origin; ethnic identification; ethnic fluidity; immigration; linked census records; American Community Survey.

Background

Population estimates, projections, and demographic profiles are central to both academic and public debates about immigration, immigrant assimilation, and minority mobility. Participants in these debates use aggregate population counts and forward-looking projections to portray both pessimistic (Coulter 2015; Krikorian 2008) and optimistic accounts (Frey 2014) of what the growing Hispanic population portends for the future of the United States.

The conclusions that analysts draw are profoundly shaped by choices that survey respondents make about how to identify themselves and other members of the household on surveys, like the decennial census (Waters 2002). Ample qualitative research shows that these choices are, especially for Hispanics, fluid, situational, and even shift over the life course (Lee and Bean 2010; Roth 2012; Smith 2014; Waters 1990). The fluidity of identity uncovered by qualitative research reflects in surveys too. Individuals may switch racial and ethnic categories (or have their identity category changed by interviewers) from one wave to the next of longitudinal surveys or between decennial censuses (Liebler et al. 2014; Saperstein and Penner 2012).

Some researchers posit that declining cohort size over time suggests that attrition from certain groups, Mexican origin in particular, appears to be related to indicators of assimilation such as inter-marriage, education and income (Alba and Islam 2009; Duncan and Trejo 2011; Emeka and Vallejo 2011). Increasingly complex ancestries also may explain inconsistent reporting of racial and ethnic identification, whether individuals trace their ancestry to multiple strands (Harris and Sim 2002; Telles and Ortiz 2008) or because the available categories may not entirely reflect the family ancestry or individuals social identity (Dowling 2014; Emeka and Vallejo 2011). These dynamics are particularly important to understand in the Mexican origin population given that its relative size and disadvantage often drive public policy debates on immigration (Bean and Stevens 2003; Telles and Ortiz 2008).

These findings point to a need to examine individual-level patterns and factors related to change in racial and ethnic identification in population survey data. But researchers are limited in doing so either because of small sample sizes inherent in longitudinal surveys or data restrictions on large government surveys like the decennial census, Current Population Survey, and American Community Survey. Instead, researchers have considered aggregate changes between survey years and infer from those changes the factors that might explain individual-level change in the selection of racial and ethnic categories. This approach has two major limitations. First, an aggregate approach limits researchers to studying net changes and likely masks a considerable amount of change in identification comprised of individuals who “join” a particular group (e.g., identify as “non-Hispanic” in one census but “Hispanic” in a subsequent census) and others who “leave” a group (e.g., identify as “Hispanic” in one census, and “non-Hispanic” in a later census) (Liebler et al. 2014). A more detailed analysis is essential for gaining a fuller picture of how individual identification choices shape the conclusions that can be drawn about racial and ethnic demographic change and about the extent and kind of group mobility (Alba 2016). Second, and relatedly, an aggregate approach does not allow for examination of individual-level factors related to changes in ethnic category selection. Identifying such factors may advance our understanding of why individual ethnic identities changes over time and in different contexts and how such changes may be related to assimilation processes.

Using similar data and methods, the present study and a forthcoming study by Fernandez, Porter, Bhaskar, and Ennis (forthcoming) address the limitations of prior research by investigating individual changes in ethnic identification over time. Whereas Fernandez et al. focus on change between Hispanic and non-Hispanic identifications, we investigate individual changes into and out of specific categories of Hispanic origin such as Mexican, Puerto Rican and Dominican. We focus in particular on the largest of the subgroups, Mexican origin, which, given the size and nature of immigration from Mexico, often draws the most media attention and drives policy debates (Bean and Stevens 2003). In this paper, we address the following research questions: How many individuals change their ethnic identification over time and to what extent do the changes contribute to population change in the Hispanic population, specifically for categories of Hispanic origin? To what extent are such changes associated with data collection and measurement issues such as imputation or changes in the Hispanic origin question? What patterns of change predominate when we disaggregate changes in identification by leavers and joiners of categories of Hispanic origin? And who is most likely to change between categories of Hispanic

identification? Our research into these questions moves beyond prior research and deepens our understanding about change in identification and its implications for social and economic outcomes in the Hispanic population.

We draw upon linked records from the 2000 and 2010 Censuses¹ that enable us to observe directly the extent to which individuals change their ethnic identification over time. Our findings affirm the fluidity of Hispanic ethnicity, and identification with Mexican origin in particular, in the United States. For example, we find that 12 percent of the net increase in the Mexican origin population between 2000 and 2010 was due to individual changes in identification. The proportions are much higher in Central and South American origin populations. We also find that changes in wording and format of the Hispanic origin question did not increase the rate at which individuals changed identification, as has been asserted previously. Rather, the question changes appear to have suppressed change out of specific origins and encouraged selection of specific origins among those who already were more prone to change. Finally, we find that the young, native born and those who experience significant life changes such as divorce or migration are more likely to change identification.

Racial and Ethnic Categorical Fluidity and Hispanics

It is a near maxim that racial and ethnic identities are fluid. Evidence for that fluidity has mostly come from historical and ethnographic research. Historical research shows that immigrant groups in particular have worked collectively and individually to be identified firmly with whiteness, which historically carried both social and legal advantages (Fox 2012; Guglielmo 2003; Haney-López 1996; Roediger 2005). Ethnographic work also highlights the fluidity of racial and ethnic identity in daily life. That fluidity is especially apparent among individuals who trace their ethnic ancestry to multiple strands (Lee and Bean 2010; Vasquez 2011; Waters 1990).

Recent research has shown that the fluidity demonstrated in historical and ethnographic research appears in survey data as well. In spite of the fact that survey responses to questions about ethnic and racial identification have the appearance of fixity, longitudinal data demonstrates quite the opposite. Changes in racial and ethnic identification across survey waves is apparent among young people who may have multiple ethnic ancestral strands (Harris and Sim 2002; Liebler et al. 2014). Inter-survey changes in identification may owe to significant life events – arrest, divorce, being laid off from a job, out of wedlock birth – that take place between surveys. These life events are predictors of identification rather than the other way around, which is more typical in models of identification (Saperstein and Penner 2012). Indeed, Saperstein (forthcoming) argues that these patterns of “racial mobility” reinforce the meaning and status of racial categories in the way that positive life events are precursors for the maintenance of a white or Asian identification, while negative life events raise the likelihood of switching to a black category.

But for Hispanics, it might be said that this fluidity is built into identity in ways that make fluctuations in their identification across surveys especially likely. Fluidity is built into the historical collective identity of Hispanic subgroups. Because populations in Latin America are made up of individuals who have Spanish and indigenous roots, there is an explicit recognition of racial mixing as part of a history that manifests in the ancestrally mixed individuals of today. For example, people of Puerto Rican and Dominican origin recognize an admixture of indigenous, Spanish, and African ancestries as important to a sense of contemporary peoplehood (Roth 2012). Some countries, like Mexico, honor their mixed origins – or “mestizaje” as part of their national self understanding (Rodriguez 2007).

If Hispanics collectively trace their origins to various Latin American countries where fluidity is asserted, Hispanics have historically encountered a U.S. context that imposes this fluidity on them. The racial identification of these groups has been historically ambiguous. In the late 19th and early 20th century, Mexicans, for example, were considered legally white, but also experienced Jim-Crow-like segregation in the Midwest and Southwest. And, their perceived racial status had a negative impact on their ability to claim the full rights of citizenship, including access to the welfare state (Fox 2012). The combination of their de-facto non-whiteness, but whiteness could be used as a tool for subjugation. For example, the legal system could invoke Mexicans’ white legal status to fend off legal challenge in regards to all-White (non-Hispanic) juries and all-Mexican schools on the grounds that Mexicans were *legally* white and therefore did in fact have a jury of their peers and all-Mexicans schools

¹ We describe these data and method of linking census records in more detail below.

were, “technically” speaking, all-white schools (Foley 2014; Gómez 2007). The ambiguous racial status in the lived experience also appeared in their official identification on the Census. For example, the 1930 census classified Mexicans as separate “Mexican” race, even though they were classified as White in all previous Censuses (Fox and Bloemraad 2015). Mexicans were reclassified as white in the 1940 Census and up until the 1970 Census, when the Hispanic category was first utilized.

Over time, the institutionalization of the Hispanic category in administrative data collection, politics, marketing, and everyday life means that individuals with ancestry in Latin America, the Spanish-speaking Caribbean, and Spain can selectively choose to identify with a larger pan-ethnic Hispanic population, or with the sub-ethnic groups (e.g., Mexican, Puerto Rican, Cuban, Spanish) that the Hispanic category encompasses (Oboler 1995; Okamoto and Mora 2014).

