

Duncan, Brian; Grogger, Jeff; Leon, Ana Sofia; Trejo, Stephen

Working Paper

New Evidence of Generational Progress for Mexican Americans

IZA Discussion Papers, No. 12704

Provided in Cooperation with:

IZA – Institute of Labor Economics

Suggested Citation: Duncan, Brian; Grogger, Jeff; Leon, Ana Sofia; Trejo, Stephen (2019) : New Evidence of Generational Progress for Mexican Americans, IZA Discussion Papers, No. 12704, Institute of Labor Economics (IZA), Bonn

This Version is available at:

<https://hdl.handle.net/10419/207528>

Standard-Nutzungsbedingungen:

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

Terms of use:

Documents in EconStor may be saved and copied for your personal and scholarly purposes.

You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.

If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.

DISCUSSION PAPER SERIES

IZA DP No. 12704

**New Evidence of Generational Progress for
Mexican Americans**

Brian Duncan
Jeffrey Grogger
Ana Sofia Leon
Stephen J. Trejo

OCTOBER 2019

DISCUSSION PAPER SERIES

IZA DP No. 12704

New Evidence of Generational Progress for Mexican Americans

Brian Duncan

University of Colorado Denver

Stephen J. Trejo

University of Texas at Austin and IZA

Jeffrey Grogger

University of Chicago and IZA

Ana Sofia Leon

Universidad Diego Portales, Chile

OCTOBER 2019

Any opinions expressed in this paper are those of the author(s) and not those of IZA. Research published in this series may include views on policy, but IZA takes no institutional policy positions. The IZA research network is committed to the IZA Guiding Principles of Research Integrity.

The IZA Institute of Labor Economics is an independent economic research institute that conducts research in labor economics and offers evidence-based policy advice on labor market issues. Supported by the Deutsche Post Foundation, IZA runs the world's largest network of economists, whose research aims to provide answers to the global labor market challenges of our time. Our key objective is to build bridges between academic research, policymakers and society.

IZA Discussion Papers often represent preliminary work and are circulated to encourage discussion. Citation of such a paper should account for its provisional character. A revised version may be available directly from the author.

ISSN: 2365-9793

IZA – Institute of Labor Economics

Schaumburg-Lippe-Straße 5–9
53113 Bonn, Germany

Phone: +49-228-3894-0
Email: publications@iza.org

www.iza.org

ABSTRACT

New Evidence of Generational Progress for Mexican Americans*

U.S.-born Mexican Americans suffer a large schooling deficit relative to other Americans, and standard data sources suggest that this deficit does not shrink between the 2nd and later generations. Standard data sources lack information on grandparents' countries of birth, however, which creates potentially serious issues for tracking the progress of later-generation Mexican Americans. Exploiting unique NLSY97 data that address these measurement issues, we find substantial educational progress between the 2nd and 3rd generations for a recent cohort of Mexican Americans. Such progress is obscured when we instead mimic the limitations inherent in standard data sources. Similar patterns emerge for cognitive test scores and for annual earnings.

JEL Classification: J15, J61, J62

Keywords: Mexican Americans, generational progress, assimilation

Corresponding author:

Stephen J. Trejo
Department of Economics
University of Texas at Austin
2225 Speedway Stop C3100
Austin, TX 78712-1690
USA
E-mail: trejo@austin.utexas.edu

* We thank the Russell Sage Foundation for support (Award Number 93-16-03). For helpful comments, we thank conference and seminar participants at the ZEW Workshop on the Assimilation and Integration of Immigrants, the Population Association of America, the Western Economic Association, the Society of Labor Economists, the Southern Economic Association, ITAM University, Banco de Mexico, the University of Chicago, the University of Texas at Austin, the IZA Annual Migration Meeting, the Austin-Bergen Applied Microeconomics Conference, and the IZA conference honoring Daniel Hamermesh.

1. Introduction

Understanding the progress that takes place across immigrant generations is crucial for assessing the long-term impact of immigration on society. In many respects, 2nd-generation immigrants—the U.S.-born children of foreign-born migrants to the United States—show signs of rapid integration. On average, the 2nd generation as a whole and 2nd-generation members from most contemporary national origin groups meet or exceed the schooling level of the typical American (Duncan and Trejo 2018). Not conforming to this pattern, however, are Mexican Americans, who because of historical continuity and demographic size constitute arguably the most significant U.S. immigrant flow.

Panel A of Table 1 illustrates the seemingly persistent educational disadvantage of Mexican Americans. These calculations of average years of schooling for men use 2003-2016 data from the Current Population Survey (CPS).¹ Mexican Americans display impressive growth in educational attainment between the 1st generation (9.5 years) and the 2nd generation (12.7 years), but no further improvement is evident for the 3rd+ generation (the grandchildren and later descendants of Mexican immigrants). As a result, even when the comparisons are confined to individuals whose families have lived in the United States for at least a couple of generations (i.e., the 3rd+ generation), Mexican Americans possess schooling deficits of more than a year relative to non-Hispanic whites and almost one-third of a year relative to African Americans.

Considering the low levels of schooling, English proficiency, and other types of human capital brought to the United States by the typical Mexican immigrant, it is not surprising that their U.S.-born children do not eliminate all of these enormous socioeconomic deficits in a single generation (Perlmann 2005; Smith 2006). Of potentially greater concern, however, is the

¹ See the note to Table 1 and Duncan and Trejo (2018) for further details about the data and calculations. Patterns for women are very similar.

evidence that progress seems to stall after the 2nd generation for Mexican Americans.² Certainly, Mexican immigrants to the United States confront obstacles that might account for slowed or stalled progress among later generations (Portes and Rumbaut 2001), including discrimination (Telles and Ortiz 2008) and widespread undocumented status (Bean *et al.* 2011).

In addition, Huntington (2004) points to several other factors that could slow the pace of generational integration by Hispanics today as compared to Europeans in the past. These factors include the large scale of current immigration flows from Mexico and other Spanish-speaking countries, the substantial (though lessening) geographic concentration of these flows within the United States, and the fact that such flows have remained sizeable over a much longer period of time than did the influx from any particular European country. In addition, the close proximity of Mexico to the United States facilitates return and repeat migration. These unique features of Hispanic immigration might foster the growth of ethnic enclaves in the United States where immigrants and their descendants could, if they so choose, live and work without being forced to learn English or to Americanize in other important ways.³ Because of these and other concerns, Mexican Americans' prospects for future upward mobility are subject to much recent debate.⁴

In evaluating such theoretical arguments for slower integration by Mexican Americans, however, it is important to consider several potentially serious limitations of the existing

² Table 1 suggests educational stagnation beyond the 2nd generation for Mexican Americans, and the same pattern has been observed for earnings. Studies reporting limited progress in education and/or earnings after the 2nd generation for Mexican Americans include Trejo (1997, 2003), Fry and Lowell (2002), Farley and Alba (2002), Grogger and Trejo (2002), Livingston and Kahn (2002), Duncan, Hotz, and Trejo (2006), Blau and Kahn (2007), Telles and Ortiz (2008), Ortiz and Telles (2017), and Duncan and Trejo (2018).

³ Contrary to Huntington's thesis, however, available evidence suggests rapid linguistic assimilation for the U.S.-born descendants of contemporary immigrant groups (Alba *et al.* 2002). This holds even for Hispanics who live in areas with high concentrations of Spanish-speaking immigrants. In Southern California, for example, 96 percent of 3rd-generation Mexican Americans prefer to speak English rather than Spanish at home, and only 17 percent of 3rd-generation Mexican Americans retain the ability to speak fluent Spanish (Rumbaut, Massey, and Bean 2006).

⁴ See, for example, Perlmann (2005), Portes (2006), Telles and Ortiz (2008), Alba *et al.* (2011), Alba, Kasinitz, and Waters (2011), Haller, Portes, and Lynch (2011a, 2011b), Perlmann (2011), Alba, Jimenez, and Marrow (2014), Park, Myers, and Jimenez (2014), Bean, Brown, and Bachmeier (2015), and Ortiz and Telles (2017).

empirical evidence. We consider, in turn, the following three measurement issues: (1) comparisons of immigrant generations in cross-sectional data, (2) ethnic attrition, and (3) the inability to distinguish the 3rd generation from higher generations.

First, as noted by Borjas (1993, 2006) and Smith (2003, 2006), generational comparisons in a single cross-section of data—like those reported in Panel A of Table 1—can be misleading because they do a poor job of matching parents and grandparents in an earlier generation with their actual descendants in later generations. If we assume that schooling is essentially complete by the age of 25 and changes little thereafter, we can use CPS data to conduct an analysis of generational changes in educational attainment similar in spirit to Smith (2003). Panel B of Table 1 reports the relevant calculations. This panel presents average schooling levels for Mexican Americans similar to those displayed in Panel A, except that now separate calculations are reported for two particular age groups: 25-34 and 50-59. By choosing age groups 25 years apart, we create a situation in which the older age group from a particular generation potentially represents the parental cohort for the younger age group in the next generation. For example, the cohort of 1st-generation men aged 50-59 includes fathers of the 2nd-generation cohort of sons aged 25-34.

Panel B reveals only slightly more progress beyond the 2nd generation for Mexican Americans than did Panel A. When we compare age/generation groups that potentially match Mexican-American fathers with their sons (by moving northeast between the connected cells with similar shading in Panel B), average schooling rises from 12.6 years for the older 2nd generation to 12.7 years for the younger 3rd+ generation, a positive but small gain.⁵ Therefore, the evidence of educational stagnation for later-generation Mexican Americans does not seem to

⁵ Note, however, that calculating schooling progress between 1st- and 2nd-generation Mexican Americans in this same way produces even bigger gains than those observed in Panel A: 4.2 years in Panel B compared with 3.2 years in Panel A.

derive largely from biases associated with comparing immigrant generations in a cross-section. Further, note in Panel B that young 3rd+ generation Mexican Americans continue to trail the average schooling of their non-Hispanic white and African-American peers by the same substantial amounts observed in Panel A.

Potential measurement bias arising from “ethnic attrition” is a second issue that might make it difficult to track progress across immigrant generations. In Table 1, 1st- and 2nd- generation Mexican Americans are identified using the relatively “objective” information collected by the CPS on the countries of birth of the respondent and his parents (e.g., a 2nd- generation Mexican American is a U.S.-born individual with at least one parent born in Mexico). Virtually no large, nationally-representative data sets, however, provide information on the countries of birth of an adult respondent’s grandparents. As a result, 3rd-and-higher-generation Mexican Americans (or the so-called 3rd+ generation) must be assigned using more “subjective” measures of racial/ethnic identification. In Table 1, we follow standard practice in defining 3rd+ generation Mexican Americans as those who are U.S.-born, have two U.S.-born parents, and identify as “Mexican” or “Mexican American” in response to the Hispanic origin question. Given data limitations, researchers seeking to study later-generation Mexican Americans seldom have a better option. Nevertheless, the problem with using subjective measures of racial/ethnic identification is that assimilation and intermarriage can cause ethnic attachments to fade across generations (Alba 1990; Waters 1990; Perlmann and Waters 2007). Consequently, subjective measures of racial/ethnic identification might miss a significant portion of the later-generation descendants of immigrants. Furthermore, if such ethnic attrition is selective on socioeconomic attainment, then it can distort assessments of integration and generational progress. For Mexican Americans, Duncan and Trejo (2007, 2011, 2017) provide evidence that ethnic attrition is

substantial and could produce significant downward bias in standard measures of attainment for later generations. In this way, measurement biases generated by ethnic attrition could create a misleading appearance of socioeconomic stagnation after the 2nd generation for Mexican Americans, similar to what is observed in Table 1.

A third but related measurement issue is that the data limitations just described also imply that, for adults, researchers typically cannot distinguish the “true” 3rd generation from higher generations. For this reason, Table 1 and the discussion so far refer to the “3rd+” generation. This is potentially a problem because Mexican Americans in generations beyond the 3rd are disproportionately descended from ancestors who came of age in places (e.g., Texas rather than California) and times (e.g., before the Civil Rights era) where Mexican Americans faced discrimination that was more severe and often institutionalized (Foley 1997; Alba 2006; Montejano 1987). The more limited opportunities for advancement experienced by these families may result in lower attainment for Mexican Americans in the 4th and higher generations compared with their 3rd-generation counterparts whose families experienced less hostile environments. Alba *et al.* (2011) and Bean, Brown, and Bachmeier (2015) provide evidence of this pattern for schooling levels, highlighting the importance of distinguishing 3rd-generation Mexican Americans from higher generations. Kosack and Ward (2018) report much slower socioeconomic progress for Mexican Americans across immigrant generations before 1940 than afterward, and for the pre-1940 period they find that Mexican Americans are particularly disadvantaged in Texas compared with California or U.S. states that do not border Mexico.

In the current paper, we exploit previously untapped information from the National Longitudinal Survey of Youth 1997 (NLSY97) that allows us to address the last two measurement issues just discussed: ethnic attrition and distinguishing the 3rd generation from

higher generations. For our purposes, a key feature of the NLSY97 is that it reports the countries of birth of respondents' grandparents. This means that we can minimize ethnic attrition by identifying 3rd-generation Mexican Americans using ancestors' countries of birth rather than subjective ethnic identification. It also means that we can distinguish 3rd-generation Mexican Americans from higher generations.

We use these data to analyze educational progress between 2nd- and 3rd-generation Mexican Americans. Once we address the measurement issues just described, we find a substantial increase in educational attainment between these generations. Moreover, we show that such progress is largely hidden when we mimic standard data sets and aggregate the 3rd and higher generations into a "3rd+" generation. Similar patterns emerge for cognitive test scores and for annual earnings. Our analysis thus provides promising evidence of generational progress for a recent cohort of Mexican-Americans. Indeed, for this birth cohort, the high school graduation rate of 3rd-generation Mexican Americans is only slightly below that of non-Hispanic whites from the 4th and higher generations.⁶

Our paper relates most closely to two important recent studies of Mexican Americans that, through ambitious data collection efforts for specific locations, are also able to distinguish the 3rd generation from higher generations and, at least in part, account for ethnic attrition. Starting with a survey conducted in 1965 of Mexican-American families living in Los Angeles and San Antonio, Telles and Ortiz (2008) re-interview in 2000 available original respondents and their U.S.-born children. They find little evidence of educational or earnings progress beyond the 2nd generation. Bean, Brown, and Bachmeier (2015) rely on survey information collected from multiple generations of Mexican-origin individuals living in the greater Los Angeles

⁶ This finding is consistent with other recent evidence of improving high school completion rates for U.S.-educated Hispanics (Murnane 2013, Gramlich 2017).

metropolitan area in 2004. Their analysis does suggest significant schooling and earnings gains for Mexican Americans between the 2nd and 3rd generations.

