CPIP Working Paper Series Paper #20215

How the 2020 Census Found No Black Disadvantage in Mexico: The Effects of State Ethnoracial Constructions on Inequality

UCI Center for Population, Inequality, and Policy

Christina Sue, University of Colorado
Fernando Riosmena, University of Colorado
Edward Telles, UC Irvine

4-1-2021

How the 2020 Census Found No Black Disadvantage in Mexico: The Effects of State Ethnoracial Constructions on Inequality

Christina Sue, Fernando Risomena and Edward Telles

DRAFT: PLEASE DO NOT CITE WITHOUT PERMISSION

Abstract

A central goal of modern ethnoracial statistics is to measure population size and inequality and an increasing number of countries are including ethnoracial questions on their censuses. Although scholarship has examined how states "make" racial categories and identities via official classification systems, much less attention has been paid in the literatures on stratification and the politics of official ethnoracial classification, to how these classifications affect portraits of population size and, especially, inequality. The Mexican government recently introduced questions on several major surveys and the 2020 Census to measure the black population, generally defining blackness and indigeneity in cultural terms but once in racial terms. We leverage these new data along with an ongoing nationally representative academic survey that has measured race-ethnicity more "neutrally" and consistently to ascertain the implications of these different framings in ethnoracial population size and inequality. After accounting for identification growth over time, we find that cultural frameworks produce a portrait of no black disadvantage, while a racial framework produces substantial black disadvantage, with similar and sizable indigenous disadvantage emerging from both racial and cultural framings. Overall, our study shows how estimates of population size and inequality can be highly dependent on states' conceptualization of ethnoracial racial categories and has broad

implications for the literatures on state ethnoracial classification, stratification, race and ethnicity, as well as for public policy.

INTRODUCTION

In recent decades, an increasing number of states across the globe have added ethnoracial classification to their official statistics to address discrimination and combat inequality (Loveman 2014). New ethnoracial data collection has spawned important conversations about how states make "race"[1] (Marx 1998) or racial categories via official classification systems. However, this literature has paid much less attention to the way official classification systems create portraits of ethnoracial inequality, focusing instead on the politics of official classification (e.g., Angosto Ferrández and Kradolfer 2012a; Loveman 2014; Mora 2014; Nobles 2000; Simon and Piché 2012; Snipp 2003) and various other consequences such as their use by racist regimes (e.g. Aly and Roth 2004) and their effects on claims for collective and individual rights (e.g., French 2009; Hooker 2005; Paschel 2016). Conversely, the literature on ethnoracial stratification generally does not approach the study of inequality through a constructivist lens (i.e., how the social process of constructing ethnoracial categories affects portraits of inequality), instead defaulting to the operationalization of race that is available in a particular dataset (James 2001; Howell and Emerson 2017; Saperstein and Penner 2013). These oversights are problematic on multiple levels. As James (2001) notes: "There are significant political and social costs associated with the continued use of race as a fixed quality in statistical analyses" (236). Moreover, official classification systems do not simply reflect social reality, but they play a central role in constructing that reality (Kertzer and Arel 2002). A significant part of that constructed reality is ethnoracial inequality.

---

[1] While we approach the study of race and ethnicity from a constructivist perspective, we generally drop the use of quotation marks for smoother reading. However, the absence of quotation marks should in no way be understood as an embrace of "groupism" or the reification of "races" (Brubaker 2004).

Self-identification has become the global standard by which states collect ethnoracial data. However, there is no standard for *how* to ask about ethnoracial self-identification (Angosto-Ferrández and Kradolfer 2014b; Del Popolo and Schkolnik 2012; Morning 2008; United Nations 2017: 205). The way in which states frame ethnoracial questions (e.g., via references to race, culture, ancestry) is oftentimes rooted in historical context, national ideology and global power relationships (Simon and Piché 2012). The United States, for example, has included a race question on its census since 1790, as racial division has been a central component of U.S. ideology and practice (Davis 1991). The most recent (2020) census asked, "What is your race?" and provided categorical options which, according to the U.S. Census Bureau, "reflect a social definition of race recognized in this country."[2] Likely because of the country's long history with official racial classification and the expectation that there is popular recognition of the term "race," the question does not define race or list criteria for categorical membership.

In Latin America, the practice of official racial categorization is newer and has been much more contested. In the first half of the twentieth century, Latin American countries gathered racial statistics (treating both blacks and indigenous as "races") to track and display the desired "whitening" of their populations. However, by the mid-twentieth century, direct questions about race were dropped from most censuses and were replaced with measures of culture such as language, food, and clothing to identify indigenous populations (Loveman 2014). This stemmed, in part, from the belief that one could not accurately capture "race" in a scientific sense. The official justification for removing the census question on race was the most extensive in Mexico (Loveman 2014), home to a particularly powerful national ideology of *mestizaje*

[2] https://www.census.gov/topics/population/race/about.html

which proclaimed that "races" had disappeared through admixture (Saldívar 2014; Saldívar and Walsh 2014; Sue 2013). A major consequence of the shift from using racial to cultural measures to understand national diversity was that indigenous populations were symbolically centralized, while black populations were rendered statistically invisible (Loveman 2014).

It has only been in the last few decades, largely in reaction to international pressure and domestic grassroots movements, that most Latin American states have re-introduced/introduced measures to enumerate their distinct ethnoracial populations (Loveman 2014). In some cases, such as the black category in Mexico,[3] this has meant officially recognizing and defining a population for the very first time. Across the region, there is great variation in how black and indigenous categories are constructed, with census questions asking about ancestry, customs, identity, group membership, physical appearance, race, and language (Loveman 2014). The consequences of these categorical constructions can be significant, as ethnoracial population estimates and socioeconomic status have been shown to vary, sometimes widely, depending on question wording and the measure being used (Bailey et al. 2013; Bailey et al. 2014; Bailey et al. 2016; Del Popolo and Schkolnik 2012; Flores et al. n.d.; Howell and Emerson 2017; Loveman et al. 2012; Saldívar et al. 2018; Telles and PERLA 2014; Telles et al. 2015). However, prior studies have been based on small-scale surveys (see Villarreal 2014 for an exception), which are generally too small to generate reliable assessments of ethnoracial variation for some "minority" populations.

[3] Throughout the article we use the term "black," (the English translation of *negro*) as it is the primary term employed in the surveys under investigation and is preferred by some scholars and local organizations such as México Negro (Saldívar and Moreno Figueroa n.d.).

The Mexican case provides a rare opportunity to assess the implications of distinct state conceptualizations of blackness and indigeneity on population estimates and ethnoracial inequality, given that it conducted multiple large surveys in a short span of time and included different formulations of the black and indigenous questions. In this article we analyze three government data sets between 2015 and the 2020 Mexican Census, which used alternative question prompts on race and culture for both the black and indigenous populations.

The 2015 data come from the Intercensus Survey, the first time in the nation's history that the National Statistical Office (INEGI) enumerated the black population, asking if individuals self-identified as black based on their culture. A year later, INEGI conducted the Intergenerational Mobility Module of the National Household Survey (MMSI), this time asking about ethnoracial identification based on race. In 2020, INEGI conducted the Population and Housing Census, asking about black identification, again referencing culture but adding ancestry in the introduction. We supplement these data with the Latin American Public Opinion (LAPOP) surveys which are conducted roughly every two years and use neutral (i.e. do not specify membership criteria) and fixed wording, thus allowing for the measurement of growth in the number of people identifying across ethnoracial categories over time.

To our knowledge, this is the first study to assess the implications of different state constructions of both black and indigenous categorical inequality, while isolating the roles of question wording vs. identification growth. We analyze unprecedented data on black identification, as well as indigenous identification, to ascertain how state classifications framed in cultural vs. racial terms affect portrayals of ethnoracial population size and inequality. After taking identification growth into account, we find that questions which reference culture produce no black disadvantage, while questions using a racial or neutral framing produce significant

black disadvantage, with indigenous disadvantage emerging regardless of question wording. Before presenting our findings, we situate our study within the literatures on the politics of state ethnoracial classification, ethnoracial stratification, and the Mexican context.

The Politics of Ethnoracial Classification

Since the mid-twentieth century, social scientists have increasingly treated "race" not as a concept rooted in biology, but instead, as a social construction. This constructionist perspective has resulted in a body of scholarship focused on ethnoracial boundaries, with a significant emphasis on categorization and classification (Brubaker et al. 2004 Wimmer 2008). Official classifications systems, and censuses in particular, have been of particular interest as they reflect the symbolic power of the modern state (Brubaker et al. 2004); census data are generally "unrivaled in their credibility as a source of knowledge about the conditions and characteristics of national populations" (Loveman 2014: 30). As a form of "racial discourse" – race as created through language and institutional practices – censuses also provide an important rationale for ethnoracial public policies (Nobles 2000).

Scholars have convincingly shown that, far from being "objective" tools used to capture "objective" social realities, official categorizations systems are highly political and ideological in nature (Angosto Ferrández and Kradolfer 2012b; Kertzer and Arel 2002; Loveman 2004; Mora 2014; Nobles 2000; Paschel 2016; Simon and Piché 2012; Skerry 2000). Questions over whether ethnoracial categories should be included in censuses, which classifications systems and categories should be included or excluded, and what criteria should be used to define categories, are contested political processes. Ideology influences this process, determining which categories

are deemed to be legitimate and important divisions in society and how recognized categories are framed.

Scholarship emphasizing the political and ideological nature of ethnoracial classification has primarily focused on the process behind the development of official ethnoracial categories and categorization (e.g., Loveman 2014; Mora 2014; Nobles 2000), with less attention being paid to the consequences and implementation of official classification systems. Studies that have addressed consequences include those on their more heinous use such as by the Nazis during the Holocaust (e.g. Aly and Roth 2004) and the South African government during Apartheid (e.g. Khalfani and Zuberi 2001), as well as their effects on legal decisions regarding immigration and nationality law (e.g. FitzGerald and Cook-Martín 2014), their positive use for enforcing civil rights legislation or distributing reparations (e.g. Berry-James et al. 2020), and their effects on shaping claims for collective and individual ethnoracial rights (e.g. French 2009; Hooker 2005; Paschel 2016). What has been generally overlooked has been the consequences of ethnoracial classification practices on stratification outcomes. However, as Loveman (2014) notes, "the politics and practices of demarcating categorical divides and naturalizing them as group boundaries is *endogenous* to the processes that generate racial inequality and injustice, not exogenous to them" (34, italics in original). In this article we explore the relationship between these two processes, bridging the literatures on the politics of ethnoracial classification and ethnoracial stratification.

Measuring Ethnoracial Stratification and Inequality

Most studies of ethnoracial stratification and inequality take ethnoracial categories for granted, either in philosophy or in practice. Undergirding the former is the belief that ethnoracial

identities and categories are static, a sentiment generally not held by social scientists. This distinction can be understood as those who study race (and understand it as situational and dynamic) and those who use race in analyses (and treat it as a fixed characteristic) (James 2001). Yet even among those who view race and ethnicity as social constructions, general practice separates the study of racial fluidity from the study of ethnoracial inequality, due to convention (James 2001; Saperstein and Penner 2013) and the difficulty of systematically assessing the effect of categorical malleability on inequality outcomes given data limitations or publication space constraints (Howell and Emerson 2017).

There is a small but growing literature on the effects of "microlevel fluidity" – classification change within an individual - on estimates of racial inequality in the U.S. (e.g., Saperstein and Penner 2013). In the Latin American context, where scholars have long contended with the high degrees of ethnoracial fluidity, the research is more extensive. Nevertheless, much of this research has focused on how question wording and measurement affects population size estimates (e.g., Del Popolo and Schkolnik 2012; Flores et al. n.d.), with less attention to their effects on inequality outcomes. Although some scholarship has addressed the effect of measurement type on inequality (e.g., Bailey et al. 2013; Bailey et al. 2014; Bailey et al. 2016; Telles and Lim 1998; Telles and PERLA 2014; Telles et al. 2015), the emphasis has been on variation across types of measures such as skin color, outside classification, and self-identification , as opposed to within a single measurement type, such as self-identification, using official data. One reason for this is likely due to data constraints - these studies are generally based on unofficial surveys with sample sizes too small for a rigorous analysis of broader ethnoracial inequality, especially for black and/or indigenous populations in nations where these

categories represent small minorities of the population. That said, the use of government data to assess the effects of question wording on inequality presents its own set of challenges.

States do not collect large-scale ethnoracial data very often and the years between government censuses preclude the isolation of question wording effects. For example, when the 1993 Colombian census asked people if they belonged to a black community, 1.5% answered affirmatively, but in 2005, when asked if they considered themselves black or mulatto based on cultural or physical features, 10.5% answered affirmatively (Paschel 2016: 133). Similarly, in 2000, 2.0% of Costa Ricans said they belonged to the culture of Afro-Costa Ricans and in 2011, 7.8% said they considered themselves to be Black or Afrodescendant (see Telles and PERLA 2014: 8). Since there was a significant time gap between these censuses, it is hard to identify the source of the shift - it could have been affected by the change in the question wording, self-reclassification and/or demographic change.

Given that ethnoracial data collection is relatively new, few Latin American countries have measured their black or indigenous populations using highly distinct question wording across multiple official surveys in a short span of time. Those that have employed multiple measures have sometimes bundled highly distinct measures into a single question. For example, the 2005 Colombian census asked: "according to your culture, people, or physical traits," making it impossible to disentangle the effects of measures of the three ethnoracial criteria invoked in the question, not to mention that mixing criteria can cause confusion for respondents (Del Popolo and Schkolnik 2012). Whereas Flores et al. (n.d.) examined the effect of changes to the Mexican census question on indigenous identification between 2000 and 2010 using a survey experiment design, such data are exceedingly rare. Our study contributes to these conversations on the effects of official question wording by leveraging the close succession of Mexican government

surveys with different question types for both black and indigenous identification, expanding the scope of Flores et al.'s study by looking at racial vs. cultural conceptions of indigenous identification, the presence of a *mestizo* option in some questions, and the impact of various question formulations on inequality outcomes.

We also build on two studies that directly assess the relationship between ethnoracial measurement type and inequality. Loveman et al. (2012) draw on a Brazilian survey to measure the effects of a binary (black/white) system versus a three-tiered system with a "mixed race" category. They find that, when the mixed-race category is removed, racial inequality is greater. This study focused on categorical options, as opposed to the framing of the question itself. We engage both aspects of question wording in this study. We also expand on the work of Villarreal (2014) who used the 2010 Mexican Census to examine how (proxy) indigenous self-identification framed in cultural terms vs. classification based on indigenous language proficiency affects the estimated size and socioeconomic status of the children of indigenous parents. He found that children of indigenous-language-speaking parents were much less likely to be classified as indigenous by the parent informant when language criteria were used compared to self-identification measures, particularly for children whose parents had higher levels of education. However, those with higher-educated parents were more likely to be classified as indigenous when (proxy) self-identification was used. Because of this, not only did Villarreal find that linguistic criteria reduced estimates of indigenous population size, but that linguistic measures revealed higher levels of socioeconomic disadvantage for children of indigenous parents. We build on this work by examining the effects of multiple measures of both indigenous and black identification on adult inequality, introducing novel data on Mexico's

black population and broadening the examination on the indigenous beyond the cultural realm and into racial framings.

Self-Identification and Official Classification Systems

Estimates of population size and inequality depend on how everyday people understand and interpret official ethnoracial questions and categories. In other words, whereas the construction of official classification systems can be understood as a state-level process (sometimes in conversation with non-state entities), the data produced from these systems reflect the interaction between official questions and popular self-classification practices. The creation of a census category does not automatically translate into the existence of a self-conscious social and political "group" organized around that category (Brubaker 2004; Mora 2014; Nobles 2000). It is therefore necessary to understand how individuals attach meanings to social categories and how those meanings interface with official measures of race and ethnicity (Hitlin et al. 2007).

Popular distinctions and understandings of ethnoracial categories are crucial to the production of population and inequality estimates, given that census and most official survey data is based on self-identification. Everyday understandings of what it means to be black in Mexico is primarily based on phenotype (Flores Dávila 2006; INEGI 2019; Lara 2014; Resano Pérez 2015; Sue forthcoming), unlike contexts such as the U.S. where ancestry plays an important role, and historically the primary role, in categorization as black (Davis 1991). Although having black/African ancestors influences black identification in Mexico to some degree, ancestry is a poor proxy for black self-identification – between a quarter and a third of black-identified respondents across three separate surveys did not claim black/African ancestry (INEGI 2019; Resano Pérez 2015; Sue forthcoming) and this percentage is much higher (~85%)

in non-black localities (Resano Pérez 2015). Regional identities also play an important role in the construction of blackness (Hoffman and Rinaudo 2014), with some individuals claiming a black-related identity because they live in a region that is associated with a black presence.

Unlike cases such as Colombia, and to a lesser extent Brazil (Paschel 2016), blacks in Mexico were not granted cultural and territorial rights as part of the multiculturalist and racial inequality reforms beginning in the 1980s, and they have not sought to distinguish themselves based on cultural difference, at least on a national scale (Hoffman 2014). Most Mexicans who identify as black do not do so based on cultural characteristics, even in regions with high concentrations of self-identified blacks (Resano Pérez 2015). Even following the extensive government campaigning and promotion of the black "cultural" questions in the 2015 Intercensus and 2020 Census, only 15.9% of respondents in INEGI pilots tests reported identifying as black based on their "culture" and 10.9% based on their "traditions and customs" (INEGI 2019). Field test reports further showed that INEGI's own interviewers struggled to explain the black question when respondents were confused, and ended up describing blackness in terms of "genetic or phenotypic" characteristics (Ruiz Ramírez 2014). Finally, in a 2019 nationally-representative survey, individuals were more likely to write in a response of "black" when asked to identify their "race" compared to when asked to identify their "ethnic group" in an open ended forma (Solís et al. 2020).