Layered on top of this historical backdrop is tremendous demographic diversity within the Hispanic population today. Not only do Hispanics trace their ethnic roots to multiple countries, the population is also marked by tremendous internal diversity with respect to language, skin color, legal status, social class, geography, and even ethnic ancestry (Alba, Jiménez, and Marrow 2014; Jiménez, Fields, and Schachter 2015). On the latter dimension of internal diversity, a high contemporary intermarriage rate – 26 percent of Hispanic newlyweds married a non-Hispanic in 2010 (Wang 2012) – contributes to a large number of Hispanics who have non-Hispanic ancestry (Alba 2016; Duncan and Trejo 2011; Macias 2006; Telles and Ortiz 2008). Ethnographic research shows that these multiethnic individuals articulate a large degree of flexibility in their lived identity (Jiménez 2004; Lee and Bean 2010). It may also be that these multiethnic individuals display similar flexibility in choosing a Hispanic or non-Hispanic category on the census forms.

It is against this historical and contemporary demographic backdrop that Hispanics choose their racial and ethnic identification on survey forms, like the U.S. Census. The patterned nature of the choices that Hispanics make about their racial and ethnic identification is best revealed not by examining their identification on any one census, but by looking at the degree of consistency in their choices across censuses. We thus examine the aggregate effect of these choices on the size and growth of the Hispanic population, and especially on the Hispanic subgroups that comprise it. We also consider the factors that predict individual-level choices about identification.

Before proceeding, it is important to bear in mind the difference between *identity* and *identification*. The former refers to a subjective sense of self that is perceived and projected in daily life. The latter has to do with the claims that people make to membership in a particular category. There is perhaps no more explicit identification claims making activity than checking a box on the census. But as a number of other scholars have warned, the logic that Hispanics use to explain their identification choices are incredibly varied, and may not necessarily correspond with popular notions of race and ethnicity (Dowling 2014; Rodriguez 2000; Roth 2012). Our examination can not examine identity in all of its complexity. Nonetheless, claiming identification on administrative forms is a regular and important part of the way that individuals express their group membership. Moreover, the research presented here is important for understanding what can be gleaned from racial and ethnic population estimates and projections. Moreover, our analysis more fully explains the individual and contextual level factors that pattern racial and ethnic identification in the United States.

DATA AND METHODS

We use restricted data from the 2000 and 2010 Censuses and 2006-2010 American Community Survey (ACS).² These data are unique in that the Census Bureau’s Center for Administrative Records Research and Applications (CARRA) uses probabilistic record linkage techniques (Fellegi and Sunter 1969) and personal information (name,

² The ACS is conducted annually and collects a variety of demographic and economic information from a representative sample of the U.S. population. The sample includes individuals from all counties and county-equivalents in the United States. The restricted ACS data that we use has not been through data perturbation, a process that ensures disclosure avoidance when data are released to the public (http://www2.census.gov/programs-surveys/acs/tech_docs/accuracy/MultiyearACSAccuracyofData2010.pdf). We ensure disclosure avoidance in any results we present by using internal Census Bureau disclosure review.

sex, date of birth, and address) to assign a unique identifier (a “Protected Identification Key” or PIK) to individuals (see Wagner and Layne 2014). Using PIKs as a match key, we first link 2000 and 2010 census records to create a longitudinal data set with responses from both censuses. To model change in ethnic identification, we then match demographic and economic variables from the 2006-2010 ACS to the linked census records.

Linked census data do not include all people who resided in the United States at the time of each census so the individuals in our analytic data set are not representative of the U.S. or Hispanic populations. Some individuals are not enumerated for various reasons (see Lujan 2014; Mule 2012; U.S. Census Bureau 2003), and others either were born, died, or had migrated out of the country after 2000. Furthermore, the Census Bureau may not be able to assign a PIK to a person in either or both censuses if that person’s personal information was incomplete or found not to be unique.³ Hereafter we refer to the population of individuals whose census records were successfully assigned a PIK and linked as the “linked population.” The linked population is comprised of 199 million person records, which includes 23 million people who identified as Hispanic or Latino in either 2000 or 2010 or in both years. To investigate our research questions we compare each individual’s answers to the ethnicity question in the two censuses to determine whether one changed their identification and if so whether the change was between Hispanic and non-Hispanic or between categories of Hispanic origin.

Exclusion Criteria

Not all identification changes observed in linked records originate from individuals explicitly choosing different responses to the Hispanic origin question. Changes related to data collection and linkage issues may be more prone to error.⁴ Data issues such as missing value imputation or proxy responses may misrepresent one’s identification and result in a false change (see Liebler et al. 2014). The extent to which observed identification changes are due to data collection and linkage issues is important because a large proportion would raise questions about the validity of measuring ethnic change over time in censuses and surveys. Alternatively, observed changes based on direct responses are less prone to error and, while not perfectly representative of ethnic identity, may help us better understand identity and its fluidity theoretically. To this end, we examine the number of and proportions of linked records that are associated with data collection and linkage issues and exclude them from our analyses of patterns and models of change.

Following Liebler et al. (2014), we use six criteria to identify changes related to data issues and PIK assignment (Table 1). We exclude responses to the 2010 Alternative Questionnaire Experiment (AQE) survey, which tested a variety of contact methods and question formats including for the race and ethnic origin questions. While we use a subset of AQE responses to understand the effects of changes in the ethnicity question, we exclude them when considering patterns and models of change so as to not confound experimental factors with individual and contextual factors. The next three exclusion criteria are related to data collection and imputation. Demographic information for residents of group quarters such as college dormitories or prisons is often obtained from administrative records, which may be out of date and not reflect a person’s ethnic identification at the time of the census. In other cases, a response may be obtained by proxy, from a neighbor for example if a person cannot not be located, or imputed if an individual provided other information on a census form but did not answer the ethnicity question.

³ A PIK could not be assigned to about 10 percent of the population in each census (Layne et al. 2014).

⁴ Observed changes based on two direct responses also may not accurately measure change in ethnic identity. For example, one person typically completes a census form and provides information for other residents in a household. If the person reporting an individual’s identification changes, they may provide an inconsistent response from the previous census. As a result, we expect to observe relatively higher rates of identification changes among individuals who are not a householder or spouse of a householder, who typically fill in a census form, and especially children.

Table 1. Excluded Linked Records, Hispanic Origin in 2000 or 2010.

	Total	Percent
Hispanic in 2000 or 2010	23,173,370	100.0
Not Excluded	18,580,005	80.2
Exclusions	4,593,365	19.8
2010 Alternative Questionnaire	164,606	0.7
Group Quarters	601,468	2.6
Proxy Response	559,725	2.4
Imputed Response	2,354,671	10.2
Imputed Age or Sex	1,100,517	4.7
Inconsistent Age or Sex	965,782	4.2

Sources: Linked 2000 and 2010 census data

Note: Exclusion criteria are not mutually exclusive so the totals and percents of the categories do not sum to total exclusions.

The last two criteria in Table 3 exclude linked records that appear to be for two different people due to an incorrect PIK assignment for one or both census records. We exclude records with imputed values for age or sex given that an inaccurate value may lead to an incorrect PIK assignment. We also exclude linked records with inconsistent values for sex or age, if the values for sex are not equal or age in 2010 is not between eight and twelve years greater than age in 2000.

Of the 23 million individuals in the linked population that identified as Hispanic in either 2000 or 2010, about 4.6 million (19.8 percent) meet at least one of the six exclusion criteria. A majority of excluded records have an imputed value for Hispanic origin (10.2 percent of the linked Hispanic population), and almost half have questionable PIK assignments, whether age or sex are imputed (4.7 percent) or they have inconsistent values across the two censuses (4.2 percent).

We find that approximately fifty percent of excluded records experienced a change in identification whereas about one quarter of non-excluded records exhibit a change. A much higher rate of change among excluded records affirms our concern that data collection issues and questionable PIK assignments may introduce additional error into identification changes. After excluding records that meet at least one of our exclusion criteria, 18.5 million linked records remain with which we investigate patterns and estimate models of identification change.

