World Inequality Lab – Working Paper N° 2022/03

Unveiling the Cosmic Race: Racial Inequalities in Latin America

L. Guillermo Woo-Mora

March 2022

WID.WORLD
THE SOURCE FOR GLOBAL INEQUALITY DATA

Unveiling the Cosmic Race: Racial Inequalities in Latin America*

L. Guillermo Woo-Mora†

For most recent version click here

March 14, 2022

Abstract

This paper uses skin tone and income information for over a hundred thousand individuals across 31 Latin American countries to study racial inequalities during the last decade. First, I estimate the welfare consequences of racial inequality. Subnational regions with higher income inequality between racial groups have worse economic development. Next, I provide evidence of a skin tone income premium. In an eleven-color palette, each darker shade in skin tone on average leads to a 3% decrease in income, with heterogeneity across countries. My analysis suggests racial discrimination is the main mechanism behind this income premium.

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

*I am especially grateful to Suanna Oh, Thierry Verdier, Thomas Piketty, Ignacio Flores, Ekaterina Zhuravskaya, Hillel Rapoport, Evan Rose, Eva Arceo-Gómez, Raymundo Campos-Vázquez, David Margolis, and Denis Cogneau. I also want to thank Mauricio Tenorio-Trillo, Sandra Ley, Edward Telles, Javier Castillo, Alice Krozer, María Montoya-Aguirre, Luis Baldomero-Quintana, Anna Barbeta Margarit, Julius Goedde, and Arely García. Thanks to all and participants at the Second World Inequality Conference, Casual Friday Development Seminar and Petit Seminaire Informel at PSE, and DEMOSAL Conference at El Colegio de México for their helpful comments. All remaining errors are my own.

†Paris School of Economics, France. guillermo.woo-mora@psemail.eu.

1 Introduction

Income inequalities are not colorblind. In Latin America, one of the most unequal regions globally (Chancel et al. 2021; De Rosa, Flores, and Morgan 2020), the mean monthly income of white individuals is at least twice higher than that of individuals with the darkest skin shades.[1] While the long-run co-evolution of culture, genes, and environment can explain differences in color pigmentation, it cannot explain contemporary disparities in economic outcomes across individuals of different skin shades (Henrich 2016; Jablonski 2021). However, color pigmentation influences racial identities which are socially constructed rather than biologically determined (Rose 2022), and differential attitudes towards racial identities can shape economic inequalities.[2]

Sociologists argue that *race* or *racialization* are physical characteristics or phenotypes that can define group membership, while *ethnicity* membership is based on cultural characteristics (Dixon and Telles 2017; Telles and Martínez Casas 2019). However, most race-related studies in economics combine both concepts and only rely on broad ethno-racial measures, as "White," "Black," "Asian," "Indigenous," or "Latino," or ethnic and linguistic groups. Such categories do not allow to isolate racial disparities due to physical characteristics from those due cultural characteristics. For example, when analyzing economic disparities and discrimination between "White" and "Latino" population: what extent is due to language? What extent is due to physical phenotype? Thus, isolating racial disparities due to phenotype from those due to other cultural or social characteristics is challenging.

Furthermore, the history and concept of race in Latin America makes it more difficult to study racial inequalities in the region. In the US *"race was clearly defined, categorical, and based primarily on descent"* (Dixon and Telles 2017). In contrast, most Latin American countries built their national identities through the 'melting pot' ethno-racial figure of *'mestizos'* or *'mulatos'*: the mixed-race descendant from European, Indigenous, and African population –also characterized as the *'Cosmic Race'* (Telles and Martínez Casas 2019). The ethno-racial category of *mestizo/mulato* is the most common one even though it identifies people with different racial phenotypes, from white skin shades to darker skin tones. However, Latin America's *"racialization has relied largely on phenotypic appearance and shades of skin color"* (Dixon and Telles 2017). Therefore, not isolating physical phenotype from other ethno-racial social identities, as the *mestizo/mulato* categories, might lead to underestimating or veiling racial inequalities in Latin America.

This paper studies racial disparities in income at the individual level and the welfare implications of racial inequalities for aggregate development in Latin America. I use a unique data set including information on skin tone, ethnicity, and income. Then, I can overcome the challenge of empirically isolating racial disparities due to differences in the phenotypic dimension from those due to other cultural characteristics. Exploiting the skin tone measures, this paper contributes with evidence for

[1] See Figure A.1.

[2] There is a growing literature in economics studying how social identities shape economic outcomes (Akerlof and Kranton 2000; Hoff and Stiglitz 2016; Oh 2022; Shayo 2020).

two overreaching questions.

First, I analyze the extent of racial disparities in Latin American countries and how they relate to disparities in development. I provide novel estimates of racial inequality –inequality between skin tone groups– at the subnational level for more than one hundred subnational regions. While racial inequality represents between 2% and 5% of total income inequality, it is far for being irrelevant for economic development. I combine the racial inequality measures with Gross Domestic Product (GDP) per capita at the regional level to compile a panel data of economic development and racial inequality for the last decade. Using cross-regional variation and controlling for time-invariant characteristics at the country level, results suggest that higher income inequality inequality between skin tone shades correlates with a lower GDP per capita. The results are robust to alternative measures of economic development, as well as to unobserved characteristics that could bias the estimates (Altonji, Elder, and Taber 2005; Oster 2019).

Could racial inequality be, in some part, directly driven by the effect of skin tone rather than through differences in educational attainment or labor occupation? In the second part of the paper, I provide evidence of a skin tone premium in Latin America. Using different specifications, including a research design that purges unobserved spatial heterogeneity –Spatial First Differences (SFD) (Druckenmiller and Hsiang 2018)–, I present evidence of an unambiguously negative effect of darker skin tone on income. On average, each darker shade in skin tone leads to 3% less income, out of a color palette with eleven tones. The results hold using a different survey for Mexico with information on skin tone, income, and the respondent's approximate geographic location. Furthermore, to test whether the gap is driven by racial discrimination, I combine SFD with an Oaxaca-Blinder decomposition for continuous variables (Ñopo 2008) and provide estimates showing that at least 80% of the racial gap can be attributed to racial discrimination, with substantial country-specific heterogeneity. Consistent with the race discrimination hypothesis, I use LAPOP data to show that individuals with darker skin tones report higher discrimination against them.

By isolating racial inequalities due to phenotype from those due to other cultural characteristics, both at the individual and at aggregate levels, this paper contributes to two strains of literature. First, the paper contributes to the growing literature studying the effects of group-based inequalities on economic development.[3] Group-based inequalities can lead to political inequality, discriminatory policies between groups, inadequate public goods provision, talent misallocation, and technological innovation losses (Alesina, Michalopoulos, and Papaioannou 2016; Cook 2014; Hsieh et al. 2019). Moreover, group-based inequalities can have persistent effects through intergenerational transmission of cultural traits, occupational segregation, spatial segregation, or historical resentment (Acharya, Blackwell, and Sen 2016; Alesina and Zhuravskaya 2011; Bisin and Verdier 2011; Bezin and Moizeau 2017; Bowles, Loury, and Sethi 2014). Previous studies of ethno-racial inequality and comparative development have relied on broad ethno-racial categories or ethnic and linguistic

[3]Closely related to the literature on the relation between culture and economic development from a historical perspective (Alesina and Giuliano 2015; Nunn 2012, 2021).

groups. I contribute to the literature by providing evidence on how income inequalities between skin tone groups systematically correlate with lower economic development.

The paper also contributes to the extensive literature studying racial disparities in economic outcomes and racial gaps in the labor market. Race-related studies have mainly focused on the United States due to the historical experience of slavery, segregation, discrimination, and the substantial racial disparities in several outcomes (Arnold, Dobbie, and Hull 2021; Brouillette, Jones, and Klenow 2021; Chetty et al. 2020; Cook and Logan 2020; Derenoncourt and Montialoux 2020; Derenoncourt 2022; Kline, Rose, and Walters 2021; Lang and Spitzer 2020). Nevertheless, racial inequalities are also salient in other regions with shared historical experiences of slavery and segregation (Baldomero-Quintana, De la Rosa-Ramos, and Woo-Mora 2022; Fujiwara, Laudares, and Valencia Caicedo 2021). Scholars have compiled extensive evidence for Latin America on how ethno-racial identities determine disparities in the labor market and wages (Arceo-Gomez and Campos-Vázquez 2014; Arceo-Gómez and Campos-Vázquez 2019; Card et al. 2018; Derenoncourt et al. 2021; Ñopo 2012), educational attainment (Botelho, Madeira, and Rangel 2015), 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 access to the financial sector (Hernández-Trillo and Martínez-Gutiérrez 2021).[4] Nonetheless, most of the previous literature focuses on a single country, with little comparability on the measures of ethnicity and race. I contribute to the literature by using a sizable sample of individuals across multiple countries and the most solid methodology available for cross-sectional data, providing further evidence that skin tone affects individual outcomes through racial discrimination.[5]

The rest of the paper is structured as follows. Section 2 presents a brief historical overview of the racial question in Latin America. 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 skin tone premium estimates. Section 6 concludes.

2 Historical background: The racial question in Latin America

Race and ethnicity are central elements in Latin America's history (Tenorio-Trillo 2017). The encounter of the original Indigenous population, European conquerors, and African populations, mostly brought as slaves for forced labor, produced an early miscegenation process throughout the

[4]Racial components also determine social norms and identities (Campos-Vázquez and Medina-Cortina 2017); electoral preferences (Campos-Vázquez and Rivas-Herrera 2021) and increase the perceptions of discrimination (Chong and Ñopo 2008; Ñopo, Chong, and Moro 2009; Trejo and Altamirano 2016).

[5]To my best knowledge, there are few studies on ethno-racial disparities with comparable data and methodology for more than one Latin American country. Ñopo (2012) studies the wage gap for Bolivia, Peru, Brazil, Guatemala, Paraguay, Chile, and Ecuador, but for broad ethno-racial groups. See Castillo J. (2022) for a thorough review from sociology on ethno-racial social stratification.

continent. The result was a complex administrative and social caste system dependent on race and ethnicity (Graham 2013). The caste system was more *flexible* than the racial hierarchies in the US –for example, one-drop rules were never consolidated–, but, overall, whiter meant closer to European and thus implied a higher social status (Dixon and Telles 2017). In the XIXth century, ethno-racial classification disappeared with the Independence processes, but the inequalities from the colonial period persisted (Loveman 2014). Moreover, racial theories justified such disparities and the *domination over "colored" populations* (Graham 1990).

Nevertheless, the narrative changed in the XXth century. In a context of salient racial segregation and anti-miscegenation in the US and racial hate by the Nazi ideology, many Latin American intellectuals promoted the formation of national identities by reinforcing the 'melting pot' ethnic identity of *mestizos* or *mulatos*: the racial mixture of Indigenous, African and European population (Dixon and Telles 2017; Martínez Casas et al. 2019). The philosopher José Vasconcelos baptized the mestizo racial mixture in the Americas as the *"Cosmic Race"*. Published in 1925 with the title *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). Given the high degree of miscegenation and racial diversity in the Americas, mestizaje identity served as the alternative to racial politics. Nowadays, most Latin Americans define themselves as *mestizos* or *mulatos*, even when there is substantial racial diversity within such ethnic categories. Thus, Latin American countries could veil racial disparities and inequalities given that everyone became *mestizo* or *mulato*.

Centuries later, the colonial and post-colonial patterns persist: whiter people are better off in many socio-economic dimensions. Whiteness is still regarded as an ideal aesthetic of beauty and wealth (Krozer and Urrutia Gómez 2021), and different forms of racism are still present in everyday life.[6] In the extreme, some countries are characterized as *Pigmentocracies* (Telles and Martínez Casas 2019). Therefore, the use of color palettes represents an improvement in studying racial disparities in Latin America given the extended *mestizo/mulato* identity. The 'colorism' research agenda might significantly benefit social science research (Dixon and Telles 2017). More importantly, learning from the Latin American experience can shed light on tackling racial inequalities in other latitudes.