With regards to indigeneity, consistent with international organizations, the Mexican government currently defines indigenous peoples under the rubric of "ethnicity."[4] At the popular

[4] https://www.inegi.org.mx/contenidos/productos/prod_serv/contenidos/espanol/bvinegi/productos/nueva_estruc/702825197520.pdf. Last accessed March 17, 2021.

level, the interpretation of indigeneity appears to be influenced by whether or not the pan-ethnoracial term "indigenous" is used, compared to specific indigenous labels (e.g. Náhuatl, Maya, Zapotec, Mixtec). The 2019 survey found that respondents were more likely to identify with a specific indigenous category when asked to identify their "ethnic group," but were more likely to identify as "Indian" when asked to identify their "race" (Solís et al. 2020). This may suggest that indigeneity, broadly speaking, is seen in both racial and cultural terms and can shift depending on whether indigeneity is operationalized as a pan-ethnoracial category or specific indigenous categories. The notion that indigeneity is understood in both racial and cultural terms is supported by findings from a regional survey, where respondents were asked to describe the characteristics associated with indigeneity - although cultural markers (e.g. language, dress) were mentioned most frequently, phenotypic markers (e.g. dark skin, straight hair) were also mentioned at high rates (Sue forthcoming).

In addition to the way in which ethnoracial questions and categories are framed and understood, various factors such as societal treatment, the status of an ethnoracial category, national ideology, cultural or identificational revitalization, and institutional incentives tied to categorical membership can affect how people choose to classify themselves on official forms. Data based on self-identification as black or mulatto, for example, does not necessarily reveal socioeconomic disadvantage (Telles et al. 2015) as they do not necessarily capture how others classify them and treat them socially (Telles and Lim 1998). Instead, self-identification may reflect racial ideology, assimilation, political and cultural attachments, and social aspirations (Telles et al. 2015). Lower status individuals socially viewed as black may opt out of the category to avoid stigma or discrimination, or as a strategy for upward mobility. In contrast, higher-status individuals may disproportionately opt into these categories, as is the case of

mulattoes in the Dominican Republic, and to some degree, blacks in Brazil, due to racial consciousness raising and institutional incentives to identify as such (Telles and Paschel 2014; Francis and Tanauri-Pianto 2013).

In the case of Mexico, black Mexicans generally do not have access to development programs and other institutional systems of support. Thus, there are currently minimal to no political, ideological, or material incentives to identifying as black in Mexico (Hoffman 2014). At the same time, there is a social stigma associated with the black category (Hernández-Cuevas 2004; INEGI 2019; Sue 2013). Villarreal and Bailey (2020) argue that these produce endogenous "self-selection" into measures of black self-identification and can underestimate black socioeconomic disadvantage. Self-identification measures may be particularly unsuitable for measuring ethnoracial inequality in contexts in which categories are new and/or hold inconsistent meanings across a population, such as the case of blacks in Mexico. That said, self-identification measures can vary widely, and some may be more problematic than others.

The Mexican Case

Historical context and racial ideology play a central role in shaping countries' approaches to ethnoracial data collection and the conceptualization of ethnoracial categories. Mexico's post-Revolutionary ideology of *mestizaje* (1921-) emphasized the Spanish and indigenous contributions to the formation of the Mexican nation, even though colonial Mexico was the destination for at least 200,000 enslaved Africans (Aguirre Beltrán 1944). It glorified the country's race mixture and proclaimed the eventual biological integration of Mexico's indigenous population resulting in a superior *mestizo* race (Vasconcelos 1925[1997]). The black population was assumed to have largely disappeared through biological and cultural integration;

continued *mestizaje* through "voluntary extinction" was expected to lead to the complete absorption of blacks into *mestizos* (Vasconcelos 1925[1997]: 32). Although racialized thinking undergirded the ideology (and continues today, see Wade 2017), national leaders oftentimes downplayed the notion of "race" in favor of culture, emphasizing language and other cultural markers to understand national diversity (Martínez et al. 2014; Saldívar 2014, 2018).

The emphasis on culture and public downplaying of race was visible in the nation's census practices. A "race" question only appeared on one census, in 1921, and included the categories of Indian, mixed, white, other, and foreigners without distinction. The question was dropped in the 1930 Census, like other Latin American countries during that time (Loveman 2014). However, Mexico produced the most elaborate justification for the question's removal, arguing race was an "antiscientific concept" and irrelevant given the country's high degree of race mixture (Loveman 2014). Officials announced that, in lieu of a race question, two additional language questions would be added to the census to acquire "precise knowledge of the process of national integration . . ." (cited in Loveman 2014: 227). This justification is illustrative of the region's mid-twentieth century emphasis on "cultural progress" (Loveman 2014) and long-standing debates over how to define indigenous peoples (Rosemblatt 2018). Subsequent censuses in Mexico included additional questions about "material culture," also reflecting national interests about ethnicity and development (Saldívar and Walsh 2014).

In the later decades of the twentieth century, the focus on indigenous populations shifted from integration to a concern over continued indigenous marginalization (Saldívar and Walsh 2014). At the turn of the 21st century, following new international norms, the Mexican Census Bureau changed the criteria for indigenous measurement to self-identification, but retained the focus on ethnicity, asking people if they considered themselves indigenous based on their

"culture" (Saldívar and Walsh 2014). The ideological centering of *mestizaje*, culture and indigeneity resulted in the symbolic and statistical erasure of Mexico's black population (Loveman 2014; Sue 2013).

The statistical invisibility of black Mexicans was broken in 2015 with the inclusion of a black self-identification question on the Intercensus (EIC). This shift occurred largely in response to pressure from international organizations, academics, government institutions such as the National Council for the Prevention of Discrimination (CONAPRED), and domestic black movement organizations. The resurgence in ethnoracial census classification across Latin America, starting in the late 1980s, was driven by the creation of new and more democratic relationships between states, ethnoracial minorities, and broader publics in Latin America, coupled with new recognition of ethnic and racial minority rights (Loveman 2014). During this period, multilateral organizations such as the United Nations, the World Bank, and the Inter-American Development Bank made the social and structural inclusion of ethnoracial categories central to their discourses and policies (Angosto-Ferrández and Kradolfer 2012b; Loveman 2014) and called for the collection of ethnoracial data to pursue their agenda (Loveman 2014). Although Mexico had consistently collected data on its indigenous population, the U.N. International Convention on the Elimination of All Forms of Racism urged the country to also collect information on its black communities, in part because of Mexico's official stance that racism did not exist in Mexico (Sue 2013).

Domestic grassroots organizing and mobilizations also drew attention to Mexico's black population. The 1990s marked "a turning point in ethnic relations in Mexico" (Saldívar and Walsh 2014: 469) not just in terms of indigenous rights but also in terms of organizing around black identification (Lara 2014; Vaughn 2013). It was thus the joining of international and

domestic pressures that ultimately led to the collection of new ethnoracial data beginning with the 2015 EIC and, eventually, the 2020 decennial census.

New Ethnoracial Classification in Mexico

Not only are countries faced with a decision of whether to count an ethnoracial population, but equally important is their decision of *how* to count (Simon and Piché 2012). Terminological choices, oftentimes rooted in political struggles, can greatly condition the results produced by official sources (Angosto-Ferrández and Kradolfer 2012b; Del Popolo and Schkolnik 2012; Skerry 2000). In the case of Mexico, when the decision to introduce a black question was finally made, debates surfaced regarding the conceptualization and formulation of the question. Various actors including census officials, national and international experts, federal, state, and municipal representatives of civil society organizations came together to discuss the issue of question wording (Resano Pérez 2015; Ruiz Ramírez 2014; Saldívar and Moreno Figueroa n.d.). INEGI also conducted extensive pilot tests of multiple versions of a black question. Based on these field tests, INEGI pointed to four challenges with identifying the black population in Mexico: the unfamiliarity with the topic of blackness in Mexico; the idea that Afrodescendants are foreigners; the fact that terms such as "negro" can be considered offensive or discriminatory; and the finding that terms such as "Afromexican" are unfamiliar, and people do not understand them (Resano Pérez 2015; see also Ruiz Ramírez 2014). Notably absent in the pilot test variations was an alternative to a cultural framing of blackness (Ruiz Ramírez 2014), suggesting that this framing was not up for debate. This could be interpreted as another instance of Mexico's particularly top-down approach to ethnoracial reform, its continued emphasis on conceptualizing diversity within an indigenous/*mestizo* framework based on cultural distinction,

and its relative reluctance to employ the language of "race" and "racism" driven in part by its longstanding attention to the country's much larger indigenous population, compared to that seen in other major countries in the region such as Brazil and Colombia (Wade 2017)

Consistent with the dominant ethnic group framework associated with the indigenous category in Mexico (Saldívar 2014; Saldívar and Walsh 2014), INEGI defined blackness in the 2015 Intercensus and 2020 Census in cultural terms. In fact, a major 2011 initiative tied to the enumeration of the black population was coordinated by the National Commission for the Development of Indigenous Communities (Resano Pérez 2015), as "Afromexican communities" are now under the Commission's purview. The emphasis on ethnic criteria is clearly articulated in INEGI's Intercensus "Conceptual Framework" document: the stated purpose of the "Afrodescendent" question is to capture "ethnic identity" using "culture as a link for said identity." The document further notes its objections to a racialized understanding of blackness: "we are careful to avoid other criterion related to phenotypical or genetic characteristics."[5]

The 2016 MMSI noticeably diverged from the cultural framework by including a question on "racial origins." Although the MMSI was administered by INEGI, the process behind the development of the MMSI question was quite distinct from that of the other surveys. While EIC and Census were initiated and designed by the Census branch of INEGI in consultation with other state entities such as CONAPRED and the National Institute of Anthropology and History (INAH), the MMSI "race" question was designed by Mexican sociologist and social stratification scholar, Patricio Solís, at *El Colegio de Mexico*. Solís modeled the question after a similar one from the Project on Ethnicity and Race in Latin

[5] https://www.inegi.org.mx/app/biblioteca/ficha.html?upc=702825098742

America, a study on ethnoracial classification and inequality (Telles and PERLA 2014)."[6] For more details of the MMSI study and its findings, see Solís et al. (2019).

The 2020 Census once again adopted a culture framework for the black question, like the EIC, while also adding a reference to ancestry. Consistent with the EIC, the 2020 Census conceptual framework and interviewer manual documents frame Afrodescendancy in terms of ethnicity.[7] They specifically caution against a phenotypic understanding of black identification: "being Afromexican, black, or Afrodescendant does not imply having a particular skin color or hair texture. For this reason, the question establishes ancestry, customs, and traditions as elements of identification instead of skin color."[8]

Although the 2015 Intercensus and 2020 Census questions were quite similar, INEGI made some important modifications to the question wording, providing the following justifications[9]: They substituted the term "ancestors" for "history" because it relates to the basic dimensions of an "ethnic group" via a reference to common origin. They replaced "culture" with "customs and traditions" because informants better associate this phrase with "situations such as form of dress, food, music, and festivities of persons who are of Afromexican or

---

[6] Personal communication with Patricio Solís

[7] https://www.inegi.org.mx/contenidos/productos/prod_serv/contenidos/espanol/bvinegi/productos/nueva_estruc/702825197520.pdf

[8] https://www.inegi.org.mx/contenidos/programas/ccpv/2020/doc/Censo2020_manual_entrevis_cuest_b.pdf

[9] https://www.inegi.org.mx/contenidos/productos/prod_serv/contenidos/espanol/bvinegi/productos/nueva_estruc/702825197520.pdf

Afrodescendants." Unlike the EIC, the 2020 Census listed "Afromexican" first, followed by *negro* and Afrodescendant, with the intent of making the question clearer for respondents, although pilot testing showed that ethnonym order did not have much impact on levels of self-identification (Ruiz Ramírez 2014).

In debates over question wording and evaluations of pilot test results, what was notably absent was a discussion of the consequences of different question formats on estimates of ethnoracial inequality. The lack of discussion is starkly at odds with publicly stated goals of ethnoracial enumeration: to identify and track inequality and develop policies to remedy existing inequality (Del Popolo and Schkolnik 2012; Loveman 2014; United Nations 2017: 205). As we will show, there are important consequences to question wording, specifically whether the black category is conceptualized in racial versus cultural terms, with regards to national portraits of ethnoracial inequality.

Black Disadvantage in Mexico?

Although there are no official data on black Mexicans prior to 2015, a few earlier studies strongly suggest black marginalization (CONAPRED 2011a, 2011b; Flores Dávila and Lézé 2007), which co-exists with the much more widely and consistently documented indigenous marginalization (e.g. Bonfil Sánchez et al. 2017; CONEVAL 2014; Villarreal 2014). There is ample evidence demonstrating that blacks suffer numerous forms of discrimination both within and outside of historically black regions of the country (CONAPRED 2011a, 2011b; Cruz Carretero 1989; Velázquez and Iturralde Nieto 2016; Sue 2013; Sue forthcoming; Vaughn 2001). Anti-black discrimination is almost certainly related to the broader phenomenon of skin color discrimination (Sue 2009, 2013) which helps explain color-based educational and socio-

economic inequality in Mexico (Arceo-Gomez and Campos-Vazquez 2014; Flores and Telles 2012; Martínez Casas et al. 2014; Telles et al 2015; Villarreal 2010; Telles et al. 2015).

A few scholars have conducted analyses on black inequality using the 2015 EIC, with mixed results. Torre-Cantalapiedra and Sánchez-Soto (2019) found a small black *advantage* in educational and occupational attainment in bivariate statistics as well as models controlling for other characteristics, including indigenous identification. However, given the historical erasure of blackness in Mexico and the newness of the black category in state statistics, black identification could be endogenous to attainment, which Villarreal and Bailey (2020) examined, also using the EIC. Leveraging state-level variation in a state-sponsored campaign to promote awareness of the black identification question as an instrumental variable, Villarreal and Bailey showed that a lack of black disadvantage from more conventional models treating ethnoracial identification as exogenous, reverses when treating it as endogenous. They argue that because an official black identification measure is new in the Mexican context, it may be particularly susceptible to selection effects related to the endogeneity between socioeconomic status and black self-identification. While their analysis provides significant insight into discussions of ethnoracial inequality in Mexico, their findings could be specific or strongly related to the EIC's cultural measure of blackness.

Our research extends Villarreal and Bailey's by examining the degree to which surveys conducted somewhat after the campaign and in the context of continued (and perhaps accelerated) awareness of blackness and ethnoracial inequality in Mexico affects estimates of ethnoracial inequality. In addition, by attempting to disentangle and estimate the contributions of growth in ethnoracial identification from question wording, our research builds on Villarreal and Bailey's by further asking whether cultural measures of blackness are likely to (underestimate)

inequality differently from other measures. While we do not assume any question is free from endogeneity, our findings suggest that endogeneity may be less severe in post-campaign estimates using racial wordings, but not in ancestry-culture ones even five years and additional campaigns after the EIC. Before presenting these findings, we provide a detailed description of the datasets we draw upon.

DATA

We analyze three major surveys conducted by Mexico's INEGI between 2015 and 2020 along with data from one continuing academic data collection effort. Importantly, the state/regional and urban-rural representation of the state data is overall good-to-excellent, which is important due to the large variation in black identification and inequality across these geographies. The EIC was a mid-Census count sampling 6.1 million households in 2015 and was designed to be representative nationally and for five different locality sizes within each of the 32 Mexican states (INEGI 2015). The MMSI, also conducted by INEGI, was collected throughout the second semester of 2016, and designed to be representative at national, state, and urban/rural levels, interviewing 32,000 households (INEGI n.d.-A). One adult age 25-64 was randomly selected per household in the MMSI survey. Because some households did not include people in this age range, the final MMSI sample size consists of 25,500 individuals. The 2020 Census long-form is a 10% sample of the Mexican population centered in mid-March and designed to be representative at national, state, and municipal levels as well as for all places with more than 50,000 inhabitants and, within each state, across four different locality sizes (INEGI 2021). Finally, to examine temporal change in identification and disadvantage using a different question

wording, we draw on the 2004-2019 Latin American Public Opinion (LAPOP 2021)[10] surveys of Mexico, which are nationally representative and draws cross-sections of 1,500 respondents roughly every two years.

Specific Ethnoracial Questions

We examine the different question wordings for the three government surveys (see Appendix III for exact wording and options) and we create three categories from each: 1) black, 2) indigenous, and 3) non-black, non-indigenous.[11] In Mexico's first ever official count of a black population on the EIC, INEGI included separate questions on black and indigenous identification using "cultural" prompts in both: "*In accordance with their culture, history, and traditions, does [name] consider themselves black, meaning Afromexican or Afrodescendant?*" and "*In accordance with their culture, does [name] consider themselves indigenous?* For the sake of simplicity, we thus refer to the EIC as "cultural" questions. Both EIC questions had two affirmative response options – "yes" and "yes, in part"– which were not read, but deduced by the interviewers based on interviewee responses (INEGI n.d.-B).[12] Roughly 30% of all blacks and 7.5% of the indigenous were classified as "in part." Because of this external attribution and the fact that our analyses revealed very small differences in the sociodemographic profile of full or

[10] LAPOP is also known as the AmericasBarometer.

[11] Another important category for the analysis of inequality is black-indigenous (Sue and Riosmena forthcoming) but we do not include that category because not all of the surveys allow for identification in both categories.