We note that the individuals who comprise the non-excluded linked population are not representative of the U.S. Hispanic origin population. Whether an individual is assigned a PIK is not random. Given the importance of address, the process favors people who move less often, which in turn reduces the rate of successful PIK assignment for men ages 20 to 50 (Rastogi and O’Hara 2012). Bond et al. (2014) find that PIK assignment rates also are lower for Hispanics and non-whites. In addition, an estimated one percent of PIKs are assigned to an incorrect person, also likely not random (Layne et al. 2014). While our exclusion criteria reduces measurement error in identification changes, we cannot adjust for bias introduced by incorrect PIK assignments. Finally, our exclusion criteria likely disproportionately affect some groups over others. For example, women respond to questions at slightly higher rates than men (Rastogi et al. 2014), a difference which is magnified when data sources are linked. Given the non-randomness of PIK assignments and exclusions, our results pertain largely to individuals who were successfully assigned a correct PIK in both censuses and provided direct responses on their census form.

Changes in the Ethnicity Question

Changes in identification with Hispanic origin may also owe to changes in the format and wording of the ethnicity question between the 2000 and 2010 Censuses. The Census Bureau changed the question in several important ways (see Figure 1). The question in 2000 did not include the word “origin” in the question or answer option whereas the 2010 version of the questionnaire included the word “origin.” Another important change was that the 2010 form listed examples of “other Spanish/Hispanic/Latino” origins. The 2000 form did not list examples.

Figure 1. Changes in Hispanic Origin Question Format, 2000 and 2010.

2000	2010
<p>7. Is Person 1 Spanish/Hispanic/Latino? Mark <input checked="" type="checkbox"/> the "No" box if not Spanish/Hispanic/Latino.</p> <p><input type="checkbox"/> No, not Spanish/Hispanic/Latino <input type="checkbox"/> Yes, Puerto Rican</p> <p><input type="checkbox"/> Yes, Mexican, Mexican Am., Chicano <input type="checkbox"/> Yes, Cuban</p> <p><input type="checkbox"/> Yes, other Spanish/Hispanic/Latino — <i>Print group.</i> ↴</p>	<p>8. Is Person 1 of Hispanic, Latino, or Spanish origin?</p> <p><input type="checkbox"/> No, not of Hispanic, Latino, or Spanish origin</p> <p><input type="checkbox"/> Yes, Mexican, Mexican Am., Chicano</p> <p><input type="checkbox"/> Yes, Puerto Rican</p> <p><input type="checkbox"/> Yes, Cuban</p> <p><input type="checkbox"/> Yes, another Hispanic, Latino, or Spanish origin — <i>Print origin, for example, Argentinian, Colombian, Dominican, Nicaraguan, Salvadoran, Spaniard, and so on.</i> ↴</p>

The Census Bureau conducted tests before and after the 2010 Census and found that the format changes likely resulted in greater proportions of respondents identifying as Hispanic and selecting more specific origins rather than a general Hispanic or Latino identity (Martin et al. 2004; Stokes et al. 2012; Compton et al. 2013). A similar shift in preferences for specific Hispanic origins occurred in data from the American Community Survey when the question format was changed between 2007 and 2008 (Ramirez 2009; Fernandez et al. 2015). Selection of specific origin categories in 2010 mirror preferences for non-Hispanic or general Hispanic origin in 2000 when the word “origin” and examples were dropped from the 1990 format (Harrison et al. 1996; Martin 2002; Logan 2002; Suro 2002; Cresce and Ramirez 2003; Cresce, Schmidley, and Ramirez 2000; Lavrankas et al. 2005).

The cross-sectional designs of these studies cannot address whether identification with specific origins resulted from changes in identification, *per se*, or that the format change interacted with other compositional changes in the Hispanic population, such as by age or nativity, that made individuals more prone to select specific origins when presented with the 2010 question format. If the latter, we would expect to observe similar rates of change in groups that received different question formats.

We use a subsample of the 2010 Alternative Questionnaire Experiment (AQE) survey to analyze whether changes in question format were related to changes in ethnic identification. To do so, we link 2000 Census records to the 2010 AQE sample and compare the probability of identification change in the AQE control group, which received the 2010 Census form, and change in an AQE treatment group that received the 2000 question format. After deleting records that satisfy other exclusion criteria listed in Table 3, our linked AQE sample includes 36,748 records that identified as Hispanic in 2000 or 2010, of which 1,937 received the 2000 question format. Although it would be ideal to analyze differences within each specific Hispanic category, we only have linked universes of sufficient sizes to analyze Mexican origin and Other Hispanic/Latino origins separately.

We calculate and compare percentages and odds of identification change in the AQE control and treatment groups. If the 2010 question changes prompted individuals to change from a general Hispanic identification to a more specific identification, then we expect the odds of change to be greater in the control group relative to the treatment group.

Patterns and Models of Identification Change

We use linked population data to assess our first and second research questions. To address the extent and patterns of identification changes, we calculate totals and rates of change into and out of eight categories of Hispanic identification. Unlike cohort studies that are limited to observing net changes, linked data allow us to observe flows of change between specific Hispanic categories and between Hispanic and non-Hispanic.

To address our third research question regarding who changes ethnic identification, we estimate logistic regression models that predict the probability of change into or out of each of seven categories of Hispanic origin. We assume with our models that joining or leaving a category of Hispanic origin indicate similar processes of ethnic fluidity, involve similar factors theoretically, and may occur multiple times over time for individuals more prone to change identification. We thus present models of individual change in general rather than modeling joining and leaving separately.

We use a variety of individual and place-related characteristics as independent variables related to change in ethnic identification. Census data are limited to race, age, sex, household relationship, and geographic location so

we match the linked census records to responses to the American Community Survey in 2006 through 2010 to obtain additional characteristics with which to model ethnic change. Of the 18.5 million linked Hispanic census records that we do not exclude, 1.1 million match to a 2006-2010 ACS record that we then use to model ethnic change.

The following general equation describes the logistic regression models we use to estimate the probability that an individual changed identification with a category of Hispanic ethnicity⁵:

$$EI_{00-10} = b + X_{00} + Y_{00-10} + Z_{06-10} + e \quad (1)$$

where

EI_{00-10} is a dummy variable indicating change into or out of a Hispanic category,

b is the model intercept,

X_{00} represents the variables age and sex from each person's 2000 census record,

Y_{00-10} are variables that represent change between 2000 and 2010 including change in household relationship status, state-to-state migration, and change in racial identification,

Z_{06-10} includes a variety of characteristics obtained from the 2006-2010 ACS microdata such as ancestry, education, and

e is a residual.

In accordance with previous research, we expect identification changes to be related to mixed origins, which we indicate as an inconsistency between an individual's ancestral and ethnic origins, and racial identification and whether it changed over the same period of time (Harris and Sim 2002; Lee and Bean 2010; Vasquez 2011; Waters 1990). The foreign-born may have stronger ties to specific national origins whether because they experience social or economic advantages in identifying with their home country or because they have fewer options in their daily lives given racial, cultural and linguistic barriers (Alba and Islam 2009; Jiménez 2008; Nagel 1995). By contrast, the native born may change identification relatively more often not only because they are more likely to be of mixed origins but also because they may have a presence for changing their identification due to intra- or inter-marriage or economic mobility (Duncan and Trejo 2011; Lee and Bean 2010). Differences by nativity may be particularly salient for the Mexican origin category given relatively high rates of unauthorized immigration status, a structural barrier to incorporation that affects this group more than other immigrant groups (Bean, Brown, and Bachmeier 2015). We expect those individuals who were not likely to self-identify on a census form, such as children and those who were not a householder or spouse, are more likely to change ethnic identification (Jiménez, Park, and Pedroza 2015; Waters 2002). Research also has shown that changes in identification are related to significant life events such as a change in marital status or migration (Saperstein and Penner 2012). And, we expect changes in identification to be associated with region given different social and cultural contexts in terms of racial and immigration histories (Jiménez 2008; Lee and Bean 2010).

Several limitations of the data and methods should be kept in mind when considering our results. Identification changes in census data do not perfectly represent change in identity. Some observed changes are due to incorrect PIK assignments and false linkages, differences in post-enumeration processing across the two censuses, different respondents and perceptions about the "correct" response for other individuals in a household, or mistakenly marking an incorrect box(es) when filling out the form. Regarding these latter two sources of error, we are not able to study those with an Hispanic identity that was not reported in the census ethnicity question. Finally, our results over-represent response stability in two ways: (a) due to our exclusion criteria and limitations of linked data, our data necessarily include individuals with stable responses – those who do not reside in group quarters and exhibit stable responses to other questions – and (b) we focus on only two measures of a person's ethnicity over a period of 10 years even though many response changes may occur.