Our analysis contributes in several important ways to the ongoing scholarly debate over Mexican-American progress after the 2nd generation. First, we employ nationally-representative data from the NLSY97. In this way, we avoid issues of selective geographic mobility that can make it difficult to interpret results from studies of particular locations (Alba, Jimenez, and Marrow 2014). Second, we are in a better position to assess and account for the effects of ethnic attrition, because roughly half of our Mexican-American respondents come from a sampling design that did *not* screen on race or ethnicity. In contrast, most of the original 1965 respondents in Telles and Ortiz (2008) and the Mexican-origin respondents in Bean, Brown, and Bachmeier (2015) had to subjectively identify as being of Mexican descent to be included in these surveys. Third, in addition to examining the educational and labor market outcomes that have been studied previously, we also analyze a commonly-used measure of cognitive ability. Finally, the recency and youth of our sample—described in greater detail below—imply that our analyses provide better information about the future trajectories of U.S.-born Mexican Americans than previous work could.⁷

2. Data

The NLSY97 provides longitudinal information for a nationally-representative sample of just fewer than 9,000 youth born in the years 1980-84 who were living in the United States when the survey began in 1997. Importantly for our purposes, there are two subsamples: a “cross-sectional sample” that is representative of all U.S. youth in the sampling universe at the time the

⁷ Our paper assesses patterns of generational progress for Mexican Americans in the contemporary period. Kosack and Ward (2018) take up similar questions for the period 1880-1940.

survey began, and a “supplemental sample” designed to oversample black and Hispanic youth. Roughly half of Mexican-origin respondents in the NLSY97 come from each of these subsamples. Note that, because Hispanic identification by the respondent (or by his parent) is used to determine inclusion in the supplemental sample but not the cross-sectional sample, the supplemental sample of Mexican Americans is subject to ethnic attrition.

We use the information available through round 17 of the NLSY97, which was conducted in 2015-16 when the respondents were between the ages of 30-36. The NLSY97 provides information on the countries of birth of the respondent, his biological parents, and his biological grandparents. Using this information, we define generations of Mexican Americans as follows:

1.5 generation: Respondent was born in Mexico and does not have a U.S.-born parent.⁸

2nd generation: Respondent was born in the United States but at least one of his parents was born in Mexico.

3rd generation: Respondent and both of his parents were born in the United States, but at least one of his grandparents was born in Mexico.

4th+ generation: Respondent, both parents, and all grandparents were born in the United States, but the respondent or one of his parents subjectively identifies as Mexican or Mexican American.

As interesting reference groups, we can also define 4th+ generation groups for non-Hispanic whites and non-Hispanic blacks. Based on these criteria, the NLSY97 data yield a sample of over 1,000 Mexican-origin respondents across the four generation categories, with sample sizes of 150 or more in each generation (see Table 2 below). These sample sizes are

⁸ Because foreign-born respondents in the NLSY97 must have been resident in the United States by the age of 12-16 to be included in the sample, we adopt the standard nomenclature of “1.5 generation” when referring to such immigrants who arrived in the destination country as children. We exclude from our sample foreign-born individuals who have any U.S.-born parents, because the definition of immigrant generation becomes murky for such individuals. There are relatively few such individuals, however, so instead retaining them in the sample has little effect on the results.

roughly similar to those employed by Telles and Ortiz (2008) and Bean, Brown, and Bachmeier (2015), but note that our samples are nationally representative, rather than coming from particular metropolitan areas. Substantially larger samples are available for the non-Hispanic white and black reference groups.

The appendix provides further details about the construction of our NLSY97 sample and how we assign individuals to immigrant generations. The appendix also discusses the representativeness of our sample and shows that key educational patterns are similar to those evident in comparable data from the CPS.

3. Generational Patterns of Educational Attainment

The primary aims of our analysis are to compare educational outcomes across generations of Mexican Americans and to make similar comparisons between later-generation Mexican Americans and the non-Hispanic white and black reference groups. We focus on education because it is a fundamental determinant of economic success, social status, health, family stability, and life opportunities (Hout 2012; Heckman, Humphries, and Veramendi 2017). In addition, information on educational attainment is available for all adults, whereas earnings data are available only for those currently working. When we can distinguish the 3rd generation from higher generations, and when we can limit the effects of ethnic attrition, do we see schooling gains for Mexican Americans between the 2nd and 3rd generations? If so, how much of a schooling gap remains between 3rd-generation Mexican Americans and other Americans?

The tabulations reported in Table 2 suggest that the answer to the first question is a resounding yes. Table 2 presents various measures of educational attainment—average years of schooling and the percent completing at least a high school degree, some college, or a bachelors

degree—for each of the Mexican-American generation groups and for the non-Hispanic white and black reference groups.⁹ Standard errors are shown in parentheses. All calculations reported in the paper employ sampling weights based on the initial sampling universe in 1997, but unweighted results show similar patterns.

For every schooling measure in Table 2, Mexican Americans exhibit steady improvement from the 1.5 to the 2nd to the 3rd generation. In most cases, this is followed by a marked decline from the 3rd to the 4th+ generation. For example, average years of schooling for Mexican Americans grow from 11.9 for the 1.5 generation to 13.0 for the 2nd generation to 13.5 for the 3rd generation, but average years of schooling then regress to 12.8 for the 4th+ generation. Similarly, the proportion of Mexican Americans with a high school diploma rises from 61.9 percent for the 1.5 generation to 77.0 percent for the 2nd generation to 84.3 percent for the 3rd generation before falling back to 68.1 percent for the 4th+ generation. The high school completion rate of 84.3 percent for 3rd-generation Mexican Americans approaches the 86.1 percent rate for 4th+ generation whites and exceeds by a considerable margin the 74.8 percent rate for 4th+ generation blacks.¹⁰ The only education measure that does not conform to this generational pattern is bachelor's degree completion, which increases slightly between the 3rd and 4th+ generations. For all education measures besides high school completion, however, large gaps ultimately remain

⁹ As described more fully in the appendix, information on educational attainment is updated each time a respondent is re-interviewed for another round of the survey. For each respondent, we construct the educational measures reported in Table 2 using the information available from the latest round of the NLSY97 that the respondent participated in. We exclude from our sample respondents who were last interviewed before they reached age 25, however, in order to focus on individuals who are likely to have completed their schooling. For the respondents in our sample, completed years of schooling ranges from a minimum of 2 to a maximum of 20. The sample sizes reported in Table 2 are for the completed years of schooling variable. Because there is less missing information regarding degree completion, the corresponding sample sizes are slightly larger for the binary measures of educational attainment.

¹⁰ In these tabulations, those with a GED (rather than a high school diploma) and no further education are counted as *not* having completed high school. If GED recipients are instead counted as high school completers, completion rates rise for all groups, but particularly so for Mexican Americans and blacks, such that the gap between 3rd-generation Mexican Americans and 4th+ generation whites almost entirely disappears (i.e., the revised rates are 94.6 percent for 3rd-generation Mexican Americans, 94.7 percent for 4th+ generation whites, and 91.4 percent for 4th+ generation blacks).

between 3rd-generation Mexican Americans and 4th+ -generation whites (i.e., deficits of 0.9 years for average schooling, 11.6 percentage points for college attendance, and 19.4 percentage points for bachelors degree completion).¹¹

In marked contrast to the CPS data in Table 1 and virtually all existing studies of Mexican-American educational progress, the NLSY97 data in Table 2 reveal substantial improvement after the 2nd generation. One crucial advantage of the NLSY97 data in Table 2 is the ability to distinguish 3rd-generation from higher-generation Mexican Americans. The final row of tabulations for Mexican Americans in Table 2 shows what happens when the 3rd and 4th+ generations are aggregated into the “3rd+ generation,” similar to what must be done in Table 1 due to limitations of CPS data. For all education measures other than bachelor’s degree completion, the NLSY97 data show little improvement after the 2nd generation and a larger remaining deficit relative to 4th+ -generation whites when 3rd- and 4th+ -generation Mexican Americans are aggregated in this way. Average years of schooling, for example, rise from 13.0 for the 2nd generation to 13.5 for the 3rd generation, whereas the corresponding increase is negligible (0.01 years) when the 3rd and higher generations are pooled together. Likewise, improvements in high school completion and college attendance between 2nd- and 3rd-generation Mexican Americans instead appear to be modest declines when the 3rd+ generation is used in place of the 3rd generation.

Table 2 reveals that, once important measurement issues are addressed, the educational attainment of 3rd-generation Mexican Americans is markedly higher and their educational deficit relative to non-Hispanic whites is considerably smaller than what would be inferred from

¹¹ Alon, Domina, and Tienda (2010) present evidence that the relatively low rates of post-secondary enrollment and degree attainment observed for U.S.-born Hispanics derive not just from having parents with lower rates of college attendance and completion, but also from those Hispanic parents who did attend college being less successful than other groups at getting their children to follow suit.

standard data sources. In this sense, socioeconomic integration is better than previously thought for the recent cohort of 3rd-generation Mexican Americans observed in the NLSY97. This is a key finding that we will explore further throughout this paper.

Also in contrast to standard data sources, Table 2 suggests that substantial educational progress takes place between 2nd- and 3rd-generation Mexican Americans, but this interpretation requires further explanation. Just to be clear, we seek to measure the educational improvement accruing to Mexican Americans as they progress across *immigrant* generations (e.g., from the 2nd-generation children of Mexican immigrants to the 3rd-generation grandchildren of Mexican immigrants). Later-generation descendants of immigrants grow up in families that have had more time to adapt to life in the United States and establish roots here. How does this process of assimilation across immigrant generations impact educational outcomes, especially for groups like Mexican Americans whose immigrant ancestors typically arrive in the United States with relatively low levels of schooling, English proficiency, and other forms of human capital? Note that we are *not* trying to measure the intergenerational progress in educational attainment that may occur between parents and their children, or the strength of the association between the schooling levels of parents and their children that is often used as an (inverse) measure of the amount of socioeconomic mobility in a society (Black and Devereux 2011). Indeed, as we show below, an important source of educational advance for 3rd-generation Mexican Americans is that their U.S.-born parents have substantially more schooling than do the Mexican-born parents of 2nd-generation Mexican Americans.

To gauge the generational progress of Mexican Americans in our NLSY97 sample, we compare levels of educational attainment across immigrant generations. Because this is a cross-sectional comparison at one point in time between Mexican Americans who are similar in age

but of different immigrant generations, the immigrant ancestors of each generation will have arrived in the United States in different time periods. Therefore, generational differences across Mexican Americans in NLSY97 data may not solely represent the impact of generational assimilation. Such differences could also reflect changes in the selectivity of Mexican immigration or other factors that are specific to an immigrant generation in cross-sectional data. Section A.5 of the appendix presents evidence from CPS data that educational differences between 2nd-generation and later-generation Mexican Americans have been stable for at least the past 20 years. This evidence suggests that the educational gains we observe in NLSY97 data between 2nd- and 3rd-generation Mexican Americans may in large part represent true generational assimilation.

Previous work used CPS data to distinguish 2nd-generation and 3rd-generation Mexican Americans (Duncan and Trejo 2007, 2011, 2017), but because of data limitations this could only be done for children living in intact families. Significantly, the NLSY97 data analyzed here allow us to observe respondents' outcomes in adulthood, and these data do not require that the sample be restricted to respondents who were raised in intact families. Moreover, with NLSY97 data we can investigate how restricting the sample to those raised in intact families affects the estimates. The NLSY97 provides information regarding whether each respondent was living with both biological parents at the initial interview in 1997 (when respondents were between the ages of 12-18) and also whether this was true at the following ages of the respondent: 2, 6, and 12. We define respondents as having been raised in an "intact" family if they were reported as living with both biological parents at *all* of these times.

For each of the Mexican-American generation groups and for the non-Hispanic white and black reference groups, Table 3 first reports the percent of respondents who were *not* raised in

intact families. By this measure, the chances that respondents experienced a single-parent household during childhood are lowest for whites (41 percent), increase across Mexican-American generations (from 43-44 percent for the 1.5 and 2nd generations to 51 percent for the 3rd generation and 63 percent for the 4th+ generation), and are highest for blacks (80 percent). The rest of the table shows how schooling patterns by race/ethnicity and immigrant generation differ for respondents who grew up in intact versus not intact families. For comparison purposes, the columns labeled as being for “all” family types simply reproduce the relevant estimates of average years of schooling from Table 2 and the implied schooling differences relative to 4th+generation whites.

Perhaps not surprisingly, for all groups average years of schooling are substantially higher for those who grew up in intact families compared with those who did not. Despite this difference in levels, however, the pattern across generations of Mexican Americans is similar for individuals from both types of families: sizeable schooling gains from the 1.5 to the 2nd to the 3rd generation followed by regress for the 4th+ generation. Moreover, this pattern is similar to what we observe in the sample that pools together individuals from both types of families. The similarity of the pattern for all individuals with that for the restricted sample of individuals from intact families suggests that the necessity of employing this sample restriction in previous analyses using CPS data may not have created large biases in the resulting estimates of schooling differences across Mexican generations.

4. Regression Analyses

In this section, we present regression analyses that illuminate and extend the basic results described so far. First, we delve further into the patterns of educational attainment discussed in

the previous section. We show that these patterns are robust to controlling for demographic and geographic variables, and we then explore the important role played by parental education.

Next, we demonstrate that similar patterns emerge for respondents' scores on the Armed Forces Qualification Test (AFQT). These scores provide a different and perhaps more discriminating measure of knowledge, aptitude, and cognitive ability than does educational attainment. Finally, we investigate whether the patterns by race/ethnicity and immigrant generation that we see for schooling and AFQT scores also show up for important labor market outcomes: annual weeks worked, hourly wages, and annual earnings.

4.1. Education

For the same samples and schooling measures introduced in Table 2, Table 4 presents least squares regressions describing how educational outcomes vary by race/ethnicity and immigrant generation. The dependent variables are the various measures of educational attainment, and the reported figures are estimated coefficients on dummy variables identifying groups defined by race/ethnicity and generation (with 4th+-generation non-Hispanic whites as the omitted reference group). Heteroskedasticity-robust standard errors are shown in parentheses. The sample sizes are 4,739 for regressions where the dependent variable is completed years of schooling and 4,780 for regressions where the dependent variables are the binary measures of educational attainment.

For each measure of educational attainment, specification (1) includes as independent variables only an intercept and the dummy variables identifying race/ethnicity and generation groups. These estimates simply reproduce, for comparison purposes, the unadjusted education differences implicit in Table 2. An advantage of regression analysis is that it allows us to

introduce control variables, the omission of which could potentially distort estimates of educational progress. Specification (2) redoes these educational comparisons while conditioning on each respondent's sex, age when last interviewed (which is the age at which the respondent's educational information was last updated), and state of birth. By comparing the estimates in specifications (1) and (2), we see that adding the control variables has little impact on the estimated coefficients and therefore on the implied schooling differences across race/ethnicity and generation groups. In particular, the striking pattern of generational gains in education for Mexican Americans through the 3rd generation followed by a substantial decline for the 4th+ generation is robust to the inclusion of the control variables, and even the magnitudes of these generational differences are altered only slightly by the controls.

Parental education is known to be one of the most important determinants of an individual's own educational attainment (Haveman and Wolfe 1994; Mulligan 1997). Parental education also is a prime candidate to help explain both the schooling deficits of Mexican Americans relative to non-Hispanic whites and the schooling differences observed for Mexican Americans across immigrant generations. In particular, schooling levels in Mexico are relatively low, so we should expect that the predominately Mexican-born parents of 1.5- and 2nd-generation Mexican Americans typically have much less education than do the U.S.-born parents of 3rd-and-higher-generation respondents.

This expectation is confirmed in Table 5, which reports average years of schooling for the mothers and fathers of the groups we that we study. These averages are calculated conditional on availability of the relevant information. It should be noted that NLSY97 data on parental education are missing with non-negligible frequency, particularly for fathers, as can be seen from the rightmost four columns of Table 5. Complete information on mother's and

father's schooling is available for nearly 90 percent of 3rd-generation Mexican Americans and 4th+ -generation non-Hispanic whites and for about 80 percent of 4th+ -generation Mexican Americans, but the corresponding rates fall to 69 percent for blacks and 72-74 percent for 1.5- and 2nd-generation Mexican Americans. Leaving this issue aside, average schooling levels of mothers and fathers are below 9 years for Mexican Americans with Mexican-born parents (the 1.5- and 2nd generations) and above 12 years for their counterparts with U.S.-born parents (the 3rd and higher generations). Finally, parental schooling is highest for later-generation non-Hispanic whites (with averages that typically exceed those of the other groups by at least a year).

