Unveiling the Cosmic Race:
Racial Inequalities in Latin America*

L. Guillermo Woo-Mora†

25 June, 2021

Abstract

Latin America is one of the regions with the highest income inequality and one of the most racially diverse. Historically, most Latin American countries build their national identities through a 'melting pot' ethnic figure: *'mestizos'* or *'mulatos'* —the mixed-race descendent from Europeans, Indigenous, and African from the colonial period. However, Latin American countries have veiled income inequalities between racial groups through the *'Cosmic Race'* or *mestizaje* identity. Using LAPOP AmericasBarometer data, I compile information on skin tone and proxies for income for nearly 150,000 individuals across Latin American countries during the last decade. The purpose of the paper is twofold. In the first part, I estimate newly racial inequality measures at the national level. Countries with higher income inequality between racial groups have worse economic development: a one percent increase in the ratio of racial over total income inequality correlates with a decrease of nearly 4 percent in GDP per capita. In the second part of the paper, I use Oaxaca-Blinder decomposition and control functions to analyze the racial income gap at the individual level. Every darker skin tone out of an eleven-color palette has at least 6 percent less monthly income per capita. Sixty percent of the effect cannot be attributed to returns to observable average characteristics of the racial groups, and most plausibly can be attributed to racial discrimination. There is substantial heterogeneity between countries. Alongside justice and reparations, progressive taxation in income or 'tagging' could decrease race-based disparities and improve economic development.

Keywords: Race, Inequality, Latin America, Discrimination, Economic Development.
JEL: *D3*, *J15*, *J71*, *O12*, *O54*, *Z13*.

1 Introduction

Inequalities lie beyond class struggle. Besides income, wealth, occupation, or educational attainment, other dimensions also explain distributive conflict: gender, race, ethnicity, religion, nation, or other social identities (Akerlof and Kranton 2000; Chakravarty 2015; Fleurbaey and Schokkaert 2011; Piketty 2020; Shayo 2009). Most economic literature studying inequalities has mainly relied on interpersonal comparisons of income or wealth, weighting little attention to the latter dimensions. Not surprisingly,

*I am grateful to Thomas Piketty and Ignacio Flores for their valuable advise and support. I am also specially grateful to Mauricio Tenorio-Trillo. Thanks to Ekaterina Zhuravskaya, Hillel Rapoport, Emiline Bezin, Sandra Poncet, Eva Arceo-Gómez, Raymundo Campos-Vázquez, Sandra Ley, Alice Krozer, María Montoya-Aguirre, Eduardo Ortiz-Juárez, Enrique de la Rosa-Ramos, Jesús Carrillo, and Federico Daverio for their helpful comments. All remaining errors are my own.

†Thesis to obtaint MSc in Analysis and Policy in Economics. Paris School of Economics, France. guillermo.woomora@pseestudent.eu.

ethnic and racial inequalities are remarkably understudied in economics compared to other social sciences, like sociology and political science: by 2020, the share of race-related publications in economics was nearly 2% (Advani et al. 2021).

Nonetheless, 2020's events showed the salience of racial inequalities: the murder of George Floyd in the U.S. was followed by mass mobilizations like Black Lives Matter denouncing racial discrimination and disparities.[1] Soon after, movements in other latitudes denounced other forms of ethnic-racial disparities: the Commission on Race and Ethnic Disparities in the United Kingdom provoked controversy;[2] the old debate on the lack of racial statistics renewed relevance in the French public debate;[3] Furthermore, movements denounced the 'invisible' racism on blacks, brown-skinned, and indigenous in many Latin American countries.[4]

This paper studies racial inequalities in Latin America. Latin America is one of the most unequal regions in the world: even when there was a decrease in income inequality in the last two decades due to higher educational attainment and better oriented social policies (Lustig, Lopez-Calva, and Ortiz-Juarez 2016), overall income and wealth inequality still high in absolute terms and relative to other regions and countries (De Rosa, Flores, and Morgan 2020). Moreover, the region is one of the most diverse, both racially and ethnically. The latter matters since, besides an extractive elite and differently from other Western experiences, race and ethnicity are also central dimensions shaping income and wealth inequalities (Sánchez-Ancochea 2021).

Using the Project on Ethnicity and Race in Latin America (PERLA) color palette and AmericasBarometer Survey data from the Latin American Public Opinion Project (LAPOP), I compile a data set of repeated cross-sections with representative information on skin tone and economic outcomes for nearly 150,000 individuals in more than 25 countries for the last decade. The purpose of this paper is twofold.

First, in the spirit of Alesina, Michalopoulos, and Papaioannou (2016), I estimate new measures of racial inequality -income inequality between racial groups- at the national level using skin tone rather than the conventional broadly defined ethnic categories. The contribution is not innocuous since there is a blurry border between ethnicity and race in Latin America due to *mestizaje* politics: namely, the idea of a shared ethnicity descendant of the miscegenation between Indigenous, Europeans, and Africans.

I use the Mean Log Deviation decomposition (Foster and Shneyerov 2000) to measure income inequality within and between components for skin tone groups. Then, I compute the ratio of the inequality between racial groups with respect to the total inequality measure. Using cross-country variation for the last decade and controlling for time-invariant characteristics and common-shocks trends, I find that higher ratios of racial inequality with respect to total income inequality are correlated to lower GDP per capita. Consistent with the ethnic-inequality results (Alesina, Michalopoulos, and Papaioannou 2016), higher racial inequality in income hinders economic development.

In the second part of the paper, I study the underlying mechanisms of racial inequalities at the individual level. More specifically, I estimate the skin tone effect on monthly household income per capita using

[1] See https://www.economist.com/leaders/2021/05/22/race-in-america.
[2] See https://www.bbc.com/news/uk-56585538.
[3] See https://www.economist.com/leaders/2020/11/21/a-lack-of-data-on-race-hampers-efforts-to-tackle-inequalities
[4] See https://elpais.com/sociedad/2020-06-08/el-racismo-invisibilizado-en-america-latina-alza-la-voz.html.

an Oaxaca-Blinder decomposition for continuous variables (Ñopo 2008b). To address measurement error concerns, I use a control function approach (Rios-Avila 2019; Wooldridge 2015): I instrument respondent's skin tone by interviewer's skin tone.

The effect is unambiguously negative: a darker skin tone has between 4 and 6 percent fewer monthly income per capita, out of a color palette with eleven tones. Around one-third of the effect can be explained by differences in average characteristics, such as lower years of schooling. Nevertheless, nearly sixty percent of the racial gap is related to unobserved factors, as race discrimination. Consistent with the race discrimination hypothesis, I find that individuals with darker skin tones report higher discrimination behavior against them. Lastly, racial inequalities are country-dependent: there is substantial country-specific heterogeneity. Consistent with historical and anecdotal evidence, countries with a high share of 'mestizo' population have higher racial gaps.

This paper is related to burgeoning literature on ethnic-racial inequalities. There is growing literature on the black-white gap on income, wealth, years of schooling, segregation, health outcomes, or economic opportunities in the United States (Chetty et al. 2020; Cook, Logan, and Parman 2016; Derenoncourt and Montialoux 2020; Logan and Parman 2017). Furthermore, besides disparities at the individual level, there is also an increasing interest in studying group-based inequalities and their potential to explain comparative economic development. For instance, recent work by Alesina, Michalopoulos, and Papaioannou (2016) uses data on nightlights intensity and the location of ethnic and language groups to show that economic inequalities between ethnic groups are correlated with lower economic development at the national level, rather than ethnic diversity by itself (Alesina and La Ferrara 2005). The correlation robust after controlling for spatial and administrative inequality, as well as geographical factors.

Nonetheless, two problems arise with most of the existing literature on ethnic-racial inequalities. The first one is that it is challenging to disentangle inequalities across a cultural dimension and physical phenotype. Sociologists argue *race* or *racialization* are physical characteristics or phenotype that can define group membership, while *ethnicity* membership is often based on cultural characteristics (Telles and Martínez Casas 2019). Most ethnic-race-related studies account for differences between broad ethnic categories as "White," "Black," "Asian," or "Latino," which can describe both phenotype and cultural characteristics simultaneously. For example, Alesina, Michalopoulos, and Papaioannou (2016) work accounts for cultural differences, but it does not account for disparities between racial groups.

The second issue is related to the scope of the study cases. Economic race-related studies mainly focus on the U.S. due to the historical experience of slavery, segregation, discrimination, and wealth inequality (Cook and Logan 2020). Nevertheless, racial inequalities are also salient in other countries and regions, some with historical experiences of slavery and segregation. For instance, when a U.S. economist says that 'racial inequality is an American tradition,' (Fryer 2011) he implicitly excludes the *Latin* part of the continent. After all, Latin America, as a historical and political category, is also a concept of race: the not-Anglo-Saxon America (Tenorio-Trillo 2017).

This study is closely related to a burgeoning interdisciplinary literature on ethnic-racial inequalities in Latin America. During the last two decades, scholars from sociology, political science, and economics have compiled extensive evidence studying how ethnic-racial components determine economic outcomes in the region. For instance, there is literature on the ethnic-racial effects on the labor market and wages (Arceo-

Gómez and Campos-Vázquez 2014, 2019; Campos-Vázquez 2020; Ñopo, Saavedra, and Torero 2007; Ñopo 2012), educational attainment (Botelho, Madeira, and Rangel 2015; Telles and Steele 2012; Telles and Martínez Casas 2019), social mobility (Campos-Vázquez and Medina-Cortina 2019; Monroy-Gómez-Franco and Vélez-Grajales 2020; Solís, Güémez Graniel, and Lorenzo Holm 2019), as well as social norms and identities (Campos-Vázquez and Medina-Cortina 2017). Racial components also determine electoral preferences (Aguilar 2011; Campos-Vázquez and Rivas-Herrera 2021) and increase the perceptions of racial discrimination (Chong and Ñopo 2008; Ñopo, Chong, and Moro 2009; Trejo and Altamirano 2016).

Besides inequality of opportunity (Ferreira and Gignoux 2011), differences in access to public goods, or occupational segregation, economic literature have argued that discrimination is the main mechanism driving ethnic-racial disparities in Latin America (Ñopo, Chong, and Moro 2009). However, economics has a specific understanding of discrimination. As explained by Altonji and Blank (1999), discrimination is *"a situation in which persons who provide labor market services and who are equally productive in a physical or material sense are treated unequally in a way that is related to an observable characteristic such as race, ethnicity, or gender."* The latter definition implies that ethnic-racial disparities are not due to discrimination as long as two individuals are *differently productive*. Therefore, as Ñopo, Chong, and Moro (2009) explain, it is more helpful to distinguish between two types of discriminating behavior: 1) people treating members of certain groups differently simply because they do not like them, or preference-based discrimination; and 2) people using group membership as a proxy measure for unobserved characteristics, statistical discrimination. The second is closely related to the concepts of stigmatization used by other social sciences like sociology.

The previous literature studying ethnic-racial disparities and discrimination has contributed significantly by showing the issue salience, but it has some relevant shortcomings. Firstly, most of the previous literature focuses on a single country, like Brazil, Mexico, or Peru. Also, there is little comparability on the measures of ethnicity and race since their meaning and understanding changes country by country. Lastly, there is no single methodology for the studies, but a mix of descriptive statistics, regression analysis, structural estimations, decomposition techniques, and some experimental and quasi-experimental designs.

To my best knowledge, there are few studies on ethnic-racial disparities with comparable data and methodology for more than one Latin American country (Cernat, Sakshaug, and Castillo 2019; Ñopo 2012; Telles and Steele 2012; Telles and Martínez Casas 2019; Zizumbo-Colunga and Flores Martínez 2017). While some present historical, anecdotal, and descriptive statistics, they do not address endogeneity concerns when studying ethnic-racial disparities. Cernat, Sakshaug, and Castillo (2019) does present evidence of measurement error concerns. Lastly, Ñopo (2012) studies the wage gap between ethnic groups for Bolivia, Peru, Brazil, Guatemala, Paraguay, Chile, and Ecuador, using an Oaxaca-Blinder decomposition with matching techniques (Ñopo 2008a). While Ñopo presents more convincing evidence on the extent of discrimination and differences in endowments to explain the wage gap, the use of broad ethnic categories rather than racial phenotype presents a challenge in countries where the majority of the population can define themselves as the broad ethnic category of *mestizo* or *mulato*.

This paper contributes to the ethnic-race-related economic literature in several dimensions. Firstly, it is one of the few studies proposing using the racial dimension rather than the ethnic one to measure the relation of inequalities across groups and economic development at aggregated levels. More specifically, in line with Alesina, Michalopoulos, and Papaioannou (2016), it is the first to use skin tone rather than broad ethnic-racial categories to compute racial inequality measures at the national level. Secondly, this is one of the few studies comparing racial gaps at the individual level using skin tone measures a comparable methodology and a sizable sample of individuals across multiple Latin American countries. Moreover, to my best knowledge, it is one of the first studies to estimate the unobserved component of racial disparities at the individual level, purging endogeneity through control functions by using the interviewers' skin tone as an instrument for the respondent's skin tone.

The rest of the paper is structured as follows. Section 2 reviews the racial question in Latin America from a historical perspective. Section 3 presents the data. Section 4 estimates the racial-inequality measures at the aggregate level and their relationship with economic development. Section 5 presents the identification strategy and results for the effect of skin tone on monthly income per capita at the individual level and its mechanisms and the heterogeneous effects. Section 6 concludes.

2 Historical review: The racial question in Latin America

Race and ethnicity have been central dimensions for Latin America, as a region, and within the countries that are part of it. As historian Tenorio-Trillo (2017) argues: *"Latin America, from its origins as a concept to this day, has fundamentally been a changing version of a single, enduring, old, and seemingly insurmountable concept: race."* Tenorio-Trillo argues that, besides notions of underdevelopment, violence, corruption, illiberal democracies, political instability, traditionalism, Catholicism, impossible utopias, or backwardness, the Latin prefix responds to the need to distinguish the not Anglo-Saxon region of the continent. *"There would be no Latin America and no Latinos/as without the United States"* (Tenorio-Trillo 2017). Moreover, within the countries fulfilling the enlisted characteristics, there has been a relatively stable racial order. This section reviews the racial question in Latin America from a historical perspective.

The encounter of the original population, later labeled as Indians or Indigenous,[5] European conquerors and populations from Africa, mostly brought as slaves or forced labor, produced an early miscegenation process throughout the continent.[6] Nonetheless, the racial mixture was not unnoticed nor unattended by the colonial empires. As Loveman (2014) argues: *"Through legal and administrative institutionalization of certain categorical divisions and not others, imperial governments in the Americas actively constituted social groupings within the colonized population that came to be perceived as natural groupings."* Figure 1 depicts one of the famous *casta* paintings, representing the everyday life of the different racial and ethnic mixtures in New Spain, nowadays most Mexico and Central America. During the colonial period, in other contexts of racial apartheid as in Spain, *casta* paintings were a precious art piece due to the exotic representation of a society where the racial mixture was not prohibited but somehow the rule.

Graham (2013) has the best historical overview of the colonial caste system across the continent. Most

[5] According to Knight (1990) *"The attribution of the Indian identity began, of course, with the Conquest."*

[6] The (re)encounter of the 'New' and 'Old' World begins after 1492, when the Spaniards, lead by Columbus, arrived to the shores of the island that nowadays is shared by Haiti and Dominican Republic. During the whole XVI century Spaniards, Portuguese, French and English dedicated to explore the unknown territories, and conquest, sometimes with local allies, pre-colonial societies throughout the continent. For a review of the early stages of the colonial period in the Americas see Taylor (2013).

Figure 1: Casta Paintings

Indigenous or Indians had a corporate status and communal properties, ensured by the king. Mestizos, the racial mixture of Indigenous and Spanish, lay between white Spaniards and Indigenous but were not accepted by both. There was also a substantial presence of Black slaves, freedmen, and freeborn population in Brazil and coastal areas of Nueva Granada (nowadays Colombia and Venezuela) and the coastal areas of nowadays Mexico, Argentina, and Peru. The mixture of African and European descent, Pardos or free mulattos, had to pay head taxes and were forbidden from marrying whites. Regarding whites or Spaniards, there were those born in Europe, *peninsulares*, and those born in the Americas, Creoles. The first had higher rankings in the most prestigious social positions.

Rather than a strict system, the caste system was a complex legal and political arrangement. As Graham (2013) also explains, in colonial Spanish-America, but also similarly in Brazil, *"the Hapsburg political and economic system had been reinforced by a social structure in which everyone was assigned a fixed position within a multilayered set of social categories (to be sure, always more flexible in practice than in theory)."* The flexibility was a matter of incentives since the caste system was also a tax and legal benefit system. For example, Loveman (2014) describes how some Indigenous wanted to pass for Mestizo to avoid paying tribute, while some Mestizos wanted to pass for Indigenous to escape Inquisition processes.

Even when each group enjoyed certain rights and prerogatives, whiter people were, on average, higher on the social ladder. In sharp contrast, Indigenous and Black were undeniably at the bottom of the caste system. Many Indigenous and Black population were used for forced labor in head-tax systems as the *encomienda*, *repartimiento*, *mita*,[7] and later in the plantations or mines as slaves (Loveman 2014). Though it is tempting to argue that racial disparities originated solely as a response to the colonial order, some scholars show evidence that labor disparities based on ethnicity and race can be tracked by pre-colonial social hierarchy and the ability of the local elites to coerce local labor (Arias and Dirod 2014).

Overall, during the colonial period, whiter meant closer to European and thus with higher social status, while there were debates on the human nature of Indigenous and Black population (Graham 1990). It is not surprising to find more historians arguing how the class-based struggle was not the only salient dimension in colonial Latin America, but instead its intersection with race and ethnicity (Anderson 1988; McCaa, Schwartz, and Grubessich 1979). However, racial disparities during the pre-colonial and colonial periods were not motivated in racial discrimination behavior, or racism, as one understands it nowadays, given racial theory had its origins until the second part of the XIX century.

Graham (1990) explains that racial theory was originated by advances in science, like biology, but mostly from bad takes of social scientists promoting notions of *"[social] evolution" as the "survival of the fittest [humans]."* Moreover, he claims:

> The idea of race as it was formulated in the nineteenth century seems to have served that function both within particular countries and in maintain or at least in justifying the economic and political power exercised by some nations over others [...] The idea of race also made possible, paradoxically, for mestizos and mulattoes –by identifying themselves with white elites as against Indian or black majorities– to accept theories that justified with domination over "colored" populations. (Graham 1990)

[7] See Dell (2010) for the long-term consequences of the *mita* system in Peru.

Thus, the complex caste system from the colonial period was used to justify social, economic, and racial disparities after the fall of the Ancient Regime.

After their independence in the early XIXth century, there are two relevant processes for the new Latin American states regarding the racial question. The first one is the nation-state building process. Without the sovereign, people in the Americas started to think about nationalities: Mexicans, Argentinians, or Venezuelans. It is not a subtle point since it was unclear who had citizenship in the new nation-states. Loveman (2014) explains that *"[s]laves, former slaves, Indian peoples, mestizos, and other castas fought on both sides of the wars for independence, as both sides held out promises of emancipation and equality."* For instance, as in the Civil War in the United States, the question about military service by slaves in Brazil caused significant controversies concerning their right to acquire freedom and citizenship (Izecksohn 2014).

Parallelly, the break with the colonial Regime implied the dissolution of the caste system. The liberalism tradition of the early XIXth century, centered on individual freedom, was incompatible with the casta corporate system. In the words of Loveman (2014), *"the end of the empire did, however, bring with explicit repudiation of the ideological rationale and legal architecture upon which the hierarchical society of castes has been created."* But it was not only a matter of ideology but also translated into formal institutions: *"across the region, newly penned republican constitutions declared the practice of official ethnoracial classification to be historically obsolete"* (Loveman 2014). The unintended consequence of the *color blindness* of the XIXth century states is that ethnoracial classification in the Americas disappeared until the late XXth century.

Between the late XIXth century and the mid-XXth century, amidst the nation-building process and the heyday of the racial theory, the racial question in Latin American countries followed two paths. Firstly, countries like Argentina, Uruguay, and Chile, following racist takes on how development was related to the whiteness of a country population, intensified their efforts to whiten their population through migration of European populations (Helg 1990). Nevertheless, other countries with a higher racial mixture from the colonial and pre-colonial periods changed their narrative. Countries as Mexico and Brazil promoted the formation of national identities by reinforcing the 'melting pot' ethnic identity of *mestizos* or *mulatos*.

Restricted by the historical experience of high racial mixture, the latter states sized the idea of promoting a common ethnicity descendant of the miscegenation between Indigenous, Europeans, and Africans. In the words of Knight (1990) regarding the aftermath of the Mexican Revolution: *"the old Indian/European thesis/antithesis had now given rise to a higher synthesis, the mestizo, who was neither Indian or European, but quintessentially Mexican."* Note how *"the process of mestizaje, sometimes seen as basically racial, is in fact social"* (Knight 1990). Then, instead of dealing with race, these states created a monolithic ethnic identity almost synonymous to their national identity.[8]

There was an active movement of Latin American intellectuals responding to the racial theory seeking the whitening of population and promoting the *mestizo* and *mulato* identity. Franz Boas in Brazil was among the first intellectuals to demystify and refute "scientific" racism (Skidmore 1990). Later on, his student, Gilberto Freyre, argued that *"Brazil's ethnic potpourri [. . .] was an immense asset"* (Skidmore 1990).

[8]Parallelly to *mestizaje* politics, other forms of racial theories and politics by the Mexican Revolutionaries were a profound sinophobia and *indigenismo*, *"yet another non-Indian formulation of the 'Indian problem'"* (Knight 1990).

In Mexico, the philosopher José Vasconcelos baptized the mestizo racial mixture in the Americas as the *'Cosmic Race'*. Published in 1925 with the title *La Raza Cósmica: Misión de la Raza Iberoamericana* (*The Cosmic Race: Mission of the Ibero-American race*), Vasconcelos argued in his essay that racial hybridism most valuable virtue was *"the ability to blend different races possessing different qualities"* (Knight 1990).

It might be understated, but Latin American mestizaje politics were an appealing alternative for racial theory in the early XXth century. Skidmore (1990) presents historical evidence. An anti-racist Brazilian Manifesto before World-Ward II declared:

> People of Indian, European, and African origins have mixed in an atmosphere of such liberality and such complete absence of legal restrictions on miscegenation that Brazil has become the ideal land for a true community of people representing very diverse origins... This Brazilian philosophy on the treatment of races is the best weapon we can offer against the monstrous Nazi philosophy, which is murdering and pillaging in the name of race.

The mestizaje identity served as the *third way* to the salient racial as segregation and anti-miscegenation in the US or racial hate from the Nazi ideology. Later on, most Latin American countries followed the mestizaje politics.