⁵ We do not use ACS weights in estimating the models because the linked population is not nationally representative. ACS weights account for survey non-response and sampling design, among other factors, but they do not adjust for record linkage and case selection so weighted estimates still would not be representative of the population.

We first focus on identification changes into and out of Mexican origin to understand how different variables relate to change alone and in combination with other variables. We then make comparisons between change into and out of Mexican origin and other categories of Hispanic origin.

RESULTS

Levels and Patterns of Identification Changes

To answer our first research question we examine how many individuals in the linked population changed their ethnic identification between 2000 and 2010 and consider the extent to which such changes affected the size of the Hispanic population overall and of specific categories of Hispanic origin. We observe that 7.2 million individuals, or approximately 31 percent of linked records that identified as Hispanic in either census, had different responses to the Hispanic origin question, whether switching between non-Hispanic and Hispanic or between categories of Hispanic origin. Table 2 shows the net effects of the 7.2 million identification changes on the size of Hispanic populations. The left panel of the table shows well-documented change in the Hispanic origin population between 2000 and 2010 (Ennis, Rios-Vargas, and Albert 2011; Passel, Cohn, and Hugo Lopez 2011). The Hispanic population grew by 15 million or about 43 percent, 11 million of which occurred in the Mexican origin population.

The right panel shows change in the linked population, which, by definition, includes only change due to individual changes in identification and excludes change due to births, deaths, and international migration. Individual changes in ethnic identification thus resulted in a net increase of the Hispanic population by 640,000. In other words, 640,000 more people switched from non-Hispanic to Hispanic than those who changed from Hispanic to non-Hispanic across the decade. This change amounts to 4.2 percent of net growth in the total Hispanic population.

Table 2. Hispanic Population Change, 2000 to 2010.

Hispanic Origin	Total Population				2000-2010 Linked Population ¹				
	2000	2010	Net Change	Percent Change	2000	2010	Net Change	Percent Change	Percent of Total Pop Change
Non-Hispanic	246,116,088	258,267,944	12,151,856	4.9	177,686,894	177,046,186	-640,708	-0.4	-5.3
Hispanic	35,305,818	50,477,594	15,171,776	43.0	21,310,964	21,951,672	640,708	3.0	4.2
Mexican	20,640,711	31,796,431	11,155,720	54.0	11,878,355	13,212,030	1,333,675	11.2	12.0
Puerto Rican	3,406,178	4,623,470	1,217,292	35.7	2,186,125	2,443,505	257,380	11.8	21.1
Cuban	1,241,685	1,785,366	543,681	43.8	864,612	1,005,962	141,350	16.3	26.0
Dominican	764,945	1,415,050	650,105	85.0	488,218	645,173	156,955	32.1	24.1
Central American	1,686,937	3,999,073	2,312,136	137.1	987,559	1,518,459	530,900	53.8	23.0
South American	1,353,511	2,770,255	1,416,744	104.7	835,538	1,199,410	363,872	43.5	25.7
Spanish	868,398	1,126,856	258,458	29.8	590,940	685,194	94,254	15.9	36.5
Other Hispanic / Latino	5,343,453	2,961,093	-2,382,360	-44.6	3,479,617	1,241,939	-2,237,678	-64.3	93.9

Sources: 2000 and 2010 censuses

¹ The linked population includes people who were enumerated in both censuses and for whom the Census Bureau was able to match their 2000 and 2010 census records.

The implications of identification changes for specific categories of Hispanic origin are more substantial. The Mexican origin population increased by 1.3 million people, or about 12 percent of total net growth, due to individual changes in identification. Changes in ethnic origin represents between one-fifth (Puerto Rican origin) and one-fourth (Cuban, Dominican, and South American origins) of net population growth in other specific Hispanic categories. In contrast to specific origins, the “Other Hispanic/Latino” category declined by 2.2 million people due to “opting out” of a general or pan-ethnic Hispanic identification.

We also observe that, of 18.5 million individuals in the linked Hispanic population, more than one quarter, 27 percent or 4.9 million, have different ethnic identifications between the two censuses. This includes those who switched between Hispanic and non-Hispanic (2 million) and between categories of Hispanic origin (2.9 million). The two million identification changes that we observe between Hispanic and non-Hispanic, or about 10 percent of the linked population, is slightly lower than the proportion that Fernandez et al. (forthcoming) find in their study (11.2 percent). They consider change across three points in time between 2000 and 2010, however. The

difference suggests that ethnic identifications likely change more often than what we observe in our data and that we underestimate the extent and frequency that changes occur.

We find a considerable amount of churning into and out of categories of Hispanic origin when we disaggregate net changes into leavers and joiners of each category (Table 3).⁶ For example, a net increase of 400,000 in the linked non-excluded Hispanic population is the result of 1.2 million individuals joining Hispanic from non-Hispanic (6.6 percent of joint Hispanic population) and 800,000 (4.4 percent) opting out of identification with Hispanic origin.

The amount of churning and rates of joining and leaving vary relatively more across specific categories of Hispanic origin. Those who identify as Mexican, Puerto Rican, or Cuban change identification least often with between 5.7 percent and 7.4 percent of the joint populations leaving these categories and between 12.5 and 15.2 percent joining between censuses. Individuals joined Dominican, Central and South American, and Spanish categories at more than twice the rates of the three categories with a separate check box (26, 38, 32, and 43 percent, respectively). Those who identified with a pan-ethnic or some other Hispanic origin exhibit the highest rate of change with about 70 percent of the joint population leaving this category and almost 20 percent joining.

How leavers identified after and joiners identified previous to switching affirm preferences for more-specific Hispanic origins in 2010. Almost all (90.9 percent) of those who switched from non-Hispanic to Hispanic (non-Hispanic “leavers”) identified with a specific origin. Likewise, 90.3 percent of the leavers from the Other Hispanic/Latino category switched to a specific Hispanic origin category. As we mention above, Census Bureau experiments and other analyses have concluded that this pattern of leaving general Hispanic identifications and joining specific categories of Hispanic origin largely was prompted by changes to the ethnicity question (we investigate this point below).

In spite of general preferences for specific Hispanic categories, leavers and joiners are not distributed evenly across Hispanic categories. For example, about half of the leavers from Mexican (41.9 percent), Puerto Rican (54.0 percent), and Cuban (59.7 percent) origins switched to non-Hispanic in 2010. Most leavers from Dominican (56.6 percent) and Central American (41.5 percent) categories changed to a general Hispanic origin. Individuals switching from South American origins appear to be more evenly split between non-Hispanic (38.2 percent) and Other Hispanic/Latino (34.9 percent).

As with leavers from Puerto Rican and Cuban identifications, most joiners of these two categories previously identified as non-Hispanic. The churning between non-Hispanic and these two categories is consistent with the options that individuals tied to each of these groups have in switching back and forth as non-Hispanic (Waters 1990). By contrast, relatively fewer joiners of Mexican (29.7 percent), Dominican (16.0 percent), and Central American (14.5 percent) and South American (26.7 percent) origins previously identified as non-Hispanic. It may be that these groups, with greater concentrations of foreign born, may be more tied to an ethnic ancestry linked to the homeland, and thus perceive themselves to have fewer options, or be more influenced by question wording and the listing of national origin examples.

⁶ Note that in this table, those joining or leaving a specific national or region-origin group are necessarily not a subset of the total number of people who joined or left Hispanic identification. Individuals who joined Mexican origin in 2010, for example, could have identified either as non-Hispanic or as some other Hispanic national origin in 2000. Those who joined a Hispanic identification in 2010, whether a general Hispanic, Mexican origin, or some other identification, could only have identified as non-Hispanic in 2000. Individual who switched between Hispanic identifications did not join Hispanic.

Table 3. Leavers and Joiners of Categories of Hispanic Origin, Non-Excluded Linked Population, 2000 – 2010.