3 Data

Measuring racial inequalities has been a challenging task due to data restrictions.[7] I exploit a rich data set that disentangles the cultural and phenotype dimensions: the Latin American Public Opinion Project's (LAPOP) AmericasBarometer survey. The AmericasBarometer is a survey conducted every two years in most countries in the Americas with stratified nationally representative

[6]See Dixon and Telles (2017) for a review of the skin-bleaching industry.

[7]In some extreme cases, there are no available measures of racial categories (The Economist 2020).

samples of voting-age adults, using a common questionnaire score and country-specific modules.[8] Since 2010, LAPOP has used the Project on Ethnicity and Race in Latin America (PERLA) palette developed by Telles and Martínez Casas (2019) and coauthors to measure skin tone.[9] The scale ranges from 1 to 11, where one is the lightest skin tone and eleven is the darkest.[10] Figure 1 panel (a) shows the PERLA color scale. Using LAPOP data, I compile all the AmericasBarometer surveys that include the PERLA palette in the core questionnaire. The sample includes more than 100 thousand individual observations from 31 countries across four waves (2012, 2014, 2016/2017, 2018/2019). Table B.1 shows the sample size by country and wave.

Racial characteristics measured by skin tone vary substantially across and within countries. Figure A.3 shows the distribution of skin tones by country. For instance, the darkest-skin tones are a majority in the Caribbean. Nevertheless, in every country there is a set of people with the darkest skin tones. 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 common in countries with little miscegenation that experienced heavy European migration during the XIXth and XXth century, like Argentina, Chile, and Uruguay. For the econometric analysis, I use a modified PERLA color scale, or *collapsed* PERLA color scale, where I top- and bottom-code skin tone for each country.[11]

Since ethnicity refers to cultural characteristics, it might differ from the racial phenotype. LAPOP works with six broad ethno-racial categories: Afro, Indigenous, Mestiza, Mulata, White, and other ethnic groups (i.e., Asian, Jew, among others). Figure A.4 shows the ethno-racial distribution for each country. The majority in Caribbean countries define themselves as Afro origin. People who define themselves as White are the majority in Argentina, Chile, Costa Rica, and Uruguay. Interesting patterns of ethnicity and racialization arise when analyzing countries with a high percentage of white and medium-dark-skinned populations. For instance, besides countries where there is an Afro or White majority, most of the population defines themselves as Mestiza or Mulata. Consistent with the historical background, such countries also happen to have a high historical miscegenation experience and a strong presence of *mestizaje* ideology.

Figure 1 panel (b) shows the distribution of skin tones by each of the ethno-racial categories. 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

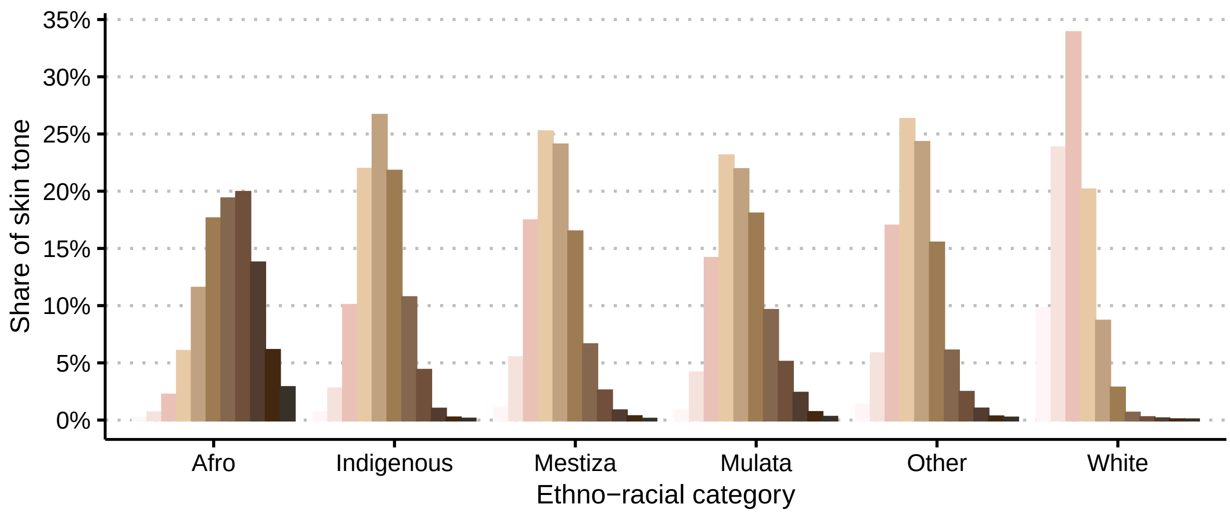
[8] See https://www.vanderbilt.edu/lapop/.

[9] See https://perla.soc.ucsb.edu/.

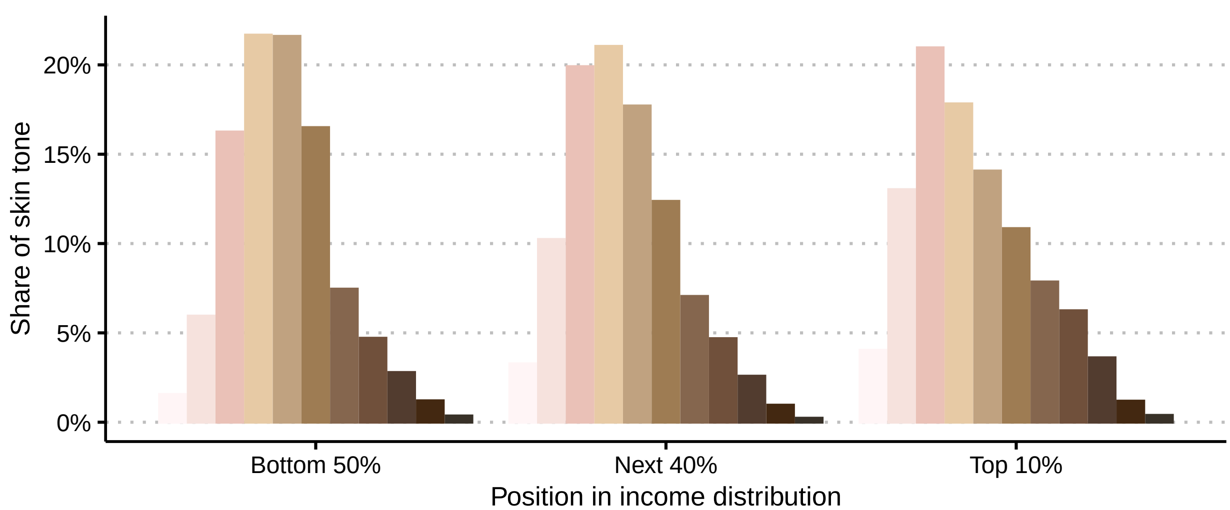
[10] In practice, interviewers are asked to discretely annotate the respondent's skin color taking as reference the PERLA palette without showing the guides to respondents (Dixon and Telles 2017). Cernat, Sakshaug, and Castillo (2019) present evidence of measurement error concerns on educational attainment. In the sensitivity checks I control for the interviewer's skin tone and include interviewer fixed effects.

[11] As Figure A.2 shows, either the whitest and the darkest PERLA colors are outliers dependent on the country. I top-code (bottom-code) skin tone replacing all values greater (smaller) than the upper (lower) limit of each country skin tone interquartile range (percentiles 25 to 75). Results are robust to both the standard PERLA color scale and the *collapsed* PERLA color scale.

(a) PERLA color palette



(b) Ethno-racial categories and skin tone



(c) Income distribution and skin tone

Figure 1: Descriptive statistics

Notes: Panel (a) shows PERLA color palette based on Telles and Martínez-Casas (2019). Panel (b) shows the share of each skin tone for each of the broad ethno-racial categories used in LAPOP's AmericasBarometer survey. Panel (c) shows the skin tone distribution for each of the broad income distribution categories constructed with LAPOP's AmercasBarometer income information.

Indigenous, Mestiza, Mulata, or other ethnic groups have mostly medium-dark skin tones, but there is considerable variation in skin tone distribution. Even when skin tones are broadly correlated to the distribution of self-reported ethnicity, there is a large diversity of racial phenotype within ethno-racial groups. Thus, using the information on skin tone has significant advantages over using self-reported ethnicity.[12]

Since the *mestizaje* ideology is strong in Latin American countries, ethno-racial identities hide racial disparities. Namely, analyzing economic disparities between ethno-racial groups might depict a general overview of the inequalities, but it might veil the disparities within broadly defined ethnic groups such as the Mestiza and Mulata populations. Figure 1 panel (c) shows the distribution of skin tones by percentiles of income. The bottom 50% of the income distribution tends to have medium- and dark skin tones. The next 40% percent have a higher share of light skin tones, and the top 10% of the income distribution has a majority of white skin tones. Therefore, the PERLA palette and the LAPOP data present an important advantage to deepen the study of racial inequalities in the region.

LAPOP data 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. The survey asks about self-reported monthly household income by brackets. 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. I use World Bank's Purchase Parity Power 2019 rates to convert local currencies. Table B.2 shows the sample descriptive statistics.

4 Racial Inequality and Economic Development

Previous studies argue that two elements shaping comparative development are ethnic diversity (Alesina and La Ferrara 2005) and income inequality (Easterly 2007; Persson and Tabellini 1994). Alesina, Michalopoulos, and Papaioannou (2016) put forward an alternative hypothesis: the economic differences between ethnic groups coexisting in the same country shapes comparative development. Using cross-country regressions, Alesina, Michalopoulos, and Papaioannou (2016) show that economic inequalities between ethnic groups are correlated with lower economic development at the national level. While the previous correlation holds globally, the authors find it is not statistically significant in Western Europe and the Americas.

A plausible hypothesis is that ethnic inequality is not as salient as racial inequality in the West. For instance, given the blurry border between ethnicity and race resulting from the *mestizaje* ideology,

[12]Skin tone is not the only relevant dimension for Latin America's racialization. Other elements as hair type, eye color, even stature, affect the conception of race. Nevertheless, skin color is the most salient dimension (Telles and Paschel 2014).

race differences could matter more than ethnic ones in Latin America. This section exploits LAPOP data to compute aggregated measures of racial inequality, or inequality between racial groups, at the subnational level. I push forward Alesina, Michalopoulos, and Papaioannou (2016) argument to test whether racial inequalities correlate with lower economic development.

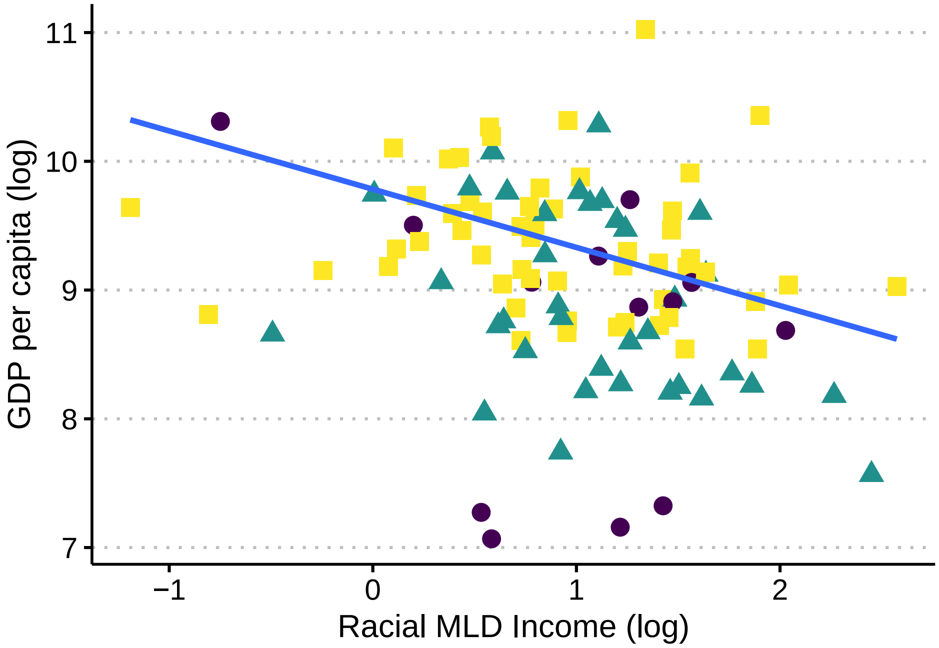
Given that cross-country analysis could mask within-country regional heterogeneity, I exploit LAPOP data at the subnational level. LAPOP data is representative at the subnational level. Each country is divided into three to eight regions, where some small countries are not stratified. First, I compute income Gini indexes for each country-region and year. Furthermore, LAPOP data allows decomposing income inequality measures by skin tone groups, as Figure 1 panel (c) show. Using the skin tone categories, I use the mean log deviation (MLD) index (Foster and Shneyerov 2000) to obtain between and within components for income and educational inequality. Besides racial inequality measures, I also use broad ethno-racial categories to compute alternative ethnic inequality measures. To obtain subnational measures of economic development, I use Kummu, Taka, and Guillaume (2018) gridded subnational data on Gross Domestic Product and Human Development Index.[13] To account for differences in population between regions, I use population grids (NASA Socioeconomic Data and Applications Center 2018).