In the early XXIth century, the *mestizo* or *mulato* identity is still in force. As the next section shows, the vast majority of the population in most Latin American countries define themselves as *mestizos* or *mulatos*, even when they belong to groups with different racial phenotype. Nevertheless, even when the racial question was not as extreme as in the US or World War II Europe, through mestizaje politics, Latin American countries could veil racial inequalities within them given that everyone became *mestizo* or *mulato*.

The patterns of the colonial and post-colonial persist: whether people are better off in many socio-economic dimensions. Moreover, different forms of racism are still present in everyday life. For example, whiteness is still regarded as an ideal aesthetic of beauty and wealth (Krozer and Urrutia Gómez 2021). As anecdotal evidence, Figure 2 shows an internet cartoon depicting the skin tone of different labor occupations.[9] The cartoon ironically depicts how white individuals have higher status or are more affluent, while labor occupations with lower status and lower-paid are usually for people with darker skin tones. Besides the historical and anecdotal evidence, it was not until recently that social sciences focused on developing methodologies to measure ethnic and racial identities and compile evidence on their disparities.[10] As the bulk of work previously mentioned studying the effects of ethnicity and race on economic outcomes, and alongside class and income, race is back in the center of attention for social researches.

In conclusion, race has returned. Tenorio-Trillo (2017) ironically states:

> Nowadays, whether one lives in Chicago or San Pedro de Macorís, the Dominican Republic, one has to be careful. The doorbell may ring at any moment, and one could carelessly open

[9]Taken from Cinismo Ilustrado (https://cinismoilustrado.com/).

[10]Contemporary sociologists were the first to study differences in socio-economic outcomes between ethnic and racial groups (Telles and Murguia 1990).

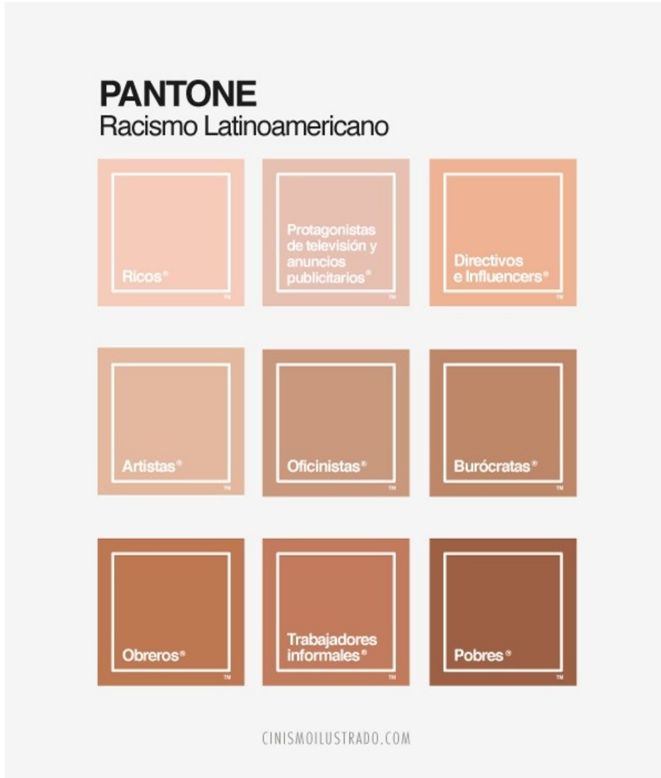

Figure 2: Cartoon Contemporary Racism in Latin America

> it to find a social scientist at the door asking how one identifies: Black? White? Mestizo? Hispanic? Latina? The pollster could be carrying a palette in her hand to contrast one's answer with the actual color of one's face (an unfair thing to do: our buttocks, not our faces, are the keepers of our true color).

Besides the best measurement method for skin or face tone, the 'colorism' fever might have important benefits for social science research. As this work will show, the use of color palettes represents an improvement to study racial disparities in Latin America given the extended *mestizo/mulato* identity, result of the historical experience previously described. More importantly, given that racial disparities and racism are global phenomena, learning from the Latin American experience, both historically and with contemporary evidence, can shed light on how to tackle racial inequalities in other latitudes.

3 Data

Measuring racial inequalities is a challenging task due to data restrictions. Few surveys and censuses register individual racial characteristics besides ethnic self-recognition, namely, whether the respondent sees herself as part of a specific ethnic group (i.e., White, Afro, Latino, Asian, among others). In other cases, the interviewer infers the ethnicity of the respondent. In some extreme cases, there are no available measures of racial categories (The Economist 2020).

Such a shortage of data is not by chance. Loveman (2014) argues that the new nation-states stopped registering race and ethnicity from national censuses from the beginning of the XIXth century. The practice responded mainly to ideological concerns: the new regimes would be *color blind* to its citizens. Even it is highly desirable to universalize citizenship and rights, the unintended consequence is that information on race and ethnicity was absent during the next two centuries.

To address the racial inequalities at the individual and national levels, I use the AmericasBarometer from the Latin American Public Opinion Project (LAPOP). The AmericasBarometer is a survey conducted every two years in 31 countries in the Americas "with stratified nationally representative samples drawn in each country, a common questionnaire score, and country-specific modules."[11] From 2010, LAPOP has used the Project on Ethnicity and Race in Latin America (PERLA) palette developed by Telles and Martínez Casas (2019) and coauthors.[12] Figure 3 shows the PERLA color scale. The scale ranges from 1 to 11, where one is the lightest skin tone and eleven is the darkest. In practice, interviewers are asked to discretely annotate the respondent's skin color taking as reference the PERLA palette without showing the guides to respondents (Dixon and Telles 2017).

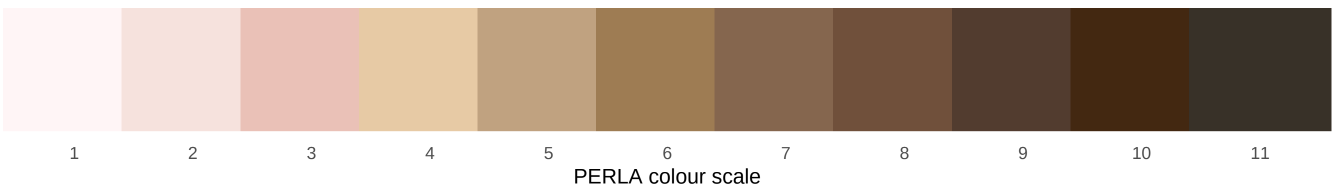

Figure 3: Perla Palette

Given that it is a public opinion survey, most of LAPOP's questions are regarding beliefs and preferences. However, the critical feature of LAPOP data is the information on race and ethnicity. Using LAPOP data, I compile all the AmericasBarometer surveys that include the PERLA palette in the core questionnaire. The sample includes more than 150,000 individual observations from 26 countries across five waves (2010, 2012, 2014, 2016/2017, 2018/2019). Most countries are present in the five waves, while some Caribbean countries had the survey just once. Table 6 in the appendix shows the sample size by country and wave.

Figure 4 shows the distribution of skin tones by country. Racial characteristics measured by skin tone vary substantially across and within countries. For instance, the darkest-skin tones are a majority in the Caribbean. Nevertheless, there is a black-skinned tone population in every country, which is usually omitted since they are a minority. Medium-dark tones are the majority in Central American countries and countries with a high miscegenation historical experience, such as Mexico, Brazil, Bolivia, Colombia, Ecuador, or Peru. Lastly, the whitest-skin tones are more usual in countries with little miscegenation and experienced high European migration during the XIXth and XXth century, like Argentina, Chile, and Uruguay.

As argued earlier, since ethnicity refers to cultural characteristics, it might differ from the racial phenotype. LAPOP works with six broad ethnic categories: Afro, Indigenous, Mestiza, Mulata, White, and other ethnic groups (i.e., Asian, Jew, among others). Figure 5 shows the ethnic distribution for each country.

[11]See https://www.vanderbilt.edu/lapop/.
[12]See https://perla.soc.ucsb.edu/.

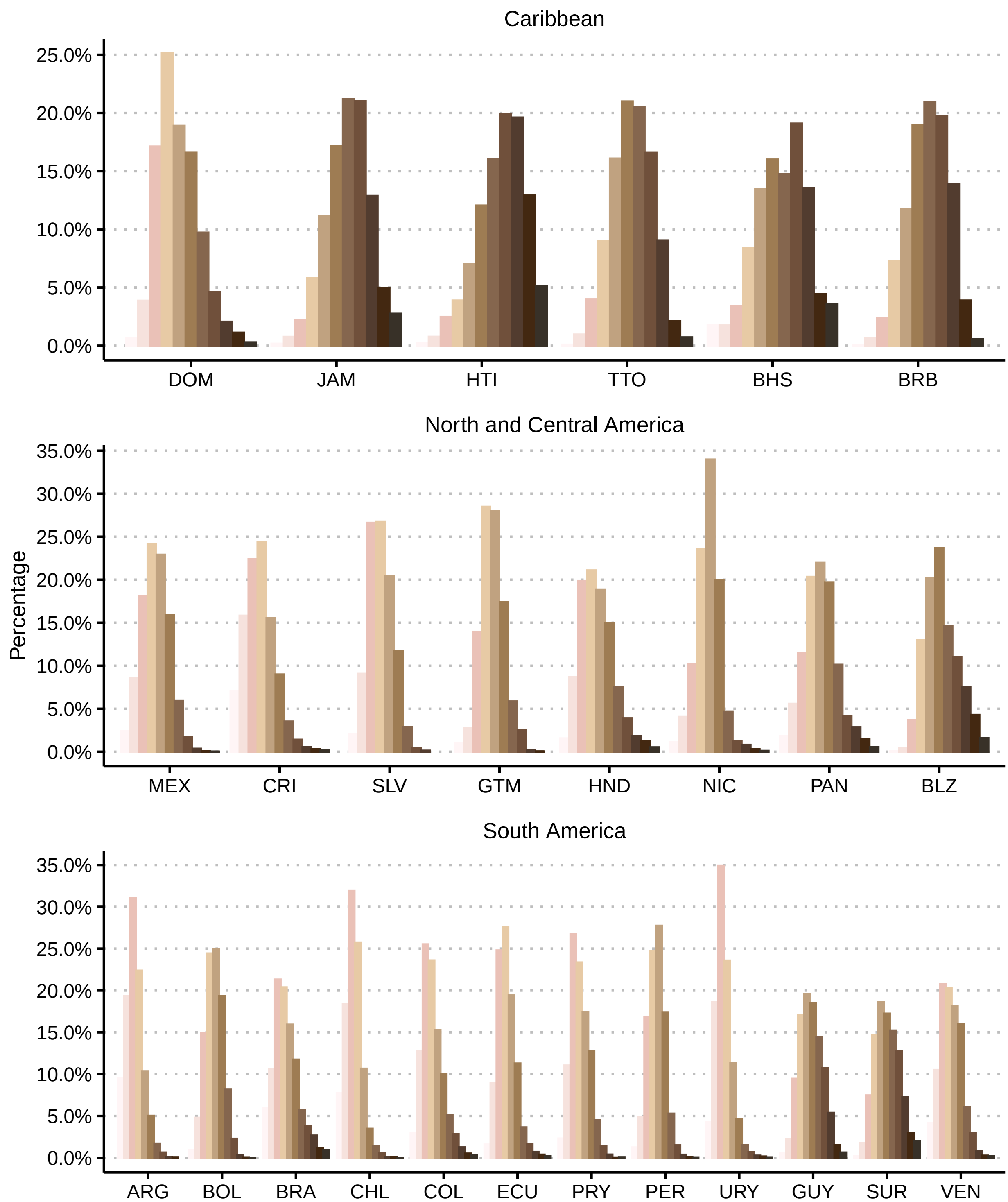

Figure 4: Skin tone distribution by country

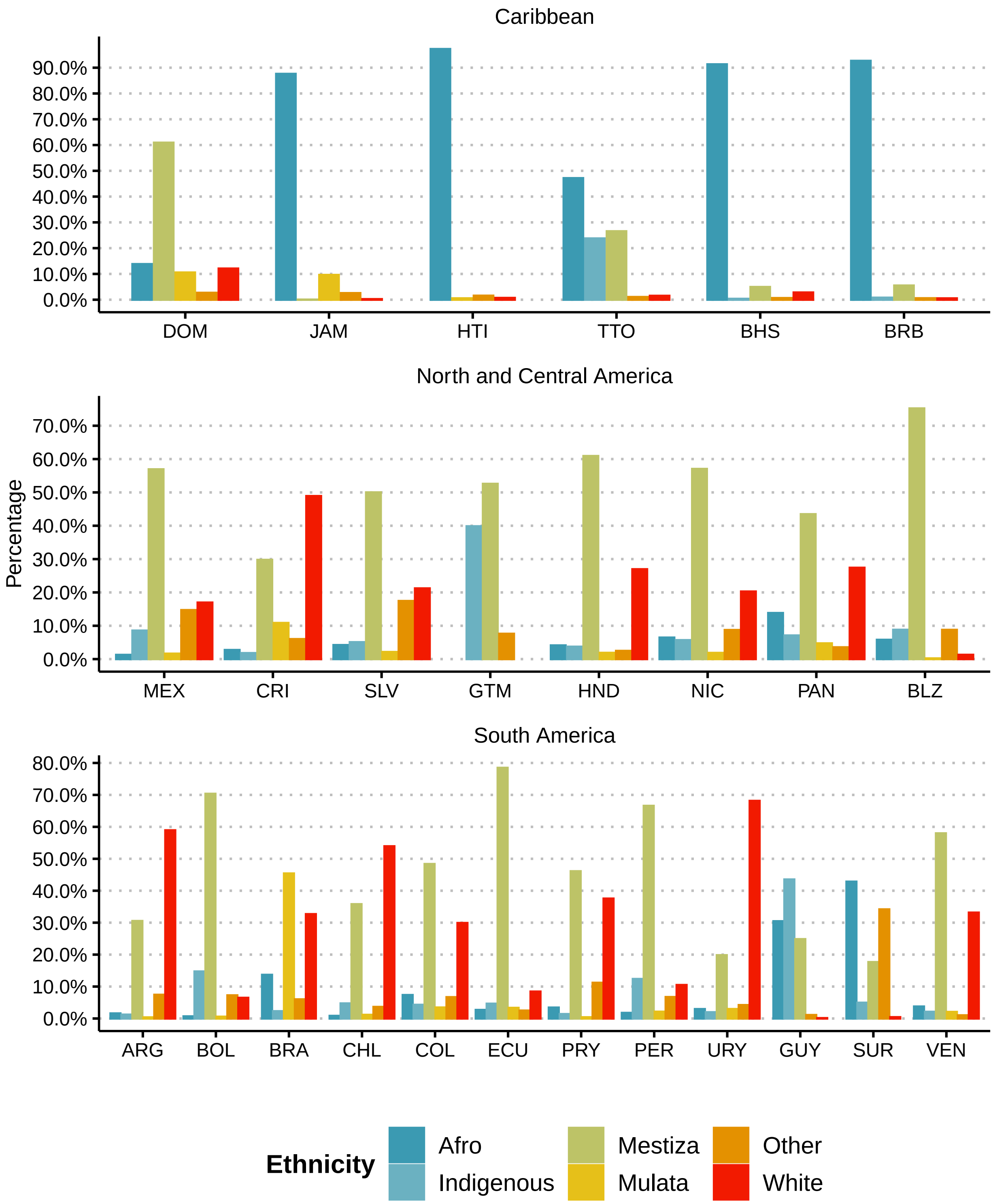

Figure 5: Ethnicity distribution by country

The majority in Caribbean countries define themselves as Afro origin, even when there is a high diversity of racial phenotype measured by skin tone. People who define themselves as White are the majority in Argentina, Chile, Costa Rica, and Uruguay. The interesting patterns of ethnicity and racialization arise analyzing countries with a high percentage of both white and medium-dark skinned populations. For instance, in every other country besides the countries where there is an Afro or White majority, most of the population defines themselves as Mestiza or Mulata. Consistent with the historical and anecdotal evidence presented in Section 2, such countries also happen to have a high historical miscegenation experience and a strong presence of *mestizaje* ideology.

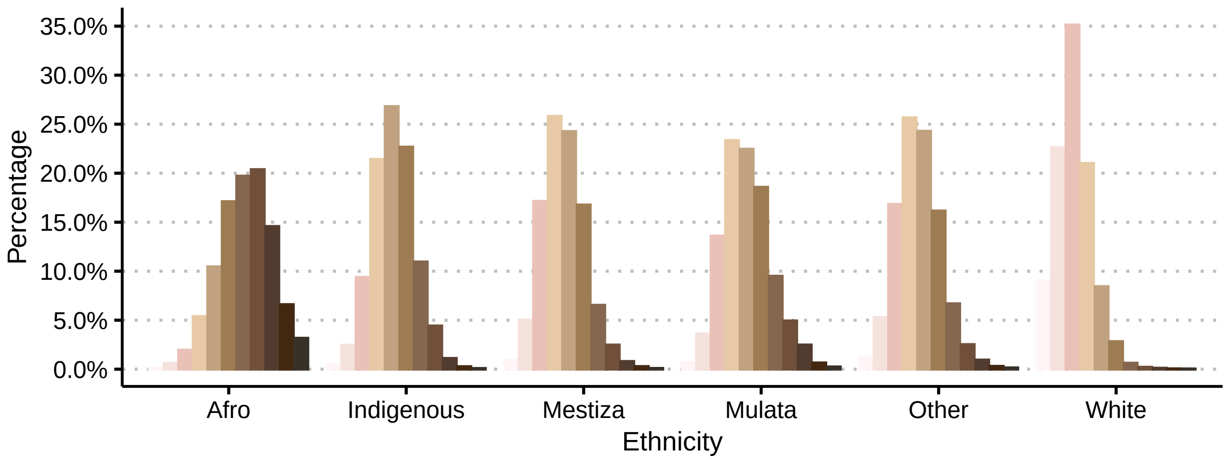

Figure 6: Ethniciy and Race

To have a better understanding of the relation between race and ethnicity, Figure 6 shows the distribution of skin tones by each of the ethnic categories for the whole sample. The patterns previously described persist: people who define themselves as Afro have darker skin tones, while those who define themselves as White have whiter skin tones. People who define themselves as Indigenous, Mestiza, Mulata, or from other ethnic groups have mostly medium-dark skin tones, but there is substantial variation in skin tone distribution. Thus, even when the distribution of skin tones is broadly correlated to the distribution of self-reported ethnicity, there is substantial diversity of racial phenotypes within an ethnic group.

Using the information on skin tone has significant advantages over using self-reported ethnicity. Since the *mestizaje* ideology is strong in Latin American countries, ethnicity might hide racial disparities. Namely, analyzing economic disparities between ethnic groups might depict a general overview of the inequalities, but it might veil the disparities within broadly define ethnic groups such as the Mestiza and Mulata populations. Therefore, the PERLA palette and the LAPOP data present an important advantage to deepen the study of racial inequalities in the region.

Besides ethnicity- and race-related questions, LAPOP also includes information on socio-demographics, such as age, gender, region, urban or rural household, years of schooling, occupational status, marital status, and household size. One shortcoming is that income is poorly measured. The survey asks on self-reported monthly household income by brackets. For each country and wave, the brackets' values change. Thus, to proxy for a continuous measure of income, I compute the bracket's median value for each monthly household income reported in the country's local currency. After, I divide the continuous measure of monthly household income between the household size to obtain a rough measure of income per capita. Lastly, to make a valid comparison for a country across time and between different countries, I use World Bank's Purchase Parity Power 2019 rates to convert local currencies. Figures 15 and 16 at the appendix shows the household income distribution by each country.[13]

As an alternative measure of household income and wealth, I construct a household asset index with information on whether the household has sewage, a bathroom in the house, television, number of vehicles, among others (Torche 2015).[14] I use Principal Component Analysis to reduce the dimensionality and use the first component to get a household asset index. Figure 17 in the appendix shows the unconditional correlation between monthly household income and the household asset index. Lastly, to have an alternative measure of income, I use Machine Learning techniques to predict monthly household income per capita with the set of variables used to construct the household asset index.[15] Descriptive statistics for the complete sample are at Table 7 in the appendix.

With information on racial phenotype, measured by skin color, and proxies of income per capita and household wealth, I can analyze racial inequalities for the region both at aggregate and individual levels.

4 Racial Inequality and Economic Development

There is an increasing interest in studying the consequences of ethnic-racial inequalities in economic development, closely related to the growing literature studying the historical and cultural patterns shaping economic development (Nunn 2020). As argued by Alesina, Michalopoulos, and Papaioannou (2016), group-based inequalities matter in comparative development: they can lead to political inequality, discriminatory policies between groups, or inadequate public goods provision. Moreover, group-based inequalities can have persistent effects through intergenerational transmission of cultural traits (Bisin and Verdier 2011), occupational segregation (Bowles, Loury, and Sethi 2014), or spatial segregation (Bezin and Moizeau 2017).

However, few studies have addressed the consequences of the ethnic-racial disparities at aggregated levels, such as the national level. A notable exception is the work of Alesina, Michalopoulos, and Papaioannou (2016), who argue that economic inequalities across ethnic groups –ethnic inequality– explains differences in economic development. Using data on the location of ethnic groups and proxying income with night light intensity at a fine grid level, the authors construct measures of inequality between ethnic groups within the same country. The authors find a robust negative correlation between GDP per capita and ethnic inequality at the national level.

Interestingly, when analyzing specific world regions, Alesina, Michalopoulos, and Papaioannou (2016) find that the negative correlation between ethnic inequality and economic development is not statistically significant for Western Europe and the Americas. In Latin America, the latter could be explained by

[13]A critical problem of the constructed proxy of income is that there is a non-trivial amount of individuals who either do not report income or report zero income.

[14]The complete asset list includes television, refrigerator, phone, cellphone, computer, number of vehicles, motorcycle, washing machine, microwave, sewage, and a bathroom in the house.

[15]I implement a linear prediction model with K-Fold Cross-Validation and sample splitting (Hastie, Tibshirani, and Friedman 2009) and use the SuperLearner package in R.

mestizaje politics and the blurry border between ethnicity and race. Thus, the authors cannot disentangle inequalities across a cultural dimension and physical phenotype. In this section, I use LAPOP individual data to compute aggregated measures of racial income inequality.

Firstly, I compute inequality measures without any interpersonal comparison besides income. Using the computed monthly income per capita, I construct a bootstrapped Gini index for each country and year in the LAPOP sample.[16] Figure 22 in the appendix shows the evolution of income inequality during the last decade for each country in Latin America using inequality measures from LAPOP data. On average, Gini indexes from LAPOP data are higher than World Bank's and lower to De Rosa, Flores, and Morgan (2020) estimates. Though noisy, my poorly aggregated measures of income inequality are positively correlated with both sets of measures. Thus, even when LAPOP data is not intended to measure income inequality, my measures of income inequality can partially describe the patterns of income inequality in Latin American countries.