Do differences in parental schooling account for some of the differences in educational attainment we see across groups? Table 6 presents regressions that shed light on this question. The dependent variable in all regressions is completed years of schooling. For comparison purposes, the regression reported in column (1) reproduces the corresponding estimates from specification (2) in Table 4 (which includes controls for the respondent's sex, age when last interviewed, and state of birth). In Table 6, the regression reported in column (2) starts with the column (1) specification and then adds as regressors the parental schooling variables, including indicators for missing information on parental schooling. The variables representing mother's and father's years of schooling are normalized to equal zero when the parent has 12 years of schooling. These variables are also set to zero when the relevant parental schooling information is missing. As a result, the coefficients of the dummy variables indicating missing parental schooling data represent differentials between individuals with the indicated missing data and otherwise similar individuals with reported parental schooling of 12 years.

The estimates in column (2) indicate that, as expected, mothers' and fathers' years of schooling have strong positive effects on the educational attainments of their adult children.

Comparing columns (1) and (2) reveals that conditioning on parental schooling substantially reduces remaining schooling gaps between other groups and non-Hispanic whites (the reference group). For 3rd-generation Mexican Americans and 4th+ -generation blacks, schooling deficits (relative to whites) shrink to less than a third of their initial levels, and for 4th+ -generation Mexican Americans the corresponding deficit is almost halved. Holding constant parental education, 1.5- and 2nd-generation Mexican Americans now exhibit schooling advantages over otherwise similar non-Hispanic whites. For the reasons discussed below, however, this latter result is misleading.

The regression specification in column (2) restricts the impact of parents' education on their child's schooling to be the same for all groups. This restriction is problematic in situations like ours where parental nativity differs across groups. In many contexts, associations between parent and child educational attainments have been shown to be much weaker for 2nd-generation individuals, whose immigrant parents typically were educated abroad, than for later-generation individuals whose parents were educated in the destination country (Gang and Zimmerman 2000; Nielsen *et al.* 2003; Dustmann 2008; Luthra and Soehl 2015). This issue is especially relevant for national origin groups, like Mexican Americans, where average educational levels in the immigrant source country are very different than those in the destination country (Feliciano 2005). Consider, for example, a respondent in our sample whose parents have ten years of schooling. This level of parental schooling is likely to mean something quite different for a 2nd-generation Mexican American whose parents were educated in Mexico than it does for a 4th+ -generation non-Hispanic white whose parents were raised in the United States. For the parent raised in the United States, ten years of schooling falls in the bottom tail of the educational distribution and may signal severe disadvantages in difficult to observe factors such as ability,

motivation, health, or family background. In contrast, this same level of schooling may not signal these things for the parent who immigrated to the United States from a country like Mexico where ten years of schooling exceeds the norm.

The regression specification reported in column (3) of Table 6 relaxes this restriction by introducing interaction terms that allow the coefficients on mother's and father's years of schooling to vary across groups.¹² Here, the main effects of the parental schooling variables represent the estimated impacts of these variables for the reference group of 4th-generation non-Hispanic whites. These main effects in column (3) are very similar to the corresponding coefficients in the restricted regression in column (2), which is not surprising given that non-Hispanic whites constitute more than half of the overall sample. In column (3), the coefficients on the interaction terms between parental schooling and the dummies for race/ethnicity and generation show how the impacts of parental schooling differ between each group and non-Hispanic whites. As expected, the estimated effects of parental schooling are much weaker (i.e., less positive) for 1.5- and 2nd-generation Mexicans, the groups with foreign-born parents.

In column (3), because of the normalization we employ for the parental schooling variables, the estimated main effects for race/ethnicity and generation represent the schooling differentials (relative to otherwise similar non-Hispanic whites) for individuals from each group whose mother and father each have 12 years of schooling. The regression specification in column (3) allows these differentials to vary with the level of parental schooling. Table 7 reports the relevant differentials when mothers and fathers are assumed to have identical years of schooling at each of the following levels: 10, 12, 14, and 16. Although the differentials vary with parental education, for most levels of parental schooling the implied deficits are

¹² For simplicity, this regression continues to impose the restriction that the coefficients on the indicators for missing parental schooling data are the same across groups. Results are similar when we also relax this restriction.

considerably smaller than the corresponding deficits that do not condition on parental education (see column (1) of Table 6). For Mexican Americans of all generations and for non-Hispanic blacks, the parental education gaps shown previously in Table 5 play an important role in generating their schooling disadvantage relative to non-Hispanic whites.

4.2. AFQT Scores

Respondents' scores on the Armed Forces Qualification Test (AFQT) provide another measure of their knowledge and abilities. During the initial round of the NLSY97, a set of military enlistment tests called the Armed Services Vocational Aptitude Battery (ASVAB) was administered to about 80 percent of NLSY97 respondents (U.S. Bureau of Labor Statistics 2019). Each respondent's score on the AFQT is created by combining his scores on the following four ASVAB tests: word knowledge, paragraph comprehension, arithmetic reasoning, and mathematics knowledge. AFQT scores are used by the U.S. military to evaluate the suitability of potential enlistees for overall service and for particular military occupations. Beginning with the NLSY79 and continuing with the NLSY97, AFQT scores have been widely used by researchers to measure cognitive ability and skill (Neal and Johnson 1996; Altonji and Pierret 2001; Cunha *et al.* 2006; Altonji, Bharadwaj, and Lange 2012; Castex and Kogan Dechter 2014).

Table 8 presents estimates from AFQT regressions similar to the education regressions shown previously in Table 4. The dependent variable in all regressions is the respondent's percentile score on the AFQT relative to others within the same three-month age group. The regression reported in column (1) controls for the respondent's sex and his age when taking the AFQT, and the column (2) regression also controls for his state of residence at the time of testing. Columns (3) and (4) present results from analogous specifications where, for

comparison purposes, 3rd-generation and 4th+ generation Mexican Americans have been aggregated into a single category denoted as the “3rd+ generation.”

AFQT scores display the same pattern across Mexican-American generations that we saw earlier for educational attainment. In column (1), the score deficit relative to the non-Hispanic white reference group shrinks from 31.8 percentile points for the 1.5 generation to 25.0 points for the 2nd generation to 17.7 points for the 3rd generation before rising back to 25.6 points for the 4th+ generation. Controlling for state of residence in column (2) yields a very similar pattern. In columns (3) and (4), failing to distinguish 3rd-generation Mexican Americans from higher generations reproduces the familiar but misleading impression of minimal gains after the 2nd generation for Mexican Americans. Consequently, the finding of hidden progress for 3rd-generation Mexican Americans demonstrated in previous sections with respect to educational attainment also holds for AFQT scores, a different and perhaps more precise and sensitive measure of attainment.

4.3. Labor Market Outcomes

We next undertake a similar type of analysis for several key labor market outcomes: annual weeks worked, hourly wages, and annual earnings. Here, we exploit the longitudinal structure of the NLSY97 and include repeated observations on outcomes for a given respondent when they are available. In order to focus on outcomes that better represent the longer-term or more “permanent” attainment of respondents, however, we exclude observations from when a respondent is younger than age 30 or is enrolled in school. For these regressions, each respondent contributes on average about two observations to the estimation sample. Standard errors are clustered at the level of the respondent.

Table 9 presents results for annual weeks worked, a measure of labor supply and attachment to the work force. The dependent variable in all regressions is the number of weeks worked during the calendar year preceding the interview date. Those with zero weeks worked are included in the estimation sample, so this measure reflects, at least in part, both the extensive and intensive margins of labor supply. The regression reported in column (1) includes as independent variables only an intercept and the dummy variables identifying race/ethnicity and generation groups, and therefore these estimates represent mean differences between each group and the reference group of non-Hispanic whites. The column (2) regression adds controls for the respondent's sex, age, and state of residence, as well as indicators for the calendar year of the observation, and the column (3) regression also conditions on the respondent's years of schooling. Columns (4), (5), and (6) present results from analogous regressions where 3rd-generation and 4th+ -generation Mexican Americans have been aggregated into a single category.

With regard to the unconditional averages, column (1) shows that annual weeks worked are lowest for 1.5- and 4th+ -generation Mexican Americans and for non-Hispanic blacks (with deficits, respectively, of 2.8, 5.7, and 4.6 weeks relative to non-Hispanic whites). In contrast, 2nd- and 3rd-generation Mexican Americans work a similar amount of weeks per year as non-Hispanic whites. Adding the demographic, geographic, and temporal control variables in column (2) shrinks the deficit for 1.5-generation Mexican Americans to 1.1 weeks but has only minor effects on the other differentials. Further conditioning on years of schooling in column (3) substantially improves for all groups the remaining differential relative to non-Hispanic whites.

In summary, the pattern of differences across Mexican-American generations for annual weeks worked does not resemble the pattern discussed previously for educational attainment and AFQT scores. What does look similar, however, is the distortion created by aggregating 3rd- and

4th+ generation Mexican Americans. Columns (4), (5), and (6) reveal that this aggregation hides the relatively strong labor force attachment of 3rd-generation Mexican Americans and creates a false impression of sharply-declining weeks worked for Mexican Americans after the 2nd generation.

Table 10 presents the corresponding results for hourly wages. The dependent variable in these regressions is the natural logarithm of the hourly wage earned at the respondent's main job as of the interview date.¹³ Once control variables are included in column (2), Mexican Americans display steady wage progress across immigrant generations. Measured in log points, hourly wage deficits relative to non-Hispanic whites fall from .47 for the 1.5 generation to .30 for the 2nd generation to .26 for the 3rd generation to .19 for the 4th+ generation. For comparison purposes, the analogous deficit for blacks is .28 log points. When schooling is held constant in column (3), all of these deficits shrink, as do the wage differences between 1.5-, 2nd-, and 3rd-generation Mexican Americans. These results indicate that substantial portions of the wage differences observed in column (2), including the wage gains between 1.5-, 2nd-, and 3rd-generation Mexican Americans, are attributable to the corresponding differences in educational attainment reported previously in Table 2. For example, the estimates in columns (2) and (3) of Table 10 suggest that the raw wage gain of about .04 log points between 2nd- and 3rd-generation Mexican Americans can be explained by the 3rd-generation's schooling advantage of half a year.¹⁴

¹³ For respondents not employed at the time of interview, the hourly wage is from the most recent job (lasting at least 13 weeks) held since the previous interview.

¹⁴ Table 10 implies this result in two ways. First, upon conditioning on schooling, the relevant wage difference transforms from a .04 log point advantage for the 3rd generation in column (2) to a slight (.01 log point) *dis*advantage in column (3). Second, the column (3) estimate of a .08 log point wage return for each year of education indicates that a half-year schooling advantage could by itself produce a wage advantage of .04 log points.

Finally, Table 11 shows comparable regression results for annual earnings. The dependent variable is the natural logarithm of earnings during the calendar year preceding the interview date. As such, variation in annual earnings reflects variation in both hourly wages and in annual hours of work. The pattern of annual earnings across Mexican-American generations resembles what we have seen for educational attainment and for AFQT scores: substantial gains from the 1.5 to the 2nd to the 3rd generation, followed by a decline from the 3rd to the 4th+ generation. In column (2), 3rd-generation Mexican Americans exhibit an earnings advantage of .10 log points relative to the 2nd generation and .04 log points relative to the 4th+ generation. Once we control for years of schooling in column (3), however, earnings differences across these generations become quite small. Evidently, the earnings advantage of 3rd-generation Mexican Americans derives primarily from their educational advantage. In column (3), also note that U.S.-born Mexican Americans (i.e., those in the 2nd generation and beyond) earn, on average, about 10 percent less than comparable non-Hispanic whites, whereas the corresponding earnings deficit for blacks exceeds 25 percent. This finding echoes other research indicating that, after controlling for observable human capital, earnings deficits (relative to non-Hispanic whites) are much smaller for Mexican Americans than for African Americans (Trejo 1997; Grogger and Trejo 2002; Duncan, Hotz, and Trejo 2006).

5. Ethnic Attrition

Biases from selective ethnic attrition are likely to be more severe in our NLSY97 sample of 4th+ generation Mexican Americans than in the corresponding sample of 3rd-generation Mexican Americans. One reason is that the 3rd generation can be identified objectively (from information on the countries of birth of the respondent, his parents, and his grandparents),

whereas inclusion in the 4th+ generation sample requires that the respondent or a parent subjectively identifies as being of Mexican descent. In addition, ethnic attachments tend to fade with more generations since immigration, and this tendency produces more extensive ethnic attrition in higher generations. Consequently, greater downward bias from ethnic attrition is one potential explanation for the relatively poor outcomes observed for 4th+ generation Mexican Americans.

Previous work by Duncan and Trejo (2007, 2011, 2017) establishes the direction and potential importance of the biases created by selective ethnic attrition, but that work suffers from significant limitations. Of particular concern is that the 3rd-generation samples in these earlier studies are confined to children living in intact families, and therefore the earlier findings can only indirectly suggest the extent and selectivity of ethnic attrition that would be observed among adults. Because the cross-sectional sample of the NLSY97 allows us to construct a sample of 3rd-generation Mexican-American adults that is free from ethnic attrition, these data offer some key advantages for further exploration of this issue. One notable disadvantage of using NLSY97 data for this purpose, however, is the small sample size: the cross-sectional sample includes only 79 3rd-generation Mexican Americans (see Table 12 below).

Table 12 reports the percentage of Mexican Americans from each generation who identify subjectively as being of Hispanic origin, based on information collected at the beginning of the survey in 1997.¹⁵ The top panel of the table shows the relevant calculations for the cross-sectional sample that is representative of all U.S. youth in the sampling universe at the time the

¹⁵ The NLSY97 also collected information about Hispanic identification in 2002 and at other times. These alternative measures of Hispanic identification display the same patterns as the 1997 measure reported in Table 12. We employ here a broad indicator of “Hispanic” identification rather than a more specific indicator for “Mexican” identification so that the resulting estimates of ethnic attrition are conservative. In addition to capturing individuals who identify as Mexican or Mexican American, Hispanic identification also captures some individuals who would not identify specifically as Mexican-origin, including those who identify with other Hispanic national origin groups (such as Puerto Rican or Cuban) as well those who identify with pan-ethnic labels such as Hispanic or Latino.

NLSY97 began. The middle panel repeats these calculations for the supplemental oversample of Hispanics, and the bottom panel does this for the combined sample that pools together observations from both the cross-sectional and supplemental samples.¹⁶

The middle panel of Table 12 reveals perfect Hispanic identification rates for every generation of Mexican Americans in the supplemental sample. This result confirms that the selection criteria for inclusion in the supplemental sample have effectively excluded from this subsample any Mexican Americans who do not identify as Hispanic. Because the supplemental sample does not provide useful information about ethnic attrition, we instead focus our attention on the cross-sectional sample in the top panel of Table 12.