	Joint 2000/2010 Non- Excluded Population ¹	Leavers ²					Joiners ³				
		Total	Percent of Leavers				Total	Percent of Joiners			
			Percent of Joint Population	To Non- Hispanic	To Specific Hispanic Origin	To Other Hispanic / Latino		Percent of Joint Population	From Non- Hispanic	Specific Hispanic Origin	Other Hispanic / Latino
Non-Hispanic	151,511,275	1,217,890	0.8	-	90.9	9.1	817,427	0.5	-	74.9	25.1
Hispanic	18,580,005	817,427	4.4	100.0	-	-	1,217,890	6.6	100.0	-	-
Mexican	11,864,685	876,547	7.4	41.9	22.2	35.9	1,807,325	15.2	29.7	9.3	61.0
Puerto Rican	2,059,447	140,477	6.8	54.0	32.7	13.3	257,777	12.5	56.8	22.9	20.2
Cuban	860,981	48,832	5.7	59.7	32.2	8.1	130,837	15.2	69.3	14.7	16.0
Dominican	541,278	39,021	7.2	18.3	25.1	56.6	141,515	26.1	16.0	17.8	66.2
Central American	1,318,781	113,495	8.6	21.6	36.9	41.5	503,740	38.2	14.5	18.3	67.2
South American	1,047,989	74,825	7.1	38.2	26.9	34.9	340,585	32.5	26.7	14.4	58.9
Spanish	808,723	299,045	37.0	26.5	59.4	14.1	349,481	43.2	42.0	26.6	31.5
Other Hispanic / Latino	2,976,932	2,123,996	71.3	9.7	90.3	-	585,441	19.7	18.9	81.1	-

Sources: Linked 2000 and 2010 Census data

Universe: U.S. residents present in both 2000 and 2010 Censuses and for whom 2000 and 2010 Census records could be linked and not excluded.

1 Non-excluded linked population that identified with category in 2000 or 2010.

2 Leavers identified with category in 2000 but did not identify with that same category in 2010.

3 Joiners identified with category in 2010 but did not identify with that same category in 2000.

The levels and patterns of identification changes observed in Tables 2 and 3 affirm that changes affiliated with direct responses (non-exclusions) comprise significant amounts of population change in categories of Hispanic ethnicity. Non-excluded identification changes account for substantial proportions of net growth in each of the specific categories of Hispanic origin, including about 12 percent of net growth in the Mexican origin population and more than 20 percent of growth in each of Cuban, Dominican, and Central and South American origin populations (Table 2). That a majority of changes in identification comes from direct responses is consistent with previous research on the fluidity of ethnic identity (Emeka and Vallejo 2011; Saperstein and Penner 2012). In other words, observed changes in census data are not simply due to data collection issues or erroneous record linkages. That said, when we consider the excluded linked records in the analysis, between one quarter and one half of *all* observed changes in ethnicity, depending on the category of Hispanic origin, come from records that are subject to a higher degree of error. This point carries important implications for estimated rates of change put forth by previous studies that could not distinguish types of data values associated with identification changes.

We expect that models of identification change will further illuminate the patterns observed in Table 3. Before we further investigate who is likely to change between which Hispanic groups, however, we first examine how changes in the format of the ethnicity question may have contributed not only to the observed patterns but to relative levels of change.

Changes in Ethnicity Question Format

We show identification changes by question format in Table 4 to understand whether changes in question wording and format prompted individual changes in identification. Previous Census Bureau studies conclude that the wording and format changes implemented in 2010 likely resulted in individuals identifying with specific Hispanic origins relatively more often than in 2000 (Martin et al. 2004; Stokes et al. 2012; Compton et al. 2013). Such studies, however, cannot tell us whether the question changes prompted relatively more people to change identification, or whether respondents receiving different question formats changed identification at similar rates and the question format simply prompted a different pattern of identification choices and resulted in relatively more people selecting specific Hispanic identifications.

Table 4 shows our analysis of linked data from Census 2000 and the 2010 AQE. For Hispanic identification in general, 23.8 percent of the AQE sample that received the 2010 question format (Control) changed identification. Among individuals who received the 2000 question format in 2010 (Treatment), 27.4 percent changed identification. The difference implies that the odds of changing identification was 20 percent greater for those who *did not* experience a change in question format. This runs contrary to the notion that changes in question wording and format prompted relatively more changes in identification.

Table 4. Differences in Likelihood of Identification Change by Question Format, 2000 to 2010.

	Control: 2010 Format		Treatment: 2000 Format		2000 to 2010 Odds Ratio ¹
	Joint 2000/2010 Sample	Percent of Origin	Joint 2000/2010 Sample	Percent of Origin	
Hispanic/Latino	34,811	100.0	1,937	100.0	
Changed Identification	8,280	23.8	530	27.4	1.21 **
Mexican Origin	21,328	100.0	1,038	100.0	
Join	2,992	14.0	139	13.4	1.04
Leave	1,303	6.1	136	13.1	2.33 **
Other Hispanic/Latino	5,406	100.0	483	100.0	
Join	881	16.3	194	40.2	1.01
Leave	4,023	74.4	179	37.1	0.20 **

Source: 2000 Census and 2010 Census Alternative Questionnaire Experiment sample

1 Ratio of odds of changing identification when given 2000 question format in 2010 to odds when given 2010 question format. Ratios for Mexican origin and Other Hispanic/Latino calculated using multinomial logistic regression with Stay/Join/Leave as dependent variable (stay as reference group) and 2000/2010 format as independent variable.

** Difference in log odds of change in identification when given 2000 format is statistically significant at 0.01 level.

We further explore implications of the question changes by assessing the likelihoods of joining and leaving the Mexican origin and Other Hispanic/Latino categories given different question formats in 2010. These are the only two Hispanic groups with sufficient sample size in the treatment group for a reliable comparison. Regarding Mexican origin, respondents who received different question formats in 2010 were just as likely to switch identification and join the Mexican origin group (14 percent versus 13.4 percent). Alternatively, those who received the same question format in both years were *more* likely to leave Mexican origin identification (13.1 percent relative to 6.1 percent of those who received 2010 question format). The net result of changes in question format and wording appears to have encouraged relatively more respondents to keep identifying as Mexican origin rather than more switching from non-Mexican to Mexican.

We find similar results in the opposite direction with regard to the Other Hispanic/Latino category. Individuals who received the 2010 question format were five times as likely to leave this category (2000-to-2010 odds ratio of 0.2) but no more likely to join. While the 2010 form resulted in fewer individuals choosing a general identification, it does not appear to have resulted in relatively more identification changes.

In sum, changes in question wording and format do not appear to have prompted more individuals to change identification relative to what would have occurred anyway without question changes. This finding affirms that the fluidity of ethnic identification is not simply a function of question wording and format but likely is associated with demographic and contextual factors.

Our analyses thus far have addressed three concerns: (1) the extent to which change in identification with Hispanic ethnicity in general and with specific Hispanic categories, occurs over time; (2) the degree to which such changes may be due to data collection and record linkage issues; and (3) the amount of influence that changes in question wording and format may have had in identification changes. Consistent with our expectations, individual change in ethnic identification is relatively common, a large majority of individual changes appear to be related more to changes in identity rather than to data collection and linkage issues, and question changes appear to affect how one changes identification but not the likelihood of changing one's identification.

Models of Change in Ethnic Identification

With such results in mind, we model identification changes for non-excluded individuals in the linked population to understand the characteristics of those who change identification. By modeling changes in general rather than joining and leaving a category separately, we assume that joining and leaving a particular identification are part and parcel of the same process of ethnic fluidity and are influenced by similar factors.⁷

We first focus in detail on identification changes related to Mexican origin to understand how our various independent variables differently relate to change in identification with this group. We then compare results for models for other Hispanic categories that include all independent variables.

In Table 5, models 1 through 4 show results for groups of independent variables separately. Model 5 includes all independent variables. Model 1 shows the association between identification change and household relationship status, distinguishing changes for those who likely self-identified (or whose ethnicity was indicated by a spouse) in both censuses from those who likely did not self identify in one census or the other, or both. Model 2 includes other types of identification including ancestry, race, and nativity. Model 3 assesses how migration and region of residence relate to change, and Model 4 includes various demographic and economic variables such as sex, age, and education.

⁷ Using the same independent variables, we modeled differences between joiners and leavers of the Mexican origin category and did not find substantive differences between the two with one exception. Those with mixed origins are much less likely to join the Mexican origin category than to leave it, which is what we would expect to the extent that those of mixed origins may be afforded and take advantage of greater options to choose their identity in their daily lives.