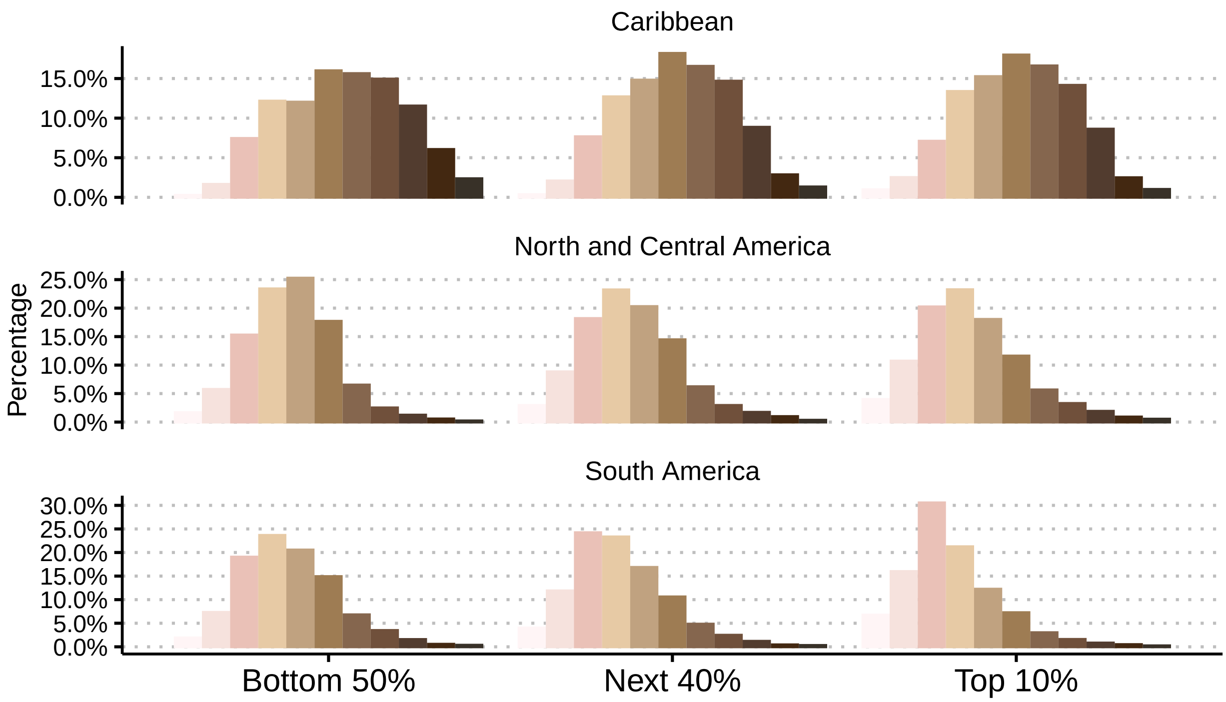

Figure 7: Skin tone distribution within quintiles of income

The attractive feature of LAPOP data is not its income information; but instead, it allows to construct the racial composition of each of the deciles of the income distribution. Figure 7 shows the distribution of skin tones for each decile of the income distribution by region in the continent. Except for the Caribbean countries, the bottom deciles have a majority of the population with skin tones within the range 4 and 6 of the PERLA scale, as upper deciles have most people with whiter skin tones. Consistent with literature in economics, sociology, and history, Latin Americans with darker skin tones have less income than their lighter skin tone peers.

Table 8 at the appendix confirms the latter stylized facts. I regress the decile's racial fractionalization

[16] I use 50 bootstrap replications to compute the index each by country and year.

index and the decile's share of the population with light skin tones with respect to the decile mean (log) income per capita. The racial fractionalization index is a Herfindahl-Hirschman index that measures the group's racial homogeneity (Alesina and La Ferrara 2005; Montalvo and Reynal-Querol 2005). Namely, it represents the probability that at a random draw of two individuals within a population belong to different racial groups or have different skin tones. Both regressions include country and year fixed effects to account for unobserved variables affecting the racial composition of the income deciles. Column 1 shows a negative correlation between the decile's mean income per capita and the fractionalization index of each decile, statistically significant at one percent. The latter implies that deciles are more racially homogeneous at the top of the distribution. Column 2 shows a positive and statistically significant correlation between a decile's mean income per capita and the share of the population with PERLA skin tones between 1 and 3, namely the lightest skin tones. The upper deciles of the income distribution are whiter than the bottom 50% or the next 40%. With the latter results in mind, I compute racial inequality measures at the national level.

4.1 Computing Racial-Inequality

Since I am interested in decomposing income inequality between racial categories, I use the mean log deviation (MLD) index to obtain the between and within components of income inequality given the PERLA scale color categories.[17] As the literature in inequality has shown, the MLD or the Generalized Entropy index GE(0) fulfills the properties of the axiomatic approach –transfer principle, population principle, decomposability principle, and scale-invariant– (Shorrocks 1984), as well as the path independent decomposability (Foster and Shneyerov 2000). Equation (1) represents the MLD and its decomposition of the between and within components (Cowell 2000; Haughton and Khandker 2009):

$$MLD = \frac{1}{N}\sum_{i=1}^{N} ln\frac{\bar{y}}{y_i} = \sum_{j}\frac{N_j}{N}MLD_j + \sum_{j}\frac{N_j}{N}ln\frac{\bar{y}}{\bar{y}_j} \tag{1}$$

For each country and year of LAPOP's data, I decompose income inequality for its between and within components for racial groups.[18] Besides the racial inequality measures, I also use the broad ethnic categories used by LAPOP data to compute alternative ethnic inequality measures. Figure's 8 left panel shows the mean MLD between-group component across the study period for each country in the analysis. Bolivia, Colombia, Jamaica, Guatemala, Dominican Republic Brazil, and El Salvador have the highest mean MLD between-group component across the last decade. In contrast, Chile, Suriname, Guyana, Venezuela, Nicaragua, and Belize, have the lowest median MLD between-group components. The mean MLD between-group component for the region is 1.70.

However, the absolute MLD between-group component is not informative about the relative weight of racial inequality. To have a more intuitive measure of racial inequality, I compute the ratio of racial over total income inequality. Namely, the ratio of the MLD between-group component over the total MLD of

[17]Earlier work by Tesei (2014) also computes racial inequality measures for US Metropolitan Areas. However, he uses the broad ethnic-racial categories used by the General Social Survey rather than measures of racial phenotype.

[18]I use 50 bootstrap replications to compute the index for each country and year.

income. The mean ratio of racial over total income inequality is 3.87 percent. Thus, in Latin American countries, 3.87 percent of the total income inequality is related to inequalities between racial groups measured by skin tone. Figure's 8 right panel shows the mean racial over total inequality ratio. With the relative measure, Bolivia, Guatemala, Colombia, Brazil, Uruguay, and Argentina have the highest mean ratios of racial over total income inequality. For the latter set of countries, nearly 5 percent of total inequality is due to differences in racial groups. Meanwhile, Chile, Nicaragua, Guyana, Belize, and Suriname, have the lowest mean ratios: around 2.5 percent. Figure 23 at the appendix charts the mean indexes values previously described, while Figure 24 at the appendix shows the evolution across time of racial over total inequality ratios for each country.

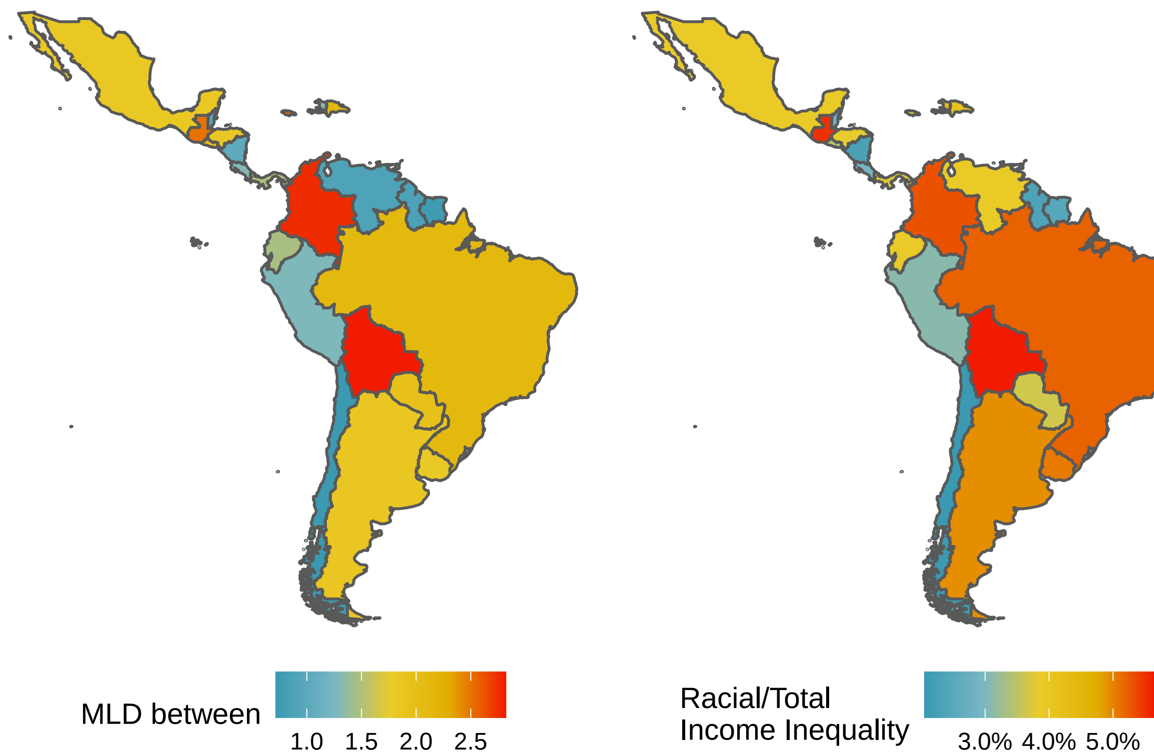

Figure 8: Racial-inequality in Latin America

To test whether *mestizaje* politics do veil racial inequalities, I regress the ratio of racial over total inequality with respect to the share of population that defines themselves as *mestizo* or *mulato*. Controlling for country and year fixed effects, Column 1 in Table 1 shows an one percent increase in the share of *mestizo* or *mulato* population correlates with an increase of 45 percent in the ratio of racial over total inequality, but only statistically significant at ten percent. Column 2 shows the correlation becomes statistically significant after the inclusion of the ethnic inequality measures using the broad ethnic categories (i.e. Mestizo, Mulato, White, Afro, Indigenous, or Other). Column 3 shows the positive correlation persist after including controls as racial and ethnic indexes of fraccionalization and segregation, as well as measures of spatial and administrative inequality (Alesina, Michalopoulos, and Papaioannou 2016). An one percent increase in the share of of *mestizo* or *mulato* population increases the ratio of racial over total inequality in nearly 60 percent. Therefore, countries with a higher *mestizo* or *mulato* identity have substantially higher racial inequality.

Table 1: Racial Inequality and *Mestizaje* Identity

	Dependent variable:		
	Racial over Total Inequality Ratio (log)		
	(1)	(2)	(3)
Share of Mestizo/Mulato (log)	0.446*	0.452**	0.601**
	(0.254)	(0.225)	(0.255)
Ethnic Inequality Ratio (log)		0.338***	0.356***
		(0.090)	(0.092)
Observations	109	109	109
R^2 within	0.039	0.213	0.296
Controls	No	No	Yes
Country FE	Yes	Yes	Yes
Year FE	Yes	Yes	Yes

Notes: Robust standard errors in parenthesis. Controls include LAPOP based racial and ethnic indexes of fraccionalization and segregation, as well as spatial and administrative inequality measures by Alesina, Michalopoulos, and Papaioannou (2016). *p<0.1; **p<0.05; ***p<0.01

4.2 Racial-Inequality and Economic Development

In this subsection, I test whether racial inequalities correlates with lower economic development. Following Alesina, Michalopoulos, and Papaioannou (2016), I use measures of GDP per capita to measure economic development at the national level. Given that I have multiple cross-sections of LAPOP data, I can construct an unbalanced panel with information on racial inequalities, ethnic inequalities, and economic development for 24 Latin American countries across the last decade.[19]

Figure 9 shows the unconditional elasticities of the pooled unbalanced panel between the MLD between racial group component and GDP per capita, and the ratio of racial over total inequality and GDP per capita. The left-sided plot shows a negative elasticity between (log) GDP per capita and the (log) MLD between racial group components. Thus, at first glance, racial inequality correlates with lower economic development. However, the right-sided plot shows no correlation between the (log) ratio of racial over total inequality and economic development.

To test more robustly the relation between racial inequality and economic development I use the following econometric specification:

$$y_{c,t} = \beta Racial\ Inequality_{c,t} + \gamma X_{c,t} + \theta_c + \eta_t + \varepsilon_{c,t} \tag{2}$$

Where $y_{c,t}$ represents (log) GDP per capita of country c at time t; $Racial\ Inequality_{c,t}$ is the (log) ratio of racial over total income inequality; $X_{c,t}$ represents a set of time varying controls. One improvement of this

[19]I exclude the countries with only one available cross-section. Results are robust to the inclusion of the latter.

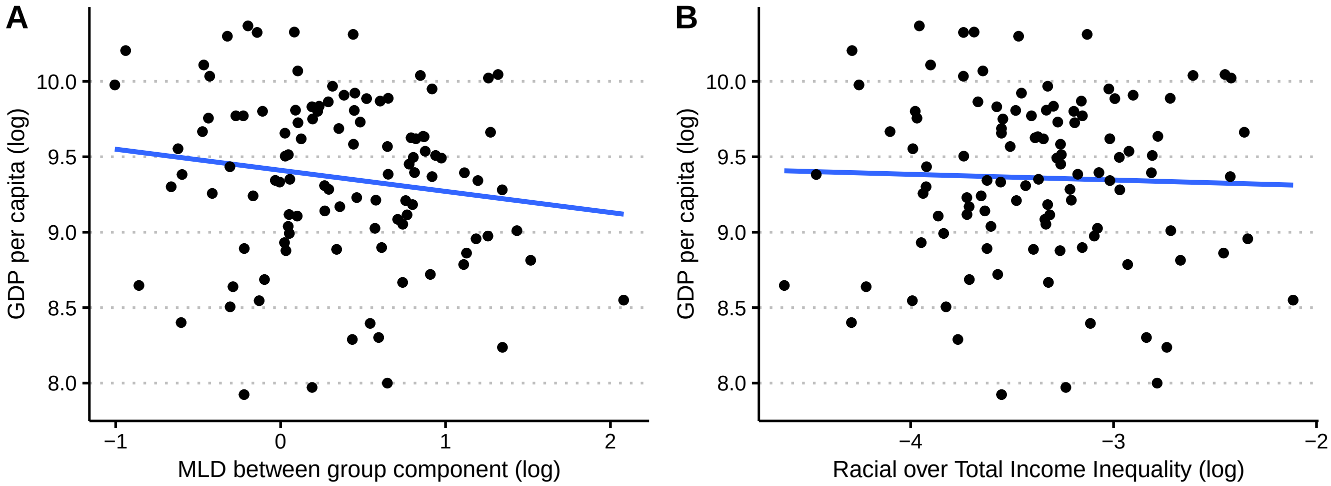

Figure 9: Racial-inequality and Economic Development

analysis with respect to Alesina, Michalopoulos, and Papaioannou (2016) is the availability of multiple observations for each country. Thus, I can include country fixed effects θ_c to control for time-invariant characteristics that affect economic development and also are correlated with racial inequality at the national level. Lastly, I include year fixed effects η_t to account for common shocks to all countries in the period of study. I use bootstrapped cluster the standard errors at the country level (Cameron, Gelbach, and Miller 2008).

Table 2 shows the results of the specifications in Equation (2). Column 1 shows that once accounting for country and year fixed effects, an one percent increase in the ratio of racial over total income inequality correlates with a decrease in 4.8 percent of GDP per capita, statistically significant at one percent. Column 2 includes my measures of ethnic inequality as a control to test whether racial-based inequalities are robust to accounting for broad ethnic-based inequalities. The elasticity decreases to 3.9 and remains statistically significant. Column 3 includes Alesina, Michalopoulos, and Papaioannou (2016) measures of spatial and administrative inequality at the national level. More specifically, given there is one single measure of spatial and administrative inequality and I include country and year fixed effects, I interact the latter measures with the continuous year variable. The elasticity of interest barely changes.

Column 4 also includes constructed measures of racial and ethnic fractionalization at the country level and racial and ethnic segregation measures using LAPOP data. Overall, the coefficient of racial over total inequality remains statistically significant and increases marginally. The most robust specification implies that an one percent increase in the ratio of racial over total inequality correlates with a decrease of 4.1 percent in GDP per capita.

To test the robustness of the results, Table 10 at the appendix includes Alesina, Michalopoulos, and Papaioannou (2016) measures of ethnic inequality. My measure of racial inequality is robust to accounting for other measures of ethnic inequalities. Table 9 in the appendix also shows the results by using only the (log) MLD between racial group component and controlling for total inequality using MLD total index. The results show that the racial group component drives the negative correlation of the ratio of racial over total inequality and economic development. Thus, racial-based inequalities hinder economic development

Table 2: Racial Inequality and Economic Development (1)

	Dependent variable:			
	GDP per capita (log)			
	(1)	(2)	(3)	(4)
Ratio Racial/Total Income Inequality (log)	-0.048^{***}	-0.039^{***}	-0.040^{***}	-0.041^{**}
	[0.016]	[0.015]	[0.015]	[0.017]
	(0.017)	(0.017)	(0.017)	(0.018)
Ratio Ethnic/Total Income Inequality (log)		-0.018	-0.018	-0.021
		[0.018]	[0.018]	[0.022]
		(0.017)	(0.017)	(0.018)
Spatial Inequality (log) × Time			0.001	0.0004
			[0.005]	[0.006]
			(0.002)	(0.002)
Administrative Inequality (log) × Time			0.003	0.004
			[0.010]	[0.010]
			(0.006)	(0.006)
Racial Fraccionalization (log)				0.042
				[0.272]
				(0.221)
Racial Segregation (log)				-0.249
				[1.263]
				(0.862)
Ethnic Fraccionalization (log)				0.048
				[0.106]
				(0.052)
Ethnic Segregation (log)				0.376
				[1.345]
				(0.928)
Observations	109	109	109	109
R^2 within	0.062	0.073	0.074	0.104
Country FE	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes

Notes: Bootstrapped standard errors clustering by country in brackets based on 500 bootstrap replications. Robust standard errors in parenthesis. Spatial and administrative inequality measures taken from Alesina, Michalopoulos, and Papaioannou (2016). *p<0.1; **p<0.05; ***p<0.01

in Latin America rather than ethnic-based inequalities.

5 Racial Disparities

In the previous section, I presented evidence that racial inequality negatively affects economic development at the national level. However, it is unclear what are the underlying mechanisms driving such a relation. In this section, I use the LAPOP data at the individual level to analyze skin tone effect on a set of socio-economic outcomes.

I follow the disparities approach, which accounts for inequalities in outcomes solely due to group membership. For instance, social status or ethnic groups might have disparities in economic outcomes. Then, economic outcomes are a function of the membership to different groups –racial, ethnic, socio-economic– and other individual characteristics –income, education–and individual needs and preferences (Fleurbaey and Schokkaert 2011). Next, I describe my identification strategy to isolate the effect of skin tone on monthly income per capita, the household asset index, and years of schooling.

Figure's 10 Panel A shows the correlation for the mean income per capita (log) for each tone of the PERLA color scale across the whole region. The unconditional linear regression depicts a negative gradient between income and darker PERLA skin tones. Using LAPOP micro data, the unconditional semi-elasticity between PERLA scale color and (log) income per capita is minus 0.093, statistically significant at one percent. Panel B shows the skin tone gradient for years of schooling. An increase in the PERLA scale correlates with a decrease of 0.05 standard deviations in years of schooling, statistically significant at one percent.

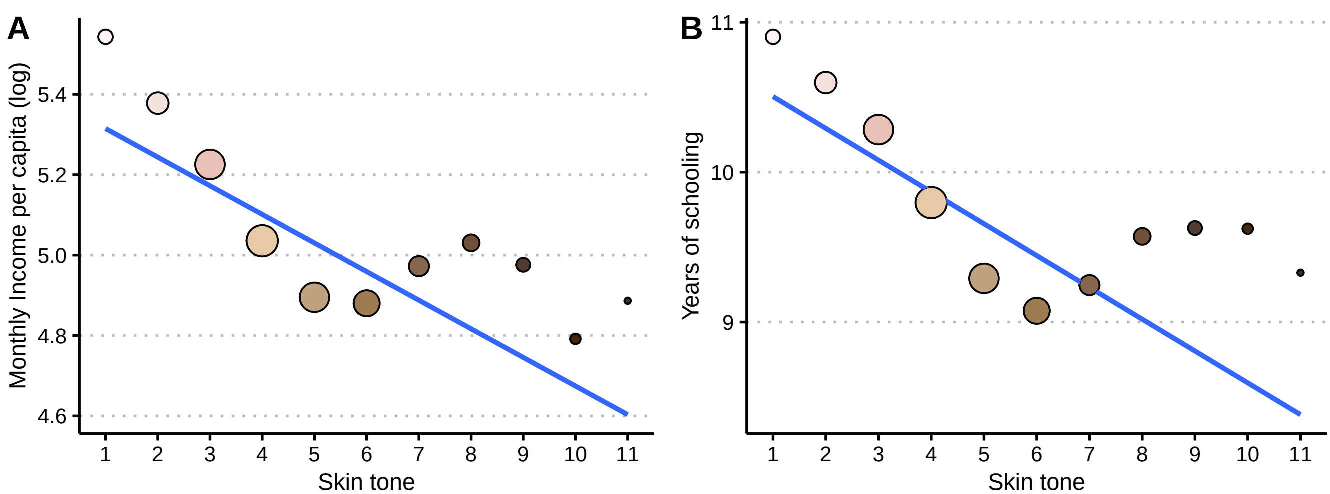

Figure 10: Skin tone gradient for mean monthly income per capita and years of schooling

Even when racial disparities follow a similar pattern throughout the region, most surely, the mechanisms vary substantially by country, or even within the country. Figures 25 and 26 in the appendix show there is substantial variation in mean income levels by country, but in most countries, the negative gradient between skin tone and income persist.[20] When using country-region and year fixed effects, the semi-

[20] Note these unconditional regressions use sample weights to diminish the weight of outlying observations.