The rates of Hispanic identification reported in the top panel of Table 12 indicate that ethnic attrition is negligible for the 1.5 and 2nd generations of Mexican Americans in the NLSY97, but it does become a significant issue by the 3rd generation. Everyone born in Mexico (i.e., the 1.5 generation) identifies as Hispanic, as do 95 percent of U.S.-born individuals with a parent born in Mexico (i.e., the 2nd generation). Among objectively-defined 3rd-generation Mexican Americans, however, only 79 percent identify as Hispanic, implying an ethnic attrition rate of 21 percent. This pattern of ethnic attrition across generations of Mexican Americans is roughly similar to what Duncan and Trejo (2016) report in recent CPS data.¹⁷

For ethnic attrition to bias estimates of socioeconomic progress, not only must it exist, but it must also be selective. Table 13, which restricts attention to 3rd-generation Mexican Americans in the cross-sectional sample of the NLSY97, provides some evidence that this is indeed the case. Among such individuals, those who do *not* identify as Hispanic average almost

¹⁶ Sample sizes in the bottom panel of Table 12 are slightly larger than the corresponding sample sizes listed in Table 2. This reflects the fact that missing data occur less frequently for Hispanic identification than for educational attainment.

¹⁷ In particular, see Appendix Tables A.1-A.3 in Duncan and Trejo (2016).

two-thirds of a year more schooling than those who do so identify. Similarly, the rate of bachelor's degree completion is higher for those not identifying as Hispanic (29 percent) than for those who do identify (23 percent). Rates of high school graduation and college attendance do not conform to this pattern, however, with slight advantages observed for 3rd-generation Mexican Americans who identify as Hispanic. Although suggestive, these estimates are imprecise because of the small samples involved (e.g., the calculations for 3rd-generation Mexican Americans who do not identify as Hispanic are based on a sample size of 11). Nonetheless, the educational selectivity of ethnic attrition among 3rd-generation Mexican Americans that we directly observe for *adults* in NLSY97 data conforms to what prior studies inferred indirectly from 3rd-generation samples that were limited to *children* (Duncan and Trejo 2007, 2011, 2017).

Table 13 and previous research suggest that selective ethnic attrition generates downward-biased estimates of socioeconomic attainment for later generations of Mexican Americans when, as is typically the case, target sample members can only be detected using subjective measures of ethnic identification. The analyses of educational attainment and other outcomes presented in previous sections largely avoid this problem by using data that can identify 1.5-, 2nd-, and 3rd-generation Mexican Americans without relying on subjective measures of ethnic identification. To preserve sample size, previous sections reported results based on the full NLSY97 sample that pooled together observations from the cross-sectional and supplemental samples. Because the supplemental sample filters out Mexican Americans who do not identify as Hispanic, the resulting ethnic attrition may generate downward-biased measures of educational attainment and other outcomes for Mexican Americans, particularly in the 3rd generation and beyond where ethnic attrition becomes non-negligible. As it turns out, however, very similar patterns of generational differences emerge when analyses for Mexican Americans

through the first three generations are reproduced using only the cross-sectional sample that does not suffer from ethnic attrition. In light of this, we will continue to report results based on the full NLSY97 sample, where appropriate, in order to increase sample sizes and improve precision.

Given that selective ethnic attrition helps explain the apparent lack of generational progress reported elsewhere, it is interesting to consider the source of selective ethnic attrition. Previous research (Duncan and Trejo 2011, 2017) indicates that the selectivity of ethnic attrition observed for Mexican Americans—i.e., the strong negative relationship between ethnic identification and socioeconomic attainment—largely reflects patterns associated with intermarriage. Mexican Americans with mixed ethnic origins are less likely to identify as Mexican or Hispanic and also display higher levels of average attainment.

Table 14 suggests that something similar occurs within our sample of 3rd-generation Mexican Americans from the NLSY97. In particular, average years of schooling are higher for 3rd-generation individuals with weaker ancestral attachments to Mexico. In the top part of Table 14, 3rd-generation Mexican Americans are distinguished by how many of their grandparents were born in Mexico. The vast majority of 3rd-generation Mexican Americans (88 percent) have only one or two grandparents born in Mexico, and such individuals average about 1.8 more years of schooling than their counterparts with stronger ethnic attachments (i.e., those with three or four Mexican-born grandparents).

The bottom part of Table 14 instead distinguishes 3rd-generation Mexican Americans according to whether their Mexican ancestry is observed on their father's side only, on their mother's side only, or on both sides of their family. In this typology, Mexican ancestry is said to be observed on the father's side of the family when at least one of the following two things is true: (1) the respondent has a paternal grandparent who was born in Mexico, or (2) the

respondent's father subjectively identifies as Mexican American. Analogously, presence of a Mexican-born maternal grandparent and/or the mother's subjective identification as Mexican American determine whether a respondent is observed to have Mexican ancestry on his mother's side of the family.¹⁸ The distribution of 3rd-generation Mexican Americans is almost evenly distributed across the three groups defined in the bottom part of Table 14. Average years of schooling are markedly higher for those with Mexican ancestry on just one side of their family (13.6 for father's side only and 14.0 for mother's side only) than for those with Mexican ancestry on both sides of their family (13.0). Once again, a substantial educational advantage is observed for those with seemingly weaker ancestral attachments to Mexico. Given previous research (Duncan and Trejo 2011, 2017) documenting that later-generation Mexican Americans with weaker ancestral attachments are much less likely to identify as Mexican or Hispanic, the schooling patterns in Table 14 are consistent with other work indicating that ethnic attrition among Mexican Americans is positively selected on socioeconomic attainment.

6. Why Is Schooling Lower for the 4th+ Generation?

In the preceding analyses, a striking and somewhat surprising finding is the low level of attainment for 4th+ generation Mexican Americans. For example, in terms of average years of schooling, high school completion, and college attendance, Mexican Americans in the 4th+ generation exhibit large deficits relative to the 3rd generation and smaller but still sizeable gaps relative to the 2nd generation. As noted in the previous section, one possible explanation for this pattern is that ethnic attrition generates greater downward bias for the 4th+ generation sample. In

¹⁸ Here, we make use of information on the subjective Mexican identification of the respondent's parents in order to construct a broader definition of Mexican ancestry. If we instead adopt a narrower but more objective definition of Mexican ancestry that is based solely on the presence of Mexican-born grandparents, similar schooling patterns emerge, but the fraction of 3rd-generation Mexican Americans defined to have Mexican ancestry on both sides of their family is cut in half (to 16 percent).

this section, we consider alternative explanations that focus on disadvantaged family and social environments for 4th+ generation Mexican Americans.

Table 15 shows how the samples of 3rd-generation and 4th+ generation Mexican Americans compare with respect to their geographic “roots.” An individual is defined as having California roots when at least one of the following three things is observed: (1) the respondent was born in California; or (2) the respondent resided in California in 1997 when the survey began; or (3) either (or both) of the respondent’s parents was born in California. An analogous procedure determines whether an individual is defined as having Texas roots.

Mexican Americans in the 3rd generation are equally distributed among those with any California roots, those with any Texas roots, and those with neither California nor Texas roots.¹⁹ Compared with the 3rd generation, the 4th+ generation displays a dramatic decline in the proportion with any California roots (from 36 percent to 10 percent) and a corresponding increase in the proportion with neither California nor Texas roots (from 36 percent to 60 percent), with relatively little change in the proportion with any Texas roots. These differences in the geographic roots of the 3rd generation versus the 4th+ generation are potentially important because Mexican Americans arguably faced less severe discrimination and enjoyed better opportunities for advancement in California than in Texas or other parts of the United States, especially prior to the civil rights reforms of the 1960s and 1970s (Foley 1997; Alba 2006; Montejano 1987).

Table 5, introduced previously, suggests that 4th+ generation Mexican Americans are also somewhat disadvantaged relative to the 3rd generation when it comes to parental education

¹⁹ Given the way we define geographic roots, it is possible for someone to have *both* California and Texas roots. In our sample of 3rd-generation Mexican Americans, 8 percent have both California and Texas roots, 27 percent have California but not Texas roots, and 28 percent have Texas but not California roots. Among the 4th+ generation, however, the proportion with overlapping California and Texas roots falls to just 1 percent of the sample.

levels. In particular, mothers' average years of schooling are 0.4 higher for the 3rd generation compared with the 4th+ generation, whereas the corresponding difference is smaller for fathers' average years of schooling (an advantage of 0.2 for the 3rd generation relative to the 4th+ generation). In addition, 4th+ generation Mexican Americans are less likely to have been raised in an intact family with both biological parents. Table 3, also introduced previously, shows that less than 38 percent of the 4th+ generation grew up in intact families, compared with almost half of the 3rd generation.

Do these deficits in California roots, parental schooling, and family cohesion for 4th+ generation Mexican Americans relative to the 3rd generation help to account for the lower educational attainment of the 4th+ generation? Table 16 presents least squares regressions which suggest that these observable differences between 3rd-generation and 4th+ generation Mexican-Americans do not explain the lion's share of the corresponding education gap. The dependent variable is completed years of schooling, and the estimation sample includes 3rd-generation and 4th+ generation Mexican Americans. All specifications include a dummy variable identifying membership in the 4th+ generation (as opposed to membership in the reference group consisting of the 3rd generation) and indicators for the respondent's sex and age when last interviewed. The regression reported in column (1) includes only these variables, and the estimated coefficient on the 4th+ generation dummy reproduces the average schooling deficit for 4th+ generation Mexican Americans (relative to the 3rd generation) of more than three-quarters of a year that was observed previously in Tables 2 and 4.

In column (2), we add to the column (1) regression an indicator for whether the respondent has any California roots, as defined previously. All else equal, later-generation Mexican Americans with California roots average three-fifths of a year more schooling than

those without California roots, and adding this variable to the regression shrinks (in absolute value) the estimated coefficient of the 4th+ generation dummy from -.83 to -.65.²⁰ Similarly, column (3) instead adds to the column (1) regression the vector of parental schooling variables (including indicators for missing information on parental schooling). Controlling for parental education changes the estimated coefficient of the 4th+ generation dummy to -.62. Column (4) includes all of these variables—the California roots dummy and the parental schooling variables—in the same regression, and the estimated coefficient of the 4th+ generation dummy becomes -.57. Finally, column (5) adds to the column (4) regression an indicator for having not grown up in an intact family, and the estimated coefficient of the 4th+ generation dummy further shrinks to -.52.

In summary, controlling for differences in geographic roots, parental education, and family structure shrinks (by up to 37 percent) but does not eliminate the substantial schooling deficit for 4th+ generation Mexican Americans relative to their 3rd-generation counterparts. For reasons discussed previously, greater downward bias from ethnic attrition in the 4th+ generation sample is another potential explanation for this schooling deficit. Indeed, more severe ethnic attrition in the 4th+ generation is a potential explanation for the differences in geographic roots, parental schooling, and family structure that we observe between 3rd-generation and 4th+ generation Mexican Americans, because previous research suggests that ethnic attrition is more prevalent among later-generation Mexican Americans with relatively advantaged family backgrounds (Duncan and Trejo 2011, 2017). As a result, although the regressions reported in Table 16 indicate that geographic roots, parental schooling, and family structure provide proximate explanations for a portion of the schooling deficit observed for 4th+ generation

²⁰ The impact on the estimated coefficient of the 4th+ generation dummy is very similar when we replace the single indicator variable for California roots with a vector of indicators distinguishing the more detailed categories of geographic roots listed in Panel A of Table 15.

Mexican Americans, it is still possible that ethnic attrition generates these proximate correlations as well as the portion of the schooling deficit not accounted for by these variables. Moreover, to the extent that ethnic attrition ultimately accounts for the schooling deficit observed for 4th+generation Mexican Americans, then this deficit is illusory rather than real. An important goal for future research, therefore, should be to better understand the role that ethnic attrition plays in generating observed schooling differences across generations of Mexican Americans.

7. Conclusion

In contrast with the descendants of almost every other contemporary immigrant group, U.S.-born Mexican Americans maintain a large schooling deficit relative to other Americans. Moreover, standard data sources suggest that this deficit does not shrink between the 2nd and later generations of Mexican Americans.²¹ The apparent generational stagnation of educational attainment for Mexican Americans raises concerns about this important group's prospects for long-term integration into American society.

Available evidence on this issue, however, suffers from some potentially serious limitations. A major problem is that data sources rarely provide information on the countries of birth of an adult respondent's grandparents. As a result, Mexican Americans beyond the 2nd generation almost always must be identified from subjective measures of ethnic identification, such as the Hispanic origin question asked in U.S. Census Bureau surveys. These data limitations create two key measurement issues for tracking the generational progress of Mexican Americans: (1) ethnic attrition, and (2) aggregation of 3rd-generation and higher-generation

²¹ See the studies cited in footnote 2.

individuals. Both of these measurement issues could lead standard analyses to understate socioeconomic improvement between 2nd-generation and 3rd-generation Mexican Americans.

Ethnic attrition takes place when U.S.-born descendants of Mexican immigrants do not subjectively identify as Mexican American or Hispanic. Previous research indicates that ethnic attrition is substantial among later-generation Mexican Americans and that such attrition typically arises in families with mixed ethnic origins (Duncan and Trejo 2007, 2011, 2017). Moreover, this research suggests that selective intermarriage and the resulting ethnic attrition produce downward bias in estimates of socioeconomic attainment that rely on subjective measures of ethnic identification to detect later-generation Mexican Americans.

The lack of information on grandparents' countries of birth also implies that analysts cannot distinguish 3rd-generation from higher-generation Mexican Americans. Instead, the only group beyond the 2nd generation available for study is an aggregated "3rd+" generation that pools together individuals from the 3rd and all higher generations. Such aggregation could hide progress for the disaggregated 3rd generation, because Mexican Americans beyond the 3rd generation may have experienced harsher family and social environments, and also because the biases from ethnic attrition are likely to be more severe for higher generations.

In this paper, we are able to address both of these measurement issues by exploiting previously untapped information from the NLSY97 on the countries of birth of respondents' grandparents. With these data, we can identify 3rd-generation Mexican Americans using ancestors' countries of birth rather than subjective ethnic identification, thereby minimizing ethnic attrition and isolating the 3rd generation from higher generations. To our knowledge, we are the first to address these measurement issues using nationally-representative data. In addition, compared to two related studies that focus on particular metropolitan areas (Telles and

Ortiz 2008; Bean, Brown, and Bachmeier 2015), our analysis is better able to account for ethnic attrition. Finally, we study a more recent cohort of Mexican Americans than others have studied, and therefore our findings provide timely insights into future trends.

Using NLSY97 data that allow us to minimize ethnic attrition and distinguish the 3rd-generation from higher generations, we find substantial educational progress between 2nd- and 3rd-generation Mexican Americans. Such progress is hidden when we instead mimic standard data sets and aggregate the 3rd and higher generations into a 3rd+ generation. Similar patterns emerge for cognitive test scores and for annual earnings. For a recent cohort of Mexican Americans, our analysis thus provides promising evidence of generational advance. In particular, for this cohort of individuals born in the years 1980-84, the high school graduation rate of 3rd-generation Mexican Americans is only slightly below that of later-generation non-Hispanic whites. Other measures of educational attainment—completed years of schooling, college attendance, and bachelors degree completion—also show sizable gains for Mexican Americans between the 2nd and 3rd generations. In contrast with high school completion, however, for these other education measures 3rd-generation Mexican Americans maintain large deficits relative to non-Hispanic whites, despite their generational gains. Further analyses document patterns of ethnic attrition among 3rd-generation Mexican-American adults in the NLSY97 that are similar to those reported previously in CPS data where the 3rd-generation samples are confined to children. Ultimately, our findings suggest that Mexican Americans do indeed experience substantial socioeconomic progress beyond the 2nd generation, and that this progress is obscured by limitations of the data sources commonly used to look for it.