Table 5. Models of Change in Mexican Origin Identification, 2000 to 2010

	Probability of Joining or Leaving Mexican Identification									
	Model 1		Model 2		Model 3		Model 4		Model 5	
	Estimate	Odds Ratio	Estimate	Odds Ratio	Estimate	Odds Ratio	Estimate	Odds Ratio	Estimate	Odds Ratio
Intercept	-1.38 ***		-0.73 ***		-1.34 ***		-1.41 ***		-0.90 ***	
Householder/Spouse Status (ref: HH/Spouse in 2000 and 2010)										
HH/Spouse in 2000 or 2010	0.31 ***	1.36							0.09 ***	1.10
Not HH/Spouse	0.22 ***	1.24							-0.18 ***	0.84
Ancestry (ref: Mexican)										
Other Specific Hispanic Origin ¹			1.85 ***	6.36					1.79 ***	6.02
Non-Specific Other Hispanic Origin			0.79 ***	2.20					0.79 ***	2.19
Non-Hispanic			1.41 ***	4.09					1.38 ***	3.99
Race in 2000 (ref: all other race categories)										
White alone			-0.64 ***	0.53					-0.63 ***	0.53
Change in Race, 2000-2010 (ref: no change)			-0.23 ***	0.79					-0.25 ***	0.78
Nativity (ref: U.S.-born)										
Foreign Born, In U.S. 10+ Years			-0.94 ***	0.39					-1.00 ***	0.37
Foreign Born, In U.S. 5-9 Years			-0.72 ***	0.49					-0.85 ***	0.43
Foreign Born, In U.S. 0-4 Years			-0.62 ***	0.54					-0.75 ***	0.47
Inter-State Migration (ref: same state)					0.47 ***	1.59			0.39 ***	1.48
Region (ref: Southwest)										
Pacific Northwest					0.30 ***	1.35			0.17 ***	1.18
Midwest					0.00	1.00			-0.10 ***	0.90
Southeast					0.43 ***	1.53			0.24 ***	1.27
Northeast					0.77 ***	2.17			0.54 ***	1.71
Female							0.07 ***	1.07	0.03 ***	1.03
Age (ref: 46+ years old)										
0 - 9							0.19 ***	1.21	0.18 ***	1.20
10 - 17							0.20 ***	1.22	0.23 ***	1.26
18 - 25							0.17 ***	1.18	0.17 ***	1.19
26 - 35							0.07 ***	1.08	0.18 ***	1.20
36 - 45							0.01	1.01	0.12 ***	1.12
Marital Status (ref: Married)										
Widowed, Divorce, Separated							0.28 ***	1.32	0.16 ***	1.18
Never Married							0.08 ***	1.08	0.00	1.00
Education (ref: Less than High School diploma)										
High School Diploma							0.12 ***	1.13	0.07 ***	1.07
Some College							0.07 ***	1.07	-0.03 **	0.97
Bachelors Degree or Higher							0.00	1.00	-0.21 ***	0.81
Speaks English Well or Very Well							-0.64 ***	0.53	0.06 ***	1.06
Sample Size and Model Fit										
N	707,875		707,875		707,875		707,875		707,875	
-2 Log Likelihood	752,414		680,765		747,793		746,599		675,503	
Likelihood Ratio	1,943 ***		73,591 ***		6,563 ***		7,757 ***		70,232 ***	

Sources: 2000 and 2010 Census, and unweighted 2006-2010 American Community Survey data. For more information on the ACS, see census.gov/acs

¹ Includes Puerto Rican, Cuban, Dominican, Central American, South American, or Spanish national origins.

*** p<0.01; ** p<0.05; * p<0.1

The results for Model 1 indicate that the odds of changing identification with Mexican origin are 24 to 36 percent greater among individuals who were not a householder or a spouse of a householder in both censuses relative to householders and their spouses. This implies that Mexican origin identification is much more fluid among those who likely do not self-identify on a census form, such as children. This may be due to shifting identities of children but also due to inconsistent responses when one individual provides responses for others in a household. The fact that this effect is much smaller in the fully specified model, Model 5, which includes age, suggests that identifications of children drive this association.

Model 2 includes factors most strongly associated with identification change, including mixed origin, racial identification and its change, and nativity. Individuals who identified with ancestries other than Mexican origin have odds of changing identification several times greater than those who indicate only Mexican ancestry (odds ratios between 2.2 and 6.4). This is consistent with previous studies that find that the identities of those with mixed origins are much more fluid (Harris and Sim 2002; Jiménez 2004; Lee and Bean 2010). The model also indicates that those who identify racially as white, which is common among those who also identify as Mexican origin, and do not change their racial identification over time are much less likely to change their ethnic identification with Mexican origin.

Also as we expected, the odds of changing identification is greater among native-born individuals relative to the foreign born (odds ratios between 0.4 and 0.5 for foreign born with various tenures in the United States). This is consistent with the literature that the native born have more options to choose and change their identities (Waters 1990, 1999). Interestingly, however, the foreign-born with more time in the United States are least likely to change and appear to hold the strongest ties to Mexican identification. Though the differences between foreign-born Mexicans of varying tenure in the United States is small, this result appears to run contrary to the notion that newcomers have fewer options to change identities because they experience greater cultural, linguistic, and structural barriers (Massey and Sánchez 2010; Nagel 1995; Waters 1999).

Results for Models 3 and 4 shows that identification changes are related to age, significant life events such as divorce and migration, and English language ability, although the effect of English ability is minimal when included with nativity in the same model. Model 3 affirms that those who migrated to a different state in the intercensal period have odds of changing identification about 60 percent greater than those who did not migrate (odds ratio of 1.59). Model 4 affirms our expectation that identification with Mexican origin changes relatively more often among children and young adults (odds ratios between 1.2 and 1.25) and that individuals who are divorced or separated also are more likely to change identification (odds ratio of 1.32). When including all variables in Model 5, the association between identification and age remains strong while the relationship with migration is somewhat moderated, perhaps because the young and native born also migrate relatively more often.

We now turn to models of identification change for other Hispanic groups. Given different histories, contexts of immigrant reception, and regional concentrations, we expect to find differences in how individual characteristics relate to identification change across groups. We investigate such differences by comparing the full Mexican origin model, Model 5 in Table 5, to models for other Hispanic groups in Table 6 (for space considerations, we exclude a model for Spanish origin given the relatively small size of this group).

One difference is in how householder/spouse status and age relate to identification change. Associations between identification and householder/spouse status and age are much stronger for Puerto Rican and Cuban origins than for other Hispanic groups. Minors have odds of changing identification with Puerto Rican or Cuban origins more than two and three times that of Mexican, Dominican, Central and South American, or Other Hispanic/Latino origins. While not as stark, a change in householder/spouse status also is more associated with changes in identification with Puerto Rican and Cuban origins.

Associations between ancestry and identification change also varies across groups. A discrepancy between one's ancestry and ethnic origin is of much greater consequence for Mexican, Puerto Rican, and Cuban origins than for Dominicans and Central and South Americans. Varying results for region is consistent with the importance of residential context for ethnic identification and reflect different regional concentrations. As we might expect, individuals residing outside the Southwest are *more* likely to change identification with Mexican origin but *less* likely to switch in relation to all other groups.

Table 6. Models of Change in Identification with Categories of Hispanic Origin, 2000 to 2010.