Table 3: Skin Tone gradient on Monthly Income per capita

	Dependent variable:				
	Income per capita (log)		Years of schooling (z-score)		Income per capita (log)
	(1)	(2)	(3)	(4)	(5)
Skin tone	-0.092^{***}	-0.062^{***}	-0.053^{***}	-0.073^{***}	
	(0.002)	(0.002)	(0.001)	(0.002)	
Skin tone (z-score)					-0.058^{***}
					(0.003)
Years of schooling (z-score)					0.364^{***}
					(0.003)
Observations	137,586	137,586	137,586	137,586	137,586
R^2 within	0.024	0.010	0.011	0.014	0.139
Country-region FE	No	Yes	No	Yes	Yes
Year FE	No	Yes	No	Yes	Yes

Notes: Robust standard errors in parenthesis. *p<0.1; **p<0.05; ***p<0.01

elasticity between skin tone and monthly income per capita equals minus 0.062. Therefore, for every darker tone in the PERLA scale, individuals earn on average 6.2 percent less on monthly household income per capita.

To gauge the magnitude of the skin tone gradient on income, I standardize skin tone by country and year and regress monthly income per capita with respect to skin tone and years of schooling, both standardized. The results in Table 3 show that a one standard deviation increase in the skin tone, namely being one standard deviation darker than the mean, correlates with a decrease of almost 6 percent in income (for the complete sample, the PERLA skin tone mean is 4.81, and standard deviation 1.99). In contrast, a one standard deviation increase in years of schooling correlates with an increase of 36 percent in income (for the complete sample, the mean years of schooling is 9.65, and its standard deviation 4.27). The skin tone gradient on income represents minus one-sixth of the years of schooling gradient. The latter results implies that the skin tone effect on income, a one standard deviation darker than the skin tone mean, is similar to the effect on income of a decrease 0.7 years of schooling.

Nevertheless, the previous coefficients can be biased for multiple reasons. As with any Mincerian equation, there are relevant observed and unobserved characteristics that are not controlled for affecting monthly household income per capita –gender, age, education, experience or ability, parental background–. Thus, the previous estimates most likely suffer from omitted variable bias. Besides, there might be measurement error bias given the survey's methodology to register respondent skin tone. The next subsections presents the identification strategy to have an unbiased effect of skin tone on income and years of schooling.

5.1 Identification strategy

5.1.1 Oaxaca-Blinder decomposition for continuous variables

To analyze more robustly the income gaps between racial groups measured by skin tone, I use an Oaxaca-Blinder (OB) decomposition for continuous variables proposed by Ñopo (2008b). As Ñopo explains, the OB decomposition is a widely used method in labor economics to study differences in earnings between two groups into two elements: the first one captures differences in observable characteristics between the two analyzed groups, and the second captures the differences in returns to those characteristics.

Following Ñopo (2008b) notation, the OB decomposition for a continuum of groups can be extended from a Mincerian regression framework for two groups. First, let t_i denote a dummy variable indicating whether individual i belongs to a given group. Thus, one have the following 'simplified' Mincer equation:

$$y_i = \alpha_0 + \alpha_1 t_i + \varepsilon_i \tag{3}$$

Where y_i is the outcome of individual i, and α_1 represents the gap between the two groups: $\alpha_1 = E[y|t = 1] - E[y|t = 0]$.

To decompose the gap between the observed characteristics and their respective returns is necessary to estimate the following 'extended' Mincer equation:

$$y_i = \beta_0 + \beta_1 x_i + \beta_2 t_i + \beta_3 t_i x_i + \varepsilon_i \tag{4}$$

Where y_i and t_i represents the same variables as in Equation (3), and x_i is a n-dimensional vector of observable characteristics. Thus, β_1 represents the rewards to the observable characteristics for group 0, and $\beta_1 + \beta_3$ are the rewards to the observable characteristics for group 1. Coefficient α_1 from Equation (3) can be expressed as:

$$\begin{aligned}
\alpha_1 &= E[(\beta_0 + \beta_2) + (\beta_1 + \beta_3)x|t = 1] - E[\beta_0 + \beta_1 x|t = 0] \\
\alpha_1 &= (\beta_0 + \beta_2) + (\beta_1 + \beta_3)E[x|t = 1] - \beta_0 - \beta_1 E[x|t = 0] \\
\alpha_1 &= \beta_1(E[x|t = 1] - E[x|t = 0]) + \beta_2 + \beta_1 E[x|t = 1] \\
\alpha_1 &= \Delta_x + \Delta_0
\end{aligned}$$

Where $\Delta_x \equiv \beta_1(E[x|t = 1] - E[x|t = 0])$, represents the differences in the outcome due to average observable characteristics of the individuals; while $\Delta_0 \equiv \beta_2 + \beta_1 E[x|t = 1]$ represents the differences in the outcome that cannot be explained by observable characteristics. Given the economic theory of discrimination, the component Δ_0 can be interpreted as the combination of discrimination and unobservable characteristics.

Ñopo (2008b) formally shows how to obtain the two components of the OB decomposition when t_i is a continuous measure. In practice, the procedure is the following:

1. Estimate the 'simplified' Mincer equation, Equation 3. The coefficient of interest is α_1.

2. Estimate the 'extended' Mincer equation, Equation 4. The coefficients of interest are β_2 and β_3.
3. Estimate the average characteristics of the sample used to estimate Equations 3 and 4, namely $E[x]$.
4. Obtain the component of the wage gap that cannot be explained by differences in average characteristics: $\Delta_0 = \beta_2 + \beta_3 E[x]$.
5. Obtain the component of the wage gap that is explained by differences in average characteristics: $\Delta_x = \alpha_1 - \Delta_0$.

As a first step to assess the racial income gap, I compute an OB decomposition using the continuous measure of skin tone by the PERLA palette using the procedure proposed by Ñopo (2008b).

5.1.2 Control function apporach

Besides the differences in observed and unobserved characteristics, one of the main concerns by estimating both the simplified and the extended Mincer equations, Equations (3) and (4) respectively, is that there is endogeneity due to measurement error bias. As previously mentioned, respondent skin tone in LAPOP surveys is registered directly by the interviewer, which could be biased due to her subjective perception of the respondent's skin tone. The latter is one of the main issues recent studies in anthropology and sociology have raised (Dixon and Telles 2017).

To better understand the bias due to measurement error in skin tone, Solís, Avitia, and Güémez (2020) uses reflectance spectrometers to measure skin tone for a Mexican representative sample. The authors find that interviewers tend to 'whiten' respondents when they register the skin tone. Namely, interviewers register respondents' skin tones in lighter skin tones than the actual skin color obtained through the reflectance spectrometers. The bias is higher when individuals have high levels of education or higher income. Cernat, Sakshaug, and Castillo (2019) also study the effect of the interviewer on the measurement bias of skin tone and socio-economic outcomes for 22 Latin American countries. The authors use the 2010 LAPOP AmericasBarometer and show that the interviewer's skin tone can explain around 20% of the respondent's skin tones. Thus, even if the Mincerian equation of interest was not bias due to unobservables, the coefficients will be either under or overestimating the effect of skin tone on monthly household income per capita due to measurement error.

Fortunately, LAPOP data has a unique feature that allows overcoming the measurement bias in the coefficient of interest. In the last part of the interview, the interviewer has to register the respondent's skin tone using the PERLA scale without showing it. The latter happens after registering the respondent's answers to report household income, years of schooling, or a series of questions on household assets. However, at the very end of the interview, the interviewer also has to register her skin tone using the same PERLA scale. If the interviewer's subjective appreciation of the respondent's skin color exists, there should be a correlation between the interviewer and the respondent's skin tones. Figure 20 at the appendix shows the mean respondent skin tone by each interviewer PERLA skin tone. Accounting for country-region and wave fixed, effects, the correlation is 0.13, statistically significant at one percent. The respondent's skin tone increases in 0.13 PERLA scale points for each PERLA skin tone of the interviewer. Thus, the interviewer's subjective perception of her skin tone might affect the respondent's skin color measurement.

With such an exogenous variation on respondent skin tones in LAPOP data, I use a control function approach to purge the measurement error of the coefficient of skin tones on household income. Wooldridge (2015) explains the control function approach for solving the problem of endogenous explanatory variables in linear and nonlinear models. To exemplify the control function approach, think in the 'simplified' Mincer equation: $y_i = \alpha_0 + \alpha_1 t_i + \varepsilon_i$. If t_i is endogenous, one can decompose it into its exogenous component $t_{i,exog}$ and its endogenous component $t_{i,endo}$, where the latter is a function of the all the exogenous variables z_i including $t_{i,exog}$:

$$t_{i,endo} = \delta_0 + \delta_1 t_{i,exog} + \delta_2 z_i + u_i$$

Now, given

$$Cov(\varepsilon_i, t_{i,endo}) = Cov(\varepsilon_i, (\delta_1 t_{i,exog} + \delta_2 z_i)) + Cov(\varepsilon_i, u_i)$$

and

$$E[(\delta_1 t_{i,exog} + \delta_2 z_i)' \varepsilon_i] = 0$$

then all the endogeneity is captured by

$$Cov(\varepsilon_i, u_i)$$

Thus, the underlying idea behind control functions is to decompose the disturbance term into the correlated and uncorrelated components, get a measure of the correlated part and add it as an extra regressor in the Mincerian equation.[21] Wooldridge (2015) describes the control function procedure as following:

1. Run the OLS regression of the endogenous variable on all the exogenous variables:

   $$t_{i,endo} = \delta_0 + \delta_1 t_{i,exog} + \delta_2 z_i + u_i$$

   Obtain the residuals: $\hat{u}_i$.

2. Plug the residuals into the structural Mincerian equation:

   $$y_i = \alpha_0 + \alpha_1 t_i + \alpha_3 \hat{u}_i + \varepsilon_i$$

Note that the control function approach follows a similar procedure than the correction for sample selection proposed earlier by Heckman (1979).[22] Furthermore, as explained by Wooldridge (2015), the control function method produces coefficients identical to the 2SLS estimates. Then, one needs a valid instrument for the control function approach: an exogenous variable correlated with the endogenous variable but uncorrelated with the error term.

I use the interviewer's skin tone as the control function or IV of the interview skin tone. Following Cunningham (2021), Figure 21 at the appendix shows a Direct Acyclical Graph (DAG) to show the intuition

[21]Example taken from David Margolis' Advanced Microeconometrics lectures at PSE.

[22]The Heckman correction model for sample selection uses a probit in the first stage and then computes the Inverse Mills Ratio as the error term of the first stage.

behind the proposed control function or IV. As previously showed, the two variables are correlated, then the relevance restriction holds. Moreover, interviewers' skin tone cannot affect the respondent's outcomes but 'only through' how the first perceives and registers the latter's skin tone. Think of any possible confounder in the relationship between skin tone and income. It is most likely implausible that an interviewer's skin tone defines and respondent self-reported long-life economic outcomes. Thus, the exclusion restriction holds.

A natural question of the identification strategy is: why using a control function approach and not an IV approach? The first reason is that the interviewer's skin tone does not literally affect the respondent's skin tone. The latter means the interviewer's skin tone is not a treatment delivered to the respondent that affects its skin tone and later its income or years of schooling. However, the control function approach allows purging the endogeneity due to measurement error. The second reason is that one can produce a Hausman test robust to heteroskedasticity by directly testing whether the residuals of the first stage are statistically different from zero in the structural equation (Wooldridge 2015).

Finally, note that as in any two-step procedure, the standard errors from the second stage, or the structural equation with the residuals from the first stage, will be incorrect for statistical inference (Murphy and Topel 1985). A simple solution is to use bootstrapped standard errors.

5.1.3 Oaxaca-Blinder decomposition with control functions

To get an unbiased effect of the skin tone on monthly household income per capita, I use an OB decomposition with control functions. Rios-Avila (2019) also proposes a procedure to obtain the OB components using control functions, but instead of using OLS, the author uses a varying coefficients semiparametric approach (Hastie and Tibshirani 1993). Moreover, Rios-Avila (2019) proposes to use control functions for sample selection using a Heckman correction (Heckman 1979).

I follow Rios-Avila's (2019) procedure but using Ñopo's (2008a) OLS OB decomposition. Given that my sample has a non-trivial amount of missing or zero reported income, I also use the Heckman correction and only use observations with strictly positive income. More specifically, I use the following procedure:

1. Run the Heckman correction first-stage for the whole sample. Run a Probit model observing a positive income, $D = 1$, against a missing or zero income, $D = 0$, on a set of observable characteristics and using the gender of the interviewer as 'instrument':

$$P(D_i = 1|Z_i) = \Phi(Z_i'\gamma + \theta_i + \eta_i)$$

   Get the Inverse Mills Ratio: IMR_i, and drop all missing and zero income observations.

2. Obtain a random paired bootstrap sample from the sample with non-missing and positive income observations.

3. Obtain the bootstrapped sample mean of the set of covariates, x_i.[23] Namely, $E[x_i]$.

[23]The set includes: sex; age; age squared; years of schooling; a dummy with marital status; a dummy with occupational status; a dummy indicating whether the individual identifies as Afro, Indigenous, Mestizo, Mulato, Other, or White; a dummy indicating whether the individual lives in the capital city, in a big size city, a medium-sized city or a rural area.

4. Run the OLS on the first stage, or regress the respondent's skin tone on the interviewer skin tone:

$$PERLA_i = \delta_0 + \delta_1 Interviewer\ PERLA_i + \theta_i + \eta_i + u_i$$

   Get the first stage residuals: $\hat{u}_i$.

5. Run the OLS 'simplified' Mincerian equation including the first stage residuals as a control function, $\hat{u}_i$, and also the Inverse Mills Ratio for sample selection, IMR_i:

$$y_i = \alpha_0 + \alpha_1 PERLA_i + \alpha_2 \hat{u}_i + \alpha_3 IMR_i + \theta_i + \eta_i + \varepsilon_i$$

6. Run the OLS 'extended' Mincerian equation including all the covariates represented in the n-dimensional vector x_i:

$$y_i = \beta_0 + \beta_1 x_i + \beta_2 PERLA_i + \beta_1 x_i \cdot PERLA_i + \theta_i + \eta_i + \varepsilon_i$$

   The first stage residuals, $\hat{u}_i$, and the Inverse Mills Ratio, IMR_i, are also included in the n-dimensional vector x_i.

7. Use Ñopo's (2008a) procedure to obtain the OB components: α_1 from Step 5; Δ_0 from Step 6, and $\Delta_x = \alpha_1 - \Delta_0$.

8. Repeat Steps 2 to 7, B times to obtain the empirical distributions of the aggregated and detailed decomposition components.

Since the data is from samples for different countries and waves, note that all regressions include country-region fixed effects, θ_i, and year fixed effects, η_i. I use LAPOP sample weights to make the results comparable across countries and waves, and also representative at the national level (Castorena 2021).

5.2 Results

Table's 4 Panel A shows the results for the Oaxaca-Blinder decomposition for continuous variables, while Panel B shows the same OB decomposition procedure using control functions.

Column 1 shows the results for my proxy of monthly income per capita. Without correcting for skin tone measurement error concerns, but correcting for sample selection, the 'simplified' Mincer equation shows an increase in one PERLA scale tone correlates with a decrease of 6.2 percent on monthly income per capita (α_1). The OB decomposition shows that out the total skin tone effect, 3.8 percent cannot be explained by differences in average characteristics. Both coefficients are statistically significant at one percent. The latter results imply that the individual with the darkest skin tone has an average income nearly 62 percent lower than the whitest individual, where differences in observed average characteristics cannot explain almost 61.3% of the gap ($\Delta_0/\alpha_1 = \frac{-0.038}{-0.062} = 61.29\%$).

Panel B shows similar results. However, I find mild evidence of endogeneity due to error measurement in the skin tone of the respondent since the coefficient for the first-stage residual, $\hat{u}$, equals to 1.5 percent,

Table 4: Skin tone effect on income: OB Decomposition with control functions

	Dependent variable:		
	Income per capita (log)	Predicted Income per capita (log)	Household Asset Index
	(1)	(2)	(3)
Panel A. ***OB decomposition***			
α_1	-0.062***	-0.039***	-0.087***
	(0.002)	(0.001)	(0.002)
Δ_0	-0.038***	-0.022***	-0.049***
	(0.002)	(0.001)	(0.002)
Δ_x	-0.024***	-0.018***	-0.039***
	(0.001)	(0.001)	(0.001)
Hausman Test: Selection			
IMR	-2.257***	-0.860***	
	(0.032)	(0.016)	
Panel B. ***OB decomposition with control functions***			
α_1	-0.075***	-0.037***	-0.072***
	(0.008)	(0.005)	(0.008)
Δ_0	-0.051***	-0.026***	-0.046***
	(0.008)	(0.004)	(0.007)
Δ_x	-0.024***	-0.010***	-0.026***
	(0.004)	(0.002)	(0.004)
Hausman Test: Selection			
IMR	-2.257***	-0.860***	
	(0.032)	(0.016)	
Hausman Test: Measurement error			
$\hat{u}$	0.015*	-0.003	-0.016**
	(0.009)	(0.005)	(0.008)
First stage			
δ	0.169***	0.174***	0.174***
	(0.003)	(0.004)	(0.003)
Observations	133,705	146,859	153,804
Individual Controls	Yes	Yes	Yes
Country-region FE	Yes	Yes	Yes
Year FE	Yes	Yes	Yes

Notes: Bootstrapped standard errors based on 500 bootstrap replications. Individual controls include sex; age; age squared; years of schooling; a dummy with marital status; a dummy with occupational status; a dummy indicating whether the individual identifies as Afro, Indigenous, Mestizo, Mulato, Other, or White; a dummy indicating whether the individual lives in the capital city, in a big size city, a medium-sized city or a rural area. *p<0.1; **p<0.05; ***p<0.01

but is not statistically different from zero at conventional significance levels. Note that without the control function the effect of skin tone on income is higher in absolute terms:

$$\alpha_1 = \alpha_{1,CF} + \hat{u} \iff \alpha_1 - \hat{u} = \alpha_{1,CF} \iff -0.062 - (0.015) \approx -0.075$$

Then, if the coefficient of the control function was statistically significant, there is endogeneity that the control function approach purges properly.

As a robustness check, Table's 4 Column 2 shows the results for the predicted monthly income per capita using the set of household assets. When correcting only for sample selection, Panel A shows that an increase in the PERLA color palette correlates with a decrease in 3.9 percent in the predicted income per capita, where 56% of the total effect cannot be attributed to differences in returns to the observed average characteristics ($\Delta_0/\alpha_1 = \frac{-0.022}{-0.039} = 56.4\%$). Nevertheless, Panel B shows that the coefficient for the first-stage residual of the regression of respondent skin tone on interviewer skin tone, $\hat{u}$, is not statistically different from zero. Not surprisingly, the coefficient for skin tone in the OB decomposition using the control function is minus 3.7 percent, also statistically significant at one percent. The results for the predicted income confirm the results using our proxy for monthly income per capita.

Lastly, Table's 4 Column 3 shows the results for the econometric specification using the (standardized) Household Asset Index. Panel B shows the control function purges endogeneity. More specifically, without the control function the skin tone coefficient is overestimated in absolute terms. The purged coefficients shows a increase in one skin tone from the PERLA palette decreases the Household Asset Index in 0.072 standard deviations, statistically significant at one percent. The latter means a mean difference of 0.72 standard deviations in household assets from the whitest individual to the same with the darkest skin tone. Similar to the previous results, 65.2% of the gap in household assets between individuals with different skin tones cannot be explained by differences in the returns to their average observed characteristics ($\Delta_0/\alpha_1 = \frac{-0.046}{-0.072} = 63.8\%$).

5.3 Mechanisms

To gain a better understanding of the factors driving the previously described results, I use the OB decomposition with control functions to analyze the income gap for individuals active in the labor market, and also analyze the racial gap in years of schooling for the whole sample. Table 5 shows the results.

The previous results could be driven by the occupational status of individuals given their skin tone. Namely, it could be that individuals with darker skin tones have less monthly income per capita because they do not work. To test the latter, I regress the skin tone gradient on the proxy for monthly income per capita just for those who reported being actively working by the time of the interview. The results in Panel A follow the pattern of those previously presented: an increase in the PERLA color palette from whither to darker skin tones correlates with a decrease of 6.1 percent on income, where only a third part of the gap can be explained for differences in average characteristics. Panel B shows the coefficient for the first-stage residuals, $\hat{u}$, is not statistically different from zero. Thus, there is no endogeneity, and the racial gap persists for individuals participating in the labor market.

Table 5: Mechanisms

	Dependent variable:				
	Workers Income per capita (log)	Years of Schooling (z-score)	Probability of Discrimination for Skin Tone	at Public Places	at School or Work
	(1)	(2)	(3)	(4)	(5)
Panel A. *OB decomposition*					
α_1	-0.061***	-0.072***	0.023***	0.012***	0.004**
	(0.003)	(0.002)	(0.002)	(0.002)	(0.002)
Δ_0	-0.039***	-0.058***	0.013***	0.006***	0.000
	(0.003)	(0.002)	(0.003)	(0.002)	(0.002)
Δ_x	-0.022***	-0.014***	0.010***	0.005***	0.005**
	(0.002)	(0.001)	(0.002)	(0.001)	(0.002)
	(0.058)		(0.035)	(0.02)	(0.028)
Panel B. *OB decomposition with control functions*					
α_1	-0.064***	-0.054***	0.013*	0.014**	0.007
	(0.011)	(0.008)	(0.007)	(0.006)	(0.008)
Δ_0	-0.049***	-0.044***	0.008***	0.01***	0.007*
	(0.011)	(0.007)	(0.003)	(0.003)	(0.004)
Δ_x	-0.015***	-0.010**	0.008***	0.01***	0.007*
	(0.005)	(0.005)	(0.003)	(0.003)	(0.004)
Hausman Test: Measurement error					
$\hat{u}$	0.003	-0.019**	0.010	-0.003	-0.003
	(0.012)	(0.008)	(0.007)	(0.007)	(0.008)
First stage					
δ	0.156***	0.169***	0.190***	0.139***	0.129***
	(0.006)	(0.003)	(0.014)	(0.009)	(0.010)
Observations	64,969	133,705	13,250	20,438	17,559
Individual Controls	Yes	Yes	Yes	Yes	Yes
Country-region FE	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes

Notes: Bootstrapped standard errors based on 500 bootstrap replications. Individual controls include: sex; age; age squared; a dummy with marital status; a dummy with occupational status; a dummy indicating whether the individual identifies as Afro, Indigenous, Mestizo, Mulato, Other, or White; a dummy indicating whether the individual lives in the capital city, in a big size city, a medium-sized city or a rural area. Column 1 includes as control years of schooling, while Column 2 includes as control (log) monthly income per capita. Columns 3 to 5 include both years of schooling and (log) monthly income per capita. *p<0.1; **p<0.05; ***p<0.01

I also analyze the racial gap in years of schooling. I regress years of schooling (standardized by country and year) using the OB decomposition with control function techniques. As Panel B shows, the coefficient for the first-stage residuals, $\hat{u}$, equals -0.02 standard deviations. Thus, the OB decomposition without control functions over-estimates the effect of skin tone on years of schooling: $\alpha_1 = \alpha_{1,CF} + \hat{u} \iff \alpha_1 - \hat{u} = \alpha_{1,CF} \iff -0.072 - (-0.020) = -0.053$. Thus, between the individual with the lightest skin tone and one with the darkest skin tone, there is, on average, a half standard deviation of the difference in years of schooling. Of the overall racial gap in years of schooling, 83% cannot be attributed to the average returns in observed characteristics.