Figure A.5 panel (a) shows the mean skin tone by each of LAPOP's stratified sampling regions. There is substantial racial diversity between country regions. Figure A.5 panel (b) shows GDP per capita (log) for each country-region. Figure A.5 panel (c) shows the mean Gini index for each region in the analysis. Figure's A.5 panel (d) shows the mean Income MLD between-group racial component. One improvement of this analysis with respect to Alesina, Michalopoulos, and Papaioannou (2016) is the availability of multiple observations for each subnational region. Figures A.6 and A.7 visualizes the data's panel structure. Most variation is between units rather than within-unit through time.[14] The ratio of the income MLD between-group racial component over the total MLD income inequality index has a mean of 3.9%, with an interquantile range between 1.8% and 4.9%. Thus, racial inequality explains relatively little income inequality. However, it has relevant consequences for economic development.

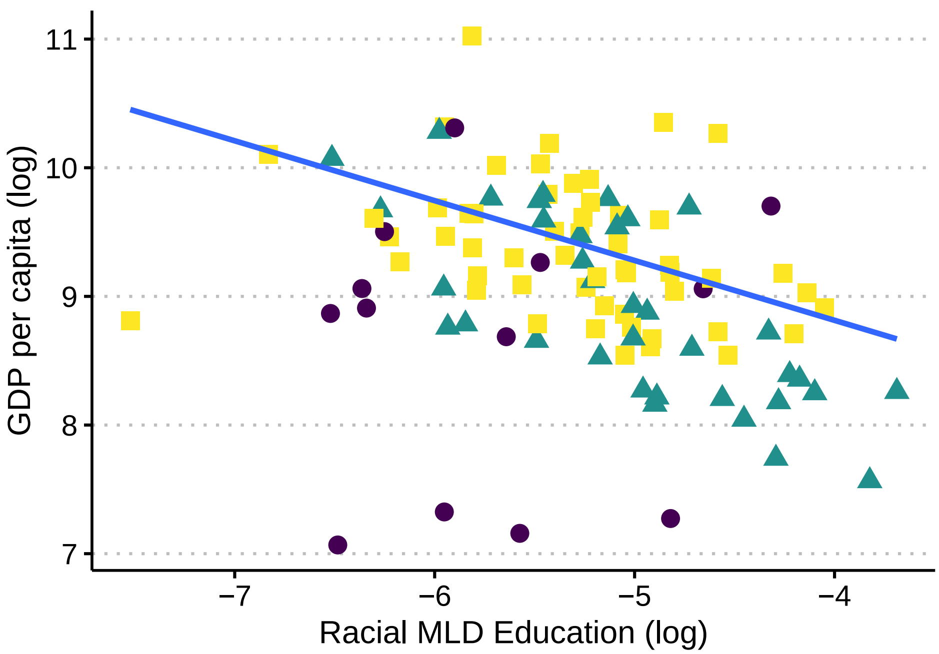
Figure 2 panel (a) shows the unconditional correlation income MLD between racial group component and mean (log) GDP per capita. Panel (b) shows the same correlation but uses years of schooling MLD between racial group component and mean (log). Consistent with the ethnic-inequality argument, there is a negative correlation between racial inequality and lower economic development at the subnational level. To test more robustly the relation between racial inequality and economic

[13]For robustness, I also use nightlight yearly data (Elvidge et al. 2021) to proxy for regional economic development.

[14]Racial inequality measures are interpolated. Given LAPOP surveys are not available every year, I interpolate with linear trends the missing values for each region. Figure A.6 shows the actual values versus the interpolation ones, Moreover, economic development data has also interpolations since Kummu, Taka, and Guillaume (2018) data is available from 1990 to 2015. In order to obtain measures for LAPOP observations between 2016 and 2019, I use linear extrapolations by subnational regions. Results are robust to the subsample without extrapolations.



(a) Income racial inequalities



(b) Educational racial inequalities

(c)

Figure 2: Racial-inequality and Economic Development

Notes: Unit of analysis is country-region. Pooled regressions weighted by population. Panel (a) shows the unconditional correlation between the income MLD between skin tone groups component (log) and GDP per capita (log). Panel (a) shows the unconditional correlation between the years of schooling MLD between skin tone groups component (log) and GDP per capita (log). GDP per capita (log) measures using Kummu, Taka, and Guillaume (2018) data grids. Racial inequality indexes constructed using LAPOP's AmericasBarometer income, skin tone, and years of schooling measures.

development I use the following specification:

$$y_{rct} = \beta Racial\ Inequality_{rct} + \gamma X_{rc} + \theta_{ct} + \varepsilon_{rct} \tag{1}$$

Where y_{crt} represents GDP per capita (log) of subnational region r in country c at time t; $Racial\ Inequality_{rct}$ is the MLD between racial groups component (log); X_{rc} represents a set of time-invariant controls at the country-region level (geographical variables and economic activity proxies). I include country times year fixed effects θ_{ct} to control for country time-invariant characteristics and common shocks at the national level. Then, the remaining variation on economic development and racial inequality is within a country for a given year. I weight all specifications by country-region population and cluster the standard errors at the country-region level.

Table 1 shows the results of the specifications in Equation (1). Column 1 shows that accounting for country fixed effects interacted with year fixed effects, an increase in 1.0% on racial income inequality correlates with a decrease in regional GDP per capita of almost 25.0%. Column 2 includes inequality in years of schooling between racial groups, Column 3 includes total income inequality and income inequality measures between broad ethno-racial groups. Column 4 includes racial and ethnic fractionalization measures, the share of each broad ethno-racial category, the number of ethnic groups by country region, and the median skin tone. Columns 5 and 6 include geographic and economic activity controls at the region level, respectively. In all specifications, even accounting for human and physical capital and geographic characteristics at the regional level, the negative elasticity between racial inequality and GDP per capita remains statistically significant. The most robust specification suggest a 1.0% increase in racial inequality correlates with a decrease of 9.0% on regional GDP per capita. In contrast, racial inequalities in education have not significant correlation with economic development. To provide evidence that the correlation is not driven by any outlying observation, I run the most robust specification –Column 6– multiple times dropping an observation of the sample each time. The jackknife elasticity of income MLD between racial groups and GDP per capita is 9 (p-value = 0.000).

Table B.3 Columns 3 and 4 replicate the specification at Table 1 Column 6, but restricting the sample to observations without extrapolation of racial inequality measures and GDP per capita for years after 2015. The coefficient remains negative and statistically significant. Thus, the results are not driven by the extrapolation of observations. Table B.3 use HDI and total nightlights as an alternative measures of economic development at the subnational region level. In both specifications, the racial inequality estimate is negative and statistically significant.

The previous results might be biased due to the omission of within-regions characteristics. To test the robustness of the results, I estimate the adjusted coefficient using the unobservable selection and coefficient stability approach (Altonji, Elder, and Taber 2005; Oster 2019). Assuming a maximum R-squared of one, the mean adjusted elasticity for the income MLD between racial groups component is 6.8 (p-value = 0.000). Then, the correlation between racial inequality and

Table 1: Racial Inequality and Economic Development

	GDP per capita (log)					
	(1)	(2)	(3)	(4)	(5)	(6)
Racial Income MLD (log)	-0.254**	-0.185**	-0.240**	-0.122*	-0.154***	-0.090**
	(0.110)	(0.076)	(0.101)	(0.067)	(0.055)	(0.035)
Racial Education MLD (log)		-0.262*	-0.271*	-0.161**	-0.123**	-0.030
		(0.140)	(0.143)	(0.077)	(0.050)	(0.034)
Income MLD (log)			0.040	-0.267	-0.226	-0.066
			(0.276)	(0.272)	(0.170)	(0.138)
Ethno-racial Income MLD (log)			0.245	0.178	0.226***	0.137***
			(0.227)	(0.125)	(0.071)	(0.045)
Racial Fractionalization index (log)				1.774	0.491	0.101
				(1.389)	(0.652)	(0.481)
Ethno-racial Fractionalization index (log)				0.360	-0.232	-0.478***
				(0.309)	(0.167)	(0.113)
Mestizo share (log)				-0.346	0.262	0.525***
				(0.361)	(0.244)	(0.164)
Indigenous share (log)				-1.392	0.328	0.265
				(1.144)	(0.970)	(0.645)
Afro share (log)				-2.833**	0.263	0.843
				(1.426)	(0.926)	(0.650)
Other ethnic groups share (log)				-2.833**	0.263	0.843
				(1.426)	(0.926)	(0.650)
No. GREG Groups				-0.005*	-0.002	-0.001
				(0.003)	(0.003)	(0.002)
Mean skin tone				0.037	-0.003	-0.124
				(0.144)	(0.138)	(0.096)
Population density (log)					0.102***	0.033
					(0.037)	(0.115)
Years of schooling (mean)						0.087
						(0.054)
No. Obs.	824	824	824	824	824	824
No. Country-regions	103	103	103	103	103	103
No. Years	8	8	8	8	8	8
R2	0.568	0.604	0.620	0.719	0.820	0.884
R2 Within	0.099	0.174	0.208	0.415	0.625	0.759
FE: Country $\times$ Year	X	X	X	X	X	X
Geographic controls					X	X
Economic controls						X

Notes: Unit of analysis is country-region. Standard errors clustering by country-region in parenthesis. All regressions are weighted by population. Geographic controls include area, longitude, latitude, altitude, ruggedness, mean temperature, mean precipitation, and mean solar radiation. Economic controls, an indicator whether the country-region host the country's capital, an indicator whether the region has coastline, an indicator whether the region shares an international border, total roads' length, the number of airports, and mean total nightlights (log). *p<0.1; **p<0.05; ***p<0.01

economic development could be mildly biased by unobserved characteristics, but remains negative and statistically significant. Even when the negative correlation is robust to omitted variables, there might be reverse causation.

While I cannot prove a casual link, the negative correlation is systematic to different specifications and robustness checks. This paper contributes by providing novel evidence that inequality between groups of different skin tone shades correlates with lower economic development at the subnational level. Thus, consistent with the group-based inequalities hypothesis literature, racial inequalities hinder economic development in Latin America.