[12] The INEGI manual instructs interviewers as follows: If the response is "Yes," circle 1. If they comment "it could be because my father is, but my mother isn't", "I'd say a little bit," or something similar, circle option 2: "Yes, in part."

partly black and indigenous persons (except for region of residence, see Appendix I), our EIC black and indigenous categories henceforth always includes those answering "yes" and "yes, in part."

In the 2020 Census, INEGI also asked about black and indigenous identification separately. While it also used a cultural question for indigenous identification (*"In accordance to their culture, Does [name] consider themselves indigenous?"*), INEGI used an "ancestry-culture" prompt (*Given their ancestors and in accordance with their customs and traditions, does [name] consider themselves Afromexican black Afrodescendant?"*) to measure black identification. The response options for both questions are yes or no.
We describe this as an "ancestry-culture" prompt but for the purposes of our conceptual framing, we treat this as a type of cultural question.

In notable contrast to the EIC and 2020 Census, the 2016 MMSI used a single question on ethno-racial identification, which referenced racial origins and offered five mutually-exclusive options: "*In our country there are people of multiple racial origins. Do you consider yourself to be… a person [read options] black or mulatto* [single option], *indigenous, mestizo, white, or other race?*" The use of a racial question was highly irregular in the Mexican context but was fortuitous for the purposes of comparing the effects of various categorical constructions. Finally, in a similar fashion, LAPOP uses a single question with a neutral introduction in every cycle (*do you consider yourself to be a white, mestizo, indigenous, black, mulatto, or other person*), with identical options as those presented here in the 2010 through 2019 surveys). We

combined black and mulatto for our afro-descendant category.[13] This stability in wording and, for the most part, options allow for an examination of temporal change. And because LAPOP includes very similar options to the MMSI, a comparison between these surveys will allow us to understand whether a racial introduction might have made a difference relative to a neutral prompt, or if the categorical options in LAPOP operate in similar ways as those offered in the MMSI. As we elaborate below, even though the LAPOP question has a neutral question format, because of the categorical options and the literature which suggests that blackness is mainly understood at the popular level as a reference to phenotype (Flores Dávila 2006; INEGI 2019; Lara 2014; Resano Pérez 2015; Sue forthcoming), we suspect that individuals likely interpret the LAPOP question within a racial framework.

Ensuring Comparability

We took several steps to ensure that our analyses across surveys were as comparable as possible. Since the MMSI only sampled individuals aged 25-64, we restrict our samples in all data sets to that age range. Second, while the MMSI only includes data on main informants, the EIC and 2020 Census data include all household members and do not identify the main informant.[14]

---

[13] Compared to mulattos, black individuals had lower socioeconomic status in terms of the two main measures we use in our analyses (schooling and an amenity index). These differences were not marginally significant at $p<0.05$ (but at $p<0.1$). However, note that mulatto-identified individuals (also) exhibited higher levels of socioeconomic status in a study in the Costa Chica of Guerrero and Oaxaca (Sue, forthcoming), suggesting the observed patterns in the LAPOP may not be due to sampling error.

[14] INEGI denied our request to access these data via a restricted-data center, stating this information is not an official part of the statistical record for the EIC.

Third, since sampling frames varied among the surveys and sampling design or other differences in data collection could be affecting our assessment of the role of wording on ethnoracial population size and inequality, we compared socio-demographic and state-level compositions across surveys. We find only very small differences (see Appendix II), with an observed higher educational attainment in 2020 likely related to increases in schooling among younger cohorts (e.g., Creighton and Park 2010). Thus, any discrepancies in ethnoracial self-identification across the three surveys are unlikely to be driven by differences in their design or coverage. We further explain our analytical strategy as we present our findings.

RESULTS

Afro-Mexican Growth and Adjusting for Changing Identification

Before proceeding to our analysis of inequality, we examine the changing size of the black population since 2004, with a focus on changes since 2015 when official data on blacks became available and when sizable variation in black identification occurred. As shown in the shaded bars in Figure 1a, 1.8% of Mexicans ages 25-64 self-identified as black with the 2015 EIC cultural yes/no question, while 2.6% did so using the 2016 MMSI racial, multiple-choice question; and 2.14% did so with the 2020 Census ancestry-culture yes/no question (all statistically significant from each other at $p<0.001$- see Appendix III). Extrapolating to the Mexican population of all ages, this means that, compared to 2015 when the EIC asked a cultural question, one million additional Mexicans self-identified as black in 2016 when asked the racial

question (a 44% increase), though this figure is lower using the 2020 Census, which shows a growth of 22% in the black population.

-FIGURE 1a ABOUT HERE-

Actual demographic growth in the short (average) 18.5 months spanning the EIC and MMSI is unlikely to account for a significant part of the increase in the estimated size of the Afro-Mexican population. As discussed in the prior section, we also ruled out differences in the demographic coverage of the surveys. Rather, the larger shares in the MMSI relative to the EIC are likely the result of re-classification, either due to changes in question wording or growth in black self-identification overall. Regarding the latter explanation, changes in the international and national arenas were occurring during this time. The United Nations declared 2015-2024 to be the International Decade for People of African Descent and Mexico engaged in awareness campaigns on Mexico's black population (see Villarreal and Bailey 2020 for a discussion of the effect of awareness campaigns on black identification).

While comparison across official surveys fielded in a relatively short time suggests that question wording may be driving the differences, an examination across LAPOP cross-sections also reveals a substantial increase in black identification between 2014 and 2017 and stability thereafter,[15] perhaps suggesting that comparisons between EIC and MMSI in particular could be clouded by growth in identification.

[15] LAPOP has eight available surveys, all with an ethnoracial question with the same (neutral) introduction and very similar options across years, and identical in wording since 2010 (see Appendix III). These data clearly show increasing black identification at three points: first, around 2008-2010 (1.4%-2.2% from 0.4% in 2004 and 0.9% in 2006, $p<0.05$ and $p<0.01$ respectively); then, circa 2014 (2.1%, from 1.4% in 2008, $p<0.001$); and, finally and perhaps most notably, around 2017 (4.6%, up from 2.1% in 2014, $p<0.001$), whereupon it remained stable in 2017-2019 ($p>0.05$).

To disentangle question wording from identification growth in the EIC vs. MMSI, we decomposed the EIC-MMSI differences into growth vs. wording contributions, leveraging the observed growth between the LAPOPs around the time of these surveys. We used 2010-2014 and 2017-2019 as the beginning and end points to estimate growth. Given that there was no growth in black identification within each of these two intervals, we pooled cross-sections for each of these periods to increase the precision of our estimates. To first estimate changes in the EIC only due to growth, we constructed a counterfactual of what the percent of black identification would have been around the time of the MMSI had interviewers used the EIC question (also shown in Figure 1a). To do this, we projected the mid-March 2015 EIC estimate to the dates of the MMSI (early-October 2016), assuming the same mean annualized growth in black identification observed between 2010-2014 and 2017-2019 LAPOPs given the stable wording in these different cross-sections.

Figure 1a presents the projected figures in gray bars. After accounting for the growth in black identification, we estimate that the EIC question would have resulted in 2.1994% and 2.493% of the Mexican population identifying as black in 2016, respectively. As shown in more detail in Table 1, these estimates suggest that a little over half of the EIC-MMSI difference (52.7%) is due to growth in identification. Having previously ruled out that differences in survey coverage could be affecting our estimates, we estimate the wording contribution as the complement/residual of this difference as the likely contribution of wording (i.e., for the MMSI-EIC $100 \cdot [1-0.527]=47.3\%$).[16] Table 1 shows these estimates along with bootstrapped 95%

[16] This figure would have been considerably larger had we used the rate of growth between 2015 EIC and 2020 Census instead of that observed in LAPOP, suggesting our estimates of the impact of question wording

confidence intervals for all Mexican adults ages 25-64 as well as for some key sociodemographic strata with this population, calculated based on strata-specific growth rates in identification in LAPOP.[17]

At the national level, our results suggest that the change from cultural to racial wording for the 2015 and 2016 surveys was just as important as expected growth for counting the black population. However, because the confidence intervals around these shares are wide for both differences, analyses of these contributions within sociodemographic strata where black identification is most common or in sociodemographic categories more closely related to socioeconomic status were more informative. Before discussing these findings, however, we discuss shifts in indigenous population size.

Indigenous Population Size

Figure 1b shows our results for the share of the population identifying as indigenous across the different surveys. First, comparing within similar question wordings/options, we estimate that indigenous identification increased substantially between 2010 and 2015 and likely declined a bit thereafter. As shown in the figure, INEGI surveys using cultural questions suggest substantial growth in indigenous identification between the 2010 Census (14.4%) and the 2015 EIC (23.5%), a 63% relative difference. Consistent with this, indigenous identification in

---

on differences in black identification between EIC and MMSI are fairly conservative and, thus, may be lower bounds.

[17] Full details on this procedure and how we generated bootstrapped confidence intervals for these contributions are available in Appendix IV. We also calculated estimates based on key sociodemographic data based on based on group-specific changes in black identification across LAPOPs (see Appendix IV)..

LAPOP increased from 7%-9% in 2006-2012 to 11%-12.5% in 2017-2019, a 40%-55% difference between the closest and most distant values in these two ranges. Finally, note that the share identifying as indigenous in the 2020 Census was lower than in the EIC (19.3%). The larger share in the EIC is likely partly explained by the unique inclusion of an "in part" option in the EIC (close to 2% of individuals in our working sample). However, even if we assume that those identifying as "in part" would have not identified as indigenous if this option were absent, there is still evidence of sizable growth in identification between 2010 and 2015. Furthermore, the slight decline in indigenous identification in the 2020 Census could potentially be attributed to the fac that these two surveys -unlike the 2010 Census and 2015 EIC- asked about indigenous language ability before self-identification, which might have increased the perceived "threshold" for indigenous identification (see Flores et al. n.d.).

Notwithstanding the growth within similar wordings, question format/options seem to be an even clearer explanation behind the estimated size of the indigenous population (see also Flores et al. n.d.). Estimates of indigenous identification from surveys taken around similar periods but with different wordings/options differ considerably. Questions on ethnoracial identification with multiple mutually exclusive options like the 2014-2019 LAPOP and the 2016 MMSI yielded similar shares identifying as indigenous (ranging from 12%-14%) that were also somewhat lower than the cultural questions in the 2015 EIC and 2020 Census (ranging at 19%-24%). This large difference is likely due to the presence of a *mestizo* option, generally understood as an indigenous-European admixture.

-FIGURE 1b ABOUT HERE-

Ethnoracial Identification within Sociodemographic Categories

In addition to estimating identification growth vs. wording contributions for the general adult population, we also did so for key sociodemographic categories where self-identification may differ, or which may translate into important differences in socioeconomic inequality (Solis et al. 2019; Telles and Paschel 2014; Torres-Cantalapiedra and Sanchez 2019; Villarreal 2014; Villarreal and Bailey 2020). Furthermore, as mentioned, since the confidence intervals in Table 1 around the contributions of wording and growth are somewhat wide, our analysis of the contributions of growth and wording for black identification within specific sociodemographic categories is perhaps more informative because, in some important cases, our analyses yielded narrower intervals due to lower variance in black identification within *some* of these strata than across strata.

Based on both levels and interval range in Table 1, wording is most clearly relevant for explaining EIC-MMSI differences in black identification among men, younger adults (ages 25-44), people with 0-11 years of schooling, and rural residents. In all these cases, the estimated contribution of wording is well over 50%, with its 95% confidence interval lower-bound of at least 46%. The relevance of wording in rural areas and among lower-educated individuals for the EIC-MMSI difference are of particular relevance because of the traditionally strong association between rural/urban residence and schooling and (other) measures of socioeconomic status.

To further illustrate how wording may produce a different sociodemographic profile, we plot the percentages of black (Figure 2a) and indigenous (Figure 2b) Mexicans, also including data from the pooled 2010-2019 LAPOP to further contextualize comparisons given some similarities between the MMSI and LAPOP questions in options. Figure 2a illustrates that while there is little sociodemographic variation in black identification for the cultural questions in the 2015 EIC and 2020 Census across many different strata, there is considerable variation in black

identification across some sociodemographic categories in the neutral and/or racial questions. Younger people are more likely to identify as racially black relative to older individuals. Perhaps more importantly, Rural residents and lower-educated adults are also more likely to identify as neutrally/racially or racially black than urban residents and higher-educated individuals. As discussed, when describing Table 1, differences in the urban-rural and schooling profile of individuals identifying as black in the EIC vs. MMSI are more likely to be due to question wording than to growth in black identification.

-FIGURE 2 ABOUT HERE-

In contrast to black identification and to the impact of question wording on the overall estimated size of the indigenous population, Figure 2b shows that that sociodemographic "gradients" in indigenous classification are similar for neutral (LAPOP), racial (MMSI) and culture (EIC and Census) questions. This indicates that, while question wording is overall influential on the extent to which people identify as indigenous and there are sharp sociodemographic differentials in who does so, question wording does not appear to produce different sociodemographic profiles. Regardless of wording, we see overall similar urban-rural and educational differentials, which foreshadows our findings related to socioeconomic inequality.

Ethnoracial Inequality

To illustrate the consequences of question type on estimates of ethnoracial inequality, we assess two outcomes: (1) schooling, as a measure of chances earlier in life and current earning potential; and (2) a normalized index of household assets/amenities, to approximate household wealth as commonly used in developing countries where income is often unreliable or

unmeasured (Filmer and Pritchet 1999; Solís et al. 2019; Telles and Torche 2019).[18] We present estimates of ethnoracial differences from OLS models for these outcomes in panels A and B in Table 2. Each column within each panel shows differences in the outcome for black and indigenous categories compared to the same outcome for non-indigenous, non-blacks in the same survey.[19] In the panel rows, we present results of ethnoracial inequality derived from different models for which we introduced relevant controls, either compositionally or as proxies for processes producing ethnoracial inequality. The first model, in the first row of each panel (a0 or b0), only includes ethnoracial identification as an "independent" variable. For the next four models in each of the subsequent rows in each panel, we controlled for only one sociodemographic variable at a time (gender, age, urban residence, and state), in addition to ethnoracial identification. In models predicting the household amenities index, we added a model that controls for years of schooling (b9). Finally, we present full models with all other control variables. In all analyses except for those with LAPOP, due to lack of consistent data availability, we also included models adding indigenous language individually and after adding all controls. We summarize these results for two of the models (no controls, all sociodemographic controls) for schooling (Figure 3a) and for household amenities (Figure 3b)

[18] We constructed this index based on whether the surveyed household had a refrigerator, washing machine, stove, radio, television, computer, telephone land line, cell phone line, internet access, and a car. Within each sample, we summed the number of amenities/assets and standardized these indices, converting them into survey-specific z-scores.

[19] Because black and indigenous identification are asked separately in the case of EIC and Census, the referent is non-black when examining black and non-indigenous when examining indigenous (though the models also controlling for indigenous and black identification, respectively).

using each of the four surveys. This shows how the ethnoracial socioeconomic hierarchy changes when controlling for different variables.

-TABLE 2 ABOUT HERE-

Most notably, there are stark differences between the socioeconomic attainment of those identifying as black according to question type. In all models, the cultural question in the EIC and ancestry-culture question in the Census (see Columns vii and ix in Table 2) exhibited small-to-moderate *advantages* in schooling (Figure 3a) and in household wealth (Figure 3b) relative to non-black individuals while also controlling for indigenous identification. In sharp contrast, the neutral LAPOP and racial MMSI questions revealed substantial black disadvantage in both outcomes (see Columns *i and iii* in Table 2; also see Figures 3a and 3b).

Neither Table 2 or Figure 3 show what happens when black and indigenous identification overlap, as it can be the case in the EIC and Census. In both surveys, the small black advantage observed in Figures 3a-3b is exclusive to individuals who self-identify as black but not indigenous (not shown). Also, note that the cultural questions in both the EIC and Census yielded similar levels of indigenous disadvantage relative to the neutral and racial question in the LAPOP and MMSI, respectively, even after adding some important sociodemographic controls. In all four surveys and across neutral, racial, and cultural wordings, indigenous identification is associated with substantially worse socioeconomic outcomes relative to other forms of ethnoracial identifications available in the data (e.g., see models a7 and b7-b8 in Table 2; see also Figures 3a and 3b).

-FIGURE 3 ABOUT HERE-

DISCUSSION

Our analyses of three official Mexican government surveys conducted over a period of five years, revealed divergent estimates of black and indigenous population size and, most notably, ethnoracial inequality. The clearest divergences were seen in estimates of the black category. After ruling out differences in coverage across surveys, we leveraged data from several cross-sections of the LAPOP study to estimate the degree of growth in black identification. By applying the degree of likely growth observed in LAPOP to forecast the level of black identification one might have observed in the EIC cultural question at the time of the MMSI (racial), we were able to decompose the differences between black identification across these surveys into growth vs. wording components. We estimate that growth and wording components were similarly relevant in explaining the increase in black identification across surveys. However, we find that wording is the most important contributor in producing differences in racial vs. cultural black identification in strata that are more likely to experience higher socioeconomic disadvantage, i.e., rural residents and people with lower schooling.

These results, paired with our findings that only questions with racial or neutral introductions produce black disadvantage in schooling and basic household wealth, lead us to conclude that different question wordings produce diverging portraits of ethnoracial inequality for black-identified individuals. While the neutral LAPOP and racial MMSI questions produced black disadvantage, the EIC and Census questions, which both included references to culture, produced no such disadvantage, and, in some cases, even black *advantage*. Although it is not clear how individuals interpret the neutral framing of the LAPOP question, given evidence that the black category is Mexico is interpreted in terms of phenotype, and thus racialized, it is quite possible that, absent of a specific framing, respondents interpreted the black option in racial terms, which would explain the similarity in results to the direct racial question of the MMSI.