So far, I have found evidence of a substantial racial gap in income and educational attainment, where more than 60% of the effect is due to unobserved characteristics. Among the unobserved characteristics, there is racial discrimination. To test whether there is discrimination for skin tone, I use the LAPOP 2014 wave, which includes questions for the individuals to self-report discrimination.[24] Columns 3 to 5 from Table 5 shows the results for the OB decomposition to the probability that the individual responded positively being discriminated: 1) because of her skin tone; 2) in a public place; 3) at school or work.

Firstly, Column 3 shows that each PERLA tone correlates with an increase of 2.2 percentage points in the probability of reporting discrimination by skin tone, where almost 60% of the effect is not related to other observable characteristics. Column 4 shows that each darker skin tone increases the probability of reporting discrimination in a public place by 1.2 percentage points. In this case, 50% of the overall effect cannot be accounted for other observable characteristics. Lastly, Column 5 shows that each darker skin tone in the PERLA palette correlates with an increase of 0.5 percentage points in the probability of reporting discrimination at the school or workplace. Interestingly, the last effect is completely driven by differences in average observable characteristics of individuals with darker skin tones. The Hausman test in Panel B shows no measurement error concerns due to skin color for each specification at conventional statistical significance levels.

5.4 Heterogeneous effects

To test for heterogeneous effects, firstly I analyze the skin tone gradient on income gradient on subsamples by gender and ethnic self-affiliation.[25] Figure 11 shows the results.[26] Panel A shows the skin tone gradient is steeper for those who define themselves as 'Other,' 'White,' or 'Mestiza,' with the semi-elasticity ranging from minus 7.3 to minus 6.6, not statistically different from each other. Meanwhile, the semi-elasticity equals minus 4.9 for those who define themselves as 'Indigenous,' and around minus 3.7 for 'Mulata' and 'Afro.' Nevertheless, for the latter ethnic categories, the $\Delta 0$ OB component represents almost all of the effect in the racial gap. For 'Afro' ethnicity, the $\Delta 0$ parameter represents 86% of the racial gap; for 'Indigenous' ethnicity, it represents 68%. Thus, even when the results suggest the skin tone effect

[24]Given that there is also a non-trivial amount of missing values for the questions on discrimination, I also use the first step Heckman correction, but to assess the probability of observing an answer to the discrimination and use the Inverse Mills Ratio's in the second stage.

[25]I use the OB decomposition with control functions, but also use a selection model to predict the probability of observing a individual with a given ethnicity or gender given a set of observable characteristics as age, urban size, country-region, among others.

[26]Confidence intervals at 95 percent significance level using 250 bootstrap replications.

is smaller for ethnic minorities, the racial gap for the latter is most likely due to unobervables as racial discrimination. Moreover, note that ‘Mestiza’ ethnic identity does veil racial inequalities: its skin tone gradient is among the steepest.

Figure’s 11 Panel B shows that the skin tone gradient on income is steeper for females. The semi-elasticity of skin tone and income for females equals minus 8.2, while for males is minus 6.3. The OB parameter Δ_0 is 50% for both genders. Then, besides the gender gap on income, females with darker skin tones are worse of than males.

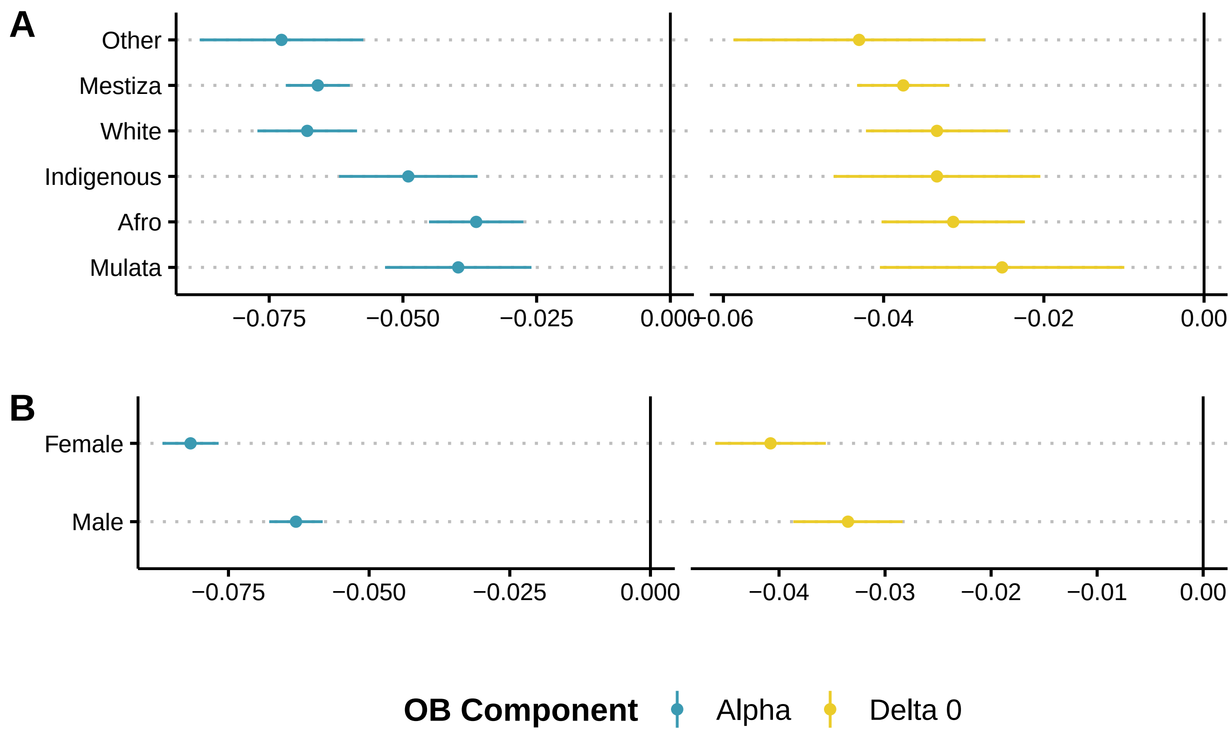

Figure 11: Skin tone effect on predicted income: OB decomposition by gender and ethnicity

If racial inequalities and discrimination are global phenomena, their expressions are local. Figure 12 shows the results for the OB decomposition using as dependent variable the proxy for monthly (log) income per capita country by country.[27] Figure 12 shows that the skin tone gradient in income is negative and statistically significant for all Latin American countries, except for Guyana and Haiti. However, there is substantial variation between countries.

The skin tone income gradient is the steepest in Uruguay: an increase in the PERLA color palette correlates with a decrease in 15 percent in monthly income per capita, where 43% of the effect cannot be accounted by differences in average observable characteristics. Interestingly, the semi-elasticities of skin tone on income is high in countries with high racial mixture and a sizable share of indigenous population. Guatemala’s and Bolivia’s skin tone elasticity with income is minus 12.5, where 39% and 47% of the total effect is not explained by differences in average observable characteristics. For Mexico, Paraguay, and El Salvador, the semi-elasticities range between minus 11 and minus 9, where 70%, 62.5% and 48% of the

[27] Confidence intervals at 95 percent significance level using 500 bootstrap replications.

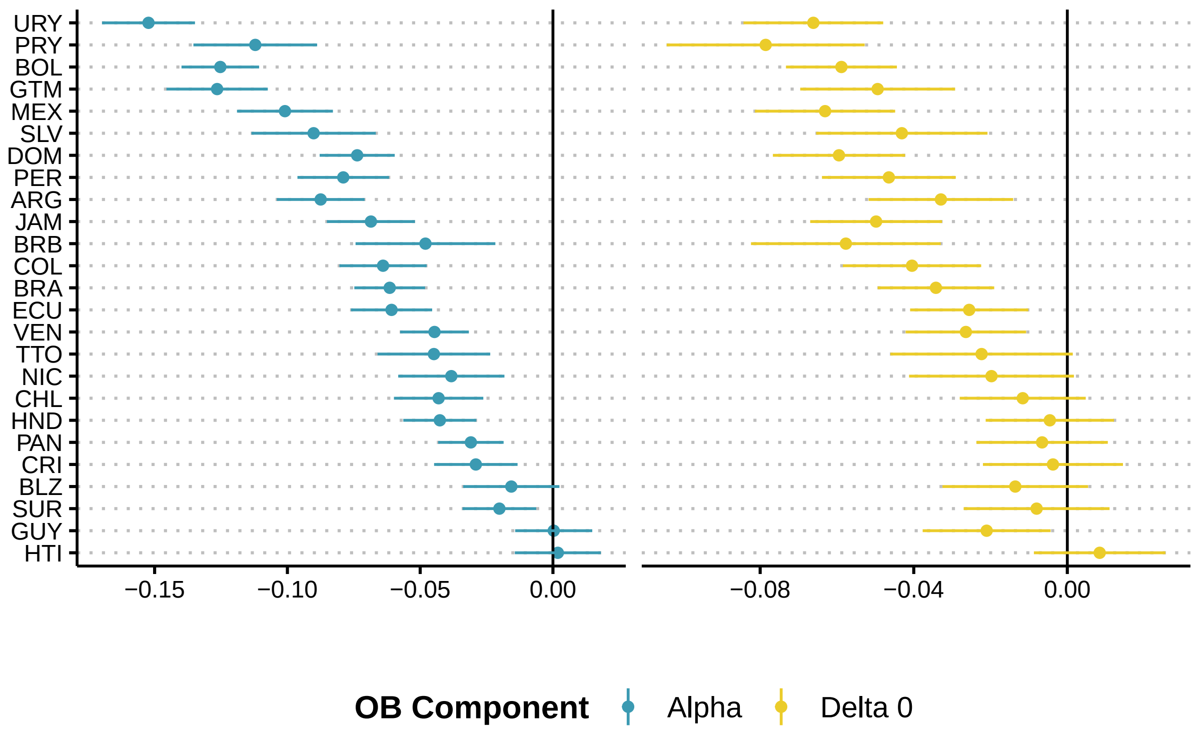

Figure 12: Skin tone effect on predicted income: OB decomposition by country

racial gap cannot be explained by differences in average observable characteristics (OB component Δ_0).

The semi-elasticity of skin tone on income for Argentina is minus 8.75, where the parameter Δ_0 represents 37% of the total effect; for Brazil, the same parameters are minus 6.14 and 55%, The parameter Δ_0 is not statistically different than zero in Belize, Chile, Costa Rica, Honduras, Haiti, Nicaragua, Panama, Suriname and Trinidad and Tobago. Thus, while these countries have the smallest skin tone gradient on income, the racial gap can be explained by differences in average characteristics of the racial groups.

The previous results do not account for measurement error concerns. Nonetheless, Figure 27 in the appendix that the only countries where the Hausman test is statistically significant, and thus there is endogeneity due to measurement error bias, are Colombia, Trinidad and Tobago, Uruguay, and El Salvador.[28] In all cases, the OB decomposition without control functions underestimates the skin tone gradient on income. Figures 28 and 29 in the appendix shows the same exercise but for racial gap on the predicted income and years of schooling (standardized), respectively.

Overall, the patterns by country are consistent with the aggregated measures of racial inequality presented in Section 4. Figure 13 shows that both the OB point estimates α and Δ_0 by country are closely correlated with the mean ratio of racial over total income inequality (MLD between groups/ Total MLD). As the racial gap measured from the micro-data is higher, racial inequality is higher from the macro-data.

[28]Note Figure 27 does not show El Salvador to keep the scale readable. El Salvador's coefficients with control functions for the skin tone gradient are -0.379, and the share of the effect that cannot be explained by returns to average observed characteristics is nearly 28%.

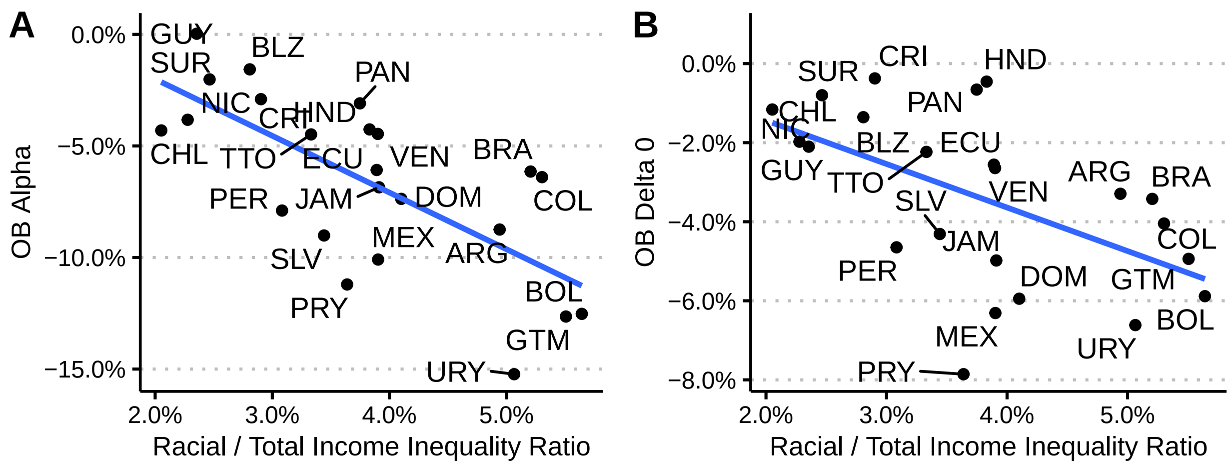

Figure 13: Micro and Macro Coefficients

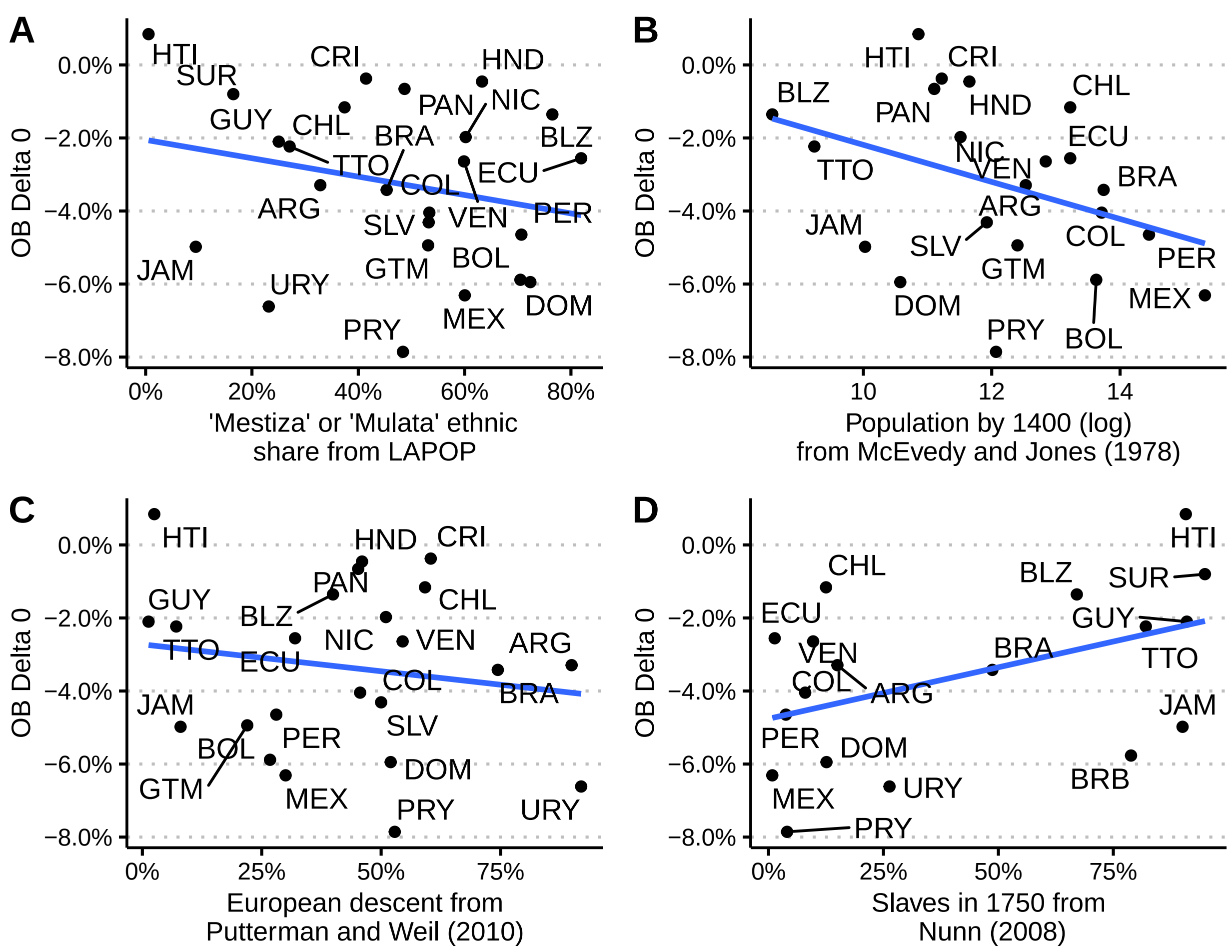

Figure 14: Historical Origins of Racial Disparities

5.5 Historical Origins of Contemporary Racial Disparities

Lastly, the patters are also broadly consistent with the historical evidence presented in Section 2. I use the OB decomposition results by country and historical data to make an exploratory analysis of the origins of contemporary racial disparities. Figure 14 shows the correlation between the OB Δ_0 point estimates and available contemporary and historical data. Panel A shows the share of population defining themselves as 'Mestiza' or 'Mulata' correlates with a higher racial income gap. Thus, the results suggest that the unexplained component of the racial gap is higher in those countries with high *mestizaje* politics. Using McEvedy and Jones (1978) data on historical population patterns,[29] Panel B shows countries with higher population density in 1400, namely places where there was pre-colonial societies, also have a higher unexplained component of racial income gap.

Panel C shows there is a mild correlation between the share of European descent population (Putterman and Weil 2010) and the OB component Δ_0. Therefore, countries with higher share of population descendant of European countries higher racial disparities due to unobserved factors, as racial discrimination. Lastly, using Nunn (2008) data on the share of slave population by 1750, the results suggest that the racial income gap is lower in countries where there was a high share of African slaves by the end of the XVIII century. The latter result is striking, since it suggest that slavery experience actually have less effects on contemporary racial disparities in Latin America. While the latter correlations are simply descriptive, future work must disentangle convincingly the historical patterns and mechanisms that explain contemporary racial inequality.

6 Conclusion

Racial disparities and discrimination are pending issues to solve globally. In this paper, I study racial inequalities at the individual and national levels in Latin America. Using the PERLA color palette instead of broad ethnic-racial categories, I complied a data set with information on skin tone and economic outcomes, as income and years of schooling, for more than 100,000 individuals in 26 Latin American countries during the last decade.

In the first part of the paper, I find that racial inequality, or income inequality between racial groups, hinders economic development at the national level. I compute new measures of racial inequality using skin tone rather than broad ethnic-racial categories. Controlling for time-invariant characteristics and common-shock across countries, an increase in one percent of the ratio of racial over total income inequality decreases GDP per capita by 4 percent. Consistent with Alesina, Michalopoulos, and Papaioannou (2016), inequalities between racial groups have adverse effects on aggregate welfare.

In the second part of the paper, I found evidence there is a substantial income gap by skin tone throughout the region: out of the 11 tones in the PERLA color palette, each increase in a darker skin tone correlates with a decrease between 4 and 6 percent in monthly income per capita. Around 60% of the gap cannot be explained by differences in observed average characteristics (i.e., years of schooling, gender, country-region, age, among others), or whether the individual participates in the labor market. Furthermore,

[29]Taken from Nunn and Puga (2012).

there is also a significant educational gap between racial groups. Each darker skin tone is associated with a decrease of 0.05 standard deviations in years of schooling. Lastly, people with darker skin tones have a higher probability of reporting discriminatory behavior against them by their skin tone in public places, schools, and workplaces.

The results show evidence of a substantial racial income gap and suggest two conclusions regarding the mechanisms. First, even I cannot isolate the effect of taste-based discrimination on the income gap given the nature of the data, results suggest that at least 60% of the gap is not explained by returns to average observable characteristics. Therefore, even if taste-based discriminatory behavior does not account for all of the latter, it is most plausible that it operates as a mechanism explaining the racial income gap. Secondly, the results suggest more convincingly that the racial income gap in Latin America is explained by statistical discrimination in labor markets due to the gap in human capital between racial groups. This hypothesis is consistent with the historical and anecdotal evidence, where darker tones have higher stigmatization. Moreover, it is also consistent with a hypothesis of occupational segregation by racial groups. Further research needs to disentangle the exact mechanisms operating and explaining the racial income gap.

There is substantial room for public policies. More specifically, and consistent with recent literature (Derenoncourt and Montialoux 2020), the results suggest that progressive income and wealth taxation is also progressive in terms of racial disparities. Furthermore, given that skin tone is fixed and there is little room for behavioral responses, taxing alternatives as ‘tagging’ (Akerlof 1978; Piketty and Saez 2013) could play an essential role in overcoming racial disparities at early year stages. Alongside its critical role in economic development, there is a historical debt in reducing racial disparities for justice and reparation.

References

Advani, Arun, Elliott Ash, David Cai, and Imran Rasul. 2021. “Race-Related Research in Economics and Other Social Sciences.”

Aguilar, Rosario. 2011. “Social and Political Consequences of Stereotypes Related to Racial Phenotypes in Mexico.”

Akerlof, George A. 1978. “The economics of ‘tagging’ as applied to the optimal income tax, welfare programs, and manpower planning.” *The American Economic Review* 68 (1): 8–19.