References

- Alba, Richard D. *Ethnic Identity: The Transformation of White America*. New Haven, CT: Yale University Press, 1990.
- Alba, Richard D. "Mexican Americans and the American Dream." *Perspectives on Politics*, June 2006, 4(2), pp. 289-96.
- Alba, Richard D.; Abdel-Hady, Dalia; Islam, Tariqul; and Marotz, Karen. "Downward Assimilation and Mexican Americans: An Examination of Intergenerational Advance and Stagnation in Educational Attainment," in Richard Alba and Mary C. Waters, eds., *The Next Generation: Immigrant Youth in a Comparative Perspective*. New York: New York University Press, 2011.
- Alba, Richard D.; Jimenez, Tomas R.; and Marrow, Helen B. "Mexican Americans as a Paradigm for Contemporary Intra-Group Heterogeneity." *Ethnic and Racial Studies*, February 2014, 37(3), pp. 446-66.
- Alba, Richard D.; Kasinitz, Philip; and Waters, Mary C. "The Kids Are (Mostly) Alright: Second Generation Assimilation; Comments on Haller, Portes and Lynch." *Social Forces*, March 2011, 89(3), pp. 733-62.
- Alba, Richard D.; Logan, John; Lutz, Amy; and Stults, John. "Only English by the Third Generation? Loss and Preservation of the Mother Tongue among the Grandchildren of Contemporary Immigrants." *Demography*, August 2002, 39(3), pp. 467-84.
- Alon, Sigal; Domina, Thurston; and Tienda, Marta. "Stymied Mobility or Temporary Lull? The Puzzle of Lagging Hispanic College Degree Attainment." *Social Forces*, June 2010, 88(4), pp. 1807-32.
- Altonji, Joseph G., and Pierret, Charles R. "Employer Learning and Statistical Discrimination." *The Quarterly Journal of Economics*, February 2001, 116(1), pp. 313-50.
- Altonji, Joseph G.; Bharadwaj, Prashant; and Lange, Fabian. "Changes in the Characteristics of American Youth: Implications for Adult Outcomes." *Journal of Labor Economics*, October 2012, 30(4), pp. 783-828.
- Bean, Frank D.; Brown, Susan K.; and Bachmeier, James D. *Parents without Papers: The Progress and Pitfalls of Mexican-American Integration*. New York: Russell Sage Foundation, 2015.
- Bean, Frank D.; Leach, Mark; Brown, Susan K.; Bachmeier, James; and Hipp, John. "The Educational Legacy of Unauthorized Migration: Comparisons across U.S.-Immigrant Groups in How Parents' Status Affects Their Offspring." *International Migration Review*, Summer 2011, 45(2), pp. 348-385.

- Black, Sandra E., and Devereux, Paul J. "Recent Developments in Intergenerational Mobility, in Orley Ashenfelter and David Card, eds., *Handbook of Labor Economics*, vol. 4B. Amsterdam: North Holland, 2011, pp. 1487-1541.
- Blau, Francine D., and Kahn, Lawrence M. "Gender and Assimilation among Mexican Americans," in George J. Borjas, ed., *Mexican Immigration to the United States*. Chicago: University of Chicago Press, 2007, pp. 57-106.
- Borjas, George J. "The Intergenerational Mobility of Immigrants." *Journal of Labor Economics*, January 1993, Part 1, 11(1), pp. 113-35.
- Borjas, George J. "Making It in America: Social Mobility in the Immigrant Population." *The Future of Children*, Fall 2006, 16(2), pp. 55-71.
- Castex, Gonzalo, and Kogan Dechter, Evgenia. "The Changing Roles of Education and Ability in Wage Determination." *Journal of Labor Economics*, October 2014, 32(4), pp. 685-710.
- Cunha, Flavio; Heckman, James J.; Lochner, Lance; and Masterov, Dimitriy V. "Interpreting the Evidence on Life Cycle Skill Formation," in Eric A. Hanushek and Finis Welch, eds., *Handbook of the Economics of Education*, vol. 1. Amsterdam: North Holland, 2006, pp. 697-812.
- Duncan, Brian; Hotz, V. Joseph; and Trejo, Stephen J. "Hispanics in the U.S. Labor Market," in Marta Tienda and Faith Mitchell, eds., *Hispanics and the Future of America*. Washington, DC: Washington, DC: National Academies Press, 2006.
- Duncan, Brian, and Trejo, Stephen J. "Ethnic Identification, Intermarriage, and Unmeasured Progress by Mexican Americans," in George J. Borjas, ed., *Mexican Immigration to the United States*. Chicago: University of Chicago Press, 2007, pp. 227-69.
- Duncan, Brian, and Trejo, Stephen J. "Intermarriage and the Intergenerational Transmission of Ethnic Identity and Human Capital for Mexican Americans." *Journal of Labor Economics*, April 2011, 29(2), pp. 195-227.
- Duncan, Brian, and Trejo, Stephen J. "The Complexity of Immigrant Generations: Implications for Assessing the Socioeconomic Integration of Hispanics and Asians." Working Paper No. 21982. Cambridge, MA: National Bureau of Economic Research, February 2016.
- Duncan, Brian, and Trejo, Stephen J. "The Complexity of Immigrant Generations: Implications for Assessing the Socioeconomic Integration of Hispanics and Asians." *Industrial and Labor Relations Review*, October 2017, 70(5), pp. 1146-75.
- Duncan, Brian, and Trejo, Stephen J. "Socioeconomic Integration of U.S. Immigrant Groups over the Long Term: The Second Generation and Beyond," in Susan Pozo, ed., *The Human and Economic Implications of Twenty-First Century Immigration Policy*,

- Kalamazoo, MI: W.E. Upjohn Institute for Employment Research, 2018, pp. 33-62.
- Dustmann, Christian. "Return Migration, Investment in Children, and Intergenerational Mobility: Comparing sons of Foreign- and Native-Born Fathers." *Journal of Human Resources*, Spring 2008, 43(2), pp. 299-324.
- Farley, Reynolds, and Alba, Richard. "The New Second Generation in the United States." *International Migration Review*, Fall 2002, 36(3), pp. 669-701.
- Feliciano, Cynthia. "Does Selective Migration Matter? Explaining Ethnic Disparities in Educational Attainment among Immigrants' Children." *International Migration Review*, Winter 2005, 39(4), pp. 841-71.
- Foley, Neil. *The White Scourge: Mexicans, Blacks and Poor Whites in Texas Cotton Culture*. Berkeley, CA: University of California Press, 1997.
- Fry, Richard, and Lowell, B. Lindsay. "Work or Study: Different Fortunes of U.S. Latino Generations." Report. Washington, DC: Pew Hispanic Center, 2002.
- Gang, Ira N., and Zimmerman, Klaus F. "Is Child like Parent? Educational Attainment and Ethnic Origin." *Journal of Human Resources*, Summer 2000, 35(3), pp. 550-69.
- Gramlich, John. "Hispanic Dropout Rate Hits New Low, College Enrollment at New High." Pew Research Center Fact Tank, September 2017. Accessed on 9/29/2017 at: <http://www.pewresearch.org/fact-tank/2017/09/29/hispanic-dropout-rate-hits-new-low-college-enrollment-at-new-high/>.
- Grogger, Jeffrey, and Trejo, Stephen J. *Falling Behind or Moving Up? The Intergenerational Progress of Mexican Americans*. San Francisco: Public Policy Institute of California, 2002.
- Haller, William; Portes, Alejandro; and Lynch, Scott M. "Dreams Fulfilled, Dreams Shattered: Determinants of Segmented Assimilation in the Second Generation." *Social Forces*, March 2011a, 89(3), pp. 733-62.
- Haller, William; Portes, Alejandro; and Lynch, Scott M. "On the Dangers of Rosy Lenses; Reply to Alba, Kasinitz, and Waters." *Social Forces*, March 2011b, 89(3), pp. 775-82.
- Haveman, Robert, and Barbara Wolfe. *Succeeding Generations: On the Effects of Investments in Children*. New York: Russell Sage Foundation, 1994.
- Heckman, James J.; Humphries, John Eric; and Veramendi, Gregory. "The Non-Market Benefits of Education and Ability." Discussion Paper No. 11047. Bonn, Germany: IZA Institute of Labor Economics, September 2017.
- Hout, Michael. "Social and Economic Returns to College Education in the United States."

- Annual Review of Sociology*, 2012, 38, pp. 379-400.
- Huntington, Samuel P. *Who Are We? The Challenges to America's National Identity*. New York: Simon and Schuster, 2004.
- Kosack, Edward, and Ward, Zachary. "El Sueno Americano? The Generational Progress of Mexican Americans Prior to World War II." Manuscript. Waco, TX: Baylor University, September 2018.
- Livingston, Gretchen, and Kahn, Joan R. "An American Dream Unfulfilled: The Limited Mobility of Mexican Americans." *Social Science Quarterly*, December 2002, 83(4), pp. 1003-12.
- Luthra, Renee R., and Soehl, Thomas. "From Parent to Child? Transmission of Educational Attainment within Immigrant Families: Methodological Considerations." *Demography*, April 2015, 52(2), pp. 543-67.
- Montejano, David. *Anglos and Mexicans in the Making of Texas: 1836-1986*. Austin, TX: University of Texas Press, 1987.
- Mulligan, Casey B. *Parental Priorities and Economic Inequality*. Chicago: University of Chicago Press, 1997.
- Murnane, Richard J. "U.S. High School Graduation Rates: Patterns and Explanations." *Journal of Economic Literature*, June 2013, 51(2), pp. 370-422.
- Neal, Derek A., and Johnson, William R. "The Role of Premarket Factors in Black-White Wage Differences." *Journal of Political Economy*, October 1996, 104(5), pp. 869-95.
- Nielsen, Helena S.; Rosholm, Michael; Smith, Nina; and Husted, Leif. "The School-to-Work Transition of 2nd Generation Immigrants in Denmark." *Journal of Population Economics*, November 2003, 16(4), pp. 755-86.
- Ortiz, Vilma, and Telles, Edward. "Third Generation Disadvantage among Mexican Americans." *Sociology of Race and Ethnicity*, October 2017, 3(4), pp. 441-57.
- Park, Julie; Myers, Dowell; and Jimenez, Tomas R. "Intergenerational Mobility of the Mexican-Origin Population in California and Texas Relative to a Changing Regional Mainstream." *International Migration Review*, Summer 2014, 48(2), pp. 442-81.
- Perlmann, Joel. *Italians Then, Mexicans Now: Immigrant Origins and Second-Generation Progress, 1890-2000*. New York: Russell Sage Foundation, 2005.
- Perlmann, Joel. "The Mexican American Second Generation in Census 2000: Education and Earnings," in Richard Alba and Mary C. Waters, eds., *The Next Generation: Immigrant Youth in a Comparative Perspective*. New York: New York University Press, 2011.

- Perlmann, Joel, and Waters, Mary C. "Intermarriage and Multiple Identities," in Mary C. Waters and Reed Ueda, eds., *The New Americans: A Guide to Immigration since 1965*. Cambridge, MA: Harvard University Press, 2007, pp. 110-23.
- Portes, Alejandro, and Rumbaut, Ruben G. *Legacies: The Story of the Immigrant Second Generation*. Berkeley, CA: University of California Press, 2001.
- Portes, Alejandro. "Review Essay: Paths of Assimilation in the Second Generation." *Sociological Forum*, September 2006, 21(3), pp. 499-504.
- Rumbaut, Rugen G.; Massey, Douglas S.; and Bean, Frank D. "Linguistic Life Expectancies: Immigrant Language Retention in the Southern California." *Population and Development Review*, September 2006, 32(3), pp. 447-60.
- Smith, James P. "Assimilation across the Latino Generations." *American Economic Review*, May 2003, 93(2), pp. 315-319.
- Smith, James P. "Immigrants and the Labor Market." *Journal of Labor Economics*, April 2006, 24(2): 203-33.
- Telles, Edward E., and Ortiz, Vilma. *Generations of Exclusion: Mexican Americans, Assimilation, and Race*. New York: Russell Sage Foundation, 2008.
- Trejo, Stephen J. "Why Do Mexican Americans Earn Low Wages?" *Journal of Political Economy*, December 1997, 105(6), pp. 1235-68.
- Trejo, Stephen J. "Intergenerational Progress of Mexican-Origin Workers in the U.S. Labor Market." *Journal of Human Resources*, Summer 2003, 38(3), pp. 467-89.
- U.S. Bureau of Labor Statistics. "Appendix 10: CAT-ASVAB Scores." *National Longitudinal Surveys of Youth, 1997, Codebook Supplement*. 2019.
<https://www.nlsinfo.org/content/cohorts/nlsy97/other-documentation/codebook-supplement/appendix-10-cat-asvab-scores>
- Waters, Mary C. *Ethnic Options: Choosing Identities in America*. Berkeley, CA: University of California Press, 1990.

**Table 1: Average Years of Schooling of Men,
by Race/Ethnicity, Age, and Immigrant Generation,
2003-2016 CPS Data**

Race/Ethnicity and Age	Immigrant Generation		
	1 st	2 nd	3 rd +
A. Ages 25-59			
Mexican American	9.49 (0.02)	12.71 (0.03)	12.65 (0.02)
Non-Hispanic White			13.80 (0.004)
Non-Hispanic Black			12.95 (0.01)
B. By Age Cohort			
Mexican American:			
Ages 25-34	9.92 (0.03)	12.68 (0.03)	12.65 (0.03)
Ages 50-59	8.50 (0.06)	12.58 (0.09)	12.51 (0.05)
Non-Hispanic White:			
Ages 25-34			13.80 (0.007)
Ages 50-59			13.79 (0.007)
Non-Hispanic Black:			
Ages 25-34			12.96 (0.02)
Ages 50-59			12.80 (0.02)

Source: 2003-2016 Current Population Survey outgoing rotation group data.

Note: Standard errors are reported in parentheses. The samples include men in the relevant racial/ethnic and age groups. For Mexican Americans, the “1st generation” consists of individuals born in Mexico, excluding those born abroad of an American parent, and the “2nd generation” consists of U.S.-born individuals who have at least one Mexican-born parent. The “3rd+ generation” (i.e., the 3rd and all higher generations) consists of U.S.-born individuals who have two U.S.-born parents, and these individuals are assigned to racial/ethnic groups based on their responses to the Hispanic origin and race questions. Sampling weights were used in the calculations.

Table 2: Educational Attainment, by Race/Ethnicity and Immigrant Generation, NLSY97 Data

Race/Ethnicity and Generation	Average Years of Schooling	Percent with at least:			Sample Size
		High School Diploma	Some College	Bachelors Degree	
Mexican American:					
1.5 generation	11.94 (0.18)	61.87 (3.52)	28.30 (3.27)	9.02 (2.08)	189
2 nd generation	13.01 (0.13)	76.95 (2.16)	47.82 (2.57)	14.16 (1.79)	378
3 rd generation	13.54 (0.23)	84.34 (2.93)	53.46 (4.02)	20.22 (3.24)	151
4 th + generation	12.78 (0.19)	68.14 (2.80)	42.12 (2.97)	20.81 (2.44)	274
3 rd + generation	13.02 (0.15)	73.42 (2.13)	45.81 (2.40)	20.62 (1.95)	425
Non-Hispanic:					
Black, 4 th + generation	13.35 (0.08)	74.80 (1.19)	51.81 (1.37)	19.85 (1.10)	1,311
White, 4 th + generation	14.47 (0.06)	86.12 (0.70)	65.08 (0.96)	39.58 (0.99)	2,436

Source: National Longitudinal Survey of Youth 1997 data through round 17 (2015-2016).

Note: Standard errors are reported in parentheses. The samples include men and women whose race/ethnicity and immigrant generation could be identified; see text for further information. The sample sizes listed above are for the completed years of schooling variable. Because of less missing information regarding degree completion, the corresponding sample sizes are slightly larger for the binary measures of educational attainment. Sampling weights were used in the calculations.