In our analyses, those identifying as black in the cultural questions have similar or slightly greater socioeconomic levels than those identifying as non-indigenous, non-black (also see Torre-Cantalapiedra and Sanchez-Soto 2019). This is consistent with findings from other Latin American countries showing that self-identification as black or mulatto does not always translate into disadvantage (e.g. Telles et al. 2015). Given evidence of a pigmentocracy in Mexico (Arceo-Gomez and Campos-Vazquez 2014; Campos-Vasquez and Medina Cortina 2019; Flores and Telles 2012; Martínez Casas et al. 2014; Telles et al 2015; Villarreal 2010; Telles et al. 2015) and of more specific forms of anti-black discrimination in Mexico (CONAPRED 2011a, 2011b; Cruz Carretero 1989; Velázquez and Iturralde Nieto 2016; Sue 2013; Sue forthcoming; Vaughn 2001), findings of a lack of black disadvantage is likely due to the socioeconomic selectivity in who self-identifies as black, mulatto, or Afrodescendant in certain question formats. Socioeconomically disadvantaged individuals may have "opted out" of both the EIC and Census, given their references to culture, but "opted into" the MMSI's racial question and the LAPOP's neutral question, for reasons we discuss in the conclusion.

One may wonder about the effect of adding a reference to ancestry in the Census. Our findings suggest that this addition does not change the picture of black socioeconomic status at all relative to a more purely cultural framing. Related to inequality, examining the 2020 Census, we saw substantial difference in black identification relative to the exclusively cultural framing of the 2015 EIC. But because this growth occurred in a roughly similar fashion at both the higher and lower ends of the social status spectrum, we observed a similar lack of disadvantage as in the EIC. Regarding the size of the black population, the total number of people opting into the black category in the 2020 Census was 2.1% (Figure 1a), which suggests that the addition of ancestry may not elicit additional black identification, or at least to a strong degree. One study in the

historically black Costa Chica region in Southern Mexico (Sue forthcoming) found that only 66% of self-identified blacks reported having African or black ancestry, which is consistent with findings from INEGI (INEGI 2019; Resano Pérez 2015) indicating that black/African ancestry may not be a good proxy for black identification.

Regardless, lower educated individuals were not particularly likely to self-identify as black in the Census or EIC in relative terms, as was the case in the LAPOP and MMSI. This confirms our interpretation of the role of question wording in explaining differences between the EIC and MMSI. As such, a question with neutral introduction or one related to racial origins may be the most inclusive of individuals in all social strata, and thus a much better fit for assessing black disadvantage in the Mexican context (Saldívar 2014).

In contrast to our findings on black classification, racial and cultural constructions of indigeneity yielded very similar portraits of indigenous socioeconomic standing. Notably, indigenous disadvantage did not vary much even though question wording - in particular, the presence of a *mestizo* option in the LAPOP and MMSI questions - produced somewhat distinct estimates of the size of the indigenous population. In prior work by Martínez et al. (2014), the highest educated individuals in Mexico were more likely to identify as *mestizo*, including over those identifying as white. While this could suggest that individuals with higher socioeconomic status could opt out of the indigenous and into the *mestizo* category if offered one, the relative stability in indigenous inequality across question wordings suggests a more generalized "movement" from indigenous to *mestizo* categories when the latter is offered.

Our work contributes to a growing body of research examining the implications of question framing on patterns of indigenous and black classification. Regarding black classification, our study builds on recent work by Villarreal and Bailey (2020), who find that

black identification is endogenous with earnings potential (and thus socioeconomic status), by suggesting that some types of questions of black identification may disproportionately appeal to higher SES individuals or "repel" lower SES individuals than others. Our findings suggest that the former dynamic may be less likely to occur in neutral or racial questions compared to cultural or ancestry-cultural questions.

Indeed, the black-awareness campaign and related conversations around ethnoracial inequality seem to have contributed to nontrivial increases in black identification, but not necessarily to changing portraits of black inequality when using cultural or ancestral-cultural framings. On one hand, the culture-based black EIC question had similar resonance as the racial MMSI question, taken only 17-21 months apart, among higher status individuals residing in urban areas, perhaps because the consciousness-raising campaigns put out by INEGI and the National Council for the Prevention of Discrimination likely reached these individuals more directly or effectively. On the other hand, the much larger relative difference in black identification among lower-educated, rural Mexicans in the 2016 MMSI's racial question relative to the 2015 EIC cultural question was more likely driven by question wording as opposed to a socially differentiated process of a greater awareness of blackness among Mexicans.

We also expand on studies focused on indigenous classification. While prior work by Villarreal (2014) shows that indigenous disadvantage is largest when using language criteria compared to (proxy or self-) identification, our findings suggest a racial framing and/or a *mestizo* option provide roughly similar estimates of higher levels of indigenous disadvantage compared to cultural measures (with no *mestizo* alternative). Our study also expands the scope of inquiry of how different measures of self-identification impact indigenous identification. In their study using a series of survey experiments, Flores et al. (n.d.) find that urban and higher-educated

Mexicans were more likely to identify as indigenous using the 2010 Census question (which they characterize as a subjective individual-level cultural question), relative to the 2000 Census question (which they characterize as an essentialist ethnic group condition criteria). Complementing this view, our study shows that sociodemographic selectivity varies little when using differing framings within the "subjective individual-level" dimensions and highlights the need for future experimental surveys to better isolate the impacts of a racial framing from a *mestizo* option on indigenous identification and disadvantage.

CONCLUSION

Latin American countries now form part of an increasingly large contingent of nations across the globe that categorize their populations by race or ethnicity (Morning 2008). Within the region, those who have been fighting for ethnoracial recognition, inclusion, and equality have viewed the addition of these categories as a victory. And while the collection of such data represents a step towards these goals, it does not guarantee them. This is not only because data collection alone may not necessarily translate into more equitable policies but, as we have shown, the way in which ethnoracial categories are conceptualized can significantly affect portraits of population size and ethnoracial inequality.

This study contributes to the literatures on the politics of state classification, stratification, and race and ethnicity. Regarding state classification, constructivist scholars, especially those working in regions such as Latin America with high degrees of racial fluidity, have placed much emphasis on the politics surrounding ethnoracial classification systems, and to a lesser extent, some of the consequences of ethnoracial categorization. However, there has been little regard for how ethnoracial question wording produces portraits of inequality. On the flip

side, stratification scholars, particularly in the U.S., have tended to treat ethnoracial data as "objective," i.e., neutral or apolitical and generally do not explore ethnoracial categorical construction as a potential mechanism in the production of inequality. In contexts of high ethnoracial fluidity which characterize Latin America, and increasingly many other parts of the world, including the U.S., this perspective is of particular relevance.

We argue that scholarship in both areas would benefit from a constructivist approach to the study of ethnoracial inequality, not just in terms of recognizing that state estimates of socioeconomic disparities are enmeshed in ethnoracial politics and ideology, but also through empirical demonstration of *how* the conceptualization of ethnoracial categories affect inequality estimates. This is particularly important given increases in ethnoracial fluidity, not just in areas if historic fluidity, but also in contexts such as the United States with historically rigid ethnoracial classification systems. Such an approach may highlight the potential problems associated with estimates of inequality when official conceptualizations and measurements are at odds with popular conceptions of ethnicity and race. We also contribute to the literatures on race and ethnicity by bringing the oft-separate study of indigenous and blacks, which are (problematically) theorized under separate rubrics of "ethnicity" and "race," illustrating how the statistical mechanisms which produce inequality operate differently across these two populations. Based on our findings on black and indigenous growth, we engage with nascent discussions regarding not only what but who is driving recent (in some cases, explosive) growth in black and indigenous identification and whether these phenomena should be understood as a form of symbolic ethnicity. We elaborate on these contributions in this concluding section.

Official Classification Systems and Their Consequences for Constructions of Inequality

As covered previously, there is a rich international literature on the politics of state ethnoracial classification systems. Within this literature, when the consequences and implementation of such classifications have been discussed, the focus has been on topics such as the use of ethnoracial data for state control (e.g., for repressive or genocidal practices), the enforcement of civil rights or reparations, their effects on claims for ethnoracial territorial and individual-level rights, and their effects on decisions regarding immigration and nationality status. However, as we have shown, the way in which states define and ask about ethnoracial identification can also matter significantly for constructions of population size and inequality.

We used the Mexican case to illustrate how state constructions of race and ethnicity can be central mechanisms driving portrayals of population size and ethnoracial inequality. Importantly, we found that these mechanisms operated differently depending on the group. For the indigenous category, we found that indigenous disadvantage surfaced across all framing types (cultural, racial, neutral), which could be explained by the fact that indigeneity is seen in both racial and cultural terms at the popular level (Solís et al. 2019; Sue forthcoming). In contrast, our analyses revealed that conceptualizing blackness in a racial versus cultural framework had a significant impact on inequality estimates, with a racial frame portraying a picture of blacks in Mexico as highly disadvantaged and a cultural frame portraying them as non-distinct from the Mexican majority, or, in some cases, even occupying a *privileged* position.

Our findings indicate that culturally based questions appear to be a poor fit for capturing black disadvantage in the Mexican context, given evidence of anti-black discrimination (CONAPRED 2011a, 2011b; Cruz Carretero 1989; Velázquez and Iturralde Nieto 2016; Sue 2013; Vaughn 2001). We posit that this is likely the case because popular understandings of blackness in Mexico are generally not associated with any sort of cultural identity (Hoffman

2014; INEGI 2019), and instead, largely defined by phenotype (Flores Dávila 2006; INEGI 2019; Lara 2014; Resano Pérez 2015; Sue forthcoming). Saldívar (2014, 2018) warns that the "uncritical deployment" of ethnicity frameworks and emphases on culture and origins can hinder understandings of racism and a focus on racial injustice, a warning not limited to the Mexican case. In Colombia, the narrow equation of blackness with culture has created tension with the recent emphases on racial inequality (Paschel 2016). However, the issue is not about whether particular categories are treated as "cultural" or "racial," but instead, the degree of alignment or misalignment between state conceptualizations and popular understandings of specific categories. When misalignment occurs, higher educated individuals, who are likely more aware of elite discourses, may align their identification accordingly, in part because these discourses may be based on their experience, whereas those in less privileged positions may tend to embrace more popular conceptions of a category, resulting in underestimates of disadvantage.

The importance of aligning official and popular identification schemes for capturing social hierarchies can be seen in other contexts. For example, Howell and Emerson (2017) tested the effects of various measurements of race, including the U.S. census model, finding the pentagon model (White, Black, Hispanic, Asian, Native American) - which has been identified as best aligning with everyday practices of categorization and social treatment (Hollinger 2006) - best captured inequality. Within the U.S. context, the disjuncture between state framings and popular identification and categorization practices is clearest in the case of Hispanics. The U.S Census Bureau continues to treat Hispanic as an ethnicity, even though many Hispanics think of their identity in racial terms or in other ways that are inconsistent with the census form of classification (Compton et al. 2012; Dowling 2014; James 2001; Hitlin et al. 2007; Rodríguez 2000; Telles 2018). Although this disjuncture has been problematized for producing high

nonresponse rates by Hispanics in the race question, or Hispanics categorizing themselves as "other" or "white" (Compton et al. 2012; Dowling 2014; Rodríguez 2000; Telles 2018), as we have argued, attention must also be paid to potential problems associated with the measurement of inequality. In the case of Hispanics, there is evidence which suggests that, when Hispanic is treated as a racial category or as an intersecting racial/ethnic category, a more nuanced and potentially more accurate picture of disadvantage emerges (Logan 2010). Ultimately, we argue that ethnoracial framings for self-identification which most closely align with how categories are understood and acted upon in everyday life, are those most likely to produce portraits of inequality which reflect lived social hierarchies.

The Social Construction of Ethnoracial Inequality

This article also contributes to the literature on ethnoracial inequality, calling for the integration of a constructivist perspective into stratification analyses. Whereas in the prior section, we made a call to address the social construction of inequality through an expansion of the literature on the politics of state classification, in this section, we argue for the fundamental recognition that measurements of inequality are a direct byproduct of survey instruments and need to be treated accordingly. Ideally, this recognition would include the assessment of inequality based on multiple measures, even within major *types* of measures (e.g., self-identification, interview classification, color palettes).

Although much of the recent literature has focused on how stratification outcomes vary across measure types or dimensions (e.g., Bailey et al. 2013; Bailey et al. 2016; Howell and Emerson 2017; Roth 2016; Telles and PERLA 2014; Telles and Lim 1998; Telles and PERLA 2014, Telles et al. 2015), we have shown that, even within a single type of measure - self-

identification - stratification outcomes can vary widely. Given that self-identification will be the global standard by which race and ethnicity are measured in the foreseeable future, it will be important to measure stratification outcomes based on multiple ethnoracial identification measures, if and when available. In the absence of such data, particular caution should be taken when interpreting results, particularly in contexts in which survey measures are at odds with popular conceptions of ethnoracial categories. Especially in these cases, scholars should consider ethnoracial question wording as a potential mechanism in the production of inequality.

Implications for the Study of Race and Ethnicity

This study bridged oftentimes separate empirical work on black and indigenous populations, and theoretical divisions on race and ethnicity, through an integrated analysis and one which measured the effects of racial and cultural approaches by the state to both blackness and indigeneity. A long-standing barrier to such analyses has been data limitations, especially in regions such as Latin America. However, the recent wave of ethnoracial data collection associated with multiculturalist and racial inequality "alignments" (Paschel 2016) presents new opportunities for empirical and theoretical advancement, as well as the exploration of identities such as black-indigenous (Sue and Riosmena forthcoming).

By leveraging unprecedented data on Mexico's black population from three surveys, we were able to provide findings on Mexico's black population (of which we know almost nothing about vis-à-vis official statistics) and compare the effects of state conceptualizations of blackness and indigeneity. While our analyses revealed that different state constructions of blackness yield widely distinct estimates of the size and socioeconomic status of the black population, for the indigenous population, racial versus cultural conceptualizations of indigeneity affected

population size, but not portraits of indigenous disadvantage. These comparative findings would have gone undetected in traditional approaches which segregate black and indigenous studies in statistical analyses, and which conceptualize indigenous populations within an "ethnic" framework, and black populations with a "racial" framework.

Our findings also contribute to broader discussions about the rise in "symbolic ethnicity" - an ethnicity that involves a nostalgic sense of pride and identification with heritage that is not accompanied by everyday practices or experiences of that ethnic group (Gans 1979; Waters 1990). Although the concept was developed based on the case of European Americans, it has been used to explain or interpret growth in the U.S. Native American population (Eschbach, Supple, and Snipp 1998), as well as in the indigenous population in Mexico (Flores et al. n.d.; Telles and Torche 2018; Villarreal 2014), particularly growth among urban, middle class, highly educated individuals. Such individuals are unlikely to be perceived as indigenous and their ethnicity emerged in response to indigenous movements and multiculturalism (Telles and Torche 2018). The dominant explanation for the growth among these so-called "New Indians" (Eschbach, Supple, and Snipp 1998; Telles and Torche 2018), as well as what could be considered "New Blacks," is that higher educated/urban/middle class people are more exposed to multicultural discourses (Flores et al. 2019; Telles and Torche 2018; Villarreal 2014; Villarreal and Bailey 2020).

While our findings support the idea that "New Indians" are contributing to identification growth, they also show that at least half of growth in indigenous identification is taking place among lower SES individuals. Based on the 2010-2019 LAPOP data, indigenous identification growth was similar among those with lower and higher levels of schooling, resulting in consistent disadvantage of those identifying as indigenous over time. As argued previously, the

existence of a *mestizo* option in the LAPOP data could help explain these patterns, although the idea that higher educated individuals would opt out of the indigenous category when given a *mestizo* option is seemingly inconsistent with the idea of symbolic ethnicity. The analysis of the 2010 and 2015 Census/Intercensus (which had very similar question wordings and no *mestizo* option) reveals disproportionate growth in indigenous identification among people with higher schooling and those living in urban areas, although there is also substantial growth among individuals with lower schooling or those living in rural areas. In terms of black growth, our analysis shows that, after accounting for question-wording effects, the rise in black identification between 2015 and 2017 is mostly occurring in urban areas and among higher-educated individuals, even if there is also substantial growth in rural areas and among individuals with less formal schooling.

These findings suggest that the recent increase in black and indigenous identification cannot be reduced to a solely "New Indian," "New Black," or "symbolic ethnicity" phenomenon. Although these concepts explain an important component of the identification growth that is occurring, other forms of growth are also occurring. Since lower SES individuals are less likely to be exposed to elite discourses, analysts need to broaden their explanations for indigenous and black growth beyond the notion of a "multicultural effect." Mechanisms for increasing identification as black and indigenous among those with less schooling may include the lessening of stigma associated with these categories, the "trickle down" effect of elite discourse (which would be consistent with the "multicultural effect" but possibly via a different mechanism), and/or the increasing influence of grassroots movements (Garza 2003).

Policy Implications

Our findings that state formulations of ethnoracial questions can significantly affect population counts and portraits of ethnoracial inequality have some direct policy implications. Whereas the racial and neutral questions produced a clear portrait of both black and indigenous inequality, Mexico adopted a cultural framing of blackness for their most important and visible surveys – the 2015 Intercensus and the 2020 Census. This decision is consistent with post-Revolutionary Mexican national ideology of *mestizaje* which has publicly eschewed the concept of "race," despite the racist underpinnings of the ideology, in favor of the dominant paradigm of "ethnic groups" and culture tied to the indigenous population. Not only has the state deemed the term "race" to be unscientific and unsuitable to the mixed-race nation of Mexico, but the ideology was developed largely in contradistinction to the race-based system of the U.S. and this distinction is deeply embedded within Mexican nationalism (Sue 2013). The omission and disavowal of the term "race," even in deeply racialized government projects, continues to be apparent today (Wade 2017). Therefore, a return to the language of "race" in the Mexican context would likely be a contentious symbolic and political battle.