5 Skin tone premium

While the previous section provides evidence of disparities between racial groups, it is unclear what shapes those disparities. For instance, do differences in education or labor stratification explain racial disparities in income? This section retakes arguments from sociology, and test if differences in income are directly linked to skin tone. Specifically, I use the LAPOP data at the individual level to analyze the skin tone premium on income following the disparities approach (Fleurbaey and Schokkaert 2011). To estimate the skin tone premium, I use the following specification:

$$y_i = \beta_1 t_i + \beta_2 t_i^2 + x_i \alpha + \theta + \varepsilon_i \tag{2}$$

Where y_i is the (log) monthly income of individual i; t_i is the continuous *collapsed* PERLA measure of skin tone; t_i^2 is the squared measure of skin tone; x_i is a vector of observable characteristics, θ is intra-municipality unit interacted with year fixed effects.[15] I use LAPOP sample weights to make the results representative at the regional level and comparable across countries and waves (Castorena 2021). I cluster the standard errors at the intra-municipality unit and year.

Table 2 panel A shows the OLS estimates.[16] The first specification demeans income and skin tone between intra-municipal units for each country and year, and then compares observations within such geographical units. Therefore, even after accounting for differences in income and skin tone within a municipality, and partially accounting for geographic selection, an increase in one skin tone correlates with a decrease in income of 10.0%, statistically significant at 1%. After accounting for gender, years of schooling, age, self-reported ethnicity, occupational status, marital status, locality size, religion, and a measure of interpersonal trust, as well as the geographic heterogeneity, Column

[15] The intra-municipality unit is a geographic unit within a municipality, thus is the smallest geographic variable.

[16] Table B.4 shows how the estimate of skin tone on income changes after accounting for non-linearities and between-geographic heterogeneity. As the R-squared statistics in Table B.4 show, a substantial part of the variation can be explained by geographical heterogeneity. Panel A uses the original PERLA skin tone palette, while Panel B uses the *collapsed* PERLA palette described in section 3. The collapsed PERLA palette has slightly smaller coefficients but is not statistically different from those using the original PERLA palette. For the remaining analysis, I only use the collapsed PERLA palette for the analysis.

Table 2: Skin tone premium on income

	Income (log)					
	(1)	(2)	(3)	(4)	(5)	(6)
Panel A: OLS						
PERLA scale	-0.100***	-0.109***	-0.085***	-0.074***	-0.076***	-0.076***
	(0.008)	(0.008)	(0.008)	(0.008)	(0.008)	(0.008)
PERLA scale (squared)	0.006***	0.006***	0.005***	0.004***	0.005***	0.005***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
1[Female]		-0.288***	-0.276***	-0.265***	-0.264***	-0.185***
		(0.006)	(0.006)	(0.006)	(0.006)	(0.006)
Years of Schooling			0.053***	0.068***	0.067***	0.062***
			(0.001)	(0.001)	(0.001)	(0.001)
Age				0.012***	0.009***	0.005***
				(0.001)	(0.000)	(0.001)
1[Afro]					-0.038**	-0.037**
					(0.015)	(0.015)
1[Indigenous]					-0.117***	-0.112***
					(0.014)	(0.014)
1[Mulata]					-0.020	-0.018
					(0.016)	(0.016)
1[Other]					-0.101***	-0.096***
					(0.014)	(0.014)
1[White]					-0.034***	-0.034***
					(0.009)	(0.009)
No. Obs.	102,092	102,092	102,092	102,092	102,092	
R2	0.554	0.569	0.593	0.605	0.605	0.620
R2 Within	0.004	0.036	0.091	0.117	0.119	0.150
Panel B: SFD						
PERLA scale	-0.033***	-0.043***	-0.030***	-0.028***	-0.028***	-0.028***
	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)
PERLA scale (squared)	-0.001	-0.001	-0.001	-0.001	-0.001	-0.001
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Female		-0.289***	-0.280***	-0.265***	-0.264***	-0.189***
		(0.008)	(0.007)	(0.007)	(0.007)	(0.008)
Years of Schooling			0.054***	0.069***	0.068***	0.063***
			(0.001)	(0.001)	(0.001)	(0.001)
Age				0.009***	0.011***	0.005***
				(0.000)	(0.001)	(0.001)
1[Afro]					-0.034*	-0.035*
					(0.019)	(0.019)
1[Indigenous]					-0.105***	-0.102***
					(0.018)	(0.018)
1[Mulata]					-0.022	-0.021
					(0.020)	(0.020)
1[Other]					-0.092***	-0.091***
					(0.018)	(0.018)
1[White]					-0.026**	-0.026**
					(0.011)	(0.011)
No. Obs.	76,747	76,747	76,747	76,747	76,747	76,747
R2	0.105	0.134	0.185	0.210	0.211	0.238
R2 Within	0.003	0.035	0.092	0.120	0.121	0.151
Specification						
FE: Intra-Mun. × Year	X	X	X	X	X	X
Socio-demographic Controls						X

Notes: Unit of analysis is the individual. Standard errors clustering by intra-municipality unit and year in parenthesis. Sociodemographic controls include: age squared, marital status, occupation status –working, taking care of the home, actively looking for a job, not working and not looking for a job, not working but have a job, studying, retired–, locality size –metro area, big city, medium city, small city, or rural area–, religion, and interpersonal trust. The ethnicity group of reference is Mestizo category. *p<0.1; **p<0.05; ***p<0.01

6 in Table 2 shows that an increase on one skin tone in the collapsed PERLA scale correlates with a decrease of 7.7% in income, statistically significant at 1%.

5.1 Purging unobserved heterogeneity

The previous estimates can be biased due to unobserved characteristics that affect income and correlate with skin tone –labor occupation, parental background, early childhood, social networks. I use the Spatial First Difference (Druckenmiller and Hsiang 2018) research design to purge common unobserved heterogeneity and obtain an educated estimate of skin tone on income. Druckenmiller and Hsiang (2018) show that if units are *densely packed across physical space*, omitted variables bias from common unobserved factors can be purged using a differencing approach where the spatial position of observations can be located and organized. Thus, exploiting the spatial dimension researchers can use first-differencing between adjacent units, such as it is commonly used in time-series contexts. See a review of the research design in Appendix C.

Given LAPOP sample methodology and survey design, units with adjacent identification number in the survey are units adjacent in space.[17] Thus, I can purge common unobserved factors between neighboring observations using SFD. For each intra-municipal unit and year, I arrange the observations in the survey by their assigned identification number. Afterwards, I use first-differencing and estimate the following specification:

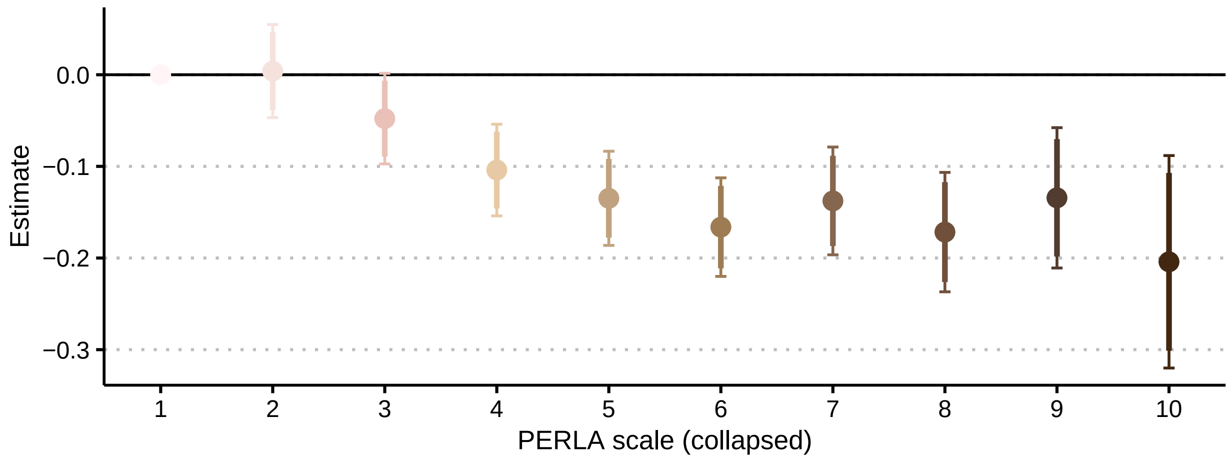
$$\Delta y_i = \beta_1 \Delta t_i + \beta_2 \Delta t_i^2 + \Delta x_i \alpha + \theta + \varepsilon_i \tag{3}$$

Where Δy_i is the spatial first difference on the outcome of interest, in this case, (log) monthly household income; Δt_i is the spatial first difference on skin tone measured by the *collapsed* PERLA palette; Δt_i^2 the spatial first difference on squared skin tone; and Δx_i is a vector of spatial first differences on socio-demographic covariates. Since the data is from samples for different countries and waves, all regressions include intra-municipal unit interacted with year fixed effects, θ. Note that the SFD specification compares adjacent units within an intra-municipal unit. Then, the intra-municipal unit times year fixed-effect controls for between-geographical unit differences, and SFD purges common unobserved heterogeneity within an intra-geographical unit. I cluster the standard errors at the intra-municipal unit and year level.

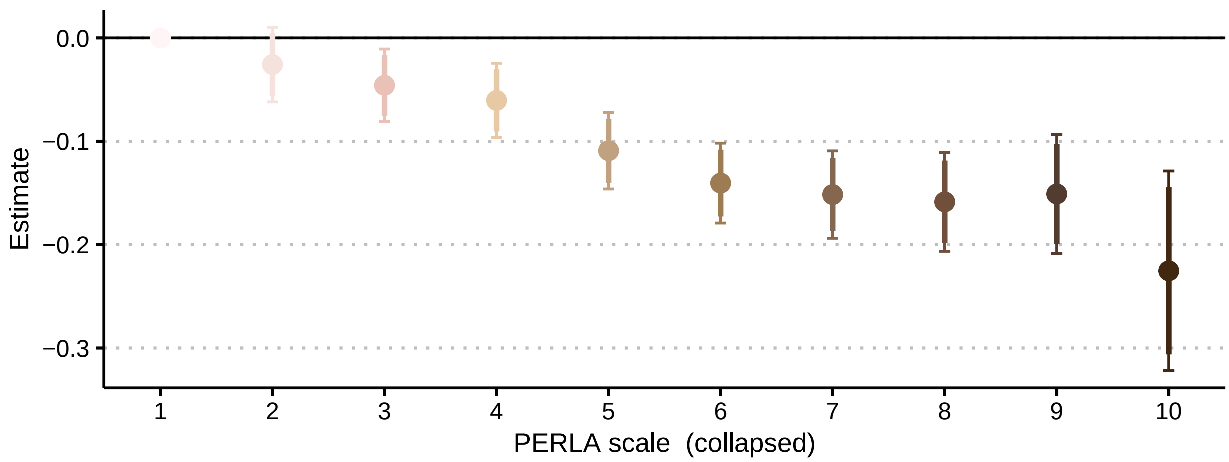
Table 2 Panel B shows the coefficients for the SFD specification in equation (3).[18] The SFD coefficients show that OLS estimates are biased from zero in absolute terms. The most robust specification in Column 6 shows that an increase on one darker skin tone correlates with a decrease of 2.8%, statistically significant at 1%. Thus, the OLS estimate is almost 3 times higher than the SFD estimate. SFD estimates also imply that the skin tone premium represents 44.0% of the

[17] See Figure C.13 and https://www.vanderbilt.edu/lapop/insights/IMN004en.pdf.