Table 3: Average Years of Schooling, by Race/Ethnicity/Immigrant Generation and Family Type

Race/Ethnicity and Generation	Percent Not from Intact Families	Average Years of Schooling, by Family Type:			Difference Relative to Whites, by Family Type:		
		All	Intact	Not Intact	All	Intact	Not Intact
Mexican American:							
1.5 generation	42.85 (3.61)	11.94 (0.18)	12.54 (0.22)	11.14 (0.28)	-2.53 (0.19)	-2.67 (0.22)	-2.27 (0.29)
2 nd generation	44.33 (2.56)	13.01 (0.13)	13.24 (0.18)	12.73 (0.21)	-1.46 (0.16)	-1.97 (0.20)	-0.69 (0.24)
3 rd generation	50.65 (4.08)	13.54 (0.23)	13.99 (0.33)	13.10 (0.31)	-0.93 (0.26)	-1.22 (0.38)	-0.31 (0.36)
4 th + generation	62.50 (2.93)	12.78 (0.19)	13.38 (0.31)	12.42 (0.23)	-1.69 (0.22)	-1.83 (0.35)	-1.00 (0.29)
Non-Hispanic:							
Black, 4 th + generation	79.99 (1.11)	13.35 (0.08)	14.50 (0.18)	13.07 (0.09)	-1.12 (0.11)	-0.72 (0.21)	-0.35 (0.14)
White, 4 th + generation	41.27 (1.00)	14.47 (0.06)	15.21 (0.07)	13.42 (0.10)			

Source: National Longitudinal Survey of Youth 1997 data through round 17 (2015-2016).

Note: Standard errors are reported in parentheses. See Table 2 and the text for further information about the sample. Respondents are defined to be from “intact” families if they grew up with both biological parents through the initial interview in 1997, when respondents were between the ages of 12-18. Sampling weights were used in the calculations.

Table 4: Education Regressions

Regressor	Dependent Variable							
	Completed Years of Schooling		Indicator for completion of at least:					
			High School Diploma		Some College		Bachelors Degree	
(1)	(2)	(1)	(2)	(1)	(2)	(1)	(2)	
Race/Ethnicity and Generation								
Mexican American:								
1.5 generation	-2.53 (.19)	-2.69 (.25)	-.243 (.039)	-.292 (.044)	-.368 (.036)	-.419 (.045)	-.306 (.022)	-.266 (.035)
2 nd generation	-1.46 (.16)	-1.52 (.19)	-.092 (.023)	-.102 (.027)	-.173 (.029)	-.191 (.034)	-.254 (.022)	-.230 (.027)
3 rd generation	-.93 (.26)	-.94 (.28)	-.018 (.032)	-.014 (.034)	-.116 (.045)	-.119 (.049)	-.194 (.037)	-.174 (.040)
4 th + generation	-1.69 (.22)	-1.66 (.22)	-.180 (.032)	-.160 (.032)	-.230 (.035)	-.216 (.035)	-.188 (.030)	-.175 (.030)
Non-Hispanic:								
Black, 4 th + generation	-1.12 (.11)	-.97 (.12)	-.113 (.015)	-.093 (.016)	-.133 (.018)	-.120 (.020)	-.197 (.016)	-.179 (.018)
White, 4 th + generation (reference group)								
Control variables included?	No	Yes	No	Yes	No	Yes	No	Yes
R ²	.05	.09	.03	.06	.03	.06	.04	.08

Source: National Longitudinal Survey of Youth 1997 data through round 17 (2015-2016).

Note: The reported figures are estimated coefficients from least squares regressions in which the dependent variables are various measures of educational attainment. Heteroskedasticity-robust standard errors are shown in parentheses. The sample sizes are 4,739 for regressions where the dependent variable is completed years of schooling and 4,780 for regressions where the dependent variables are the binary measures of educational attainment. See Table 2 and the text for further information about the sample. The “control variables” included in specification (2) are indicators for the respondent’s sex, age (when last interviewed), and state of birth. Sampling weights were used in the calculations.

Table 5: Average Parental Years of Schooling, by Race/Ethnicity and Immigrant Generation

Race/Ethnicity and Generation	Average Years of Schooling of:		Percent with Missing Parental Schooling Data for:			
	Mother	Father	Mother Only	Father Only	Both Parents	Neither Parent
Mexican American:						
1.5 generation	7.02 (0.25)	6.89 (0.27)	4.41 (1.50)	17.59 (2.78)	5.92 (1.72)	72.07 (3.27)
2 nd generation	8.79 (0.20)	8.17 (0.25)	1.99 (0.72)	19.29 (2.03)	5.20 (1.14)	73.53 (2.27)
3 rd generation	12.56 (0.21)	12.34 (0.25)	2.17 (1.19)	5.55 (1.87)	2.72 (1.33)	89.56 (2.50)
4 th + generation	12.15 (0.15)	12.17 (0.19)	2.33 (0.91)	14.69 (2.14)	3.53 (1.12)	79.45 (2.45)
Non-Hispanic:						
Black, 4 th + generation	12.43 (0.06)	12.31 (0.07)	3.15 (0.48)	23.06 (1.16)	5.29 (0.62)	68.50 (1.28)
White, 4 th + generation	13.44 (0.05)	13.51 (0.06)	1.65 (0.26)	7.12 (0.52)	2.46 (0.31)	88.78 (0.64)

Source: National Longitudinal Survey of Youth 1997 data through round 17 (2015-2016).

Note: Standard errors are reported in parentheses. See Table 2 and the text for further information about the sample. Sampling weights were used in the calculations.

Table 6: Years of Schooling Regressions

Regressor	(1)	(2)	(3)
Race/Ethnicity and Generation			
1.5-generation Mexican American	-2.69 (.25)	.65 (.27)	-.71 (.41)
2 nd -generation Mexican American	-1.52 (.19)	1.08 (.20)	.14 (.23)
3 rd -generation Mexican American	-.94 (.28)	-.31 (.28)	-.15 (.28)
4 th +generation Mexican American	-1.66 (.22)	-.87 (.19)	-.86 (.19)
4 th +generation non-Hispanic Black	-.97 (.12)	-.28 (.11)	-.29 (.12)
Parental years of schooling – 12:			
Mother		.28 (.02)	.29 (.03)
Father		.28 (.02)	.31 (.02)
Missing parental schooling data for:			
Mother only		-.75 (.28)	-.67 (.28)
Father only		-.81 (.14)	-.68 (.14)
Both parents		-.69 (.30)	-.48 (.29)
Interactions with parental schooling			
(1.5-generation Mexican American) × (Mother’s schooling – 12)			-.17 (.06)
(1.5-generation Mexican American) × (Father’s schooling – 12)			-.24 (.06)
(2nd-generation Mexican American) × (Mother’s schooling – 12)			-.26 (.05)
(2nd-generation Mexican American) × (Father’s schooling – 12)			-.14 (.04)
(3rd-generation Mexican American) × (Mother’s schooling – 12)			-.04 (.12)
(3rd-generation Mexican American) × (Father’s schooling – 12)			-.20 (.13)
(4 th +generation Mexican American) × (Mother’s schooling – 12)			.15 (.09)
(4 th +generation Mexican American) × (Father’s schooling – 12)			.10 (.06)
(4 th +generation Black) × (Mother’s schooling – 12)			.14 (.05)
(4 th +generation Black) × (Father’s schooling – 12)			-.12 (.06)
Control variables included?	Yes	Yes	Yes
R ²	.09	.27	.28

Source: National Longitudinal Survey of Youth 1997 data through round 17 (2015-2016).

Note: The reported figures are estimated coefficients from least squares regressions in which the dependent variable is completed years of schooling. Heteroskedasticity-robust standard errors are shown in parentheses. The sample size is 4,739 for all regressions. See Table 2 and the text for further information about the sample. The reference group for the race/ethnicity and generation dummy variables is 4th+ generation non-Hispanic whites. The variables representing mother's and father's years of schooling are normalized to equal zero when the parent has 12 years of schooling. These variables are also set to zero when the relevant parental schooling information is missing. As a result, the coefficients of the dummy variables indicating missing parental schooling data represent differentials between individuals with the indicated missing data and otherwise similar individuals with reported parental schooling of 12 years. The "control variables" included in all specifications are indicators for the respondent's sex, age (when last interviewed), and state of birth. Sampling weights were used in the calculations.

**Table 7: Predicted Years of Schooling Differentials (Relative to Non-Hispanic Whites),
by Parental Years of Schooling**

Race/Ethnicity and Generation	Mother's and Father's Years of Schooling:			
	10	12	14	16
Mexican American:				
1.5 generation	.11 (.32)	-.71 (.41)	-1.53 (.54)	-2.35 (.67)
2 nd generation	.93 (.20)	.14 (.23)	-.66 (.30)	-1.46 (.38)
3 rd generation	.32 (.33)	-.15 (.28)	-.63 (.32)	-1.10 (.43)
4 th + generation	-1.36 (.25)	-.86 (.19)	-.36 (.23)	.13 (.35)
Non-Hispanic:				
Black, 4 th + generation	-.33 (.18)	-.29 (.12)	-.24 (.13)	-.20 (.21)

Source: National Longitudinal Survey of Youth 1997 data through round 17 (2015-2016).

Note: The reported figures are predicted years of schooling differentials between each group and 4th+ generation non-Hispanic whites. Heteroskedasticity-robust standard errors are shown in parentheses. The predicted differentials derive from the least squares regression reported in the column labeled (3) in Table 6. See Table 6 for further information about the regression. The predicted differentials compare individuals with the same characteristics (sex, age at last interview, and state of birth) whose mother and father each have the years of schooling indicated above.

Table 8: AFQT Score Regressions

Regressor	(1)	(2)	(3)	(4)
Race/Ethnicity and Generation				
Mexican American:				
1.5 generation	-31.75 (2.54)	-33.75 (3.11)	-31.75 (2.54)	-34.18 (3.10)
2 nd generation	-24.98 (1.76)	-25.89 (2.05)	-24.98 (1.76)	-26.12 (2.04)
3 rd generation	-17.72 (2.53)	-18.86 (2.69)		
4 th + generation	-25.63 (1.91)	-24.79 (1.98)		
3 rd + generation			-23.21 (1.60)	-23.02 (1.69)
Non-Hispanic:				
Black, 4 th + generation	-27.94 (1.01)	-26.00 (1.17)	-27.94 (1.01)	-25.96 (1.17)
White, 4 th + generation (reference group)				
Control variables included?				
Sex	Yes	Yes	Yes	Yes
Age at testing	Yes	Yes	Yes	Yes
State of residence	No	Yes	No	Yes
R²	.172	.203	.170	.202

Source: National Longitudinal Survey of Youth 1997 data through round 17 (2015-2016).

Note: The reported figures are estimated coefficients from least squares regressions in which the dependent variable is the percentile score on the Armed Forces Qualification Test (AFQT). The sample size is 3,833 for all regressions. Heteroskedasticity-robust standard errors are shown in parentheses. Sampling weights were used in the calculations.

Table 9: Annual Weeks Worked Regressions

Regressor	(1)	(2)	(3)	(4)	(5)	(6)
Race/Ethnicity and Generation						
Mexican American:						
1.5 generation	-2.758 (1.716)	-1.050 (1.760)	3.886 (1.754)	-2.758 (1.716)	-1.213 (1.759)	3.773 (1.753)
2 nd generation	0.067 (1.029)	0.358 (1.123)	3.487 (1.133)	0.067 (1.029)	0.231 (1.121)	3.397 (1.132)
3 rd generation	-0.430 (1.828)	-0.546 (1.840)	1.565 (1.838)			
4 th + generation	-5.703 (1.548)	-5.836 (1.584)	-2.567 (1.515)			
3 rd + generation				-4.007 (1.228)	-4.175 (1.274)	-1.261 (1.235)
Non-Hispanic:						
Black, 4 th + generation	-4.625 (0.736)	-4.233 (0.782)	-2.374 (0.745)	-4.625 (0.736)	-4.215 (0.782)	-2.355 (0.745)
White, 4 th + generation (reference group)						
Years of schooling			1.723 (0.099)			1.728 (0.099)
Control variables included?	No	Yes	Yes	No	Yes	Yes
R ²	.011	.064	.133	.010	.063	.132

Source: National Longitudinal Survey of Youth 1997 data through round 17 (2015-2016).

Note: The reported figures are estimated coefficients from least squares regressions in which the dependent variable is the number of weeks worked during the calendar year preceding the interview date. The sample size is 8,836 for all regressions. The sample excludes weeks worked observations from when a respondent is younger than age 30 or is enrolled in school. Standard errors clustered at the level of the respondent are shown in parentheses. The “control variables” included in some specifications are indicators for the respondent’s sex, age, and state of residence, as well as for the year of the weeks worked observation. Sampling weights were used in the calculations.

Table 10: Hourly Wage Regressions

Regressor	(1)	(2)	(3)	(4)	(5)	(6)
Race/Ethnicity and Generation						
Mexican American:						
1.5 generation	-0.314 (0.060)	-0.470 (0.061)	-0.225 (0.056)	-0.314 (0.060)	-0.467 (0.061)	-0.222 (0.056)
2 nd generation	-0.164 (0.038)	-0.301 (0.043)	-0.141 (0.040)	-0.164 (0.038)	-0.299 (0.043)	-0.139 (0.040)
3 rd generation	-0.121 (0.055)	-0.258 (0.055)	-0.152 (0.052)			
4 th + generation	-0.156 (0.050)	-0.187 (0.050)	-0.040 (0.044)			
3 rd + generation				-0.144 (0.039)	-0.211 (0.040)	-0.077 (0.036)
Non-Hispanic:						
Black, 4 th + generation	-0.295 (0.024)	-0.279 (0.026)	-0.200 (0.024)	-0.295 (0.024)	-0.279 (0.026)	-0.201 (0.024)
White, 4 th + generation (reference group)						
Years of schooling			0.080 (0.003)			0.080 (0.003)
Control variables included?	No	Yes	Yes	No	Yes	Yes
R ²	.033	.100	.232	.033	.100	.232

Source: National Longitudinal Survey of Youth 1997 data through round 17 (2015-2016).

Note: The reported figures are estimated coefficients from least squares regressions in which the dependent variable is the natural logarithm of the hourly wage at the time of interview. The sample size is 7,478 for all regressions.

The sample excludes wage observations from when a respondent is younger than age 30 or is enrolled in school.

Standard errors clustered at the level of the respondent are shown in parentheses. The “control variables” included in some specifications are indicators for the respondent’s sex, age, and state of residence, as well as for the year of the wage observation. Sampling weights were used in the calculations.

Table 11: Annual Earnings Regressions

Regressor	(1)	(2)	(3)	(4)	(5)	(6)
Race/Ethnicity and Generation						
Mexican American:						
1.5 generation	-0.479 (0.105)	-0.612 (0.102)	-0.294 (0.091)	-0.479 (0.105)	-0.613 (0.102)	-0.293 (0.091)
2 nd generation	-0.225 (0.070)	-0.334 (0.075)	-0.123 (0.071)	-0.225 (0.070)	-0.335 (0.075)	-0.122 (0.071)
3 rd generation	-0.078 (0.074)	-0.229 (0.075)	-0.104 (0.074)			
4 th + generation	-0.197 (0.084)	-0.270 (0.084)	-0.079 (0.075)			
3 rd + generation				-0.156 (0.062)	-0.256 (0.064)	-0.087 (0.058)
Non-Hispanic:						
Black, 4 th + generation	-0.382 (0.040)	-0.366 (0.042)	-0.268 (0.038)	-0.382 (0.040)	-0.366 (0.042)	-0.268 (0.038)
White, 4 th + generation (reference group)						
Years of schooling			0.107 (0.005)			0.107 (0.005)
Control variables included?	No	Yes	Yes	No	Yes	Yes
R ²	.024	.088	.187	.024	.088	.187

Source: National Longitudinal Survey of Youth 1997 data through round 17 (2015-2016).