A much more politically viable alternative would be to introduce a neutral question wording, which, as we showed, produces similar levels of black disadvantage as the "race" question. We posit that the neutral framing is indirectly measuring a racialized understanding of blackness since the black category is currently understood in phenotypic terms at the popular level. Not only is the neutral wording a politically feasible alternative to the cultural framing, but it allows popular interpretations of ethnoracial categories, as opposed to top-down impositions, to guide responses. As we argued above, for self-identification measures, we believe that when state and popular conceptions of ethnoracial categories most closely align, portraits of inequality will most closely reflect social categorization practices.

Even though self-identification measures are imperfect proxies for observed race and exposure to discrimination (Telles and Lim 1998), they are the global standard set by organizations such as the United Nations. Our findings demonstrate that, within self-identification measures, some question formats are better suited to measuring ethnoracial inequality and population size than others, specifically those which most closely align with popular understandings and practices of categorization. This alignment, however, is not the focus of international guidance on ethnoracial data collection. In their recommendations for collecting data on ethnicity, the U.N. states that:

> "Ethnicity can be measured using a variety of concepts, including ethnic ancestry or origin, ethnic identity, cultural origins, nationality, race, colour, minority status, tribe, language, religion or various combinations of these concepts. Because of the interpretative difficulties that may occur with measuring ethnicity in a census, it is important that, where such an investigation is undertaken, the basic criteria used to measure the concept are clearly explained to respondents and in the dissemination of the resulting data." (205).

Although the U.N. seemingly encourages a "top-down" definitional process (as long as countries clearly explain their criteria), their use of "self-declaration" and the "subjective nature" of ethnicity represents a tension point surrounding the question of who defines the meaning of ethnicity – the state or the respondent? The above recommendations seem to suggest the former, while the emphasis on self-identification could imply the latter. We believe that the use of a neutral question framing could help assuage this tension.

Also at the global scale, our findings provide a cautionary tale for the policy diffusion that has taken place surrounding policies related to ethnoracial thinking and policies (Loveman

2014; Wade 2017) and their connection to immigration policy (FitzGerald and Cook-Martín 2014). This diffusion process is clearly implicated in the politics and decisions surrounding the conceptualization of blackness in Mexico. For example, Mexico adopted recommendations by international organizations regarding the use of the term "Afrodescendant" in the 2020 Census, despite the fact that extensive pilot testing by its own Census Bureau showed that the term is unfamiliar in the Mexican context and causes substantial confusion (INEGI 2019). There are tradeoffs to implementing global models of race and ethnicity, especially when those models are at odds with popular understandings of ethnoracial categories. As we have shown, not only are there substantial political consequences involved in decisions about if and how to engage in ethnoracial classification, but there are significant empirical consequences, such as official constructions of inequality, which inform and shape state priorities and policies to remedy discrimination and inequality and determine who are the rightful benefactors of state resources and support.

References

Angosto Ferrández, Luis Fernando, and Sabine Kradolfer, Eds. 2012a. *Everlasting Countdowns: Race, Ethnicity and National Censuses in Latin American States*. Newcastle, Cambridge Scholars Publishing.

Angosto Ferrández, Luis Fernando, and Sabine Kradolfer. 2012b. "Race, Ethnicity and National Censuses in Latin American States: Comparative Perspectives." Pp. 1-40 in *Everlasting Countdowns: Race, Ethnicity and National Censuses in Latin American States*, edited by Luis Fernando Angosto Ferrández and Sabine Kradolfer. Newcastle upon Tyne: Cambridge Scholars Publishing.

Aguirre Beltran, Gonzalo. 1944. "The Slave Trade in Mexico." *The Hispanic American Historical Review* 24(3):412-30.

Arceo-Gomez, E. O., & Campos-Vazquez, R. M. 2014. “Race and marriage in the labor market: A discrimination correspondence study in a developing country.” *American Economic Review*, 104(5), 376-80.

Bailey, Stanley, M. Fialho Fabrício and M. Penner Andrew. 2016. "Interrogating Race: Color, Racial Categories, and Class across the Americas." *American Behavioral Scientist* 60(4):538-55. doi: 10.1177/0002764215613400.

Bailey, Stanley, Aliya Saperstein, and Andrew Penner. 2014. "Race, color, and income inequality across the Americas." *Demographic Research* 31(24): 735-56.

Bailey, Stanley, Mara Loveman, and Jeronimo Muniz. 2013. "Measures of "Race" and the analysis of racial inequality in Brazil." *Social Science Research* 42(1):106-19.

Berry-James, RaJade E., Susan T. Gooden, and Richard Greggory Johnson III. 2020. “Civil Rights, Social Equity, and Census 2020.” *Public Administration Review* 80(6): 1100-1108.

Bonfil Sánchez, Paloma, Natalia De Marinis, Bertha Patricia Rosete Xotlanihua, and Raúl Martínez Navarro. 2017. "Violencia de Género Contra Mujeres en Zonas Indígenas en México." SEGOB.

Brubaker, Rogers. 2004. *Ethnicity without Groups*. Cambridge: Harvard University Press.

Brubaker, Rogers, Mara Loveman and Peter Stamatov. 2004. "Ethnicity as Cognition." *Theory and Society* 33:31-64.

Campos-Vazquez, Raymundo and Eduardo Medina-Cortina. 2019. "Skin Color and Social Mobility: Evidence from Mexico." *Demography* 56: 321–343.

Compton, Elizabeth, Michael Bentley, Sharon Ennis and Sonya Rastogi. 2013. "2010 Census Race and Hispanic Origin Alternative Questionnaire Experiment." U.S. Census Bureau.

https://www.census.gov/2010census/pdf/2010_Census_Race_HO_AQE.pdf

CONAPRED (Consejo Nacional para Prevenir la Discriminación). 2011a. "Documento Informativo sobre Discriminación Racial en México." México, D.F.

CONAPRED (Consejo Nacional para Prevenir la Discriminación). 2011b. "Guía para la acción pública: Afrodescendencia. Población Afrodescendiente en México." México, D.F., México: Consejo Nacional para Prevenir la Discriminación.

CONEVAL. 2014. "La pobreza en la población indígena en México, 2012". Mexico City.

Creighton, M., and Park, H. (2010). Closing the gender gap: Six decades of reform in Mexican education. *Comparative Education Review* 54(4):513-537.

Davis, James F. 1991. *Who Is Black? One Nation's Definition*. University Park: Pennsylvania.

Del Popolo, Fabiana, and Susana Schkolnik. 2012. "Indigenous Peoples and Black-Descendants: The Difficult Art of Counting." Pp. 304-34 in *Everlasting Countdowns: Race, Ethnicity and National Censuses in Latin American States*, edited by Luis Fernando Angosto Ferrández and Sabine Kradolfer. Newcastle: Cambridge Scholars Publishing.

Dowling, Julie A. 2014. *Mexican Americans and the Question of Race*. Austin: University of Texas Press.

Eschbach, Karl, Khalil Supple and C. Matthew Snipp. 1998. "Changes in Racial Identification and the Educational Attainment of American Indians, 1970–1990." *Demography* 35(1):35-43.

Filmer, Deon, and Lant Pritchett. "The effect of household wealth on educational attainment: evidence from 35 countries." *Population and development review* 25.1 (1999): 85-120.

FitzGerald, David and David Cook-Martín. 2014. *Culling the Masses: The Democratic Origins of Racist Immigration Policy in the Americas*. Cambridge: Harvard University Press.

Flores, René D., María Vignau Loria, and Regina Martínez-Casas. 2019. "What Explains the Indigenous Population Boom in Contemporary Mexico?" American Sociological Association Annual Conference, New York, New York, August 10-13.

Flores Dávila, Julia. 2006. *Afrodescendientes en México; reconocimiento y propuestas antidiscriminación*. México DF, CONAPRED.

Flores Dávila, Julia and Florence Lézé. 2007. "Procesos De Construcción De Identidad, Condiciones De Vida Y Discriminación: Un Estudio Comparativo De Comunidades De Los Afrodescendientes En México." Paper presented at the XXVI Congreso de la Asociación Latinoamericana de Sociología, Guadalajara, Mexico.

Francis, Andrew and Maria Tannuri-Pianto. 2013. "Endogenous Race in Brazil: Affirmative Action and the Construction of Racial Identity among Young Adults." *Economic Development and Cultural Change* 61:731-53. doi: 10.1086/670375.

French, Jan Hoffman. 2009. *Legalizing Identities: Becoming Black or Indian in Brazil's Northeast.* University of North Carolina Press.

Gans, Herbert J. 1979. "Symbolic Ethnicity: The Future of Ethnic Groups and Cultures in America." *Ethnic and Racial Studies* 2:1-20.

González Estrella, Jorge. 2012. "The (Re)Construction of Blackness in Costa Chica, Oaxaca: Ngos and the Making of an Afro-Mexican Ethnic Group." M.A., Latin American and Iberian Studies, University of California, Santa Barbara, Santa Barbara, CA.

Hernández Cuevas, Marco Polo. 2004. *African Mexicans and the Discourse on Modern Nation*. Dallas: University Press of America.

Hitlin, Steven, J. Scott Brown and Glen H. Elder. 2007. "Measuring Latinos: Racial Vs. Ethnic Classification and Self-Understandings." *Social Forces* 86(2):587-611.

Hoffman, Odile. 2014. "The Renaissance of Afro-Mexican Studies." in *Blackness and Mestizaje in Mexico and Central America*, edited by Elisabeth Cunin and Odile Hoffman. Trenton, New Jersey: Africa World Press.

Hoffman, Odile and Christian Rinaudo. 2014. "The Issue of Blackness and Mestizaje in Two Distinct Mexican Contexts: Veracruz and Costa Chica." *Latin American and Caribbean Ethnic Studies* 9(2):138-55.

Hooker, Juliet. 2005. "Indigenous Inclusion/Black Exclusion: Race, Ethnicity and Multicultural Citizenship in Latin America." *Journal of Latin American Studies* 37:285-310.

Howell, Junia and Michael O. Emerson. 2017. "So What "Should" We Use? Evaluating the Impact of Five Racial Measures on Markers of Social Inequality." *Sociology of Race and Ethnicity* 3(1):14-30.

Hollinger, David A. 2006. *Postethnic America: Beyond Multiculturalism*. New York: Basic Books.

INEGI. 2021. Censo 2020 de Población y Vivienda: Síntesis Metodológica y Conceptual. Available at https://www.inegi.org.mx/contenidos/productos/prod_serv/contenidos/espanol/bvinegi/productos/nueva_estruc/702825197537.pdf. Last accessed March 17, 2021.

INEGI. 2019. Proceso del Diseño del la Pregunta de Afrodescendencia.

INEGI. 2015. “Encuesta Intercensal 2015: Síntesis Metodológica y Conceptual.” Retrieved from http://www3.inegi.org.mx/contenidos/proyectos/enchogares/especiales/intercensal/2015/doc/eic2015_sintesis.pdf. Last accessed September 1, 2017.

INEGI. N.D.-A “Módulo de Movilidad Social Intergeneracional 2016 MMSI Diseño Estadístico.” Retrieved from http://internet.contenidos.inegi.org.mx/contenidos/productos/prod_serv/contenidos/espanol/bvinegi/productos/nueva_estruc/promo/diseno_estadistico_mmsi_2016.pdf. Last accessed September 4, 2017

INEGI. N.D.-B. Encuesta Intercensal 2015: Manual del Entrevistador. Available at https://www.inegi.org.mx/contenidos/programas/intercensal/2015/doc/eic2015_manual_entrevistador.pdf. Last accessed November 22, 2020.

James, Angela. 2001. "Making Sense of Race and Racial Classification." *Race and Society* 4:235-47.

Kertzer, D. and D. Arel, Eds. 2002. *Census and Identity: The Politics of Race, Ethnicity, and Language in National Censuses*. Cambridge, Cambridge University Press.

Khalfani, Akil Kokayi and Tukufu Zuberi. 2001. "Racial Classification and the Modern Census in South Africa, 1911–1996." *Race and Society* 4 (2): 161-76.

Latin America Public Opinion Project (LAPOP). 2021. *Survey Design*. Available at https://www.vanderbilt.edu/lapop/survey-designs.php.

Lara, Gloria. 2014. "An Ethno-Political Trend on the Costa Chica, Mexico (1980-2000)." Pp. 117-38 in *Blackness and Mestizaje in Mexico and Central America*, edited by Elisabeth Cunin and Odile Hoffman. Trenton: Africa World Press.

Logan, John R. 2010. "How race counts for Hispanic Americans." In *The Afro-Latino reader: History and culture in the United States*, eds. Miriam Jiménez Román and Juan Flores, 471–84. Durham, NC: Duke University Press.

Loveman, Mara, Jeronimo O. Muniz and Stanley R. Bailey. 2012. *Brazil in black and white? Race categories, the census, and the study of inequality, Ethnic and Racial Studies*. 35(8): 1466-1483.

Loveman, Mara. 2014. *National Colors: Racial Classification and the State in Latin America*: Oxford University Press.

Martínez Casas, Regina, Emiko Saldívar, René D. Flores, and Christina A. Sue. 2014. "The Different Faces of Mestizaje: Ethnicity and Race in Mexico." Pp. 36-80 in *Pigmentocracies: Ethnicity, Race, and Color in Latin America*, edited by Edward E. Telles. Chapel Hill: University of North Carolina Press.

Marx, Anthony. 1998. *Making Race and Nation: A Comparison of the United States, South Africa, and Brazil*. Cambridge, Cambridge University Press.

Mora, Cristina. 2014. *Making Hispanics: How Activists, Bureaucrats, and Media Constructed a New American*. Chicago, University of Chicago Press.

Morning, Ann. 2008. "Ethnic Classification in Global Perspective: A Cross-National Survey of the 2000 Census Round." *Population Research and Policy Review,* 27(2): 239-272.

Nobles, Melissa. 2000. *Shades of Citizenship: Race and the Census in Modern Politics*. Stanford: Stanford University Press.

Omi, Michael and Howard Winant. 2015. *Racial Formation in the United States from the 1960s to the 1990s*. New York: Routledge.

Paschel, Tianna S. 2016. *Becoming Black Political Subjects: Movements and Ethno-Racial Rights in Colombia and Brazil*. Princeton University Press: New Jersey.

Resano Pérez, Elsa. 2015. "El Tema De Etnicidad En Los Censos De Población En México." INEGI.

Rodríguez, Clara E. *Changing race: Latinos, the census, and the history of ethnicity in the United States*. NYU Press, 2000.

Rosemblatt, Karin Alejandra. 2018. *The Science and Politics of Race in Mexico and the United States, 1910–1950*. Chapel Hill: University of North Carolina Press.

Roth, Wendy D. 2016. “The Multiple Dimensions of Race.” *Ethnic and Racial Studies*, 39(8): 1310-1338.

Ruiz Ramírez, Roberto. 2014. “Resultados de las Pruebas del Tema Afrodescendencia.” INEGI.

Saldívar, Emiko. 2018. "Uses and Abuses of Culture: Mestizaje in the Era of Multiculturalism." *Cultural Studies* 32(3):438-59.

Saldívar Tanaka, Emiko, Erika Arenas, Fernando Riosmena, and Christina Sue. 2018. *Consideraciones Teóricas y Metodológicas para el Conteo de la Población Afrodescendiente en el Censo 2020*. University of California, Santa Barbara and Kellog Foundation Report

Saldívar Tanaka, Emiko, and Mónica G. Moreno Figueroa. n.d. "‘Soy Afro: Me Reconozco y Cuento;” Race, Racism and Representation of Blackness in Mexico."

Saldívar Tanaka, Emiko. 2014. "‘It’s Not Race, It’s Culture’: Untangling Racial Politics in Mexico." *Latin American and Caribbean Ethnic Studies* 9(1):89-108.

Saldívar, Emiko, and Casey Walsh. 2014. "Racial and Ethnic Identities in Mexican Statistics." *Journal of Iberian and Latin American Research* 20(3):455-75.

Simon, Patrick and Victor Piché. 2012. Accounting for ethnic and racial diversity: the challenge of enumeration, *Ethnic and Racial Studies,* 35:8, 1357-1365

Skerry, P. 2000. Counting on the Census?: Race, Group Identity, and the Evasion of Politics. Washington, D.C.: Brookings Institution Press.

Snipp, C. Matthew. 2003. "Racial Measurement in the American Census: Past Practices and Implications for the Future." *Annual Review of Sociology* 29(1):563-88. doi: 10.1146/annurev.soc.29.010202.100006.

Solís, P., Güémez Graniel, B., & Lorenzo Holmes, V. 2019. "Por mi raza hablará la desigualdad. Efectos de las características étnico-raciales en la desigualdad de oportunidades en México." Mexico City: Oxfam México

Sue, Christina. Forthcoming *The Erosion of Mexico's Hegemonic National Ideology of Mestizaje?: Blackness, Race Mixture and Discrimination*

Sue, Christina and Riosmena, Fernando. Forthcoming. "Black and Indigenous Inequality in Mexico: Implications for Multiracialism and Intersectionality Research." *Sociology of Race and Ethnicity.*

Sue, Christina. 2013. *Land of the Cosmic Race: Race Mixture, Racism, and Blackness in Mexico*. New York: Oxford University Press.