Akerlof, George A., and Rachel E. Kranton. 2000. “Economics and identity.” *Quarterly Journal of Economics* 115 (3): 715–53. https://doi.org/10.1162/003355300554881.

Alesina, Alberto, and Eliana La Ferrara. 2005. “Ethnic Diversity and Economic Performance.” *Journal of Economic Literature* 43 (3): 762–800. https://doi.org/10.1257/002205105774431243.

Alesina, Alberto, Stelios Michalopoulos, and Elias Papaioannou. 2016. “Ethnic Inequality.” *Journal of Political Economy* 124 (2): 428–88. https://doi.org/10.1086/685300.

Altonji, Joseph G., and Rebecca M. Blank. 1999. “Chapter 48 Race and gender in the labor market.” In *Handbook of Labor Economics*, edited by Orley C. Ashenfelter and David Card, 3 PART:3143–3259.

3. Elsevier. https://doi.org/10.1016/S1573-4463(99)30039-0.

Anderson, Rodney D. 1988. "Race and Social Stratification: A Comparison of Working-Class Spaniards, Indians, and Castas in Guadalajara, Mexico in 1821." *The Hispanic American Historical Review* 68 (2): 209. https://doi.org/10.2307/2515513.

Arceo-Gómez, Eva O., and Raymundo M. Campos-Vázquez. 2014. "Race and Marriage in the Labor Market: A Discrimination Correspondence Study in a Developing Country." *American Economic Review* 104 (5): 376–80. https://doi.org/10.1257/aer.104.5.376.

———. 2019. "Double Discrimination: Is Discrimination in Job Ads Accompanied by Discrimination in Callbacks?" *Journal of Economics, Race, and Policy* 2 (4): 257–68. https://doi.org/10.1007/s41996-019-00031-3.

Arias, Luz Marina, and Desha M. Dirod. 2014. "Indigenous Origins of Colonial Institutions." *Quarterly Journal of Political Science* 9 (3): 371–406. https://doi.org/10.1561/100.00013135.

Bezin, Emeline, and Fabien Moizeau. 2017. "Cultural dynamics, social mobility and urban segregation." *Journal of Urban Economics* 99 (May): 173–87. https://doi.org/10.1016/j.jue.2017.02.004.

Bisin, Alberto, and Thierry Verdier. 2011. "The Economics of Cultural Transmission and Socialization." In *Handbook of Social Economics*, 339–416. https://doi.org/10.1016/B978-0-444-53187-2.00009-7.

Botelho, Fernando, Ricardo A. Madeira, and Marcos A. Rangel. 2015. "Racial Discrimination in Grading: Evidence from Brazil." *American Economic Journal: Applied Economics* 7 (4): 37–52. https://doi.org/10.1257/app.20140352.

Bowles, Samuel, Glenn C. Loury, and Rajiv Sethi. 2014. "Group Inequality." *Journal of the European Economic Association* 12 (1): 129–52. https://doi.org/10.1111/jeea.12037.

Cameron, A. Colin, Jonah B. Gelbach, and Douglas L. Miller. 2008. "Bootstrap-Based Improvements for Inference with Clustered Errors." *Review of Economics and Statistics* 90 (3): 414–27. https://doi.org/10.1162/rest.90.3.414.

Campos-Vázquez, Raymundo M., and Carolina Rivas-Herrera. 2021. "The Color of Electoral Success: Estimating the Effect of Skin Tone on Winning Elections in Mexico." *Social Science Quarterly* 102 (2): 844–64. https://doi.org/10.1111/ssqu.12933.

Campos-Vázquez, Raymundo M. 2020. "The higher price of whiter skin: an analysis of escort services." *Applied Economics Letters*. https://doi.org/10.1080/13504851.2020.1725229.

Campos-Vázquez, Raymundo M., and Eduardo M. Medina-Cortina. 2017. "Identidad social y estereotipos por color de piel. Aspiraciones y desempeño en jóvenes mexicanos." *El Trimestre Económico* 85 (337): 53. https://doi.org/10.20430/ete.v85i337.659.

———. 2019. "Skin Color and Social Mobility: Evidence From Mexico." *Demography* 56 (1): 321–43. https://doi.org/10.1007/s13524-018-0734-z.

Castorena, Oscar. 2021. "Factores de expansión en los datos del Barómetro de las Américas." https://www.vanderbilt.edu/lapop/insights/IMN007es.pdf.

Cernat, Alexandru, Joseph W. Sakshaug, and Javier Castillo. 2019. "The Impact of Interviewer Effects on Skin Color Assessment in a Cross-National Context." *International Journal of Public Opinion Research* 31 (4): 779–93. https://doi.org/10.1093/ijpor/edy030.

Chakravarty, Satya R. 2015. *Inequality, Polarization and Conflict*. Vol. 12. Economic Studies in Inequality, Social Exclusion and Well-Being. New Delhi: Springer India. https://doi.org/10.1007/978-81-322-2166-1.

Chetty, Raj, Nathaniel Hendren, Maggie R. Jones, and Sonya R. Porter. 2020. "Race and Economic Opportunity in the United States: an Intergenerational Perspective." *The Quarterly Journal of Economics* 135 (2): 711–83. https://doi.org/10.1093/qje/qjz042.

Chong, Alberto E., and Hugo Ñopo. 2008. "The Mystery of Discrimination in Latin America." *Economía* 8 (2): 79–115.

Cook, Lisa D., and Trevon D. Logan. 2020. "Racial Inequality." Economics for Inclusive Prosperity.

Cook, Lisa D., Trevon D. Logan, and John M. Parman. 2016. "The mortality consequences of distinctively black names." *Explorations in Economic History* 59 (January): 114–25. https://doi.org/10.1016/j.eeh.2015.10.001.

Cowell, F. A. 2000. "Measurement of inequality." In *Handbook of Income Distribution*, edited by Anthony B. Atkinson and François Bourguignon, 87–166. https://doi.org/10.1016/S1574-0056(00)80005-6.

Cunningham, Scott. 2021. *Causal Inference. The Mixtape.* Yale University Press.

De Rosa, Mauricio, Ignacio Flores, and Marc Morgan. 2020. "Inequality in Latin America Revisited: Insights from Distributional National Accounts." World Inequality Lab.

Dell, Melissa. 2010. "The Persistent Effects of Peru's Mining Mita." *Econometrica* 78 (6): 1863–903. https://doi.org/10.3982/ECTA8121.

Derenoncourt, Ellora, and Claire Montialoux. 2020. "Minimum Wages and Racial Inequality." *The Quarterly Journal of Economics* 136 (1): 169–228. https://doi.org/10.1093/qje/qjaa031.

Dixon, Angela R., and Edward E. Telles. 2017. "Skin color and colorism: Global research, concepts, and measurement." *Annual Review of Sociology.* https://doi.org/10.1146/annurev-soc-060116-053315.

Ferreira, Francisco H. G., and Jérémie Gignoux. 2011. "The Measurement of inequality of opportunity: Theory and an Application to Latin America." *Review of Income and Wealth* 57 (4): 622–57. https://doi.org/10.1111/j.1475-4991.2011.00467.x.

Fleurbaey, Marc, and Erik Schokkaert. 2011. "Equity in Health and Health Care." In, edited by Mark V. Pauly, Thomas G. Mcguire, and Pedro P. Barros, 1003–92. https://doi.org/10.1016/B978-0-444-53592-4.00016-5.

Foster, James E., and Artyom A. Shneyerov. 2000. "Path Independent Inequality Measures." *Journal of Economic Theory* 91 (2): 199–222. https://doi.org/10.1006/jeth.1999.2565.

Fryer, Roland G. 2011. *Racial inequality in the 21st century: The declining significance of discrimination.* Vol. 4. PART B. https://doi.org/10.1016/S0169-7218(11)02408-7.

Graham, Richard. 1990. "Introduction." In *The Idea of Race in Latin America. 1870-1940*, edited by Richard Graham, 1–6. Austin, TX: University of Texas Press.

———. 2013. *Independence in Latin America. Contrast and comparisons.* University of Texas Press.

Hastie, Trevor, and Robert Tibshirani. 1993. "Varying-Coefficient Models." *Journal of the Royal Statistical Society. Series B (Methodological)* 55 (4): 757–96.

Hastie, Trevor, Robert Tibshirani, and Jerome Friedman. 2009. *The Elements of Statistical Learning.* Springer Series in Statistics. New York, NY: Springer New York. https://doi.org/10.1007/978-0-387-84858-7.

Haughton, Jonathan, and Shahidur R Khandker. 2009. *Handbook on Poverty and Inequality.* Edited by World Bank. Washington, DC: World Bank.

Heckman, James J. 1979. "Sample Selection Bias as a Specification Error." *Econometrica* 47 (1): 153. https://doi.org/10.2307/1912352.

Helg, Aline. 1990. "Race in Argentina and Cuba, 1880-1930: Theory, Policies, and Popular Reaction." In *The Idea of Race in Latin America. 1870-1940*, edited by Richard Graham, 37–70. University of Texas Press.

Izecksohn, Vitor. 2014. *Slavery and War in the Americas: Race, Citizenship, and State Building in the United States and Brazil, 1861-1870.* University of Virginia Press.

Knight, Alan. 1990. "Racism, Revolution, and Indigenismo: Mexico, 1910-1940." In *The Idea of Race in Latin America. 1870-1940*, edited by Richard Graham, 71–114. University of Texas Press.

Krozer, Alice, and Andrea Urrutia Gómez. 2021. "Not in the eyes of the beholder: Racialisation, Whiteness and Beauty Standards in Mexico." Project on Racial and Ethnic Discrimination in Mexico. El Colegio de México.

Logan, Trevon D., and John M. Parman. 2017. "The National Rise in Residential Segregation." *The Journal of Economic History* 77 (1): 127–70. https://doi.org/10.1017/S0022050717000079.

Loveman, Mara. 2014. *National Colors. Racial Classification and the State in Latin America.* Oxford University Press. https://doi.org/10.1093/acprof:oso/9780199337354.001.0001.

Lustig, Nora, Luis F. Lopez-Calva, and Eduardo Ortiz-Juarez. 2016. "Deconstructing the Decline in Inequality in Latin America." In *Inequality and Growth: Patterns and Policy*, edited by Kaushik Basu and Joseph E. Stiglitz, 212–47. London: Palgrave Macmillan UK. https://doi.org/10.1057/9781137554598_7.

McCaa, Robert, Stuart B. Schwartz, and Arturo Grubessich. 1979. "Race and Class in Colonial Latin America: A Critique." *Comparative Studies in Society and History* 21 (3): 421–33. https://doi.org/10.1017/S0010417500013074.

McEvedy, Cohn, and Richard Jones. 1978. *Atlas of world population history.* Penguin Books Ltd.

Monroy-Gómez-Franco, Luis, and Roberto Vélez-Grajales. 2020. "Skin Tone Differences in Social Mobility in Mexico: Are We Forgetting Regional Variance?" *Journal of Economics, Race, and Policy*, July. https://doi.org/10.1007/s41996-020-00062-1.

Montalvo, José G., and Marta Reynal-Querol. 2005. "Ethnic Polarization, Potential Conflict, and Civil Wars." *American Economic Review* 95 (3): 796–816. https://doi.org/10.1257/0002828054201468.

Murphy, Kevin M., and Robert H. Topel. 1985. "Estimation and Inference in Two-Step Econometric Models." *Journal of Business & Economic Statistics* 3 (4): 370. https://doi.org/10.2307/1391724.

Nunn, Nathan. 2008. "Slavery, Inequality, and Economic Development in the Americas: An Examination of the Engerman-Sokoloff Hypothesis." In *Institutions and Economic Performance*, edited by E. Helpman, 148–80. Harvard University Press.

———. 2020. "The historical roots of economic development." *Science* 367 (6485). https://doi.org/10.1126/science.aaz9986.

Nunn, Nathan, and Diego Puga. 2012. "Ruggedness: The Blessing of Bad Geography in Africa." *Review of Economics and Statistics* 94 (1): 20–36. https://doi.org/10.1162/REST_a_00161.

Ñopo, Hugo. 2008a. "Matching as a Tool to Decompose Wage Gaps." *The Review of Economics and Statistics* 90 (2): 290–99.

———. 2008b. "An extension of the Blinder–Oaxaca decomposition to a continuum of comparison groups." *Economics Letters* 100 (2): 292–96. https://doi.org/10.1016/j.econlet.2008.02.011.

———. 2012. *New Century, Old Disparities : Gender and Ethnic Earnings Gaps in Latin America and the Caribbean*. The World Bank.

Ñopo, Hugo, Alberto Chong, and Andrea Moro, eds. 2009. *Discrimination in Latin America: an economic perspective*. Washington DC: The Inter-American Development Bank.

Ñopo, Hugo, Jaime Saavedra, and Máximo Torero. 2007. "Ethnicity and earnings in a mixed-race labor market." *Economic Development and Cultural Change* 55 (4): 709–34. https://doi.org/10.1086/516762.

Piketty, Thomas. 2020. *Capital e Ideología*. Grano de Sal.

Piketty, Thomas, and Emmanuel Saez. 2013. "Optimal Labor Income Taxation." In *Handbook of Public Economics*, 391–474. https://doi.org/10.1016/B978-0-444-53759-1.00007-8.

Putterman, Louis, and David N. Weil. 2010. "Post-1500 population flows and the long-run determinants of economic growth and inequality." *Quarterly Journal of Economics* 125 (4): 1627–82. https://doi.org/10.1162/qjec.2010.125.4.1627.

Rios-Avila, Fernando. 2019. "A Semi-Parametric Approach to the Oaxaca–Blinder Decomposition with Continuous Group Variable and Self-Selection." *Econometrics* 7 (2): 28. https://doi.org/10.3390/econometrics7020028.

Sánchez-Ancochea, Diego. 2021. "All about ideology? Reading Piketty´s with Latin American lenses." *British Journal of Sociology* 72 (1): 125–38. https://doi.org/10.1111/1468-4446.12803.

Shayo, Moses. 2009. "A model of social identity with an application to political economy: Nation, class, and redistribution." *American Political Science Review*. https://doi.org/10.1017/S0003055409090194.

Shorrocks, Anthony F. 1984. "Inequality Decomposition by Population Subgroups." *Econometrica* 52 (6): 1369. https://doi.org/10.2307/1913511.

Skidmore, Thomas E. 1990. "Racial Ideas and Social Policy in Brazil, 1870-1940." In *The Idea of Race in Latin America. 1870-1940*, edited by Richard Graham, 7–36. University of Texas Press.

Solís, Patricio, Marcela Avitia, and Braulio Güémez. 2020. "Tono de piel y desigualdad socioeconómica en México. Reporte de la Encuesta Proder # 1." El Colegio de México. https://discriminacion.colmex.mx/wp-content/uploads/2020/07/info1.pdf.

Solís, Patricio, Braulio Güémez Graniel, and Virginia Lorenzo Holm. 2019. "Por mi raza hablará la desigualdad." Ciudad de México: Oxfam México.

Taylor, Alan. 2013. *Colonial America. A very Short Introduction.* Oxford University Press.

Telles, Edward, and Regina Martínez Casas. 2019. *Pigmentocracias: Color, etnicidad y raza en América Latina.* FCE. https://doi.org/10.25100/sye.v0i29.3925.

Telles, Edward, and Edward Murguia. 1990. "Phenotypic discrimination and income differences among Mexican Americans." *Social Science Quarterly* 71 (4): 682–96.

Telles, Edward, and Liza Steele. 2012. "Pigmentocracia en las Américas: ¿cómo se relaciona el logro educativo con el color de piel?" Perspectivas Desde El Barómetro de Las AMéricas.

Tenorio-Trillo, Mauricio. 2017. *Latin America. The Allure and Power of an Idea.* The University of Chicago Press.

Tesei, Andrea. 2014. "Trust, Racial Fragmentation and Income Inequality: New Evidence from the U.S." https://papers.ssrn.com/abstract=2425582.

The Economist. 2020. "A lack of data on race hampers efforts to tackle inequalities." *The Economist.* https://www.economist.com/leaders/2020/11/21/a-lack-of-data-on-race-hampers-efforts-to-tackle-inequalities.

Torche, Florencia. 2015. "Intergenerational Mobility and Gender in Mexico." *Social Forces* 94 (2): 563–87. https://doi.org/10.1093/sf/sov082.

Trejo, Guillermo, and Melina Altamirano. 2016. "The Mexican Color Hierarchy: How Race and Skin Tone Still Define Life Chances 200 Years after Independence." In *Racial and Class Inequalities in the Americas*, edited by Juliet Hooker and Alvin Tillery. Washington, DC: American Political Science Association.

Wooldridge, Jeffrey M. 2015. "Control Function Methods in Applied Econometrics." *Journal of Human Resources* 50 (2): 420–45. https://doi.org/10.3368/jhr.50.2.420.

Zizumbo-Colunga, Daniel, and Iván Flores Martínez. 2017. "Is Mexico a Post-Racial Country? Inequality and Skin Tone across the Americas." Latin American Public Opinion Project.

Appendix

Table 6: LAPOP AmericasBarometer: Sample size by country and wave.

Country	ISO Code	2010	2012	2014	2016/17	2018/19	Total
Mexico	MEX	1353	1215	1143	1337	1345	6393
Argentina	ARG	1089	1038	896	1191	1319	5533
Bolivia	BOL	2492	2450	2383	1443	1456	10224
Brazil	BRA	2262	1071	1441	1391	1349	7514
Chile	CHL	1629	1331	1139	1418	1392	6909
Colombia	COL	1341	1215	1376	1283	1313	6528
Costa Rica	CRI	1102	1039	1109	1276	1330	5856
Dominican Republic	DOM	1293	1280	1316	1166	1296	6351
Ecuador	ECU	2777	1351	1315	1322	1407	8172
El Salvador	SLV	1460	1240	1276	1331	1227	6534
Guatemala	GTM	1207	1117	1301	1075	1184	5884
Honduras	HND	1491	1317	1240	1123	1099	6270
Jamaica	JAM	1116	884	961	1013	1016	4990
Nicaragua	NIC	1418	1540	1421	1232	1365	6976
Panama	PAN	1469	1332	1366	1327	1351	6845
Paraguay	PRY	1106	1305	1101	1081	1317	5910
Peru	PER	1368	1306	1162	2319	1327	7482
Uruguay	URY	1396	1336	1394	1378	1456	6960
Haiti	HTI		1141	873			2014
Belize	BLZ	1205	850	994			3049
Guyana	GUY	1109	1120	1102			3331
Suriname	SUR	1197	863	1819			3879
Trinidad and Tobago	TTO	926	796	1589			3311
Venezuela	VEN	1270	942	1178			3390
Bahamas	BHS			1711			1711
Barbados	BRB			1726			1726
Total by wave		33076	29079	34332	23706	23549	143742

Table 7: LAPOP AmericasBarometer: Descriptive Statistics

Statistic	Mean	St. Dev.	Min	Pctl(25)	Pctl(75)	Max
Age	39.54	15.74	16	26	50	89
Skin tone (PERLA palette)	4.61	1.89	1	3	6	11
Years of schooling	9.72	4.25	0	6	12	18
Income per capita (PPP 2019)	276.68	361.43	1.22	78.74	333.34	4,297.46
Predicted Income per capita (PPP 2019)	245.22	163.09	1.05	130.60	325.13	1,164.18
Household size	4.27	2.23	1	3	5	17
Gender						
Female	0.50	0.50				
Male	0.50	0.50				
Ethnicity						
Afro	0.12	0.32				
Indigenous	0.07	0.26				
Mestiza	0.46	0.50				
Mulata	0.05	0.22				
Other	0.07	0.25				
White	0.23	0.42				
Occupation						
Actively looking for a job	0.08	0.27				
Not working and not looking for a job	0.03	0.16				
Not Working but have job	0.04	0.20				
Retired	0.08	0.27				
Studying	0.07	0.26				
Taking care of the home	0.19	0.39				
Working	0.51	0.50				
Marital Status						
Divorced or Separated	0.06	0.24				
Living together	0.25	0.43				
Married	0.34	0.47				
Single	0.31	0.46				
Widowed	0.04	0.20				
Urbanization						
Big City	0.17	0.37				
Medium City	0.16	0.37				
Small City	0.15	0.35				
Metropolitan area	0.22	0.41				
Rural area	0.30	0.46				

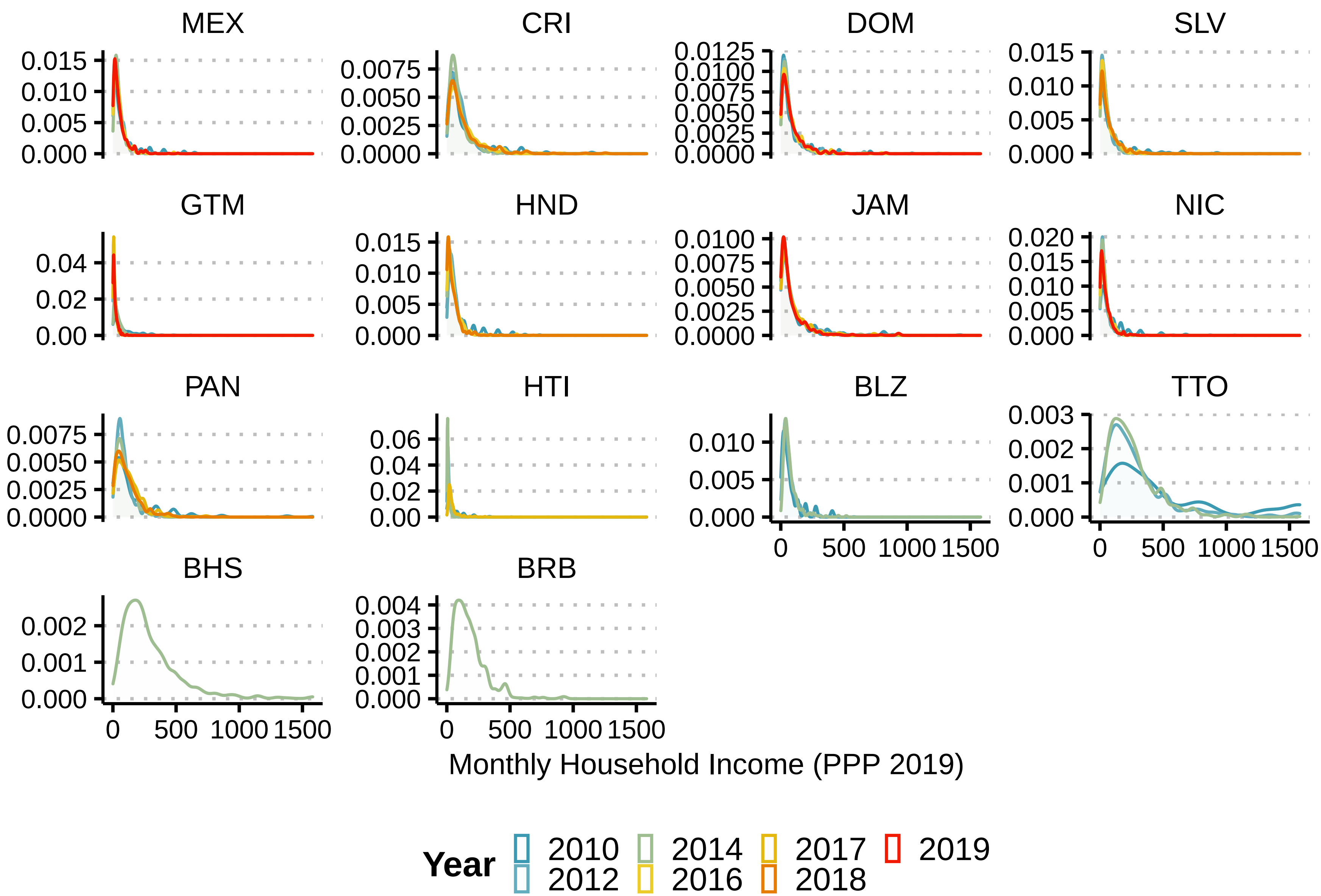

Figure 15: Monthly Household Income distributions by country (1)