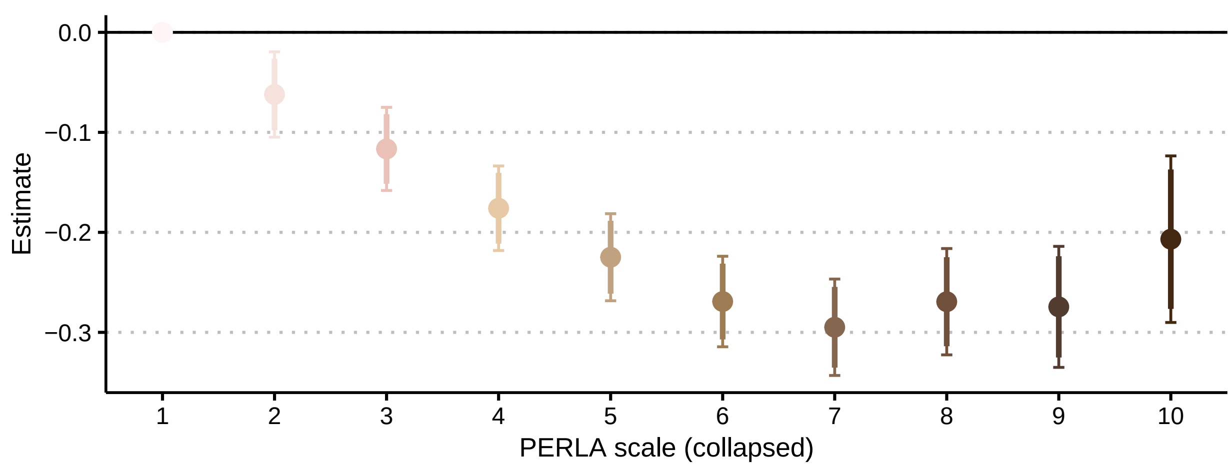
[18] Note that the number of observations falls drastically for the SFD estimates since the first-differencing operation of adjacent units by each cluster-municipality and year drops the first unit.



(a) Monthly Income (log)



(b) Household asset index (z-score)



(c) Years of schooling (z-score)

Figure 3: SFD estimates by skin tone

Notes: Coefficients using specification 3 and dummy variables for each skin tone. The skin tone of reference is the whitest skin tone. Error bars represent confidence intervals at 95% statistical significance. Solid bars represent confidence intervals at 90% statistical significance. Coefficients in Panel (a) are expressed in elasticities. Coefficients in Panels (b) and (c) are expressed in standard deviations. Covariates include intra-municipal unit interacted with year fixed effects, gender, years of schooling, ethno-racial self-identification, age, age squared, marital status, occupation status , locality size, religion, interpersonal trust, and interviewer self-reported skin tone. Panel (c) includes income as control.

return of years of schooling. When comparing the skin tone premium SFD estimates, note that only the estimates from Columns 2 and 6 are statistically different from the one in Column 1. More specifically, the difference between the the benchmark and the most robust SFD specification is 0.06 percentage points (p-value = 0.031). Overall, SFD estimates are stable across the different specifications.

To discretely see the premium associated with each shade, I reproduce specification in Column 6 using dummy variables for each of the eleven skin tones using as reference the whitest skin tone. Figure 3 panel (a) plots the SFD coefficients for each of the collapsed PERLA scale skin tones with respect to the whitest skin tone. There is a non-linear skin tone premium in income. The coefficient for the second skin tone in the PERLA palette is not statistically different from the first skin tone. The third tone has 5.0% less income than the second (p-value = 0.075). The fourth has 5.6% less than the third (p-value = 0.058). However, coefficients for the fifth to the tenth tone are not statistically different than the fourth, while they are statistically different than the whitest skin tone.

5.2 Robustness

I perform several sensitivity checks to test the robustness of the results. First, I use the SFD specification and two alternative measures of individual economic welfare on LAPOP's data: years of schooling and a household asset index (Torche 2015). Figure 3 panel (b) and (c) shows the results for each PERLA skin tone. Consistent with the previous results, the three panels show that darker skin tones have lower economic outcomes than whiter skin tones.

Table B.5 shows specifications with further robustness checks. First, I test whether the results are biased given I use observations with strictly positive income. Column 1 uses specification (2) with a Poisson regression. To account for measurement error concerns given the survey's methodology to register respondent skin tone, I use interviewers information included in waves 2016/17 and 2018/17. Column 2 uses SFD specification, or specification (3), and includes interviewer's skin tone as control and interviewer fixed effects. Another source of concern is if the result hold after accounting for labor stratification. Only LAPOP's wave 2018/19 includes information on the respondents' type of labor occupation. Column 3 uses the subsample including wave 2018/19 and includes interviewer fixed effects and occupational controls. Finally, I use Spatial Second Differences (SSD) (Druckenmiller and Hsiang 2018) design which, instead of using a one-spatial lag difference between neighboring units, it uses two-spatial lag differences. Column 4 uses SSD for the subsample including waves 2016/17 and 2018/17, and interviewer fixed effects. In all specifications, the coefficient of skin tone on income remains negative and statstically significant.

As any observational study with no exogenous source of variation, a relevant concern is omitted variable bias. I use the SFD specifications to compute the adjusted coefficients á la Altonji, Elder, and Taber (2005) and Oster (2019) for different values of the maximum R-squared using the complete sample and the subsample with information on labor occupations.[19] I construct confidence intervals using one thousand bootstrap repetitions. Figure A.9 shows the adjusted coefficients and their confidence intervals. For the complete sample, if I could account for most of the remaining unobserved heterogeneity, the adjusted coefficients would not be statistically different from the coefficient of the most robust SFD specification – -0.031 – and would remain statistically different from zero. The same pattern appears for the subsample for wave 2018/19 with information on labor occupations. Thus, there are few concerns for omitted variable bias on the skin tone effect on income.

Finally, I test if the results hold using a different sample and whether SFD design actually purges unobserved spatial heterogeneity. To do so, I use ESRU Social Mobility Survey 2017 (ESRU-EMOVI) data which is representative for Mexico's population, Mexico City's population, and also measures skin tone using the PERLA palette.[20] Furthermore, it includes information on income, labor occupation, parental education and early childhood household wealth, and the respondent's approximate location, in longitude and latitude. Thus, I can replicate my main specification but exploiting a finer level of geolocation. Following a motivating example used by Druckenmiller and Hsiang (2018), I also use to subsamples of observations within a five kilometers ratio from two of the most important avenues in Mexico City. Figure A.10 shows the spatial distribution of respondents' skin tone and income for Mexico City, as well as the subsamples near each avenue. Table B.6 shows the OLS and SFD skin tone premium coefficients. The SFD does purge substantial unobserved heterogeneity biasing the OLS estimates. For the sample representative of Mexico's population, An increase in one PERLA skin tone correlates with a decrease in income of 4.0%. For Mexico City, the elasticities lie between 2% and 9%. All coefficients are statistically significant at one percent, using Conley standard errors to account for spatial autocorrelation (Conley 1999).

5.3 Mechanisms and heterogeneous effects

What drives the skin tone premium? Besides occupational segregation, differences in access to public goods, or inequality of opportunity (Ferreira and Gignoux 2011), economic literature argues that discrimination is the main mechanism driving ethnic-racial disparities in Latin America (Ñopo, Chong, and Moro 2009). Nevertheless, causal inference of racial discrimination is not straightforward for experimental and observational data (Rose 2022). To overcome the problem, I combine an Oaxaca-Blinder (OB) decomposition for continuous variables proposed by Ñopo (2008) with the SFD design to test whether racial discrimination explains the racial gap. See Appendix D for a detailed exposition of the methodology.

Figure A.11 shows the results for the Oaxaca-Blinder decomposition for continuous variables with SFD. The estimates for the complete sample imply that once purging unobserved heterogeneity

[19] González and Miguel (2015) argue that the maximum R-squared is way below one in contexts where there is measurement error.

[20] See https://ceey.org.mx/contenido/que-hacemos/emovi/.

through SFD, an increase in one darker skin tone decreases income by 3.2%. The OB discrimination estimate implies that 3.0% is due to discriminatory traits. Both coefficients are statistically significant at 1%. The latter results imply that differences in observed average characteristics cannot explain more than 90.0% of the gap.[21] I also compute the OBD-SFD for the subsample for wave 2018/19 with info on labor occupations. The racial gap coefficient is 7.6%, where discrimination can explain almost 80.0% of the latter.[22] Thus, racial discrimination would explain between 80.0% and 90.0% of the skin tone premium, accounting for unobserved heterogeneity.

Figure A.12 shows the results for the SFD-OB decomposition per capita country by country. Skin tone gradient in income is negative and statistically significant for most Latin American countries, with a relevant between-country variation. The skin tone premium is above the mean estimate for the whole region for countries such as Uruguay, Guatemala, Paraguay, Bolivia, Argentina, and Brazil. There is no statistically significant effect of skin tone premium in Costa Rica, Nicaragua, Belize, Guyana, Panama, and Suriname. The OB discrimination component is statistically different from zero for a few countries, but not different for the regional mean. If racial inequalities and discrimination are global phenomena, the heterogeneous effects show that their expressions are local.

LAPOP's wave 2012 includes questions on self-reported discrimination at work or school, at public places, or by governmental institutions. To provide further evidence of the racial discrimination hypothesis, I use the SFD research design and LAPOP questions on self-reported experiences of discrimination. Table B.7 shows that darker skin tones have a higher probability of reporting experiences of discrimination on the school or work, in public places, and by governmental institutions.

6 Conclusion

Racial disparities and discrimination are pressing issues to solve globally. This paper provides evidence of a skin tone premium and its adverse effects on economic development for Latin American countries. A novelty of the paper is the use of data allowing to disentangle inequalities due to phenotype. First, consistent with previous evidence on group-based inequalities, I provide evidence that inequalities between skin tone groups correlate with lower economic development at the subnational level. Next, I use the most robust methodology to provide evidence of a skin tone premium in income. Consistent with the historical and anecdotal evidence, racial discrimination drives the significant racial gap in Latin America.

Thus, contrary to the belief that only class or occupational status shapes individual income differences, results suggest a non-trivial white-skin tone premium in Latin America. Nevertheless, I

21 $\frac{\Delta^0_{OBD}}{\alpha_1} = \frac{-0.030}{-0.032} = 93.75\%$

22 $\frac{\Delta^0_{OBD}}{\alpha_1} = \frac{-0.060}{-0.076} = 78.94\%$

cannot disentangle the type of discrimination driving the skin tone premium: due to *taste-based* or *statistical discrimination* mechanisms. If discrimination, as Rose (2022) argues, is the result of acting based on perceived social identities, the mechanisms must be contingent on specific local, historical and cultural factors. Given that there is significant heterogeneity country by country, further research needs to disentangle the exact mechanisms operating and explain the skin tone premium at the local level. Moreover, future research have to disentangle the historical or institutional origins of racial inequalities.

On a purely theoretical perspective, the results suggest that, since skin tone is fixed and there is little room for behavioral responses, taxing alternatives as 'tagging' (Akerlof 1978; Alesina, Ichino, and Karabarbounis 2011; Piketty and Saez 2013) could play an essential role in overcoming racial disparities at early year stages. On a more practical stand, the results also suggest that progressive income and wealth taxation is also progressive in racial disparities (Derenoncourt and Montialoux 2020; Derenoncourt et al. 2021). Moreover, the results highlight the relevance of attenuating racial inequalities to improve aggregate welfare. Alongside its critical role in economic development, there is a historical debt in reducing racial disparities for justice and reparation.

References

Acharya, Avidit, Matthew Blackwell, and Maya Sen. 2016. "The Political Legacy of American Slavery." *Https://Doi.org/10.1086/686631* 78 (3): 621–41. https://doi.org/10.1086/686631.

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 Paola Giuliano. 2015. "Culture and institutions." *Journal of Economic Literature* 53 (4): 898–944. https://doi.org/10.1257/jel.53.4.898.

Alesina, Alberto, Andrea Ichino, and Loukas Karabarbounis. 2011. "Gender-based taxation and the division of family chores." *American Economic Journal: Economic Policy* 3 (2): 1–40. https://doi.org/10.1257/pol.3.2.1.

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.

Alesina, Alberto, and Ekaterina Zhuravskaya. 2011. "Segregation and the Quality of Government in a Cross Section of Countries." *American Economic Review* 101 (5): 1872–1911. https://doi.org/10.1257/aer.101.5.1872.

Altonji, Joseph G., Todd E. Elder, and Christopher R. Taber. 2005. "Selection on observed and unobserved variables: Assessing the effectiveness of Catholic schools." *Journal of Political Economy* 113 (1): 151–84. https://doi.org/10.1086/426036.

Arceo-Gomez, 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.