Sue, Christina A. 2009. "The Dynamics of Color: *Mestizaje*, Racism, and Blackness in Veracruz, Mexico." Pp. 114-29 in *Shades of Difference: Why Skin Color Matters*, edited by Evelyn Nakano Glenn. Stanford: Stanford University Press.

Telles, Edward. 2018. "Latinos, Race, and the U.S. Census." *The ANNALS of the American Academy of Political and Social Scienc*e 677(1):153-64.

Telles, Edward and Florencia Torche. 2018. "Varieties of Indigeneity in the Americas." *Social Forces* 97(4): 1543-1570.

Telles, Edward, René D Flores, and Fernando Urrea-Giraldo. 2015. "Pigmentocracies: Educational inequality, skin color and census ethnoracial identification in eight Latin American countries." Research in Social Stratification and Mobility 40:39-58.

Telles, Edward E., and The Project on Race and Ethnicity in Latin America (PERLA). 2014. *Pigmentocracies: Ethnicity, Race, and Color in Latin America.* Chapel Hill: University of North Carolina Press.

Telles, Edward and Tianna Paschel. 2014. “Who Is Black, White, or Mixed Race? How Skin Color, Status, and Nation Shape Racial Classification in Latin America.” *American Journal of Sociology*, 120 (3): 864-907.

Telles, Edward E. and Nelson Lim. 1998. "Does It Matter Who Answers the Race Question? Racial Classification and Income Inequality in Brazil." *Demography* 35(4):465-74.

Torre-Cantalapiedra, Eduardo, and Gabriela Sánchez-Soto. 2019. “Afrodescendientes y estratificación social en México. Nueva evidencia con datos de la Encuesta Intercensal 2015.” Papeles de Población 25(100), 273-302.

Treitler, Vilna Bashi. 2013. *The Ethnic Project: Transforming Racial Fiction into Ethnic Factions*. Stanford: Stanford University Press.

United Nations. 2017. "Principles and Recommendations for Population and Housing Censuses." New York: United Nations.

Vasconcelos, Jose. 1925. The Cosmic Race, La Raza Cosmica. Translated by D. Jaen. Baltimore: The John Hopkins University Press.

Vaughn, Bobby. 2001. "Race and Nation: A Study of Blackness in Mexico." in A*nthropology*. Stanford: Stanford University.

Vaughn, B. 2013. "México Negro: From the Shadows of Nationalist Mestizaje to New Possibilities in Afro-Mexican Identity." *The Journal of Pan African Studies* 6(1): 227-240.

Velázquez, María Elisa e Iturralde Gabriela. 2016. *Afrodescendientes en México. Una historia de silencio y discriminación*, CONAPRED, INAH

Villarreal, Andrés, and Stan Bailey. 2020. "The Endogeneity of Race: Black Racial Identification and Men's Earnings in Mexico." *Social Forces*, soz096, https://doi.org/10.1093/sf/soz096.

Villarreal, Andrés. 2014. "Ethnic Identification and its Consequences for Measuring Inequality in Mexico." *American Sociological Review* 79(4): 775-806.

Villarreal, Andrés. 2010. "Stratification by Skin Color in Contemporary Mexico." *American Sociological Review* 75(5):652-78.

Wade, Peter. 2010. *Race and Ethnicity in Latin America*. London: Pluto Press.

Wade, Peter. 2018. "Afro-Indigenous Interactions, Relations, and Comparisons." Pp. 92-129 in Afro-Latin American Studies: An Introduction, edited by A. De la Fuente and G. R. Andrews. Cambridge: Cambridge University Press.

Waters, Mary. 1990. *Ethnic Options: Choosing Identities in America*. Berkeley: University of California Press.

Wimmer, Andreas. 2008. "The Making and Unmaking of Ethnic Boundaries: A Multilevel Process Theory." *American Journal of Sociology* 113(4):970–1022.

Figure 1a. Percent of Mexican adults ages 25-64 self-identifying as black by question type and year, various surveys.

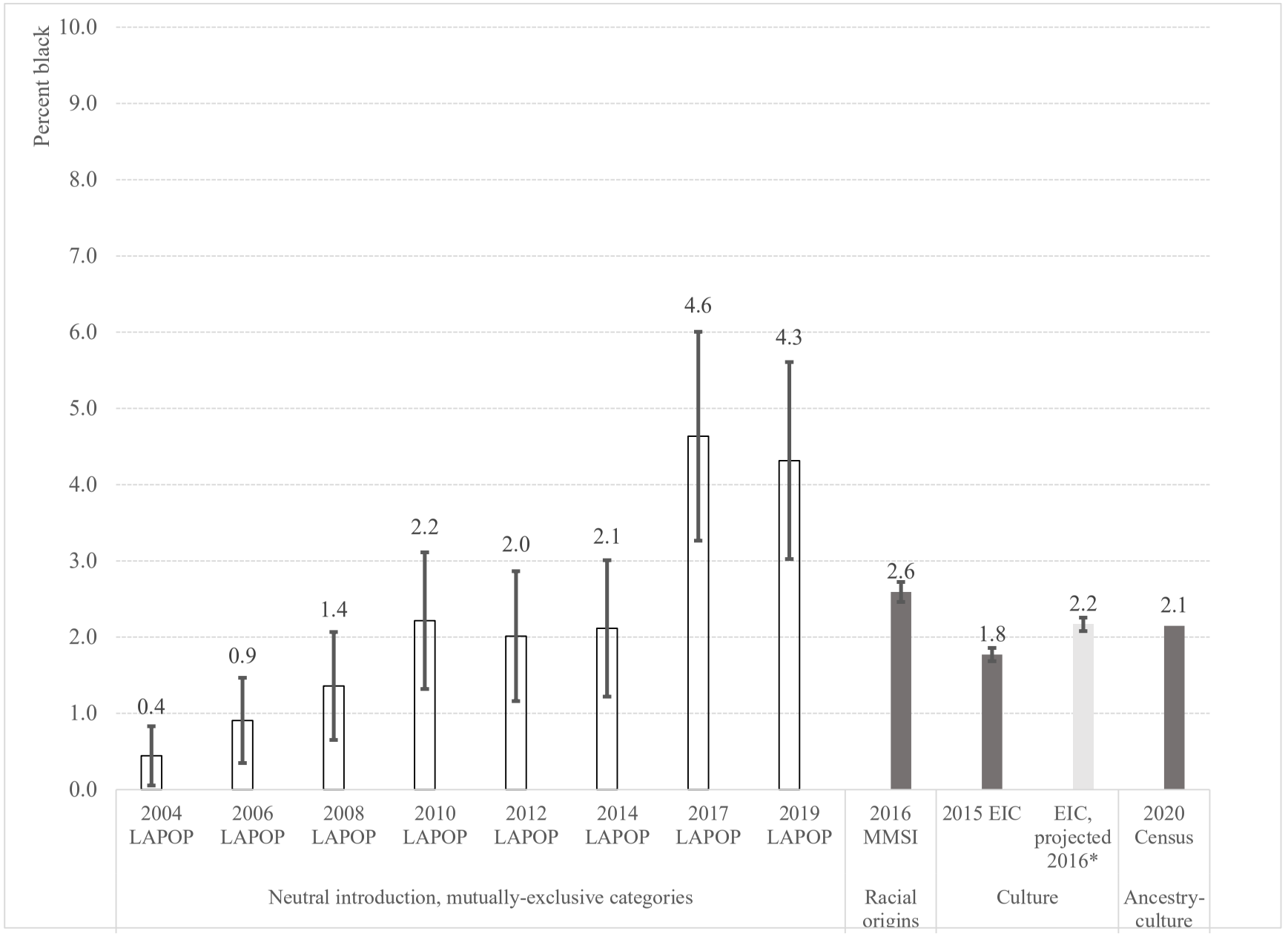


Figure 1b. Percent of Mexican adults ages 25-64 self-identifying as indigenous by question type and year, various surveys.

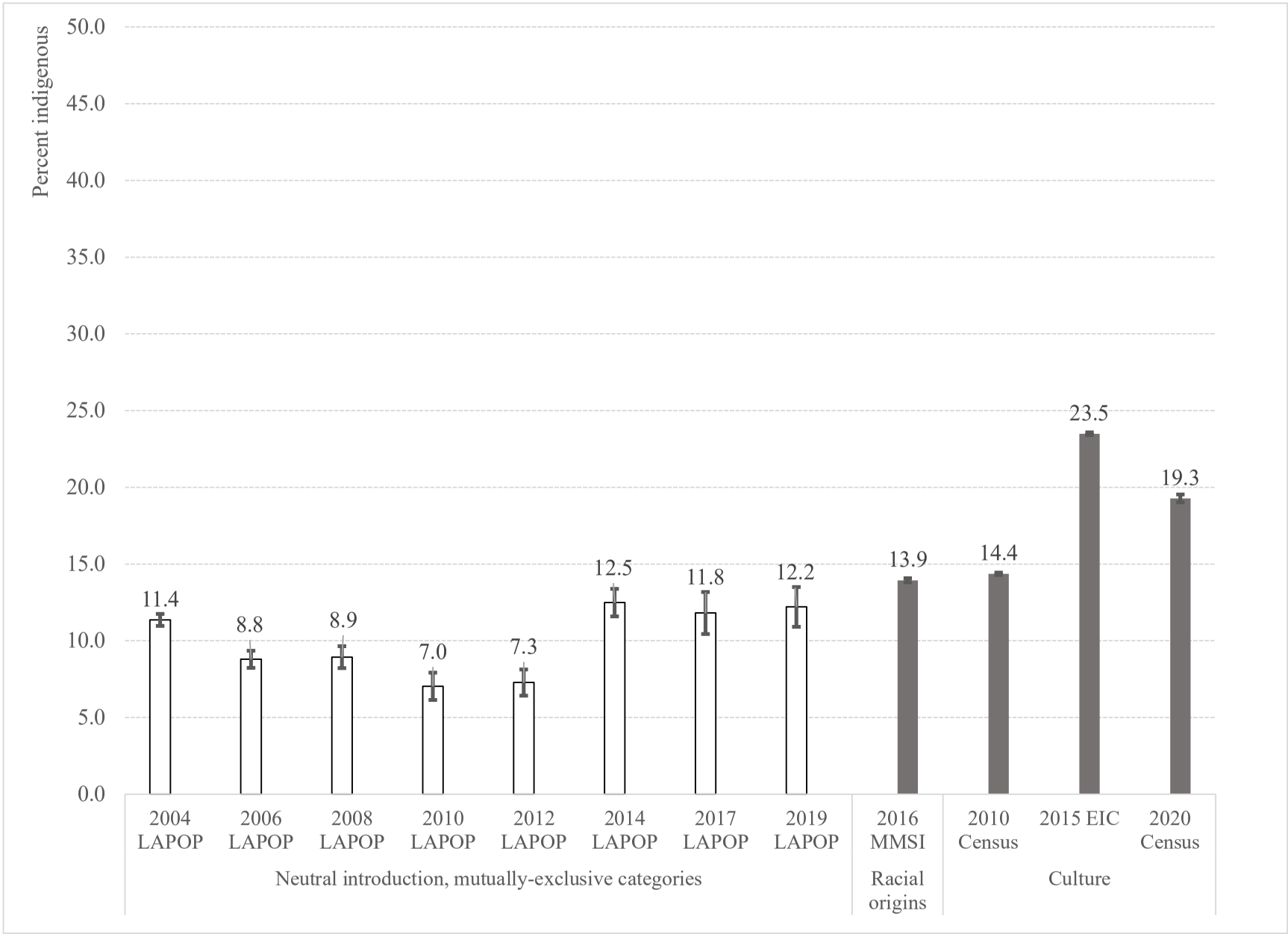


Table 1. Estimated contributions of secular growth in black identification and question wording on the difference between the share of Mexican adults ages 25-64 in 2015 EIC and 2016 MMSI, mean estimates and bootstrapped 95% confidence intervals.

		Est. growth rate in black identification	2015 EIC		2016 MMSI	Est. contributions of… to observed MMSI-EIC difference.			
						…growth		…wording[4]	
			Actual	Projected, 2016[2]	Actual	Mean	95% confidence interval	Mean	95% confidence interval
Nationwide		13.0%	1.80%	2.20%	2.60%	52.7%	(34.8% , 81.9%)	47.3%	(18.1% , 65.2%)
Gender	Women	21.2%	1.76%	2.45%	2.63%	87.1%	(48.7% , 162.0%)	12.9%	(-62.0% , 51.3%)
	Men	7.1%	1.84%	2.05%	2.57%	31.9%	(20.2% , 53.0%)	68.1%	(47.0% , 79.8%)
Age	25-44 years	12.6%	1.78%	2.16%	3.07%	30.8%	(20.8% , 47.1%)	69.2%	(52.9% , 79.2%)
	45-64 years	16.1%	1.83%	2.35%	1.91%	561%	(-1559% , 2613%)	-461%	(-2513% , 1659%)
Schooling	0-11 years	14.0%	1.81%	2.25%	3.10%	35.3%	(24.1% , 53.3%)	64.7%	(46.7% , 75.9%)
	12+ years	13.0%	1.77%	2.17%	1.71%	-42.7%	(-1165% , 1119%)	142.7%	(-1019% , 1265%)
Locality size	Rural	17.8%	1.86%	2.45%	4.27%	25.5%	(18.4% , 37.8%)	74.5%	(62.2% , 81.6%)
	Urban	11.9%	1.78%	2.14%	2.16%	121.6%	(51.1% , 293%)	-21.6%	(-193% , 48.9%)

[1] Mean annualized growth rate in black identification, estimated between 2010-2014 and 2017-2019 LAPOP (see Eq. 4 in Appendix IV).

[2] Estimated share culturally-black at time of MMSI survey based on projecting 2015 EIC value using growth rate in (1). See Eq. 2 in

[3] Estimated share culturally-black at time of MMSI survey based on projecting 2015 EIC value using growth rate in (1). See Eq. 2 in

[4] Estimated as complement of growth component (see Eq. 5 in Appendix IV). Value of wording confidence intervals were flipped to reflect minimum and maximum values of range in cases where growth contribution estimates are higher than 100% for the 95th percentile

Figure 2a. Percent of Mexican adults ages 25-64 self-identifying as black by sociodemographic characteristics, 2015 EIC, 2016 MMSI, 2020 Census, & 2010-2019 LAPOP.

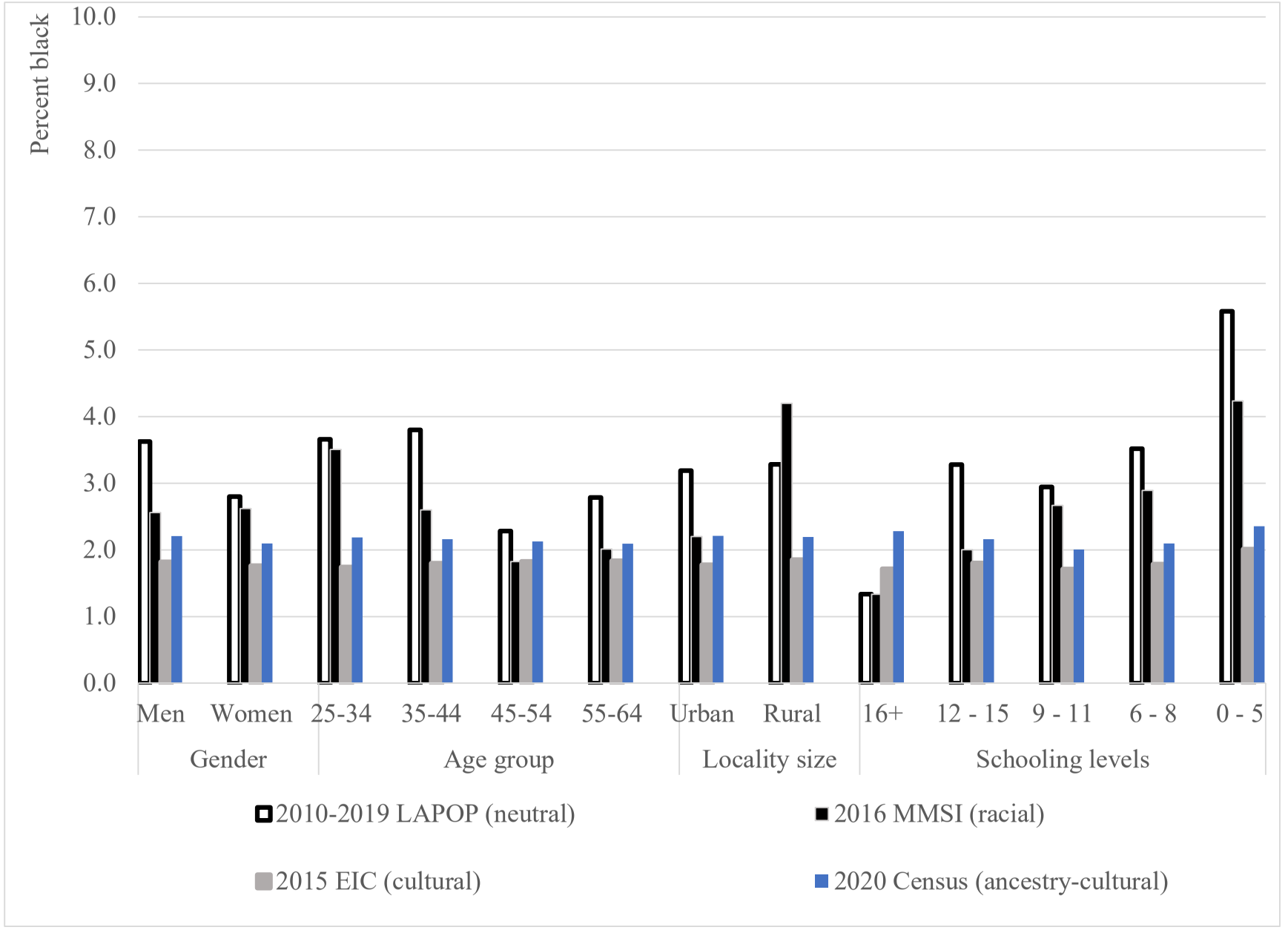


Figure 2b. Percent of Mexican adults ages 25-64 self-identifying as indigenous by sociodemographic characteristics, 2015 EIC, 2016 MMSI, 2020 Census, & 2010-2019 LAPOP.