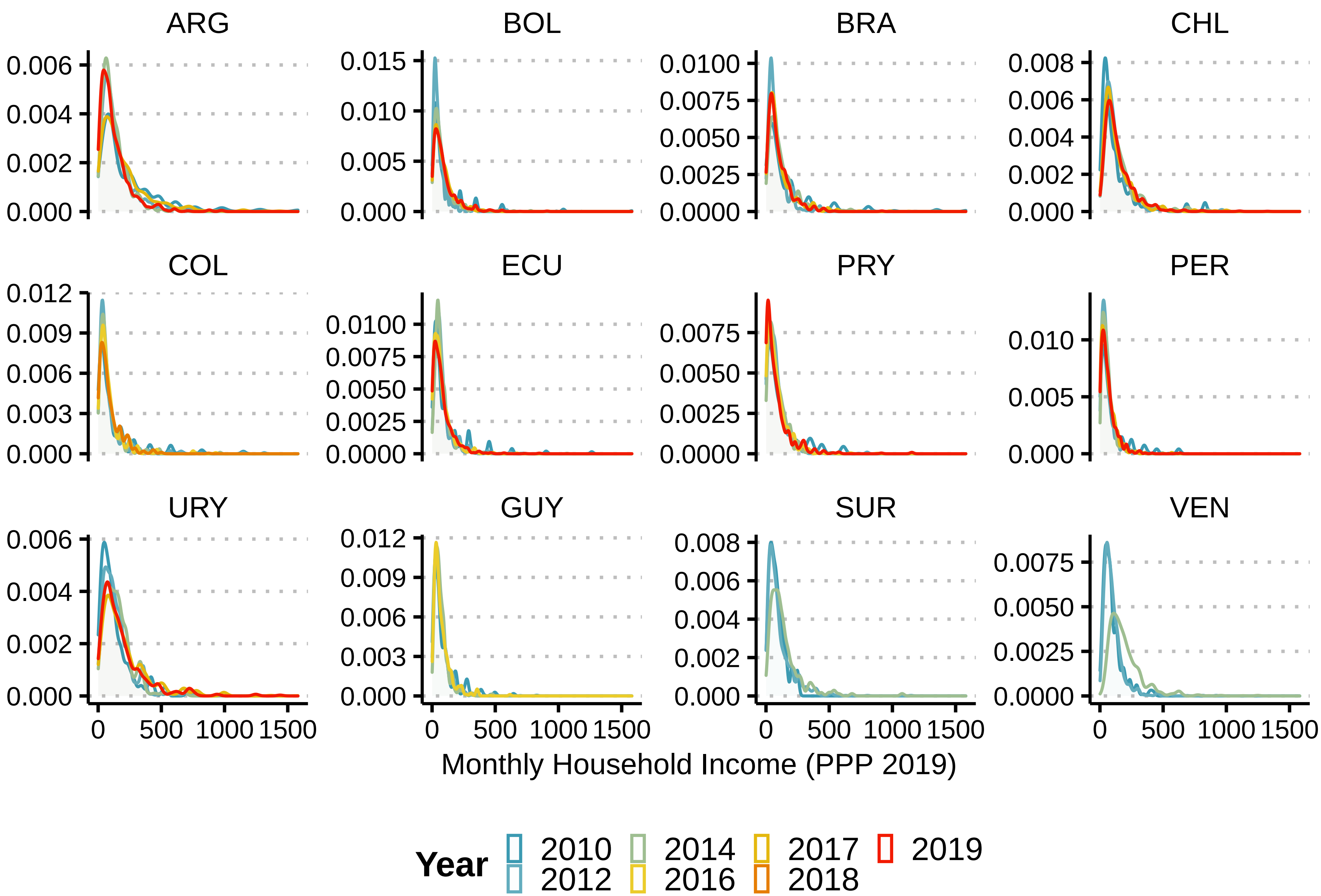

Figure 16: Monthly Household Income distributions by country (2)

Figure 17: Income and Household Asset Index

Table 8: Racial composition within income deciles

	Dependent variable:	
	Fractionalization	Share of lightest PERLA colors
	(1)	(2)
Decile's mean income per capita (log)	-0.667^{***}	0.043^{***}
	(0.214)	(0.002)
Observations	1,100	1,100
R^2 within	0.009	0.251
Country FE	Yes	Yes
Year FE	Yes	Yes

Notes: Robust standard errors in parenthesis. Deciles are weighted by number of people in each decile. *p<0.1; **p<0.05; ***p<0.01

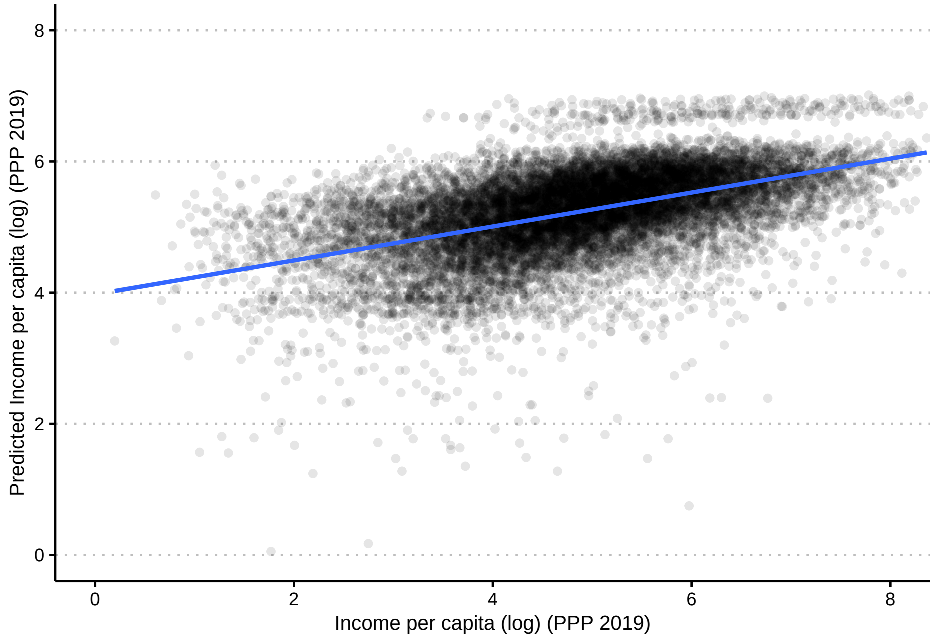

Figure 18: Income and predicted Income

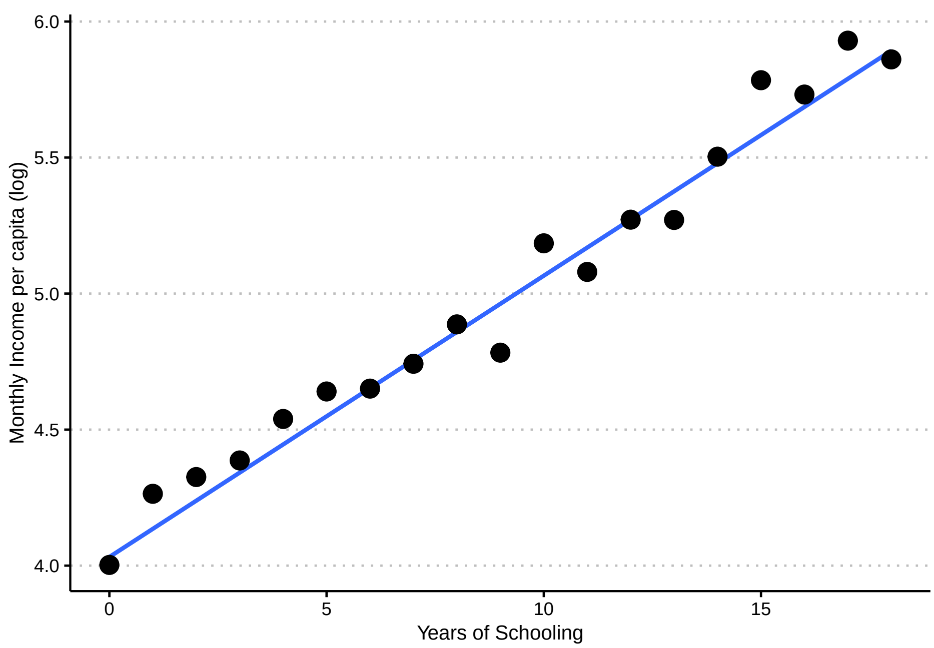

Figure 19: Income per capita and Years of Schooling

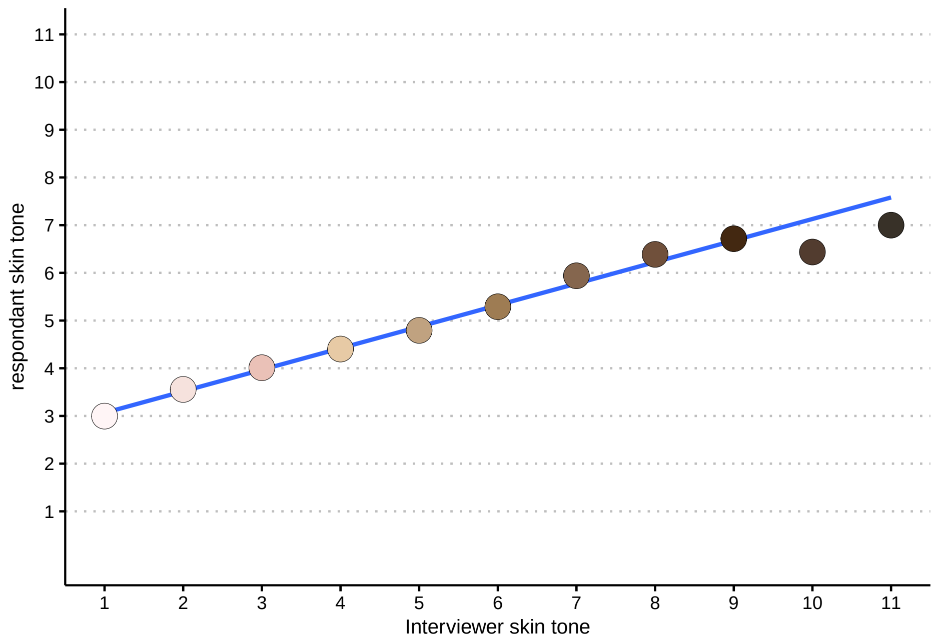

Figure 20: Interviewer and respondant skin tone

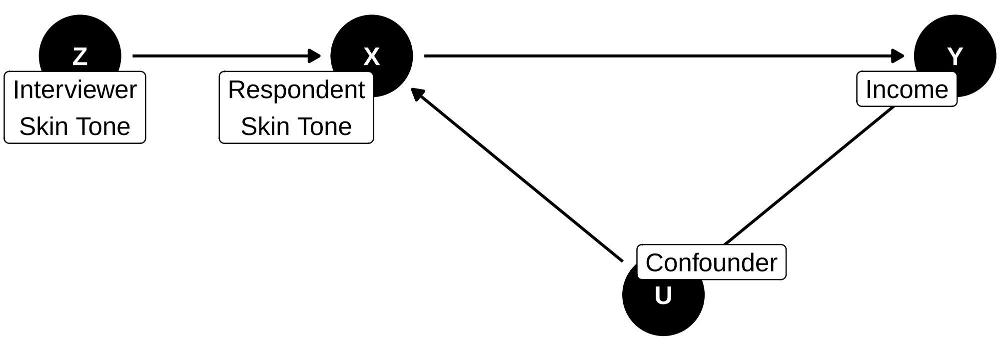

Figure 21: DAG for interviewer skin tone CF or IV

Table 9: Racial Inequality and Economic Development (2)

	Dependent variable:			
	GDP per capita (log)			
	(1)	(2)	(3)	(4)
MLD Total (log)	0.040	0.049	0.044	0.052
	[0.069]	[0.076]	[0.078]	[0.086]
	(0.057)	(0.061)	(0.062)	(0.069)
MLD Between Racial groups (log)	-0.048^{***}	-0.040^{***}	-0.040^{***}	-0.041^{**}
	[0.017]	[0.014]	[0.015]	[0.018]
	(0.018)	(0.017)	(0.017)	(0.018)
MLD Between Ethnic groups (log)		-0.018	-0.018	-0.021
		[0.018]	[0.017]	[0.023]
		(0.014)	(0.014)	(0.016)
Spatial Inequality (log) × Time			0.001	0.0004
			[0.006]	[0.005]
			(0.002)	(0.002)
Administrative Inequality (log) × Time			0.004	0.004
			[0.010]	[.011]
			(0.007)	(0.007)
Racial Fraccionalization (log)				0.038
				[0.259]
				(0.226)
Racial Segregation (log)				-0.230
				[1.253]
				(0.890)
Ethnic Fraccionalization (log)				0.048
				[0.102]
				(0.055)
Ethnic Segregation (log)				0.365
				[1.375]
				(0.947)
Observations	109	109	109	109
R^2 within	0.089	0.104	0.116	0.125
Country FE	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes

Notes: Bootstrapped standard errors clustering by country in brackets based on 500 bootstrap replications. Robust standard errors in parenthesis. Spatial and administrative inequality measures taken from Alesina, Michalopoulos, and Papaioannou (2016). *p<0.1; **p<0.05; ***p<0.01
*p<0.1; **p<0.05; ***p<0.01