Arceo-Gómez, Eva O., and Raymundo M. Campos-Vázquez. 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.

Arnold, David, Will Dobbie, and Peter Hull. 2021. "Measuring Racial Discrimination in Algorithms." *AEA Papers and Proceedings* 111 (May): 49–54. https://doi.org/10.1257/pandp.20211080.

Baldomero-Quintana, Luis, Enrique De la Rosa-Ramos, and L. Guillermo Woo-Mora. 2022. "Infrastructures of Race? Colonial Indigenous Zoning and Contemporaneous Urban Segregation."

William & Mary. https://scholarworks.wm.edu/cgi/viewcontent.cgi?article=2947%7B/&%7D context=aspubs.

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.

Brouillette, Jean-Felix, Charles I. Jones, and Peter J. Klenow. 2021. "Race and Economic Well-Being in the United States." Cambridge, MA: National Bureau of Economic Research. https://doi.org/10.3386/W29539.

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., 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.

Card, David, Francois Gerard, Lorenzo Lagos, and Edson Severnini. 2018. "Assortative Matching or Exclusionary Hiring? The Impact of Firm Policies on Racial Wage Differences in Brazil." *National Bureau of Economic Research*. https://doi.org/10.3386/w25176.

Castillo J., Javier. 2022. "Inequality and Stratification in Latin America." In *Oxford Handbook of Social Stratification*, edited by Markus Gangl, Lucinda Platt, and Herman van de Werfhorst.

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.

Chancel, L., T. Piketty, E. Saez, and G. Zucman. 2021. "World Inequality Report 2022." World Inequality Lab.

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.

Conley, T. G. (1999). "GMM estimation with cross sectional dependence." *Journal of Econometrics*, 92(1), 1–45. https://doi.org/10.1016/S0304-4076(98)00084-0.

Cook, Lisa D. 2014. "Violence and economic activity: Evidence from African American patents, 1870-1940." *Journal of Economic Growth* 19 (2): 221–57. https://doi.org/10.1007/S10887-014-9102-Z/FIGURES/4.

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

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

Derenoncourt, Ellora. 2022. "Can You Move to Opportunity? Evidence from the Great Migration." *American Economic Review* 112 (2): 369–408. https://doi.org/10.1257/AER.20200002.

Derenoncourt, Ellora, François Gérard, Lorenzo Lagos, and Claire Montialoux. 2021. "Racial Inequality, Minimum Wage Spillovers, and the Informal Sector."

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.

Druckenmiller, Hannah, and Solomon Hsiang. 2018. "Accounting for Unobservable Heterogeneity in Cross Section Using Spatial First Differences." NBER Working Paper. https://doi.org/10.3386/w25177.

Easterly, William. 2007. "Inequality does cause underdevelopment: Insights from a new instrument." *Journal of Development Economics* 84 (2): 755–76. https://doi.org/10.1016/j.jdeveco.2006.11.002.

Elvidge, Christopher D., Mikhail Zhizhin, Tilottama Ghosh, Feng-Chi Hsu, and Jay Taneja. 2021. "Annual Time Series of Global VIIRS Nighttime Lights Derived from Monthly Averages: 2012 to 2019." *Remote Sensing* 13 (5): 922. https://doi.org/10.3390/rs13050922.

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.

Fortin, Nicole, Thomas Lemieux, and Sergio Firpo. 2011. *Decomposition Methods in Economics.* Vol. 4. PART A. Elsevier. https://doi.org/10.1016/S0169-7218(11)00407-2.

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.

Fujiwara, Thomas, Humberto Laudares, and Felipe Valencia Caicedo. 2021. "Tordesillas, Slavery and the Origins of Brazilian Inequality."

González, Felipe, and Edward Miguel. 2015. "War and local collective action in Sierra Leone: A comment on the use of coefficient stability approaches." *Journal of Public Economics* 128 (August): 30–33. https://doi.org/10.1016/j.jpubeco.2015.05.004.

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.

Henrich, Joseph. 2016. *The Secret of Our Success. How Culture is Driving Human Evolution, Domesticating Our Species, and Making Us Smarter.* Princeton University Press.

Hernández-Trillo, Fausto, and Ana Laura Martínez-Gutiérrez. 2021. "The Dark Road to Credit Applications: The Small-Business Case of Mexico." *Journal of Financial Services Research*, July, 1–25. https://doi.org/10.1007/s10693-021-00356-x.

Hoff, Karla, and Joseph E. Stiglitz. 2016. "Striving for balance in economics: Towards a theory of the social determination of behavior." *Journal of Economic Behavior & Organization* 126 (June): 25–57. https://doi.org/10.1016/j.jebo.2016.01.005.

Hsieh, Chang-Tai, Erik Hurst, Charles I. Jones, and Peter J. Klenow. 2019. "The Allocation of Talent and U.S. Economic Growth." *Econometrica* 87 (5): 1439–74. https://doi.org/10.3982/ECTA11427.

Jablonski, Nina G. 2021. "Skin color and race." *American Journal of Physical Anthropology* 175 (2): 437–47. https://doi.org/10.1002/ajpa.24200.

Jann, Ben. 2008. "The Blinder–Oaxaca Decomposition for Linear Regression Models." *The Stata Journal: Promoting Communications on Statistics and Stata* 8 (4): 453–79. https:

//doi.org/10.1177/1536867X0800800401.

Kelly, Morgan. 2020. “Understanding Persistence.” *SSRN*.

Kline, Patrick M., Evan K. Rose, and Christopher R. Walters. 2021. “Systemic Discrimination Among Large U.S. Employers.” Cambridge, MA: National Bureau of Economic Research. https://doi.org/10.3386/W29053.

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.

Kummu, Matti, Maija Taka, and Joseph H. A. Guillaume. 2018. “Gridded global datasets for Gross Domestic Product and Human Development Index over 1990–2015.” *Scientific Data* 5 (1): 180004. https://doi.org/10.1038/sdata.2018.4.

Lang, Kevin, and Ariella Kahn Lang Spitzer. 2020. “Race discrimination: An economic perspective.” American Economic Association. https://doi.org/10.1257/JEP.34.2.68.

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.

Martínez Casas, Regina, Emiko Saldívar, René D. Flores, and Christina A Sue. 2019. “Las múltiples caras del mestizaje.” In *Pigmentocracias. Color, Etnicidad y Raza En América Latina.*, 56–107. FCE.

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.

NASA Socioeconomic Data and Applications Center. 2018. “Documentation for the Gridded Population of the World, Version 4 (GPWv4), Revision 11 Data Sets.” Center for International Earth Science Information Network (CIESIN), Columbia University. https://doi.org/https://doi.org/10.7927/H49C6VHW.

Nunn, Nathan. 2012. “Culture and the Historical Process.” *Economic History of Developing Regions* 27 (sup1): S108–26. https://doi.org/10.1080/20780389.2012.664864.

———. 2021. “History as evolution.” In *The Handbook of Historical Economics*, 41–91. Academic Press. https://doi.org/10.1016/b978-0-12-815874-6.00010-1.

Ñopo, Hugo. 2008. “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.

Oh, Suanna. 2022. "Does Identity Affect Labor Supply?" *SSRN Electronic Journal.* https://doi.org/10.2139/ssrn.3998025.

Oster, Emily. 2019. "Unobservable Selection and Coefficient Stability: Theory and Evidence." *Journal of Business & Economic Statistics* 37 (2): 187–204. https://doi.org/10.1080/07350015.2016.1227711.

Persson, Torsten, and Guido Tabellini. 1994. "Is Inequality Harmful for Growth?" *The American Economic Review* 84 (3): 600–621. https://doi.org/https://www.jstor.org/stable/2118070.

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.

Rose, Evan K. 2022. "A Constructivist Perspective on Empirical Discrimination Research."

Shayo, Moses. 2020. "Social identity and economic policy." Annual Reviews Inc. https://doi.org/10.1146/annurev-economics-082019-110313.

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

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 Tianna Paschel. 2014. "Who is black, white, or mixed race? How skin color, status, and nation shape racial classification in latin america." *American Journal of Sociology* 120 (3): 864–907. https://doi.org/10.1086/679252.

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

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.

A Figures

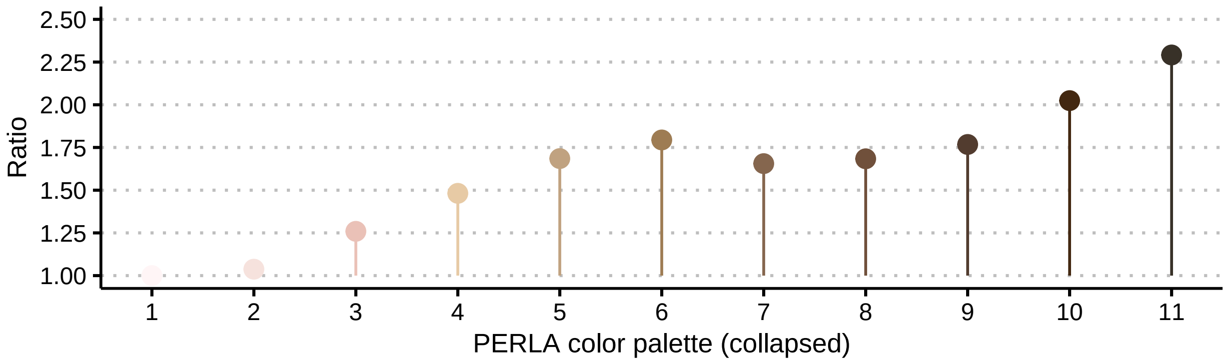


Figure A.1: Relative income

Notes: Mean income of whitest skin tone relative to mean income for each skin tone. Constructed using LAPOP's AmericasBarometer survey.

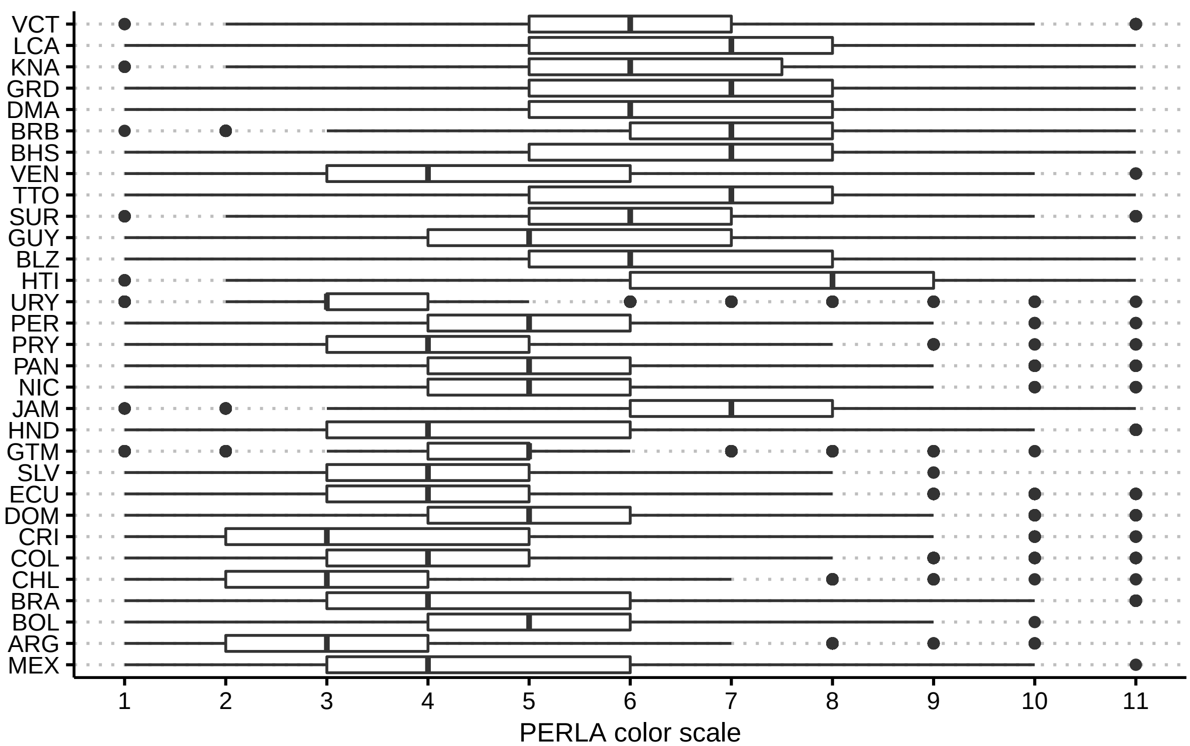


Figure A.2: PERLA skin tone boxplot by country

Notes: Skin tone boxplots by country. The line within the box represents the median. The lower and upper lines closing the box represents the interquantile range (percentiles 25 and 75, respectively). The lines outside the box, or whiskers, represent the observations outside the interquantile range. Dots outside the whiskers represent outlying observations. Constructed using LAPOP's AmericasBarometer survey.