Note: The reported figures are estimated coefficients from least squares regressions in which the dependent variable is the natural logarithm of earnings during the calendar year preceding the interview date. The sample size is 6,473 for all regressions. The sample excludes earnings observations from when a respondent is younger than age 30 or is enrolled in school. Standard errors clustered at the level of the respondent are shown in parentheses. The “control variables” included in some specifications are indicators for the respondent’s sex, age, and state of residence, as well as for the year of the earnings observation. Sampling weights were used in the calculations.

Table 12: Rates of Hispanic Identification (%) for Mexican Americans, by Sample Type and Immigrant Generation

Sample Type and Generation	Percent Identified as Hispanic	Sample Size
<i>Cross-Sectional Sample</i>		
Mexican American:		
1.5 generation	100.00 (0.00)	87
2 nd generation	94.63 (1.82)	155
3 rd generation	79.38 (4.58)	79
<i>Supplemental Sample</i>		
Mexican American:		
1.5 generation	100.00 (0.00)	104
2 nd generation	100.00 (0.00)	229
3 rd generation	100.00 (0.00)	76
<i>Both Samples Combined</i>		
Mexican American:		
1.5 generation	100.00 (0.00)	191
2 nd generation	97.37 (0.82)	384
3 rd generation	86.97 (2.71)	155

Source: National Longitudinal Survey of Youth 1997 data through round 17 (2015-2016).

Note: Standard errors are reported in parentheses. The samples include men and women who could be identified as 1.5-, 2nd-, or 3rd-generation Mexican Americans based on the countries of birth reported for each respondent, his parents, and his grandparents; see text for further information. The “sample type” indicates if a given observation is part of the “cross-sectional” sample that is representative of all U.S. youth in the sampling universe when the survey began in 1997, or if the observation instead comes from the “supplemental” oversample of blacks and Hispanics. Hispanic identification is based on information collected at the beginning of the survey in 1997. Sampling weights were used in the calculations.

Table 13: Educational Attainment of 3rd-Generation Mexican Americans from the Cross-Sectional Sample, by Hispanic Identification

Sample Type and Hispanic Identification	Average Years of Schooling	Percent with at least:			Sample Size
		High School Diploma	Some College	Bachelors Degree	
<i>Cross-Sectional Sample</i>					
Identified as Hispanic	13.58 (0.35)	85.50 (4.33)	51.96 (6.15)	23.01 (5.18)	65
Not identified as Hispanic	14.22 (1.16)	82.07 (11.57)	49.26 (15.07)	29.17 (13.70)	11
All	13.70 (0.35)	84.79 (4.07)	51.41 (5.66)	24.28 (4.85)	76

Source: National Longitudinal Survey of Youth 1997 data through round 17 (2015-2016).

Note: Standard errors are reported in parentheses. The sample includes men and women who could be identified as 3rd-generation Mexican Americans based on the countries of birth reported for each respondent, his parents, and his grandparents; see text for further information. Hispanic identification is based on information collected at the beginning of the survey in 1997. The sample sizes listed above are for the completed years of schooling variable. Because of less missing information regarding degree completion, the corresponding sample sizes are slightly larger for the binary measures of educational attainment. Sampling weights were used in the calculations.

Table 14: Average Years of Schooling of 3rd-Generation Mexican Americans, by Source of Mexican Ancestry

Source of Mexican Ancestry	Percent of Sample	Average Years of Schooling	Sample Size
Number of Mexican-born grandparents:			
1	61.0	13.80 (0.29)	91
2	27.1	13.70 (0.47)	39
3	4.7	12.35 (1.02)	8
4	7.2	11.67 (0.58)	13
All 3 rd -generation Mexican Americans	100.0	13.54 (0.23)	151
Mexican ancestry observed on:			
Father's side only	36.5	13.61 (0.37)	50
Mother's side only	32.0	14.00 (0.42)	48
Both sides of family	31.5	13.00 (0.39)	53
All 3 rd -generation Mexican Americans	100.0	13.54 (0.23)	151

Source: National Longitudinal Survey of Youth 1997 data through round 17 (2015-2016).

Note: Standard errors are reported in parentheses. The sample includes men and women who could be identified as 3rd-generation Mexican Americans based on the countries of birth reported for each respondent, his parents, and his grandparents; see text for further information. Mexican ancestry is said to be observed on the father's side of the family when at least one of the following two things is true: (1) the respondent has a paternal grandparent who was born in Mexico, or (2) the respondent's father subjectively identifies as Mexican American. Analogously, presence of a Mexican-born maternal grandparent and/or the mother's subjective identification as Mexican American determine whether a respondent is observed to have Mexican ancestry on his mother's side of the family. Sampling weights were used in the calculations.

Table 15: Differences between 3rd- and 4th+ -Generation Mexican Americans in Geographic Roots

	Immigrant Generation	
	3 rd	4 th +
A. Percent of generation with roots in:		
California but not Texas	27.3	8.4
Texas but not California	28.4	30.7
Both California and Texas	8.4	1.1
Neither California nor Texas	35.9	59.7
Total for generation	100.0	100.0
Sample size for generation	155	279
B. Percent of generation with:		
Any California roots	35.7	9.6
Any Texas roots	36.8	31.9

Source: National Longitudinal Survey of Youth 1997 data through round 17 (2015-2016).

Note: The samples include men and women who could be identified as 3rd-generation or 4th+ -generation Mexican Americans; see text for further information. An individual is defined as having California “roots” when at least one of the following things is observed: (1) the respondent was born in California; or (2) the respondent resided in California in 1997 when the survey began; or (3) either (or both) of the respondent’s parents was born in California. An analogous procedure determines whether an individual is defined as having Texas roots. Sampling weights were used in the calculations.

Table 16: Education Regressions for 3rd- and 4th+ -Generation Mexican Americans

Regressor	(1)	(2)	(3)	(4)	(5)
Mexican American: 3 rd generation (reference group)					
4 th + generation	-.83 (.34)	-.65 (.35)	-.62 (.28)	-.57 (.34)	-.52 (.35)
Any California roots		.63 (.41)		.17 (.37)	.14 (.38)
Parental years of schooling – 12:					
Mother			.35 (.06)	.35 (.08)	.36 (.08)
Father			.34 (.06)	.34 (.06)	.31 (.06)
Missing parental schooling data for:					
Mother only			-.71 (.87)	-.73 (1.03)	-.58 (.99)
Father only			-.54 (.41)	-.54 (.48)	-.27 (.49)
Both parents			.51 (.73)	.52 (1.00)	.72 (.99)
Not from intact family					-.63 (.31)
Control variables included?	Yes	Yes	Yes	Yes	Yes
R ²	.05	.06	.27	.27	.28

Source: National Longitudinal Survey of Youth 1997 data through round 17 (2015-2016).

Note: The reported figures are estimated coefficients from least squares regressions in which the dependent variable is completed years of schooling. Heteroskedasticity-robust standard errors are shown in parentheses. The sample includes men and women who could be identified as 3rd-generation or 4th+ -generation Mexican Americans; see text for further information. The sample size is 425 for all regressions. An individual is defined as having California “roots” when at least one of the following things is observed: (1) the respondent was born in California; or (2) the respondent resided in California in 1997 when the survey began; or (3) either (or both) of the respondent’s parents was born in California. The variables representing mother’s and father’s years of schooling are normalized to equal zero when the parent has 12 years of schooling. These variables are also set to zero when the relevant parental schooling information is missing. As a result, the coefficients of the dummy variables indicating missing parental schooling data represent differentials between individuals with the indicated missing data and otherwise similar individuals with reported parental schooling of 12 years. Respondents are defined to be from “intact” families if they grew up with both biological parents through the initial interview in 1997, when respondents were between the ages of 12-18. The “control variables” included in all specifications are indicators for the respondent’s sex and age (when last interviewed). Sampling weights were used in the calculations.

Appendix

A1. Summary

In this appendix, we discuss the construction of our analysis sample from National Longitudinal Survey of Youth 1997 (NLSY97) data and show how the representativeness of our sample is affected by panel attrition and by sample selection due to missing information regarding immigrant generation. We also investigate how the educational patterns by race/ethnicity and immigrant generation in our NLSY97 sample compare to those in a sample of similar individuals constructed from Current Population Survey (CPS) data. Finally, we use CPS data to explore the stability over time of these educational patterns by race/ethnicity and immigrant generation.

A2. Constructing the NLSY97 Sample

The NLSY97 is a nationally-representative longitudinal survey of 8,984 youth who were ages 12-16 on December 31, 1996. Surveys were conducted annually from 1997 through 2011 and every other year thereafter. We use information from round 1 (1997) through round 15 (2015-2016) to construct our sample and create our key variables. By round 15, respondents were between the ages of 30-36. To construct our measures of educational outcomes, we use information on the respondent's "highest grade completed" and "highest degree received" collected up through their most recent survey. There is some panel attrition, however, and so to make it probable that we observe completed schooling, we drop 699 individuals who were last

interviewed before they reached age 25. Given the ages of the respondents in the initial round, this implies that all remaining individuals were interviewed in 2005 (round 9) or later.²²

As described in Section 2 of the main text, we assign each respondent to an immigrant generation using information on the respondent's birthplace and his parents' and grandparents' birthplaces. This information was collected in different rounds of the NLSY97. For instance, the respondent's birthplace was first asked in 2001 (round 5). Mother's and father's birthplaces were asked in the original 1997 parent interview, and the question "was mother/father born in the U.S.?" was first asked in 2002 (round 6) if the respondent did not have a 1997 parent interview. Grandparents' regions of birth were first asked in 2006 (round 10). As a consequence, we drop an additional 14 individuals for whom we are unable to determine their immigrant generation because they were not interviewed in 2006 (round 10) or later. This brings the total number of respondents lost to panel attrition to 713 (or 7.9% of the original NLSY97 sample of 8,984).

We identify foreign-born respondents using the NLSY97 geocode variables for state of birth and country of birth. We drop 16 individuals for whom we cannot identify immigrant generation because they did not report their own country of birth. These individuals were not lost to panel attrition, but rather they were lost because their birthplace response was "don't know" or "refused to answer." The birthplace information for parents and grandparents necessary to assign a U.S.-born individual to an immigrant generation varies somewhat by generation.²³ For example, 2nd generation individuals are those born in the United States with at

²² Respondents may have been interviewed before or after their birthdate in a given survey year, and interviews for a given year were sometimes conducted the following year. As a result, there was one respondent who was age 25 in round 8 (2004) and did not have a round 9 or later interview. However, this individual was excluded from our sample due to other sample restrictions.

²³ To construct parents' and grandparents' places of birth, we start with the NLSY97 geocode variables for biological parents' and grandparents' regions of birth, where Mexico is separately identified. We then fill in missing values as U.S.-born or unknown foreign-born using the NLSY97 public use variables indicating if the biological parents and grandparents were born in the U.S.

least one foreign-born parent. Therefore, we can assign an individual to the 2nd generation even if one parent's birthplace is missing, provided that the other parent is known to be foreign-born. However, we cannot assign a U.S.-born person to an immigrant generation if one parent's birthplace is unknown and the other parent's birthplace is U.S.-born, or if both parents' birthplaces are unknown. As a result, we drop an additional 426 individuals with one or both parents' birthplaces missing because we cannot distinguish whether these individuals are 2nd generation versus higher generations using the available information. Similarly, individuals can be identified as 3rd generation even if one or more of their grandparents' birthplaces are unknown, provided that they are known to have at least one foreign-born grandparent. For a person to be identified as 4th+ generation, however, all four grandparents must be known to be U.S.-born. As a consequence, we drop an additional 1,005 individuals with missing information on some or all grandparents' places of birth because we cannot distinguish whether these individuals are 3rd generation versus higher generations with available information. This brings the total number of respondents lost due to missing birthplace information to 1,447 (16.1% of the original NLSY97 sample), leaving a sample of 6,824 individuals for whom we have information on both immigrant generation and educational attainment at age 25 or later.

The NLSY97 collected information about the respondent's race and Hispanic origin in 1997 and at various other times. We use the information from the baseline 1997 survey to determine whether respondents identify subjectively as being of Hispanic origin.²⁴ We also use the 1997 race and Hispanic origin information to identify respondents who are non-Hispanic white or non-Hispanic black. In identifying non-Hispanic whites and blacks, we do not include individuals who report any indication of Hispanic origin or some other race in later rounds of the

²⁴ Alternate measures of Hispanic identification that incorporate information from later rounds of the NLSY97 or from supplemental surveys display similar patterns as the 1997 measure. The 1997 measure is available for nearly all individuals at baseline, however, and so it is the preferred measure for investigating issues of panel attrition.

NLSY97. The remaining sample now includes all respondents with available information on race/ethnicity, immigrant generation, and educational attainment at age 25 or later. From this sample, we limit our analysis to the following groups: Mexican Americans (1.5, 2nd, 3rd, or 4th+ generation), 4th+ generation non-Hispanic whites, and 4th+ generation non-Hispanic blacks.²⁵ Finally, we also exclude 28 individuals who were born abroad and have a U.S.-born parent. In total, 2,025 individuals (22.5% of the original NLSY97 sample) were excluded from our sample because they are not members of the particular groups that we target for study. The end result is an analysis sample that includes 4,799 respondents.²⁶

A3. Panel Attrition and Missing Immigrant Generation

The preceding section shows that, starting from the original NLSY97 sample of 8,984 respondents, our analysis sample excludes 713 respondents (7.9%) because of panel attrition and another 1,447 respondents (16.1%) due to missing information about immigrant generation.²⁷ To investigate how such attrition and missing data impact the representativeness of our analysis sample, Table A1 reports averages of various respondent characteristics in the baseline 1997 survey for three different samples: the full NLSY97 sample of 8,984 respondents (in columns (1) and (4) of the table), the sample of 8,271 non-attriters that survive the panel attrition relevant for our analysis sample (in columns (2) and (5)), and the sample of 6,824 non-attriters for whom immigrant generation can be assigned (in columns (3) and (6)). For the characteristics not measured as binary indicators, standard errors are reported in parentheses. Columns (1)-(3)

²⁵ See Section 2 of the main text for further information about how these groups are defined.

²⁶ Sixty of these individuals have missing information regarding their completed years of schooling. Excluding these individuals yields the overall sample size of 4,739 reported in Table 2.

²⁷ As indicated previously, our analysis sample also excludes an additional 2,025 respondents (22.5%) because they are not members of the groups of interest, but this intentional exclusion is not relevant for the sample representativeness analysis conducted in this section.

present unweighted averages, and columns (4)-(6) replicate these calculations using sampling weights. Very similar patterns emerge from the unweighted and the weighted results.