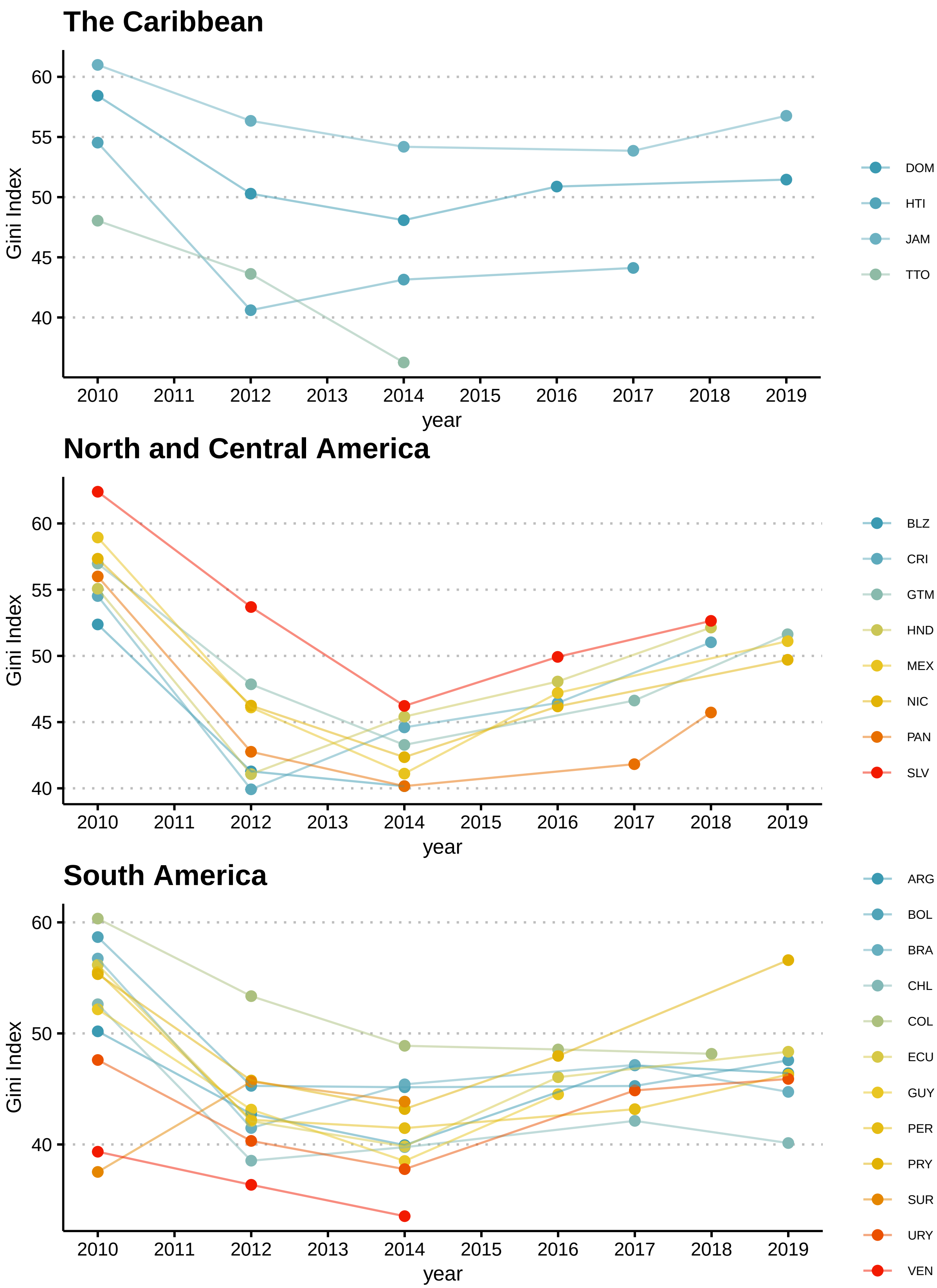

Figure 22: Evolution of income inequality in Latin America

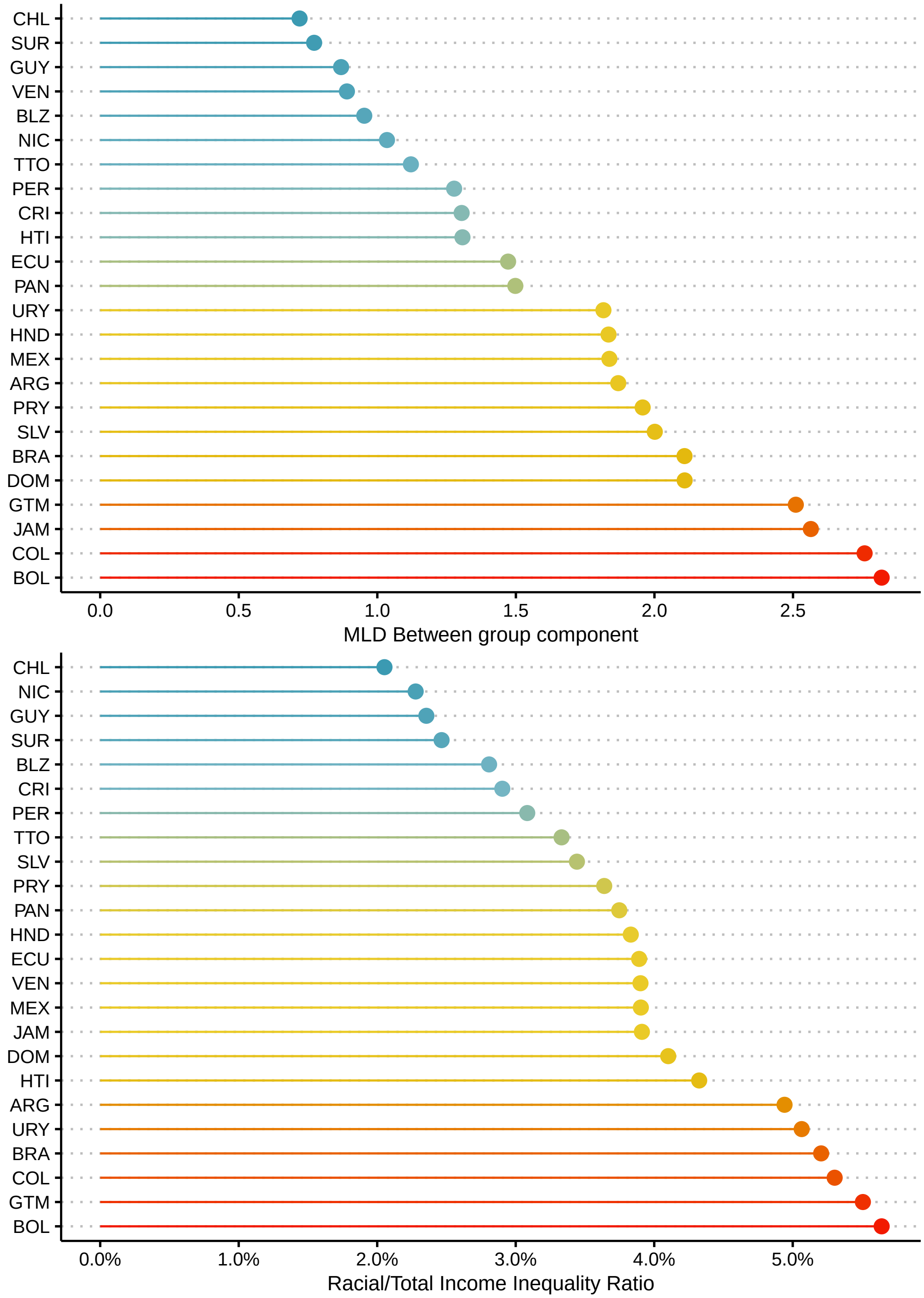

Figure 23: Mean Racial inequality measures

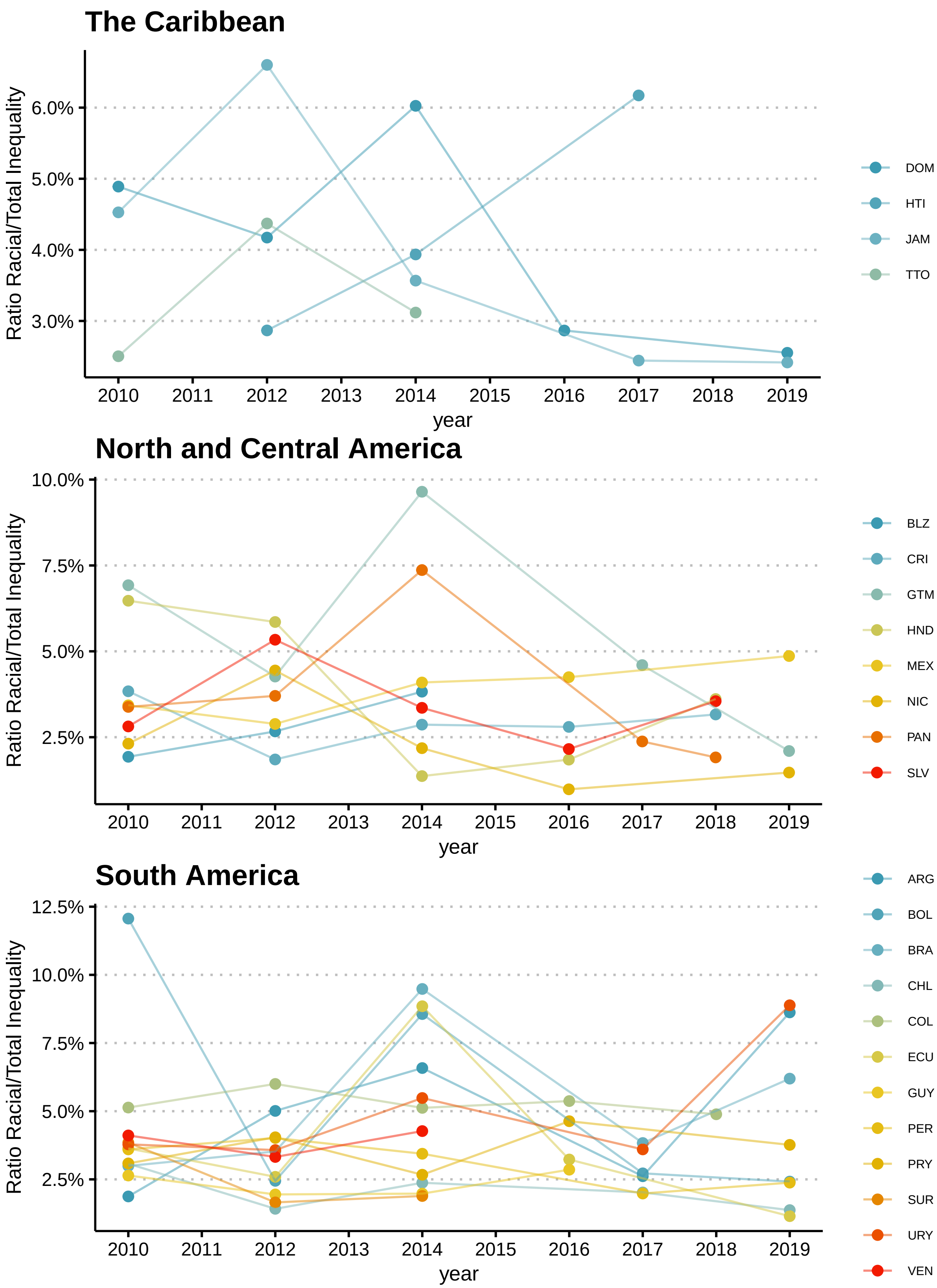

Figure 24: Evolution of racial-inequality over total inequality in Latin America

Table 10: Racial Inequality and Economic Development (3)

	Dependent variable:	
	GDP per capita (log)	
	(1)	(2)
Racial/Total Inequality Ratio (log)	-0.049^{**}	-0.051^{**}
	[0.021]	[0.021]
	(0.020)	(0.021)
GREG Ethnic Inequality (log) × Time	0.007	
	[0.016]	
	(0.006)	
Ethnologue Ethnic Inequality (log) × Time		-0.0003
		[0.006]
		(0.002)
Observations	109	109
R^2 within	0.128	0.107
Controls	Yes	Yes
Country FE	Yes	Yes
Year FE	Yes	Yes

Notes: Bootstrapped standard errors clustering by country in brackets based on 500 bootstrap replications. Robust standard errors in parenthesis. Controls include spatial and administrative inequality, and LAPOP based indexes of racial and ethnic fraccionalization and polarization, and spatial and administrative inequality measures taken from Alesina, Michalopoulos, and Papaioannou (2016).
*p<0.1; **p<0.05; ***p<0.01
*p<0.1; **p<0.05; ***p<0.01

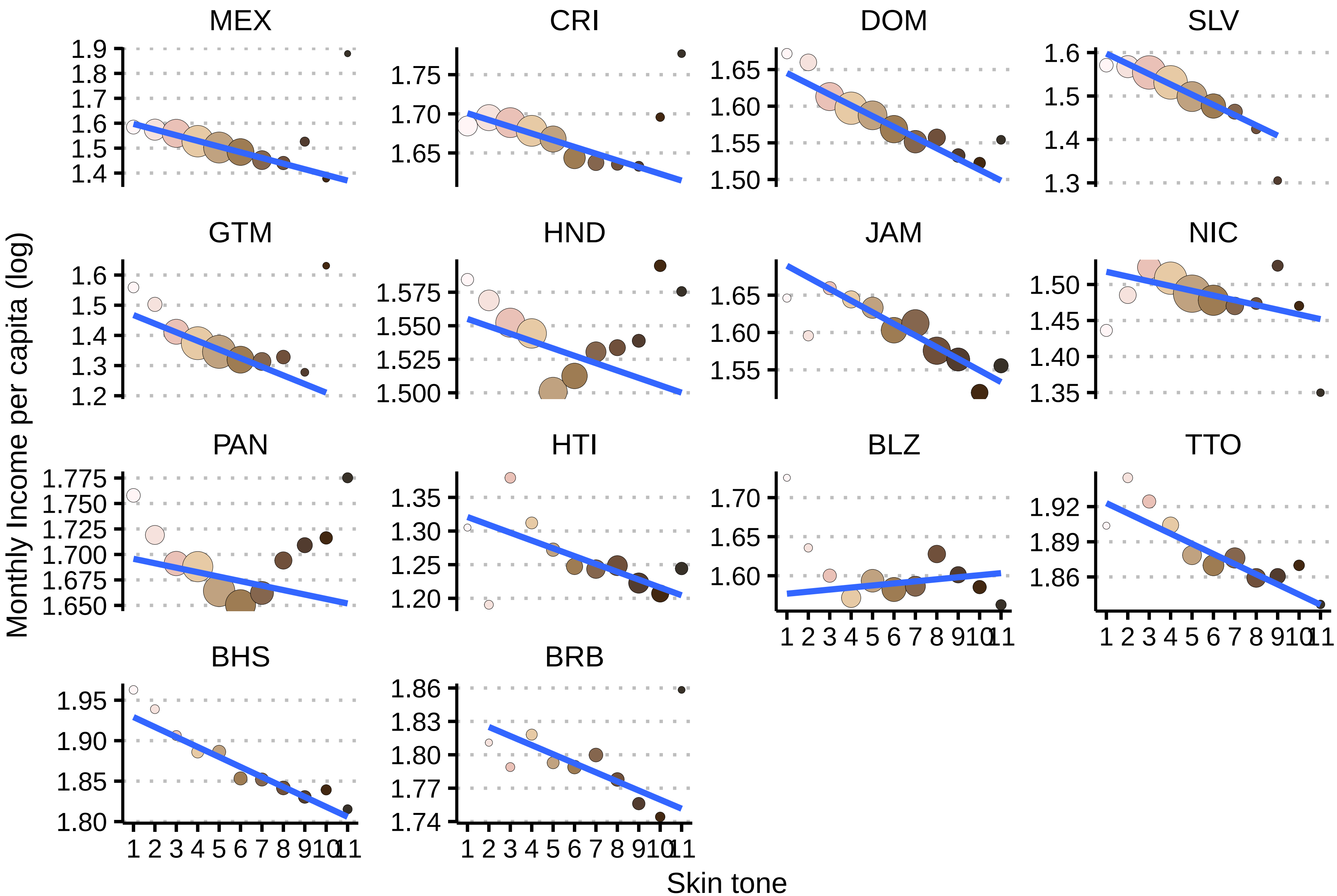

Figure 25: Skin tone gradient for mean monthly income per capita by country (1)

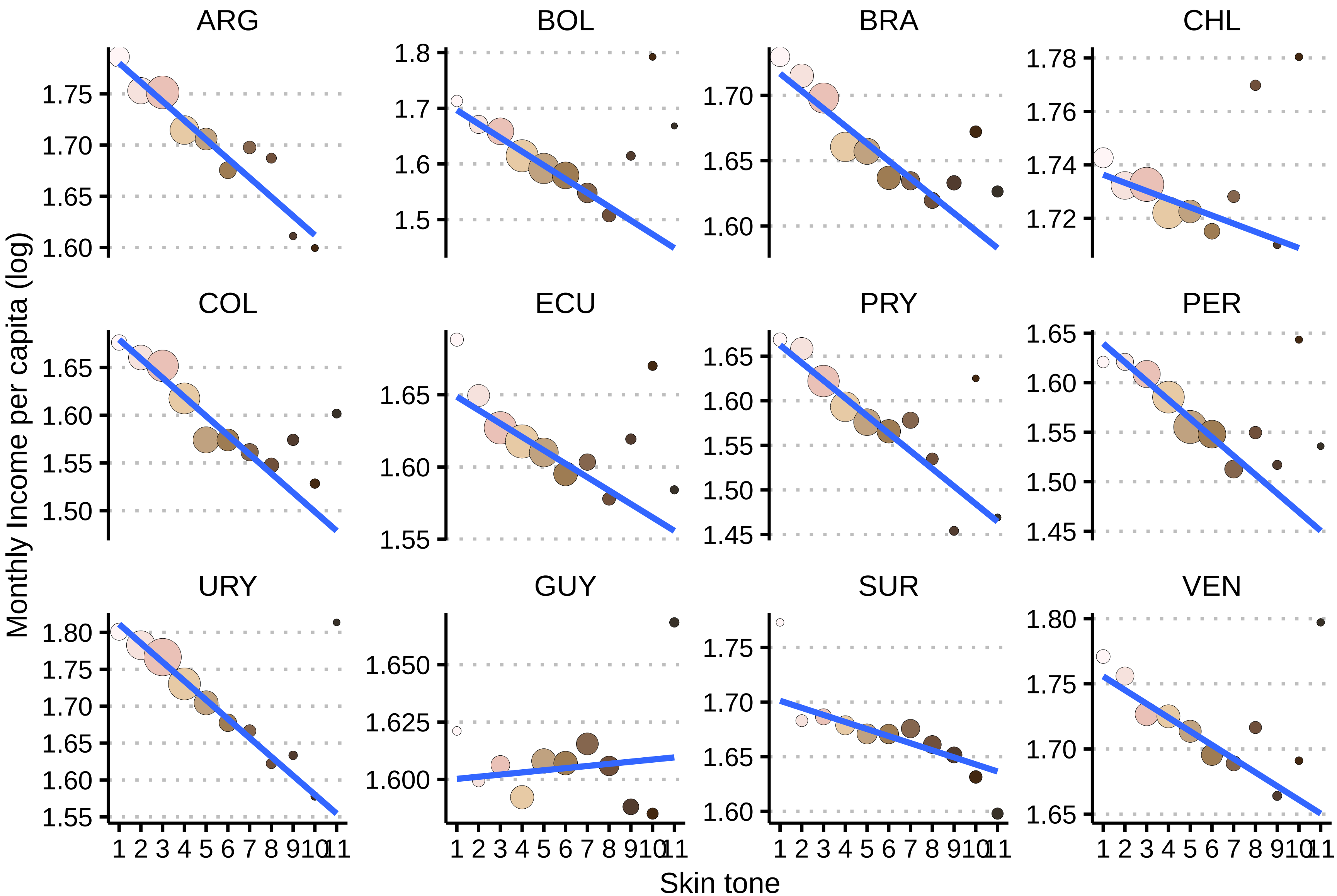

Figure 26: Skin tone gradient for mean monthly income per capita by country (2)

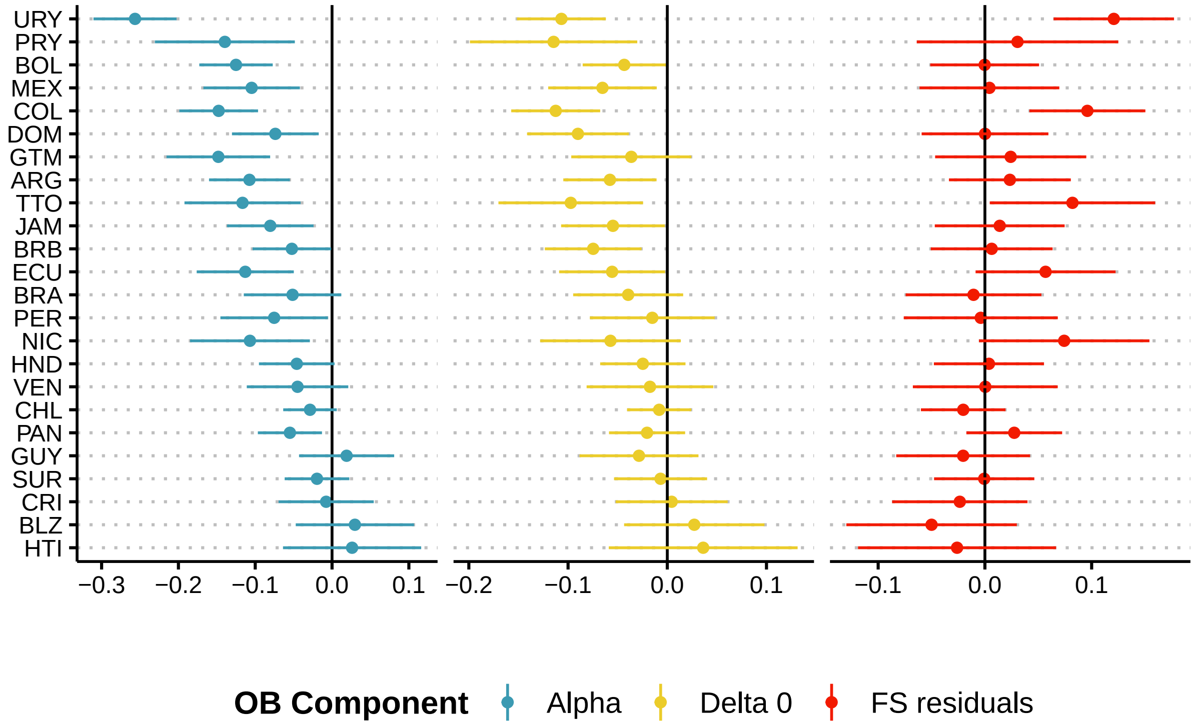

Figure 27: Skin tone effect on predicted income: OB decomposition by country (with control functions)

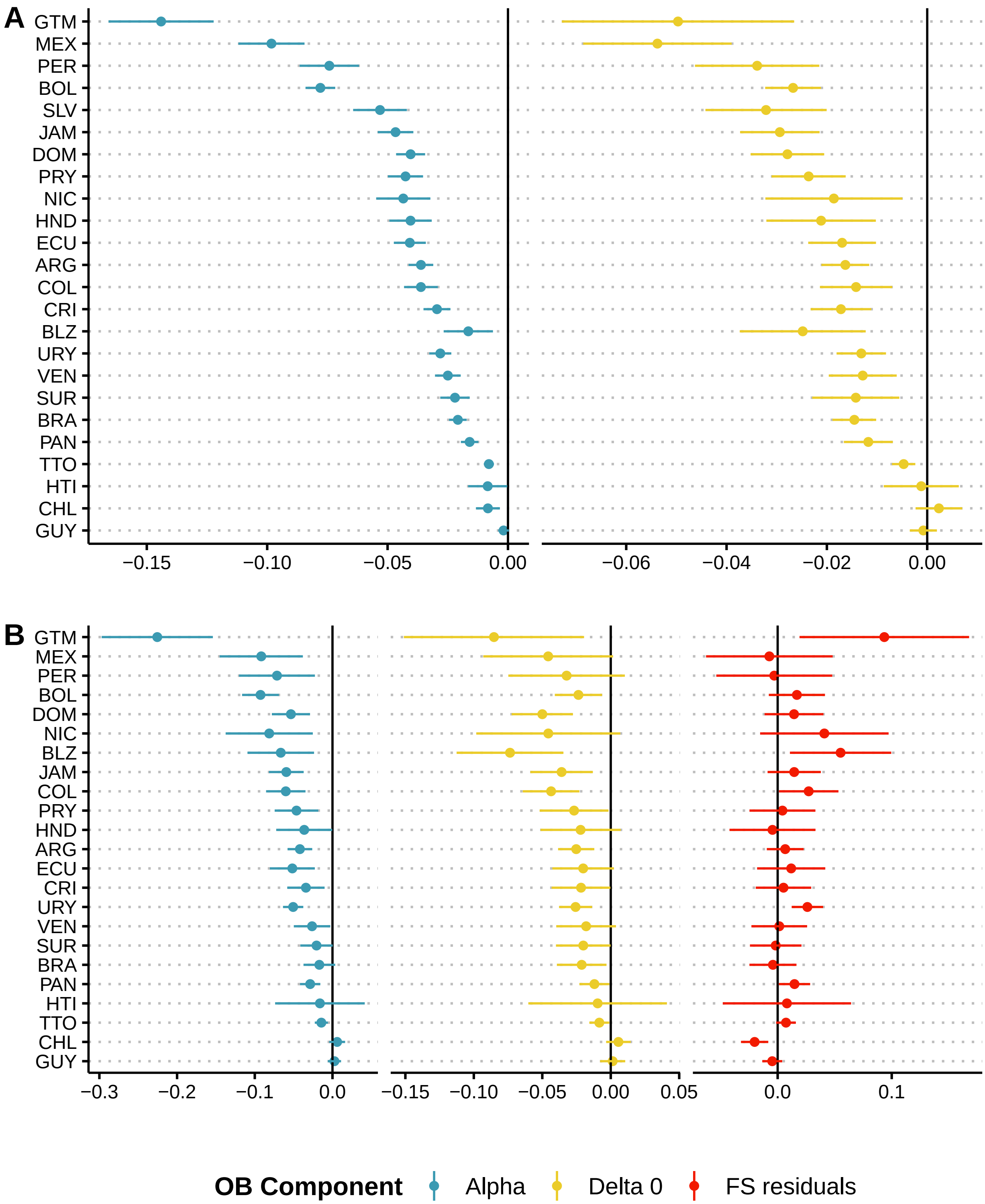

Figure 28: Skin tone effect on years of schooling: OB decomposition by country

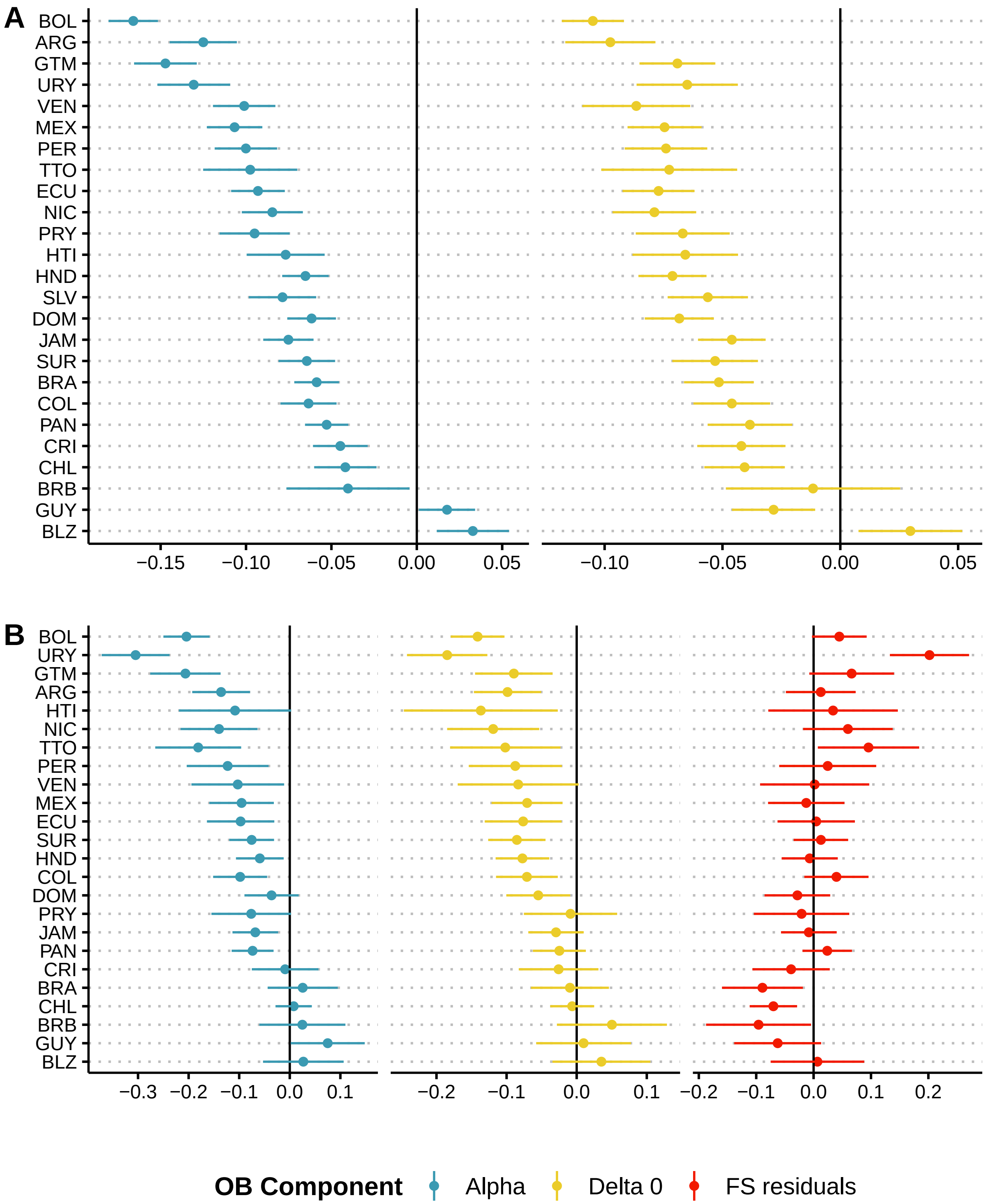

Figure 29: Skin tone effect on years of schooling: OB decomposition by country