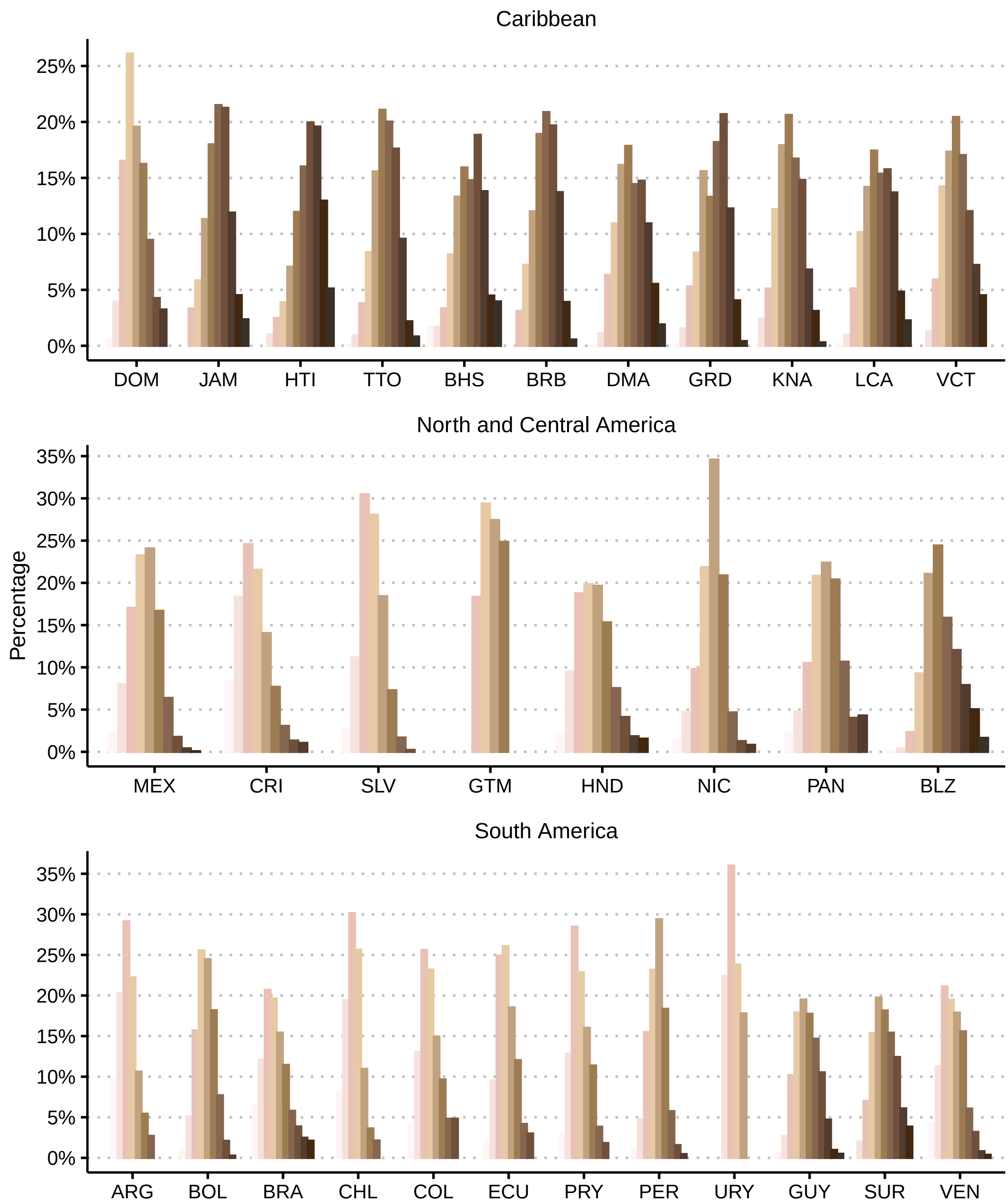
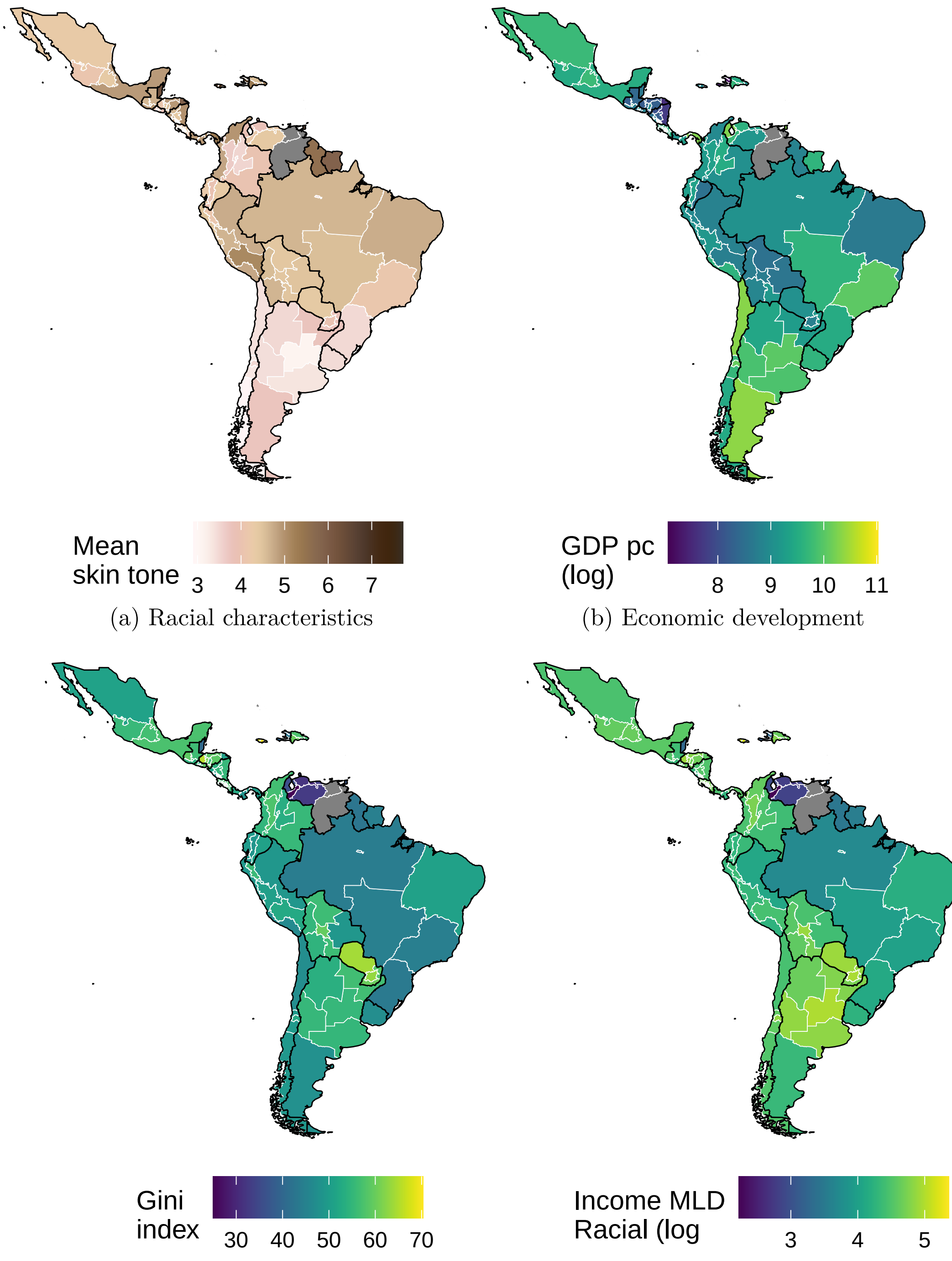


Figure A.3: Skin tone distribution by country

Notes: Histogram of skin tone share for each country in LAPOP's AmericasBarometer survey using the collapsed PERLA scale.

Figure A.4: Ethnicity distribution by country

Notes: Histogram of ethno-racial share for each country in LAPOP's AmericasBarometer survey.



(a) Racial characteristics (b) Economic development

(c) Income Inequality (d) Racial inequality

Figure A.5: LAPOP Stratified Regions

Notes: Stratified regions constructed using LAPOP's survey methodology and technical reports. All panels show each measure mean by region across the time period 2012-2019. Panel (a) shows skin tone mean constructed using LAPOP's AmericasBarometer skin tone measures. Panel (b) shows GDP per capita (log) mean constructed using Kummu, Taka, and Guillaume (2018) data grids. Panel (c) shows Gini index mean using LAPOP's AmericasBarometer income measures. Panel (d) shows income MLD between skin tone shades component (log) mean using LAPOP's AmericasBarometer skin tone and income measures.

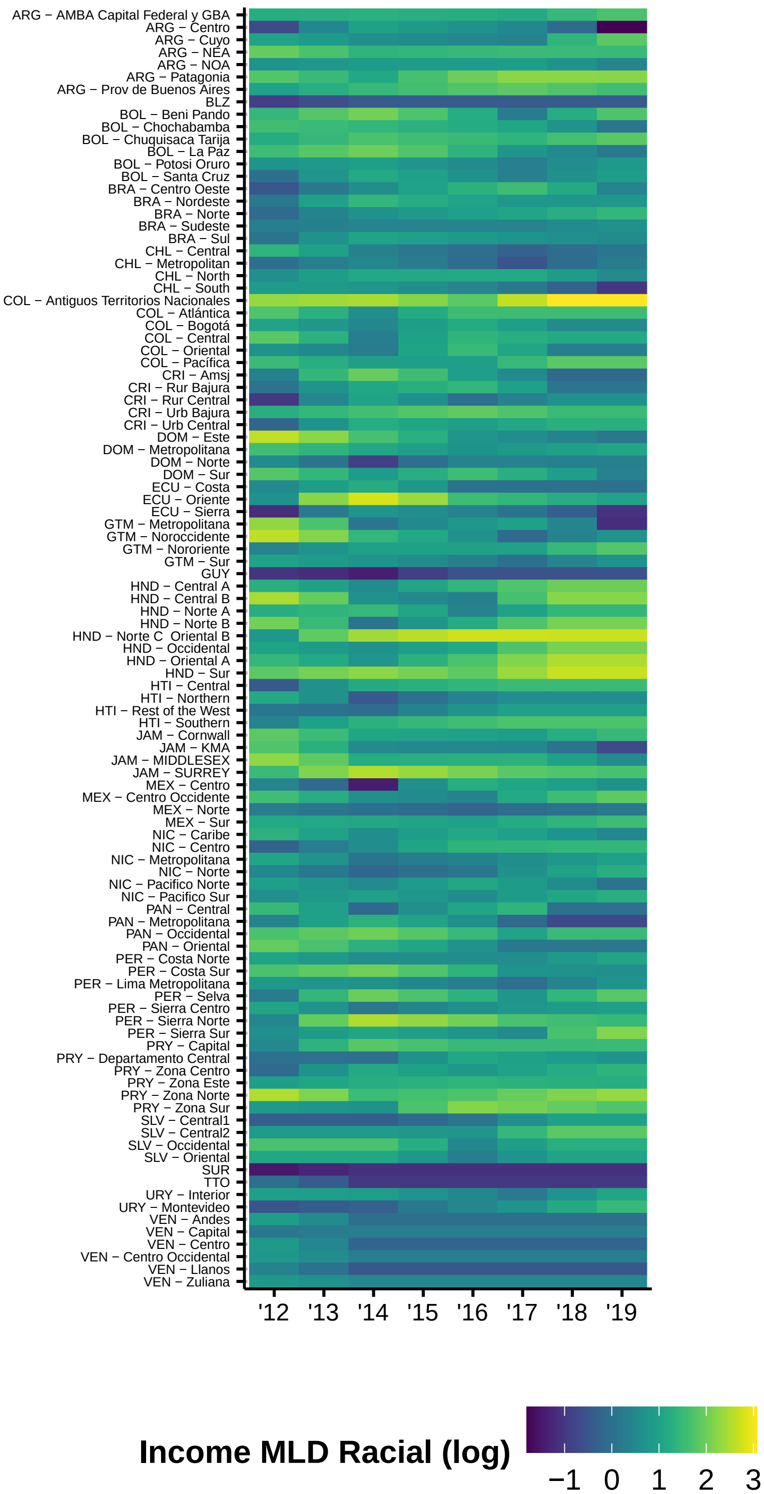


Figure A.6: Aggregate data by region and year: Income MLD Racial (log)

Notes: Stratified regions constructed using LAPOP's survey methodology and technical reports. Each row represents a subnational region. Each column represents a year. Income MLD between skin tone shades component (log) constructed using LAPOP's AmericasBarometer skin tone and income measures.