Table A1 suggests that panel attrition and missing information on immigrant generation have little impact on the representativeness of our analysis sample. In general, average characteristics are very similar across the three samples. Comparisons between columns (1) and (3) (or between columns (4) and (6)) reveal the joint impact of panel attrition and missing generation data on our analysis sample. Men, whites, and blacks are underrepresented among those who remain in our sample, whereas women and Hispanics are overrepresented, but all of these differences are quite small (i.e., differences between the samples of less than one percentage point in the relevant shares, with the exception of a slightly larger difference for the share of Hispanics in the unweighted calculations). Similarly, birth year, region, parental education, and mother's age at her first birth and at the respondent's birth vary little across samples. The largest observed differences are for the share of respondents living with both biological parents when the survey began in 1997 and for the share of respondents with missing schooling data for their biological father. These differences emerge between columns (2) and (3) (and between columns (4) and (5)), indicating that they are created by the selection on being able to assign immigrant generation, rather than by selective panel attrition. In the unweighted averages, for example, 48.1 percent of the non-attriter sample lived with both biological parents in 1997, whereas the corresponding share increases to 52.4 percent among our analysis sample. Between the same samples there is a decline from 21.0 to 17.9 percent in the unweighted share of respondents with missing schooling data for their father. These patterns probably reflect the fact that respondents who grew up in intact families and have greater knowledge about both of

their parents are also better able to report the information regarding the countries of birth of their parents and grandparents that is needed to assign immigrant generation.

A4. Comparing the NLSY97 and the CPS

The U.S. Current Population Survey (CPS) provides nationally-representative data for large samples of individuals. The CPS surveys residences rather than individuals, so panel attrition is not an issue, and rates of missing information in the CPS are low for the variables of interest here. Therefore, the concerns about representativeness discussed above for our NLSY97 analysis sample would not be expected to have much effect on the representativeness of a comparable sample of individuals from the CPS. In this section, we show that educational patterns by race/ethnicity and immigrant generation in our NLSY97 sample are quite similar to those evident in CPS data. This finding lends additional support to the argument made in the preceding section that the basic patterns highlighted in our paper are unlikely to be artifacts generated by non-representativeness of our NLSY97 sample.

We proceed as follows to create a CPS sample comparable to our NLSY97 sample. We use CPS outgoing rotation group data from the years 2015 and 2016, in order to match the survey years of the round 17 data that we employ from the NLSY97.²⁸ We limit the CPS sample to men and women ages 30-36, in order to match the age range of NLSY97 respondents in round 17. Among foreign-born individuals, we limit the CPS sample to those who arrived in the United States after 1995, in order to approximate the NLSY97 sampling frame.²⁹ We exclude from the

²⁸ To avoid repeated observations on a given individual, we only use data from the first time a household appears in an outgoing rotation group (i.e., we only use data from the fourth month that a household appears in the CPS sample).

²⁹ To be included in the NLSY97 sampling frame, foreign-born individuals had to arrive in the United States before January 1, 1997. Ideally, therefore, we would like to use 1996, rather than 1995, as the year of arrival threshold for determining inclusion of foreign-born individuals in the CPS sample. The CPS, however, reports immigrant year of arrival in intervals, and 1994-1995 and 1996-1997 are the relevant intervals reported in the data.

CPS sample those born abroad of an American parent. Using the information reported in the CPS on race, Hispanic origin, own country of birth, and parents' countries of birth, we can identify three generations of Mexican Americans: 1.5, 2nd, and 3rd+. We can also identify the comparison groups of 3rd+-generation non-Hispanic whites and 3rd+-generation non-Hispanic blacks.³⁰

Table A2 reports average years of schooling by race/ethnicity and immigrant generation for our NLSY97 sample and for the comparable sample we constructed from the CPS.³¹ Schooling levels and the patterns across groups are similar in the two data sets. In particular, both data sets show for Mexican Americans substantial schooling gains between the 1.5 and 2nd generations but no further progress between the 2nd and 3rd+ generations, and both data sets indicate that a large schooling gap remains between 3rd+-generation Mexican Americans and later-generation non-Hispanic whites.

A5. Schooling Patterns in CPS Data across Survey Years

In this section, we use CPS data to examine how educational patterns by race/ethnicity and immigrant generation have varied over time. With this analysis, we hope to gain some insight into whether the cross-sectional comparisons we make between Mexican-American generations in NLSY97 reflect forces beyond generational assimilation.

The CPS began collecting information about the countries of birth of each respondent and his parents in 1994, so we use data from that year until 2016, which is the last survey year

³⁰ See Table 1 of the main text for further information about how these groups are defined.

³¹ The NLSY97 schooling averages shown in Table A2 are taken from Table 2 of the main text. For non-Hispanic whites and blacks, note that the NLSY97 calculations are for 4th+-generation individuals whereas the CPS calculations, by necessity, are for 3rd+-generation individuals (because it is impossible in CPS data to determine the countries of birth of an adult respondent's grandparents). This difference has a negligible impact on the comparisons, however. For non-Hispanic whites and blacks in NLSY97 data, schooling averages calculated for 3rd+-generation individuals are almost identical to those shown in the table for 4th+-generation individuals.

relevant for our NLSY97 data from round 17. We focus on two age groups: a younger group ages 25-34 that roughly corresponds to the ages of our NLSY97 sample, and an older group ages 50-59. As described in the preceding section, we define three generations of Mexican Americans (1st, 2nd, and 3rd+) as well as the comparison groups of 3rd+ -generation non-Hispanic whites and 3rd+ -generation non-Hispanic blacks.³² To improve the precision of the estimates, we split the 1994-2016 sample period into three-year intervals that are five years apart: 1994-1996, 1999-2001, ..., 2014-2016.

For groups defined by race/ethnicity, immigrant generation, and age, the left-side columns of Table A3 show how average schooling varies across the survey-year intervals. For all groups, educational attainment has been rising over time. To account for this secular trend of increasing education, the right-side columns of Table A3 report the corresponding schooling differences between each group and 3rd+ -generation non-Hispanic whites.

Among those ages 25-34 (i.e., the top panel of the table), average schooling levels are almost identical for 2nd- and 3rd+ -generation Mexican Americans, and the rate of increase over time for these groups has been very similar to that for non-Hispanic whites. As a result, the schooling gaps between younger U.S.-born Mexican Americans and their non-Hispanic white counterparts have been quite stable over the 20-year span for which the relevant CPS data are available. The NLSY97 sample that we study corresponds roughly to the younger CPS cohort in 2014-2016. The temporal stability of the schooling differences between 2nd- and 3rd+ -generation Mexican Americans and between these groups and non-Hispanic whites suggest that the analogous differences we study in NLSY97 data may largely reflect more permanent forces, such

³² Note that here, as in Table 1 of the main text, the immigrant sample includes foreign-born individuals who arrived in the United States at any age, and so we refer to them as the 1st generation (rather than the 1.5 generation).

as generational assimilation, rather than factors idiosyncratic to particular birth cohorts and immigrant generations.

Some of the temporal stability described above does not hold for the older age group shown in the bottom panel of Table A3. As was also the case for the younger age group, schooling differences between 2nd- and 3rd+ -generation Mexican Americans are small and vary little over time. Schooling gaps between older U.S.-born Mexican Americans and non-Hispanic whites, however, steadily widen as we go farther back in time. Schooling gaps between older blacks and whites also become larger in earlier periods, although less dramatically. The contrasting temporal patterns for the younger and older age groups could reflect civil rights and educational reforms that took place in the United States in the 1960s and early 1970s. Given the time periods examined in Table A3, the younger age groups shown in the top panel have birth years of 1960 and later, whereas almost all members of the older age groups shown in the bottom panel were born prior to 1960.

For both younger and older age groups, Table A3 also shows that the schooling levels of Mexican immigrants have been rising steadily over time, both in absolute terms and relative to the schooling levels of 3rd+ -generation non-Hispanic whites. If we interpret this trend to indicate that Mexican immigrants to the United States have become more favorably selected over time, this would suggest that our cross-sectional NLSY97 comparisons between 2nd- and 3rd- generation Mexican Americans may *understate* the educational gains from generational assimilation. In our NLSY97 sample, the immigrant ancestors of the 2nd generation arrived in the United States much more recently than did the immigrant ancestors of the 3rd generation. If Mexican immigrants indeed have become more favorably selected over time, this would inflate the educational outcomes of the 2nd generation relative to the 3rd generation and bias downward

our estimates of the improvement across generations. In other words, correcting for this type of cross-sectional bias would only strengthen our finding of generational progress in education for Mexican Americans. Alternatively, the rising schooling levels of Mexican immigrants to the United States may only reflect secular increases in average education levels in the source country, rather than changes in the selectivity of immigrants.

Table A1: Average Respondent Characteristics by Attrition and Selection Status, NLSY97 Data

Respondent Characteristic	Unweighted Averages			Weighted Averages		
	Full Sample	Non-Attriters		Full Sample	Non-Attriters	
		All	Known Generation		All	Known Generation
	(1)	(2)	(3)	(4)	(5)	(6)
Male	51.19	50.83	50.40	51.32	50.98	50.67
Female	48.81	49.17	49.60	48.68	49.02	49.33
White	58.24	57.47	57.78	72.41	71.91	71.81
Black	26.58	27.36	26.14	15.71	16.25	15.61
Hispanic	21.19	21.39	22.56	12.87	13.01	13.64
Birth year:						
1980	18.82	18.69	18.99	20.08	19.94	20.35
1981	20.86	20.77	21.13	19.94	19.90	20.20
1982	20.49	20.75	20.27	20.27	20.48	19.86
1983	20.11	20.20	20.08	19.44	19.53	19.52
1984	19.71	19.59	19.53	20.26	20.15	20.07
Family background:						
Lives with both biological parents	48.92	48.11	52.40	52.94	52.12	56.30
Mother's age at first birth	22.85	22.80	22.96	23.28	23.24	23.42
	(0.05)	(0.05)	(0.06)	(0.05)	(0.05)	(0.06)
Mother's age at respondent birth	25.43	25.37	25.56	25.71	25.64	25.86
	(0.06)	(0.06)	(0.07)	(0.06)	(0.06)	(0.06)
Mother's years of schooling	12.44	12.43	12.46	12.85	12.86	12.92
	(0.03)	(0.03)	(0.04)	(0.03)	(0.03)	(0.04)
Father's years of schooling	12.56	12.53	12.56	12.99	12.96	13.03
	(0.04)	(0.04)	(0.04)	(0.04)	(0.04)	(0.04)
Missing schooling data for:						
Mother	7.72	7.74	7.28	6.99	7.02	6.35
Father	20.75	20.98	17.92	16.97	17.15	14.39
Region:						
Northeast	17.64	17.28	16.87	18.45	18.07	17.61
Midwest	22.82	22.69	22.57	26.34	26.25	26.31
South	37.39	38.01	38.19	34.22	34.79	35.23
West	22.15	22.02	22.38	20.99	20.88	20.85
Sample size	8,984	8,271	6,824	8,984	8,271	6,824

Source: National Longitudinal Survey of Youth 1997 data through round 17 (2015-2016).

Note: Standard errors are reported in parentheses. All characteristics are taken from the baseline 1997 survey. The full sample shown in columns (1) and (4) includes all respondents in the 1997 survey. The non-attriters sample shown in columns (2) and (5) excludes 713 respondents lost to panel attrition. The sample for columns (3) and (6) excludes the respondents lost to panel attrition and also excludes an additional 1,447 respondents with missing information on immigrant generation. Sampling weights were used in the calculations where indicated.

Table A2: Average Years of Schooling, by Race/Ethnicity and Immigrant Generation, NLSY97 and CPS Data

Race/Ethnicity and Generation	Average Years of Schooling		Difference Relative to Whites	
	NLSY97	CPS	NLSY97	CPS
Mexican American:				
1.5 generation	11.94 (0.18)	11.70 (0.14)	-2.53 (0.19)	-2.60 (0.15)
2 nd generation	13.01 (0.13)	13.12 (0.07)	-1.46 (0.16)	-1.18 (0.08)
3 rd + generation	13.02 (0.15)	13.01 (0.07)	-1.45 (0.18)	-1.29 (0.08)
Non-Hispanic:				
Black, 3 rd + generation		13.46 (0.04)		-0.84 (0.05)
Black, 4 th + generation	13.35 (0.08)		-1.12 (0.11)	
White, 3 rd + generation		14.30 (0.02)		
White, 4 th + generation	14.47 (0.06)			

Source: National Longitudinal Survey of Youth 1997 data through round 17 (2015-2016) and 2015-2016 Current Population Survey outgoing rotation group data.

Note: Standard errors are reported in parentheses. The samples include men and women whose race/ethnicity and immigrant generation could be identified; see text for further information. Sampling weights were used in the calculations.

Table A3: Average Years of Schooling, by Race/Ethnicity, Immigrant Generation, Age, and Survey Year, CPS Data

Race/Ethnicity, Generation, and Age Cohort	Average Years of Schooling, by Survey Year Interval:					Difference Relative to Whites, by Survey Year Interval:				
	1994- 1996	1999- 2001	2004- 2006	2009- 2011	2014- 2016	1994- 1996	1999- 2001	2004- 2006	2009- 2011	2014- 2016
A. Ages 25-34										
Mexican American:										
1 st generation	9.07 (0.07)	9.45 (0.06)	9.70 (0.05)	9.98 (0.05)	10.71 (0.05)	-4.38 (0.04)	-4.28 (0.04)	-4.16 (0.03)	-4.01 (0.03)	-3.51 (0.04)
2 nd generation	12.45 (0.08)	12.51 (0.07)	12.69 (0.06)	12.71 (0.05)	13.05 (0.04)	-1.00 (0.06)	-1.22 (0.06)	-1.17 (0.05)	-1.28 (0.05)	-1.17 (0.04)
3 rd + generation	12.31 (0.05)	12.52 (0.05)	12.66 (0.05)	12.80 (0.04)	13.06 (0.04)	-1.14 (0.05)	-1.22 (0.05)	-1.20 (0.05)	-1.19 (0.04)	-1.17 (0.04)
Non-Hispanic:										
Black, 3 rd + generation	12.72 (0.02)	12.91 (0.02)	13.03 (0.02)	13.11 (0.02)	13.29 (0.02)	-0.73 (0.02)	-0.82 (0.03)	-0.83 (0.03)	-0.87 (0.03)	-0.93 (0.03)
White, 3 rd + generation	13.45 (0.01)	13.74 (0.01)	13.86 (0.01)	13.99 (0.01)	14.22 (0.01)					
B. Ages 50-59										
Mexican American:										
1 st generation	6.49 (0.16)	7.33 (0.14)	7.70 (0.12)	8.30 (0.09)	9.25 (0.08)	-6.57 (0.08)	-6.23 (0.07)	-6.10 (0.06)	-5.47 (0.05)	-4.59 (0.04)
2 nd generation	10.83 (0.18)	11.44 (0.20)	12.09 (0.16)	12.28 (0.16)	12.85 (0.10)	-2.23 (0.11)	-2.12 (0.12)	-1.72 (0.10)	-1.50 (0.10)	-0.99 (0.08)
3 rd + generation	10.70 (0.17)	11.40 (0.12)	12.11 (0.09)	12.43 (0.07)	12.67 (0.06)	-2.37 (0.11)	-2.16 (0.08)	-1.70 (0.06)	-1.34 (0.06)	-1.17 (0.06)
Non-Hispanic:										
Black, 3 rd + generation	12.00 (0.04)	12.61 (0.03)	12.88 (0.03)	13.01 (0.03)	13.14 (0.02)	-1.06 (0.04)	-0.95 (0.03)	-0.92 (0.03)	-0.77 (0.03)	-0.71 (0.03)
White, 3 rd + generation	13.06 (0.01)	13.56 (0.01)	13.81 (0.01)	13.78 (0.01)	13.85 (0.01)					

Source: 1994-2016 Current Population Survey outgoing rotation group data.

Note: Standard errors are reported in parentheses. The samples include men and women whose race/ethnicity and immigrant generation could be identified; see text for further information. Sampling weights were used in the calculations.