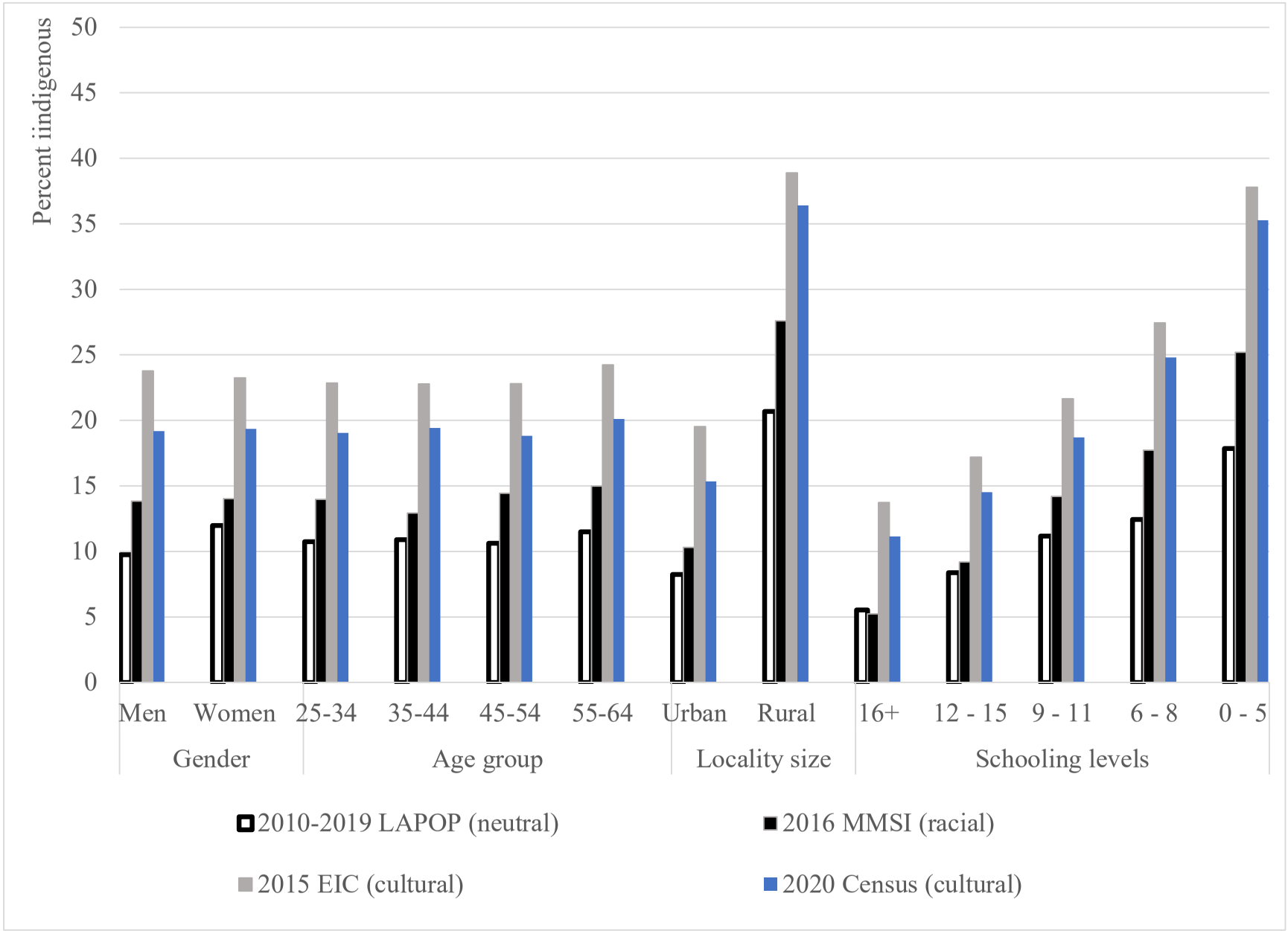


Table 2. Predicted differences in (a) years of schooling and (b) household amenities/assets index (z-scores) between Mexican adults 25-65 identifying as black or indigenous relative to non-indigenous, non-black counterparts by controls added, 2015 EIC, 2016 MMSI, and 2020 Census long-form surveys.

	Neutral introduction & neutral/racial options, 2010-2014 LAPOP		Neutral introduction & neutral/racial options, 2017-2019 LAPOP		Racial question: 2016 MMSI - sampled informant ages 25-64		Cultural questions: 2015 EIC - all household members ages 25-64		Ancestry-cultural (black) & cultural (indigenous) questions: 2020 Census - all household members ages 25-64	
	i. Black	ii. Indigenous	iii. Black	iv. Indigenous	v. Black	vi. Indigenous	vii. Black	viii. Indigenous	ix. Black	x. Indigenous
a. Schooling (years-difference vs. NINB[1]).										
	Est. Sig.	Est. Sig.	Est. Sig.	Est. Sig.	Est. Sig.	Est. Sig.	Est. Sig.	Est. Sig.	Est. Sig.	Est. Sig.
a0. No additional controls.	-1.388 *	-1.557 ***	-1.819 ***	-1.536 ***	-1.907 ***	-2.435 ***	0.664 ***	-2.043 ***	0.551 ***	-2.101 ***
a1. Race/ethnicity + sex.	-1.557 **	-1.501 ***	-1.809 ***	-1.507 ***	-2.248 ***	-2.807 ***	0.662 ***	-2.045 ***	0.549 ***	-2.101 ***
a2. Race/ethnicity + age.	-1.663 **	-1.519 ***	-2.024 ***	-1.514 ***	-2.522 ***	-2.791 ***	0.678 ***	-2.020 ***	0.529 ***	-2.083 ***
a3. Race/ethnicity + urban/rural.	-1.387 *	-0.8871 *	-1.749 ***	-1.346 ***	-1.716 ***	-2.070 ***	0.441 ***	-1.459 ***	0.408 ***	-1.463 ***
a4. Race/ethnicity + state.	-1.515 *	-1.5013 ***	-1.773 ***	-1.429 ***	-2.019 ***	-2.513 ***	0.629 ***	-1.730 ***	0.552 ***	-1.846 ***
a5. All sociodemg. controls.[2]	-2.018 **	-0.8886 **	-1.947 ***	-1.250 ***	-1.574 ***	-1.545 ***	0.412 ***	-1.315 ***	0.385 ***	-1.389 ***
a6. Race/ethnicity + indig. lang.	N/A	N/A	N/A	N/A	-2.166 ***	-2.013 ***	0.431 ***	-1.221 ***	0.387 ***	-1.155 ***
a7. All sociodemg. + indig. lang.	N/A	N/A	N/A	N/A	-1.575 ***	-1.037 ***	0.255 ***	-0.857 ***	0.271 ***	-0.812 ***
b. Asset index ownership (difference in z-scores, vs. NINBs[1]).										
	Est. Sig.	Est. Sig.	Est. Sig.	Est. Sig.	Est. Sig.	Est. Sig.	Est. Sig.	Est. Sig.	Est. Sig.	Est. Sig.
b0. No additional controls.	-0.647 **	-0.943 ***	-0.509 ***	-0.664 ***	-0.483 ***	-0.778 ***	0.152 ***	-0.593 ***	0.106 ***	-0.662 ***
b1. Race/ethnicity + sex.	-0.690 **	-0.927 ***	-0.496 ***	-0.653 ***	-0.482 ***	-0.778 ***	0.152 ***	-0.593 ***	0.106 ***	-0.662 ***
b2. Race/ethnicity + age.	-0.645 **	-0.943 ***	-0.513 ***	-0.653 ***	-0.473 ***	-0.779 ***	0.151 ***	-0.595 ***	0.107 ***	-0.663 ***
b3. Race/ethnicity + urban/rural.	-0.649 **	-0.783 ***	-0.490 ***	-0.608 ***	-0.340 ***	-0.576 ***	0.092 ***	-0.436 ***	0.068 ***	-0.491 ***
b4. Race/ethnicity + state.	-0.734 **	-0.753 ***	-0.459 ***	-0.465 ***	-0.307 ***	-0.560 ***	0.176 ***	-0.436 ***	0.128 ***	-0.496 ***
b5. Race/ethnicity + schooling.	-0.455 *	-0.754 ***	-0.328 **	-0.527 ***	-0.271 ***	-0.495 ***	0.086 ***	-0.394 ***	0.057 ***	-0.477 ***
b6. All sociodemg. controls.[2]	-0.535 **	-0.467 ***	-0.267 **	-0.329 ***	-0.048 N.S.	-0.276 ***	0.088 ***	-0.224 ***	0.065 ***	-0.283 ***
b7. Race/ethnicity + indig. lang.	N/A	N/A	N/A	N/A	-0.456 ***	-0.494 ***	0.082 ***	-0.348 ***	0.058 ***	-0.383 ***
b8. Model 7 + indig. lang.	N/A	N/A	N/A	N/A	-0.214 ***	-0.281 ***	0.080 ***	-0.209 ***	0.063 ***	-0.230 ***
b9. Model b8 + schooling.	N/A	N/A	N/A	N/A	-0.053 N.S.	-0.171 ***	0.058 ***	-0.135 ***	***	***

Notes: all estimates were weighted and standard errors adjusted for complex sampling.

[1] Comparison pertains to "group" in column relative to NINB = Non-indigenous, non-black-identified individuals.

2 For both outcomes, model with all "sociodemographic" controls include race/ethnicity, sex, age (years), urban/rural locality, and state fixed-effects. For analyses of the assets/amenities index, model also includes controls for schooling (years).

*** $p < 0.01$ ** $p < 0.01$ * $p < 0.05$ N.S. Not significant at 0.05 or higher level.

Figure 3a. Predicted difference in years of schooling and 95% confidence intervals for black- and indigenous-identified individuals ages 25-64 relative to non-indigenous, non-blacks, by model adjusting for different controls, 2010-2019 LAPOP, 2016 MMSI, 2015 EIC, & 2020 Census long form surveys.

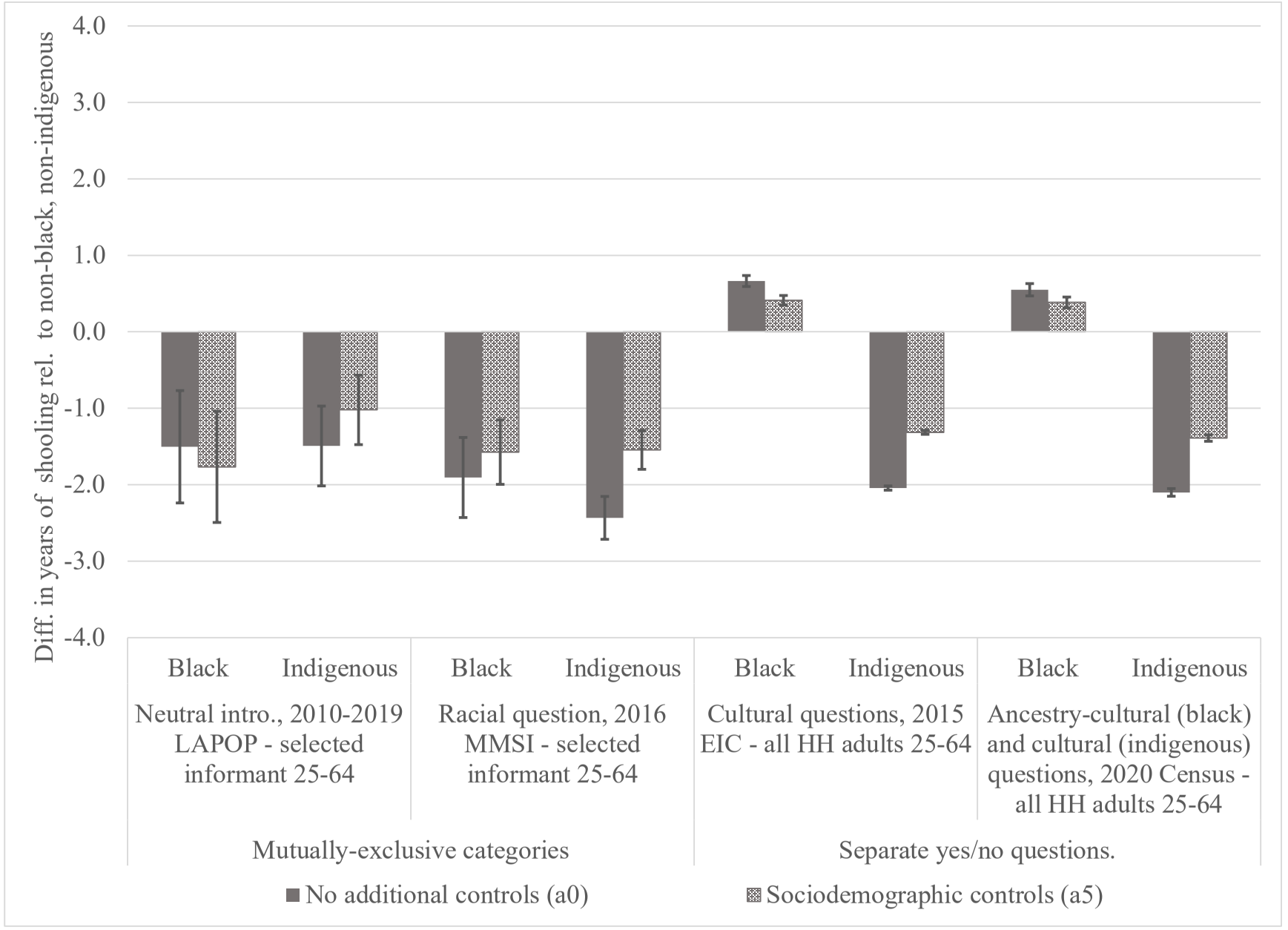
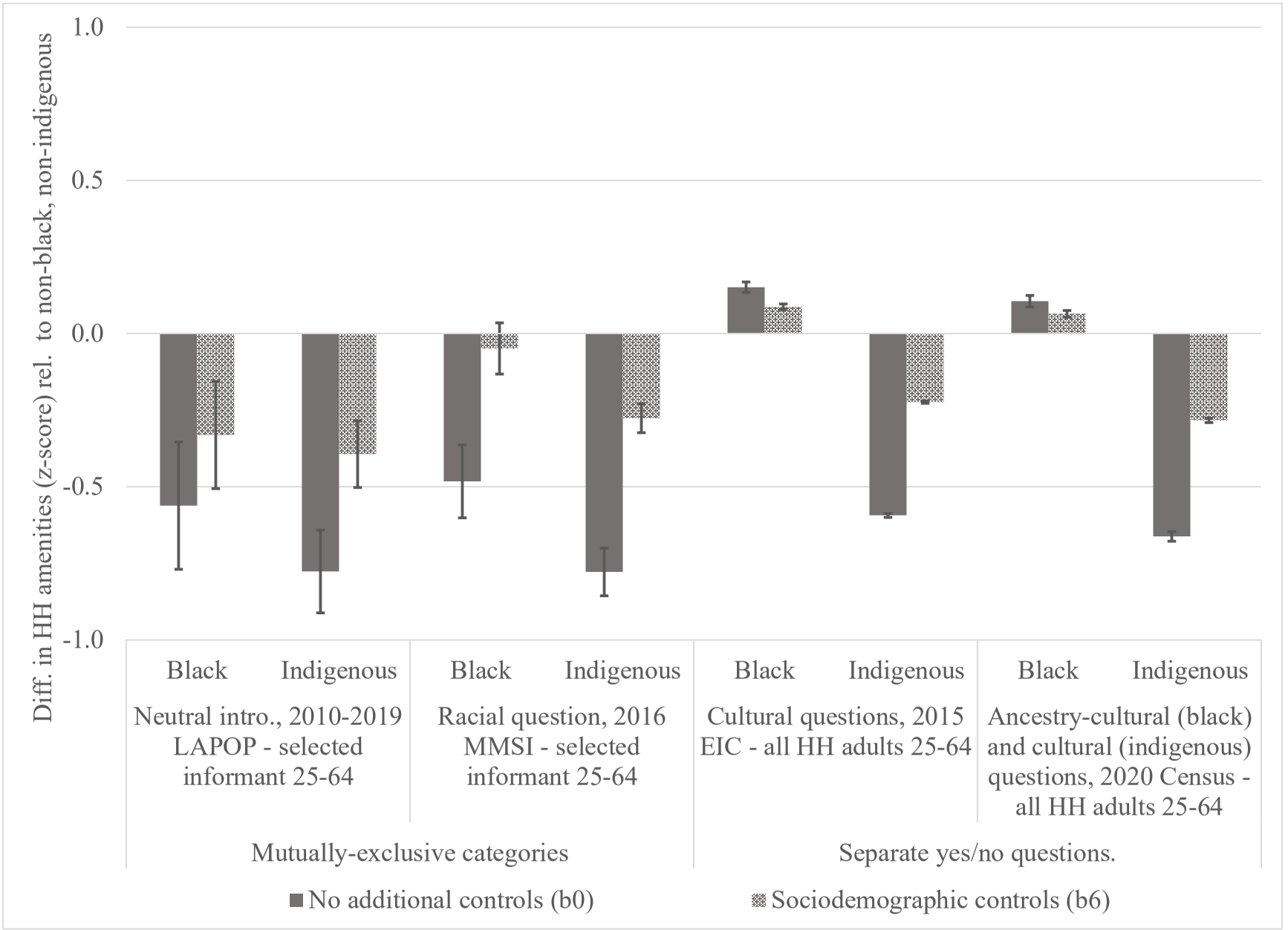


Figure 3b. Predicted difference in household amenities index (z-score) and 95% confidence intervals for black- and indigenous-identified individuals ages 25-64 relative to non-indigenous, non-blacks, by model adjusting for different controls, 2010-2019 LAPOP, 2016 MMSI, 2015 EIC, & 2020 Census long-form surveys.