	Puerto Rican		Cuban		Dominican		Central American		South American		Other Hispanic	
	Estimate	Odds Ratio	Estimate	Odds Ratio	Estimate	Odds Ratio	Estimate	Odds Ratio	Estimate	Odds Ratio	Estimate	Odds Ratio
Intercept	-1.88 ***		-0.77 ***		-0.13		0.09 **		0.11 **		1.89 ***	0.00
Householder/Spouse Status (ref: HH/Spouse in 2000 and 2010)												
HH/Spouse in 2000 or 2010	0.38 ***	1.46	0.45 ***	1.57	0.24 ***	1.27	0.12 ***	1.13	0.21 ***	1.23	0.28 ***	1.33
Not HH/Spouse	0.24 ***	1.27	0.26 ***	1.30	0.15 ***	1.16	0.04	1.04	0.19 ***	1.21	0.05	1.05
Ancestry (ref: Modeled Group)												
Other Specific Hispanic Origin ¹	1.09 ***	2.98	0.92 ***	2.50	0.48 ***	1.61	0.54 ***	1.71	0.39 ***	1.48	0.54 ***	1.71
Non-Specific Other Hispanic Origin	0.41 ***	1.50	0.56 ***	1.75	0.19 ***	1.20	0.20 ***	1.23	0.24 ***	1.27	-	-
Non-Hispanic	1.37 ***	3.93	1.26 ***	3.52	0.54 ***	1.71	0.42 ***	1.52	0.53 ***	1.70	0.72 ***	2.04
Race in 2000 (ref: all other race categories)												
White alone	-0.28 ***	0.76	-0.50 ***	0.61	0.17 ***	1.19	-0.07 ***	0.94	-0.19 ***	0.83	-0.32 ***	0.72
Change in Race, 2000-2010 (ref: no change)	-0.39 ***	0.68	0.28 ***	1.32	-0.06 **	0.95	-0.01	0.99	-0.02	0.98	-0.35 ***	0.70
Nativity (ref: U.S.-born)												
Foreign Born, In U.S. 10+ Years	0.67 ***	1.96	-2.18 ***	0.11	-0.52 ***	0.59	-0.68 ***	0.51	-0.77 ***	0.47	1.08 ***	2.93
Foreign Born, In U.S. 5-9 Years	0.42 ***	1.52	-2.48 ***	0.08	-0.44 ***	0.64	-0.67 ***	0.51	-0.83 ***	0.44	1.00 ***	2.73
Foreign Born, In U.S. 0-4 Years	-0.04	0.96	-1.93 ***	0.15	-0.63 ***	0.53	-0.63 ***	0.53	-0.84 ***	0.43	0.65 ***	1.91
Inter-State Migration (ref: same state)	0.44 ***	1.56	0.62 ***	1.86	0.31 ***	1.37	0.39 ***	1.47	0.24 ***	1.27	0.30 ***	1.36
Region (ref: Southwest)												
Pacific Northwest	-0.09	0.92	0.02	1.02	-1.21 ***	0.30	-0.20 ***	0.82	-0.13 *	0.87	0.65 ***	1.92
Midwest	-0.75 ***	0.47	-0.20 ***	0.82	-0.58 ***	0.56	-0.44 ***	0.64	-0.35 ***	0.70	0.71 ***	2.04
Southeast	-0.74 ***	0.48	-0.34 ***	0.71	-0.68 ***	0.51	-0.22 ***	0.80	-0.17 ***	0.84	0.26 ***	1.29
Northeast	-1.02 ***	0.36	-0.16 ***	0.85	-0.94 ***	0.39	-0.37 ***	0.69	-0.27 ***	0.76	0.18 ***	1.20
Female	-0.02	0.99	-0.01	0.99	0.02	1.02	-0.03 *	0.97	-0.05 ***	0.95	-0.05 ***	0.95
Age (ref: 46+ years old)												
0 - 9	1.30 ***	3.68	0.89 ***	2.43	0.55 ***	1.74	0.45 ***	1.57	0.39 ***	1.47	0.06	1.06
10 - 17	1.18 ***	3.25	0.82 ***	2.27	0.36 ***	1.43	0.41 ***	1.51	0.27 ***	1.31	0.13 ***	1.14
18 - 25	0.89 ***	2.44	0.51 ***	1.67	0.06	1.06	0.10 ***	1.11	-0.02	0.98	0.19 ***	1.21
26 - 35	0.79 ***	2.21	0.30 ***	1.35	0.02	1.02	0.02	1.02	-0.17 ***	0.85	0.15 ***	1.16
36 - 45	0.56 ***	1.75	0.33 ***	1.39	0.07	1.07	0.03	1.03	-0.09 ***	0.91	0.09 ***	1.09
Marital Status (ref: Married)												
Widowed, Divorce, Separated	0.07 **	1.08	0.07	1.07	0.05	1.05	0.07 ***	1.08	0.04 *	1.05	-0.06 **	0.94
Never Married	0.07 **	1.07	-0.03	0.97	0.10 **	1.10	0.06 **	1.06	0.02	1.02	-0.10 ***	0.91
Education (ref: Less than High School diploma)												
High School Diploma	-0.01	0.99	0.17 ***	1.18	-0.07 *	0.93	0.03	1.03	0.08 ***	1.09	-0.01	0.99
Some College	-0.01	0.99	0.13 ***	1.14	-0.07 *	0.93	-0.11 ***	0.90	0.09 ***	1.09	0.08 ***	1.08
Bachelors Degree or Higher	-0.08 ***	0.92	-0.09 **	0.92	-0.28 ***	0.75	-0.20 ***	0.82	-0.06 **	0.95	0.21 ***	1.23
Speaks English Well or Very Well	-0.43 ***	0.65	-0.10 *	0.91	0.12 ***	1.13	0.19 ***	1.21	0.16 ***	1.17	0.05	1.06
Sample Size and Model Fit												
N	117,669		50,214		28,224		73,273		64,001		172,647	
-2 Log Likelihood	97,989		36,375		33,513		95,402		80,052		104,394	
Likelihood Ratio	15,960 ***		14,824 ***		2,284 ***		5,699 ***		5,618 ***		4,548 ***	

Sources: 2000 and 2010 Census, and unweighted 2006-2010 American Community Survey data. For more information on the ACS, see census.gov/acs

1 Includes specific national origins other than modeled group. Includes all specific groups for Other Hispanic model.

*** p<0.01; ** p<0.05; * p<0.1

Discussion and Conclusion

Our results offer greater understanding of patterns of change in Hispanic ethnic populations. Scholars and pundits like to make much of America's changing demographics. On one hand, some see the growing Hispanic population, and especially the Mexicans among them, as a threat to American national identity (Coulter 2015). On the other hand, others see the growth of these populations as the basis for a call for greater representation in multiple aspects of American life, especially in politics and government (Barreto and Segura 2014). Whatever differences in views about the implications of a growing Hispanic population, analysts implicitly assume that the trinity of demographic change – birth, death, migration – drives the growth of the Hispanic population. Among these three factors, migration, or, more precisely, immigration, plays a central role on contemporary debates about the implications of a growing Hispanic population. Our results show that immigration and natural growth are not the exclusive drivers of the rising numbers of Hispanics. Our findings show that such changes are influenced by the choices that individuals make about how to identify themselves on the census. Indeed, nearly four percent of the overall 2000-2010 growth in the Hispanic population was driven by individuals choosing to identify as

Hispanic themselves (or other in the household) in 2010, while choosing not to just a decade earlier. For Mexicans, individual choice was an even bigger factor in rising inter-census numbers: about 12 percent of the rise in the number of Mexicans between 2000 and 2010 owed to individuals choosing to identify as Mexican who did not do so in 2000. Those choices – whether to “join” or “leave” a group – ultimately shape what analysts can conclude from basic population profiles and population projections. Theoretically, our findings offer further evidence of the fluidity of ethnicity, and particularly Hispanicity, in the United States. But, unlike most studies, our results affirm a “two-way” fluidity that involves not just exit from Hispanic categories, as an assimilation might suggest; fluidity also involves joining a category (Liebler et al. 2014). To more fully flush out the theoretical implications of this two-way fluidity, future research should examine the contextual and demographic factors that determine whether an individual leaves or joins various Hispanic categories.