Appendix I. Sociodemographic characteristics of Mexican adults ages 25-64 according to whether they were classified as black and/or indigenous "fully" or "in part", 2015 EIC.

			Black		Indigenous		
			Non-indigenous black		Black & indigenous		Non-black
	All groups	NINB*	("yes")	("in part")	("yes")	("in part")	indigenous
	Mean	Mean	Mean	Mean	Mean	Mean	Mean
	(S.E.)	(S.E.)	(S.E.)	(S.E.)	(S.E.)	(S.E.)	(S.E.)
Speaks an indigenous lang.	6.7						
	(0.036)						
Female	52.7	53.1	50.4	51.7	53.0	51.7	52.4
	(0.019)	(0.024)	(0.433)	(0.640)	(0.242)	(0.477)	(0.040)
Age	41.5	41.4	42.2	41.5	41.5	41.6	41.6
	(0.011)	(0.012)	(0.118)	(0.149)	(0.090)	(0.111)	(0.016)
Schooling (years)	9.4	9.9	10.1	10.0	8.8	9.0	7.9
	(0.010)	(0.011)	(0.065)	(0.076)	(0.055)	(0.069)	(0.012)
0 - 5 years	15.3	12.3	12.9	12.8	19.6	18.8	25.2
	(0.045)	(0.044)	(0.347)	(0.478)	(0.423)	(0.456)	(0.085)
6 - 8 years	19.6	18.4	17.1	17.5	20.7	20.6	23.4
	(0.044)	(0.050)	(0.434)	(0.533)	(0.316)	(0.436)	(0.060)
9 - 11 years	28.8	29.4	26.6	29.0	27.5	28.1	27.0
	(0.054)	(0.063)	(0.479)	(0.660)	(0.396)	(0.581)	(0.067)
12 - 15 years	20.1	21.8	23.3	22.1	19.4	18.0	14.9
	(0.050)	(0.057)	(0.504)	(0.611)	(0.369)	(0.516)	(0.067)
16+ years	16.1	18.2	20.2	18.6	12.8	14.6	9.5
	(0.082)	(0.096)	(0.577)	(0.612)	(0.355)	(0.504)	(0.068)
Urban locality	79.5	83.7	84.0	83.2	76.5	75.8	65.7
	(0.092)	(0.092)	(0.594)	(0.718)	(0.828)	(0.764)	(0.177)
Mun. marginalization (z)	-1.1	-1.3	-1.1	-1.2	-0.8	-0.8	-0.6
	(0.002)	(0.002)	(0.015)	(0.011)	(0.020)	(0.019)	(0.004)
Region of residence							
Guerrero, Costa Chica	0.1	0.0	3.2	0.8	1.2	0.5	0.1
	(0.001)	(0.001)	(0.127)	(0.079)	(0.053)	(0.037)	(0.002)
Oaxaca, Costa Chica	0.2	0.1	5.0	0.6	3.3	1.2	0.5
	(0.005)	(0.003)	(0.210)	(0.055)	(0.172)	(0.096)	(0.013)
Guerrero	2.6	2.1	12.6	4.7	13.8	5.8	3.3
	(0.034)	(0.035)	(0.798)	(0.362)	(0.590)	(0.444)	(0.060)
Oaxaca	2.8	1.1	4.8	2.2	11.5	6.4	8.5
	(0.023)	(0.015)	(0.239)	(0.170)	(0.352)	(0.338)	(0.071)
Veracruz	6.8	6.0	14.2	9.3	21.7	11.6	8.5
	(0.054)	(0.059)	(0.542)	(0.719)	(0.877)	(0.539)	(0.090)
Mexico City metro	19.3	21.7	37.7	25.2	30.5	17.9	10.6
	(0.168)	(0.192)	(1.086)	(0.976)	(0.991)	(0.929)	(0.159)
Mexico State[1]	3.5	3.1	3.8	2.2	3.6	1.9	5.0
	(0.082)	(0.091)	(0.458)	(0.224)	(0.260)	(0.178)	(0.117)
All others[1]	64.7	65.9	20.0	55.1	15.3	55.0	63.7
	(0.168)	(0.193)	(0.971)	(1.012)	(0.831)	(1.039)	(0.198)
N (thousands)	9,937	6,057	45.611	17.586	105.576	41.185	3,264

Notes: all figures are percentages unless noted otherwise; all means were weighted and standard errors were adjusted for complex sampling design.

[1] Excludes municipalities that are part of Mexico City metro area.

*** $p < 0.001$ ** $p < 0.01$ * $p < 0.05$ N.S. $p > 0.05$.

Appendix II. Descriptive statistics, adults 25-64, 2015 EIC, 2016 MMSI & 2020 Census long-form surveys.

	2015 EIC - all household members ages 25-64	2016 MMSI - sampled informant ages 25-64	2020 Census - all household members 25-64
	Mean (S.E.)	Mean (S.E.)	Mean (S.E.)
Speaks indig. language	6.7 (0.04)	7.0 (0.48)	6.5 (0.09)
Woman	52.7 (0.02)	52.6 (0.46)	52.4 (0.03)
Age (years)	41.5 (0.01)	42.0 (0.11)	42.0 (0.02)
Schooling (years)	9.5 (0.01)	9.6 (0.05)	10.0 (0.02)
0 - 5	15.3 (0.04)	14.6 (0.37)	11.6 (0.07)
6 - 8	19.5 (0.04)	18.0 (0.36)	16.8 (0.07)
9 - 11	28.8 (0.05)	31.3 (0.49)	30.3 (0.09)
12 - 15	20.2 (0.05)	19.7 (0.41)	23.4 (0.08)
16+	16.3 (0.08)	16.5 (0.38)	17.9 (0.14)
Refrigerator	88.1 (0.05)	88.6 (0.41)	90.1 (0.08)
Stove	85.5 (0.06)	91.6 (0.40)	N/A
Washing machine	74.0 (0.07)	74.8 (0.49)	77.2 (0.12)
Radio	76.1 (0.06)	68.7 (0.45)	71.7 (0.10)
TV	95.1 (0.02)	71.8 (0.45)	93.5 (0.06)
Land line	40.2 (0.11)	40.9 (0.53)	40.0 (0.18)
Cell phone	82.3 (0.05)	91.2 (0.32)	91.1 (0.07)
PC	36.1 (0.10)	37.9 (0.48)	41.3 (0.18)
Internet access	36.7 (0.11)	45.7 (0.50)	58.0 (0.18)
Car*	48.0 (0.10)	43.3 (0.51)	51.3 (0.16)
Lives in urban locality	79.6 (0.09)	79.0 (0.28)	81.3 (0.22)
State of residence			
Aguascalientes	1.1 (0.02)	1.0 (0.03)	1.1 (0.05)
Baja California	2.9 (0.06)	3.0 (0.10)	3.1 (0.12)
Baja California Sur	0.6 (0.02)	0.7 (0.03)	0.7 (0.02)
Campeche	0.8 (0.02)	0.8 (0.02)	0.7 (0.03)
Coahuila	2.5 (0.04)	2.4 (0.08)	2.5 (0.05)
Colima	0.6 (0.01)	0.6 (0.02)	0.6 (0.02)
Chiapas	3.8 (0.04)	3.8 (0.13)	3.8 (0.10)
Chihuahua	3.0 (0.05)	3.1 (0.10)	3.0 (0.09)
Distrito Federal	8.6 (0.12)	8.3 (0.24)	8.1 (0.21)
Durango	1.4 (0.03)	1.4 (0.05)	1.4 (0.04)
Guanajuato	4.7 (0.07)	4.5 (0.14)	4.7 (0.12)
Guerrero	2.6 (0.03)	2.6 (0.09)	2.5 (0.06)
Hidalgo	2.4 (0.02)	2.3 (0.09)	2.4 (0.06)
Jalisco	6.5 (0.08)	6.3 (0.18)	6.6 (0.13)
México	14.1 (0.15)	14.5 (0.35)	14.0 (0.21)
Michoacán	3.6 (0.04)	3.6 (0.12)	3.6 (0.09)
Morelos	1.6 (0.03)	1.6 (0.05)	1.6 (0.05)
Nayarit	1.0 (0.02)	1.0 (0.03)	0.9 (0.04)
Nuevo León	4.5 (0.06)	4.2 (0.11)	4.7 (0.07)
Oaxaca	3.1 (0.02)	3.1 (0.10)	3.1 (0.04)
Puebla	4.9 (0.07)	4.8 (0.15)	5.0 (0.14)
Querétaro	1.7 (0.04)	1.7 (0.05)	1.9 (0.07)
Quintana Roo	1.3 (0.04)	1.4 (0.04)	1.5 (0.13)
San Luis Potosí	2.2 (0.03)	2.2 (0.07)	2.2 (0.06)
Sinaloa	2.5 (0.04)	2.5 (0.07)	2.4 (0.07)
Sonora	2.4 (0.04)	2.5 (0.07)	2.4 (0.05)
Tabasco	2.0 (0.08)	1.9 (0.06)	1.9 (0.11)
Tamaulipas	2.9 (0.05)	3.0 (0.10)	2.9 (0.06)
Tlaxcala	1.0 (0.02)	1.1 (0.03)	1.0 (0.03)
Veracruz	6.8 (0.05)	6.8 (0.18)	6.5 (0.10)
Yucatán	1.8 (0.03)	1.8 (0.04)	1.9 (0.05)
Zacatecas	1.2 (0.02)	1.2 (0.04)	1.2 (0.03)
Sample size	10,019,727	25,587	6,801,283

Notes: weighted estimates and design-adjusted standard errors.
Unless otherwise noted, all figures are expressed as percentages.

Appendix III. Percent of the Mexican population ages 25-64 identifying as black or indigenous, various surveys and years.

Row	Data source (institution)	Dates	N (people)	Pct. black Est.	Pct. black (S.E.)	Pct. black Sig.[1]	Pct. black Sig.[2]	Pct. indigenous Est.	Pct. indigenous (S.E.)	Pct. indigenous Sig.[1]	Pct. indigenous Sig.[2]	Question wording; prompt and response options.
A. Measures with neutral introduction.												
1	LAPOP project Mexico samples (Vanderbilt) - sampled informants	2004 (Mar 17)	1,556	0.4	(0.1981)	***	***	11.4	(0.9456)	***	**	Do you consider yourself white, mestizo, indigenous, or black? (White/mestizo/indigenous/black/other/DK or non-response [NR]).
2		2006 (June 6-29)	1,560	0.9	(0.2855)	**	***	8.8	(0.8531)	***	***	Do you consider yourself to be a: white, mestizo, indigenous, Afromexican (black), mulatto, or other person? (White/mestizo/indigenous/afro-mexican (black)/mulatto/other/DK or NR).
3		2008 (Jan 27-Feb 26)	1,560	1.4	(0.3610)	N.S.	***	8.9	(0.8891)	***	***	Do you consider yourself to be a: white, mestizo, indigenous, black or Afromexican, mulatto, or other person? (White/mestizo/indigenous/black or afro-mexican/mulatto/other/DK or NR).
4		2010 (Jan 17-Feb 19)	1,562	2.2	(0.4571)	N.S.	N.S.	7.0	(0.7940)	***	***	Do you consider yourself to be a: white, mestizo, indigenous, black, mulatto, or other person? (White/mestizo/indigenous/black/mulatto/other/DK/NR).
5		2012 (Jan 25-Feb 19)	1,560	2.0	(0.4347)	N.S.	N.S.	7.3	(0.8045)	***	***	Same as in 2010.
6		2014 (Jan 24-Feb 24)	1,535	2.1	(0.4568)	N.S.	N.S.	12.5	(1.0496)	***	N.S.	Same as in 2010.
7		2017 (Jan 28-Mar 23)	1,563	4.6	(0.6989)	***	**	11.8	(1.0728)	***	*	Same as in 2010.
8		2019 (Jan 30-Mar 27)	1,580	4.3	(0.6597)	***	**	12.6	(1.0905)	***	N.S.	Same as in 2010.
B. Measures with culture-based introduction.												
9	Decennial Census long-form sample (INEGI) - all household members ages 25-64	2010 (May 31-Jun 25)	1.95 million	N/A.	N/A.	N/A.	N/A.	14.4	(0.0333)	***	N.S.	According to [NAME]'s culture, does s/he consider her/himself indigenous? (Yes/No).
10	EIC (INEGI) - all household members ages 25-64	2015 (Mar 2-27)	14.4 million	1.8	(0.0227)	N/A	***	23.5	(0.0828)	N/A	***	Black: In accordance with their culture, history, and traditions, does [name] consider themselves black, meaning Afromexican or Afrodescendant? (Yes/Yes in part/No) Separately, indigenous: In accordance with your their culture, does [name] consider themselves indigenous? (Yes/Yes in part/No).
11	Decennial Census long-form sample (INEGI) - all household members 25-64	2020 (March 2-27)	26,216	See row 13.				19.3	(0.1351)	***	***	Does [name] consider themselves indigenous, based on their traditions or customs? (Yes/No/Don't know). For separate question on black identification, see row 13.
C. Measures with race-based introduction.												
12	MMSI (INEGI) - sampled informants	2016 (Jul-Dec)	25,634	2.6	(0.1579)	***	N/A	13.9	(0.4330)	***	N/A	In our country, there are people with multiple racial origins. Do you consider yourself a … person? (Black or mulata/indigenous/mestiza/white/other [e.g., Asian, Eurodescendant] {Don't know option available but not mentioned})
D. Measures with ancestry-culture-based introduction.												
13	Decennial Census long-form sample (INEGI) - all household members 25-64	2020 (March 2-27)	6.84 million	2.1	(0.0311)	***	**	See row 11.				Given your ancestors and in accordance with your customs, does [name] consider themselves black, meaning Afromexican (Afrodescendant)? (Yes/No/[Don't know]). For separate question on indigenous identification, see row 11.

Notes: EIC estimates for black (indigenous) include individuals also identifying as indigenous (black). N = full sample size. In LAPOP surveys after 2012, question includes an instruction to interviewers that, if person identifies as "Afromexican", interviewer should code response option as "Black."

[1] Diff. bet. estimate relative to corresponding one for 2015 EIC ("yes" + "in part") is statistically significant at...***0.001 **0.01 *0.05 … level (otherwise, N.S. = not sig. at 0.05 or lower level).

[2] Diff. bet. estimate relative to corresponding one for 2016 MMSI is statistically significant at...***0.001 **0.01 *0.05 ...level (otherwise, N.S. = not sig. at 0.05 or lower level).

[3] Question includes "yes" & "in part" response options, chosen by interviewer based on interviewee's response (see text for explanation).

Appendix IV. Description of procedure to estimate growth vs. wording components in difference between 2015 EIC and 2016 MMSI.

As described in the text, the goal is to decompose the difference between estimates of black identification between 2015 EIC and 2016 MMSI into growth and wording components. The procedure is depicted in Equations 1 through 5, explained below:

$$Share_{growth} = \frac{\left(\widehat{EIC}_{2016.75} - \widehat{EIC}_{2015.2083}\right)}{\left(\widehat{MMSI}_{2016.75} - \widehat{EIC}_{2015.2083}\right)} \quad (1)$$

$$\widehat{EIC}_{2016.75} = \widehat{EIC}_{2015.2083} \cdot e^{t \cdot r_{LAPOP}} \quad (2)$$

$$t = 2016.75 - 2015.2083 = 1.541667 \quad (3)$$

$$r_{LAPOP} = \frac{ln\left({LAPOP_{2017-2019}}/{LAPOP_{2010-2014}}\right)}{2018.125 - 2012.125} \quad (4)$$

$$Share_{wording} = 1 - Share_{growth} \quad (5)$$

Meaning, we estimated the growth contribution (Eq. 1) by dividing the difference between the estimated growth in cultural identification between the EIC and MMSI survey periods (centered in mid-March 2015 and early October 2016, respectively) by the difference between actual EIC and MMSI estimates. All figures in Eq. 1 are observed from the data, with the exception of $\widehat{EIC}_{2016.75}$, i.e., the share in black cultural identification one might have observed using the EIC cultural question at the time of the MMSI survey, which we estimated by projecting the actual

2015 estimates ($\widehat{EIC}_{2015.2083}$) for the period between surveys (t, see Eq. 3) using the mean annualized growth rate in black identification between 2010-2014 and 2017-2019 LAPOPs, centered in mid-February 2012 and 2018, respectively (see Eq. 4). Finally, we estimated the wording contribution to be the complement of the growth contribution (Eq. 5).

To estimate the 95% confidence interval around the contributions of growth vs. wording, we bootstrapped standard errors by simultaneously randomly drawing estimates off the 95% confidence interval distributions of the share black in 2010-2014 and 2017-2019 LAPOPs, 2015 EIC, and 2016 MMSI, doing 2,000 of these "simulations" to obtain the 5th and 95th percentiles of their distribution along the mean.