Works Cited

- Alba, Richard. 2016. “The Likely Persistence of a White Majority.” *The American Prospect*. Retrieved March 15, 2016 (<http://prospect.org/article/likely-persistence-white-majority-0>).
- Alba, Richard and Tariq Islam. 2009. “The Case of the Disappearing Mexican Americans: An Ethnic-Identity Mystery.” *Population Research and Policy Review* 28(2):109–21.
- Alba, Richard, Tomás R. Jiménez, and Helen B. Marrow. 2014. “Mexican Americans as a Paradigm for Contemporary Intra-Group Heterogeneity.” *Ethnic and Racial Studies* 37(3):446–66.
- Barreto, Matt and Gary Segura. 2014. *Latino America: How America’s Most Dynamic Population Is Poised to Transform the Politics of the Nation*. New York: PublicAffairs.
- Bean, Frank D., Susan K. Brown, and James D. Bachmeier. 2015. *Parents Without Papers: The Progress and Pitfalls of Mexican American Integration*. New York: Russell Sage Foundation.
- Bean, Frank D. and Gillian Stevens. 2003. *America’s Newcomers and the Dynamics of Diversity*. New York, NY: Russell Sage Foundation.
- Coulter, Ann. 2015. *Adios, America!: The Left’s Plan to Turn Our Country Into a Third World Hellhole*. Washington, DC: Regnery Publishing.
- Cresce, Arthur R., Audrey Dianne Schmidley and Roberto R. Ramirez. 2000. “Identification of Hispanic Ethnicity in Census 2000: Analysis of Data Quality for the Question on Hispanic Origin.” Population Division Working Paper No. 75. Washington D.C.: U.S. Census Bureau.
- Cresce, Arthur R. and Roberto R. Ramirez. 2003. “Analysis of General Hispanic Responses in Census 2000.” Population Division Working Paper No. 72. Washington, D.C.: U.S. Census Bureau.
- Dowling, Julie A. 2014. *Mexican Americans and the Question of Race*. Austin: University of Texas Press.
- Duncan, Brian and Stephen J. Trejo. 2011. “Intermarriage and the Intergenerational Transmission of Ethnic Identity and Human Capital for Mexican Americans.” *Journal of labor economics* 29(2):195–227.
- Emeka, Amon and Jody Agius Vallejo. 2011. “Non-Hispanics with Latin American Ancestry: Assimilation, Race, and Identity among Latin American Descendants in the US.” *Social Science Research* 40(6):1547–63.
- Ennis, Sharon R., Merarys Ríos-Vargas, and Nora G. Albert. 2011. “The Hispanic Population: 2010.” *Census Bureau Briefs*. Washington, D.C.: U.S. Census Bureau.
- Fellegi, Ivan P. and Alan B. Sunter. 1969. “A Theory for Record Linkage.” *Journal of the American Statistical Association* 64(328): 1183-1210.
- Fernandez, Leticia, Sonya R. Porter, Renuka Bhaskar, and Sharon R. Ennis. Forthcoming. “Factors that Influence Change in Hispanic Identification: Evidence from Linked Decennial Census and American Community Survey Data.” Unpublished manuscript.
- Foley, Neil. 2014. *Mexicans in the Making of America*. Cambridge, MA: Harvard University Press.
- Fox, Cybelle. 2012. *Three Worlds of Relief: Race, Immigration, and the American Welfare State from the Progressive Era to the New Deal*. Princeton University Press.
- Fox, Cybelle and Irene Bloemraad. 2015. “Beyond ‘White by Law’: Explaining the Gulf in Citizenship Acquisition between Mexican and European Immigrants, 1930.” *Social Forces* 94(1): 181-208.
- Frey, William H. 2014. *Diversity Explosion: How New Racial Demographics Are Remaking America*. Washington, DC: Brookings Institution Press.
- Gómez, Laura E. 2007. *Manifest Destinies: The Making of the Mexican American Race*. New York: New York University Press.

- Guglielmo, Thomas. 2003. *White on Arrival: Italians, Race, Color and Power in Chicago, 1890-1945*. New York, NY: Oxford University Press.
- Haney-López, Ian. 1996. *White by Law: The Legal Construction of Race*. New York: New York University Press.
- Harris, David R. and Jeremiah Joseph Sim. 2002. "Who Is Multiracial? Assessing the Complexity of Lived Race." *American Sociological Review* 67(4):614–27.
- Jiménez, Tomás R. 2004. "Multiethnic Mexican Americans and Ethnic Identity in the United States." *Ethnicities* 4(1):75–97.
- Jiménez, Tomás R. 2008. "Mexican-Immigrant Replenishment and the Continuing Significance of Ethnicity and Race." *American Journal of Sociology* 113(6):1527–67.
- Jiménez, Tomás R., Corey Fields, and Ariela Schachter. 2015. "How Ethnoraciality Matters: Looking Insider Ethnoracial Groups." *Social Currents* 2(2):107–15.
- Jiménez, Tomás R., Julie Park, and Juan Pedroza. 2015. "The New Third Generation: Post-1965 Immigration and the Next Stage in the Long March of Assimilation." in *Annual Meetings of the Population Association of America*. San Diego, CA.
- Krikorian, Mark. 2008. *The New Case Against Immigration: Both Legal and Illegal*. New York: Penguin.
- Lee, Jennifer and Frank D. Bean. 2010. *The Diversity Paradox: Immigration and the Color Line in Twenty-First Century America*. New York: Russell Sage Foundation Publications.
- Liebler, Carolyn A., Sonya Rastogi, Leticia E. Fernandez, James M. Noon, and Sharon R. Ennis. 2014. *America's Churning Races: Race and Ethnic Response Changes between Census 2000 and the 2010 Census*. United working paper #2014-09 States Bureau of the Census. Washington, DC.
- Lujan, Carol Chiago. 2014. "American Indians and Alaska Natives Count: The U.S. Census Bureau's Efforts to Enumerate the Native Population." *The American Indian Quarterly* 38(3):319-41.
- Macias, Thomas. 2006. *Mestizo in America: Generations of Mexican Ethnicity in the Suburban Southwest*. Tucson, : University of Arizona Press.
- Massey, Douglas S. and Magaly R. Sánchez. 2010. *Brokered Boundaries: Creating Immigrant Identity in Anti-Immigrant Times*. New York: Russell Sage Foundation Publications.
- Mora, G. Cristina. 2014. *Making Hispanics: How Activists, Bureaucrats, and Media Constructed a New American*. Chicago: University of Chicago Press.
- Mule, Thomas. 2012. "Census Coverage Measurement Estimation Report: Summary of Estimates of Coverage for Persons in the United States." Washington, D.C.: U.S. Census Bureau.
- Nagel, Joane. 1995. "American Indian Ethnic Renewal: Politics and the Resurgence of Identity." *American Sociological Review* 60(6):947–65.
- Passel, Jeffery S., D'Vera Cohn, and Mark Hugo Lopez. 2011. "Census 2010: 50 Million Latinos Hispanics Account for More Than Half of Nation's Growth in Past Decade." Washington, D.C.: Pew Hispanic Center.
- Oboler, Suzanne. 1995. *Ethnic Labels, Latino Lives: Identity and the Politics of (re)presentation in the United States*. Minneapolis: University of Minnesota Press.
- Okamoto, Dina and G. Cristina Mora. 2014. "Panethnicity." *Annual Review of Sociology* 40:219–39.
- Ramirez, Roberto R. 2009. "Changes to the American Community Survey Between 2007 and 2008 and Their Potential Effect on the Estimates of Hispanic Origin Type, Nativity, Race, and Language." Unpublished internal U.S. Census Bureau research note.
- Rodriguez, Clara E. 2000. *Changing Race: Latinos, the Census and the History of Ethnicity*. New York: New York University Press.
- Rodriguez, Gregory. 2007. *Mongrels, Bastards, Orphans and Vagabonds: Mexican Immigration and the Future of Race in America*. New York: Pantheon.
- Roediger, David R. 2005. *Working Toward Whiteness: How America's Immigrants Become White. The Strange Journey from Ellis Island to the Suburbs*. New York: Basic Books.
- Roth, Wendy. 2012. *Race Migrations: Latinos and the Cultural Transformation of Race*. Stanford, CA: Stanford University Press.
- Saperstein, Aliya. forthcoming. *Racial Mobility: How Selective Fluidity Reinforces Social Inequality*. New York: Russell Sage Foundation.
- Saperstein, Aliya and Andrew M. Penner. 2012. "Racial Fluidity and Inequality in the United States." *American Journal of Sociology* 118(3):676–727.
- Smith, Robert Courtney. 2014. "Black Mexicans and Beyond: Black Mexicans and Beyond: Long Term Ethnographic Analysis of Evolving Ethnic and Racial Identities in Contemporary America." *American*

- Sociological Review* 79(3):517–48.
- Telles, Edward E. and Vilma Ortiz. 2008. *Generations of Exclusion: Mexican Americans, Assimilation, and Race*. New York: Russell Sage Foundation.
- Vasquez, Jessica M. 2011. *Mexican Americans across Generations : Immigrant Families, Racial Realities*. New York: New York University Press.
- Wagner, Deborah, and Mary Layne. 2014. "The Person Identification Validation System (PVS): Applying the Center for Administrative Records Research and Applications' (CARRA) Record Linkage Software." *CARRA Working Paper #2014-01*. Washington, D.C.: U.S. Census Bureau.
- Wang, Wendy. 2012. *The Rise of Inter-marriage Rates, Characteristics Vary by Race and Gender*. Pew Research Center, Washington, DC.
- Waters, Mary C. 1990. *Ethnic Options: Choosing Identities in America*. Berkeley: University of California Press.
- Waters, Mary C. 1999. *Black Identities: West Indian Immigrant Dreams and American Realities*. New York and Cambridge, MA: Russell Sage Foundation; Harvard University Press.
- Waters, Mary C. 2002. "The Social Construction of Race and Ethnicity: Some Examples from Demography." Pp. 25–49 in *American Diversity: A Demographic Challenge for the Twenty-first Century*, edited by N. A. Denton and S. E. Tolnay. Albany, NY: State University of New York Press.