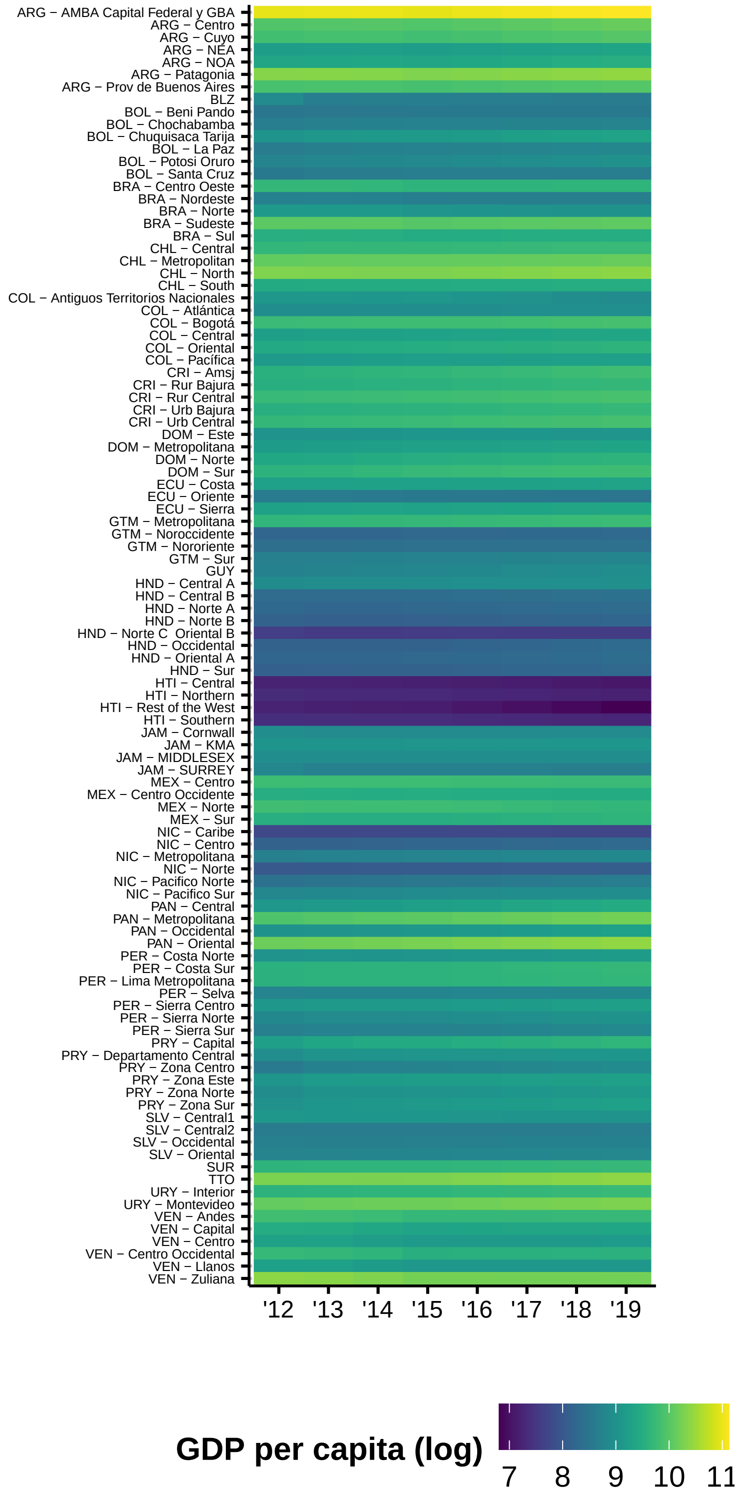


Figure A.7: Aggregate data by region and year: GDP per capita (log)

Notes: Stratified regions constructed using LAPOP's survey methodology and technical reports. Each row represents a subnational region. Each column represents a year. GDP per capita measures based on Kummu, Taka, and Guillaume (2018) data grids.

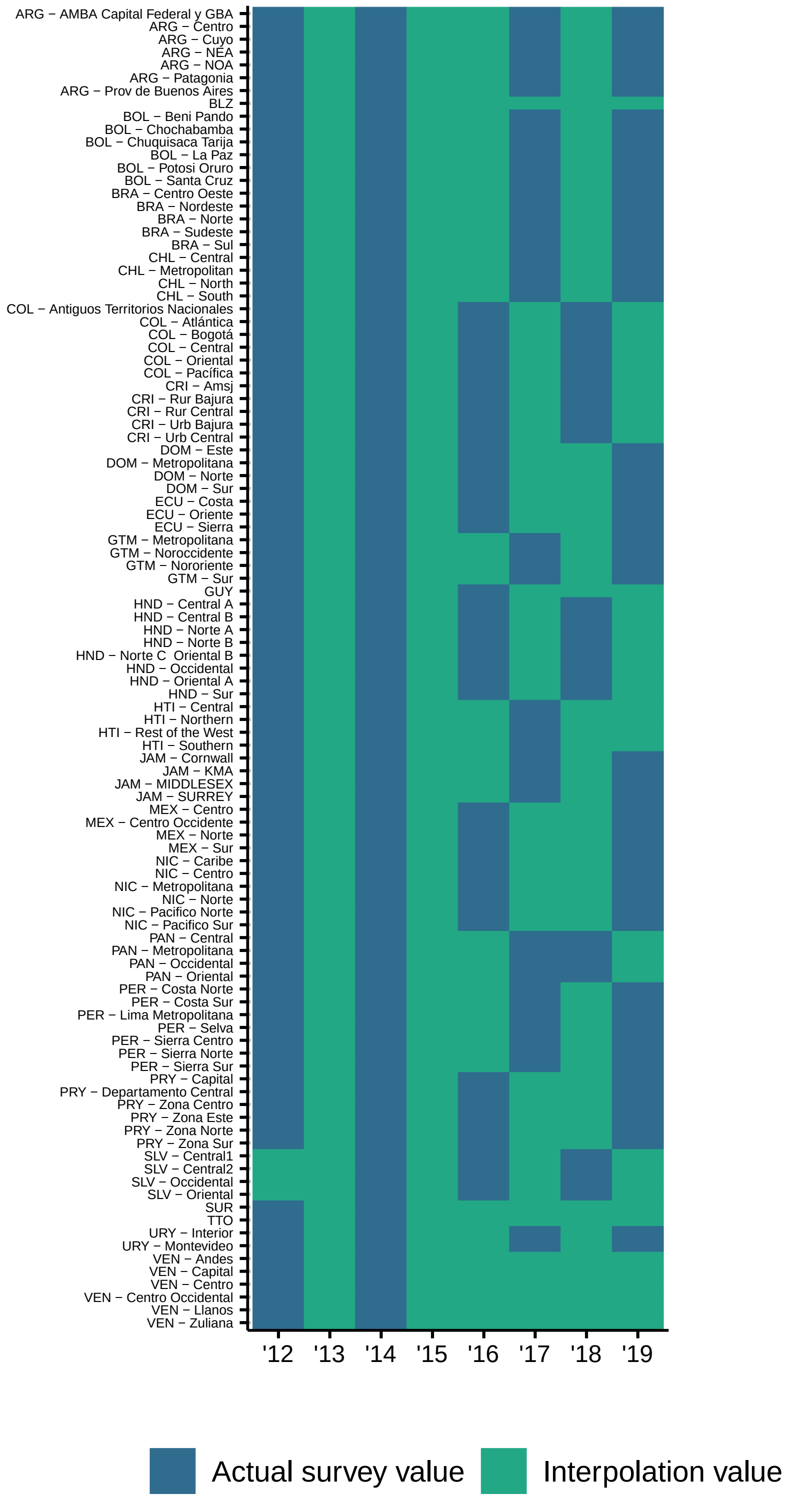
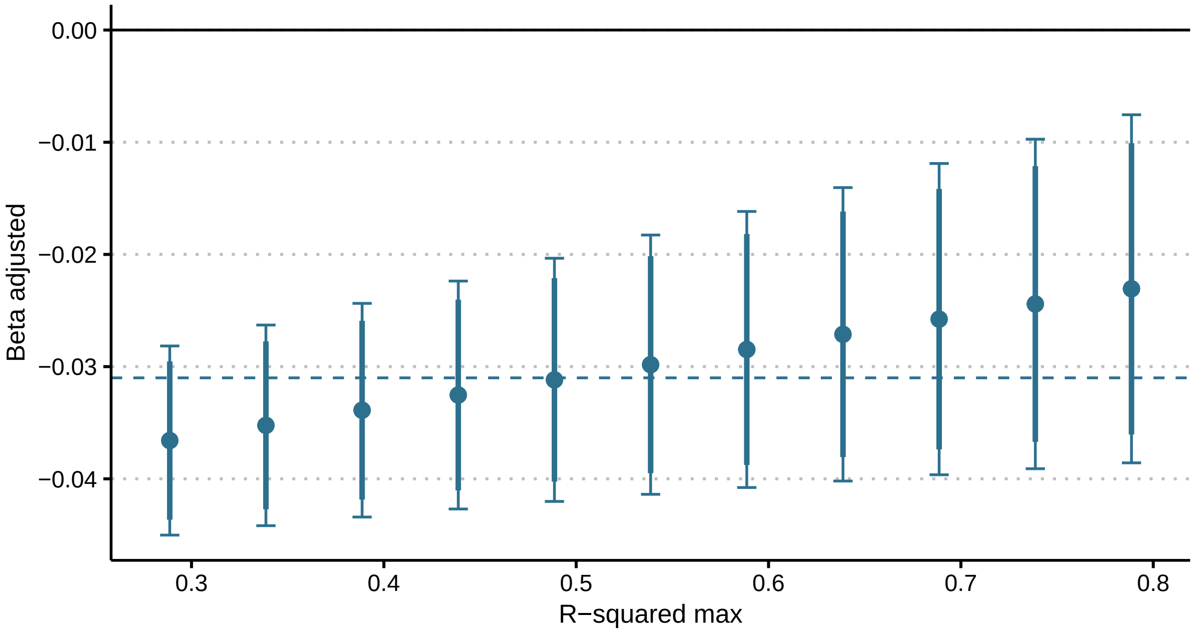
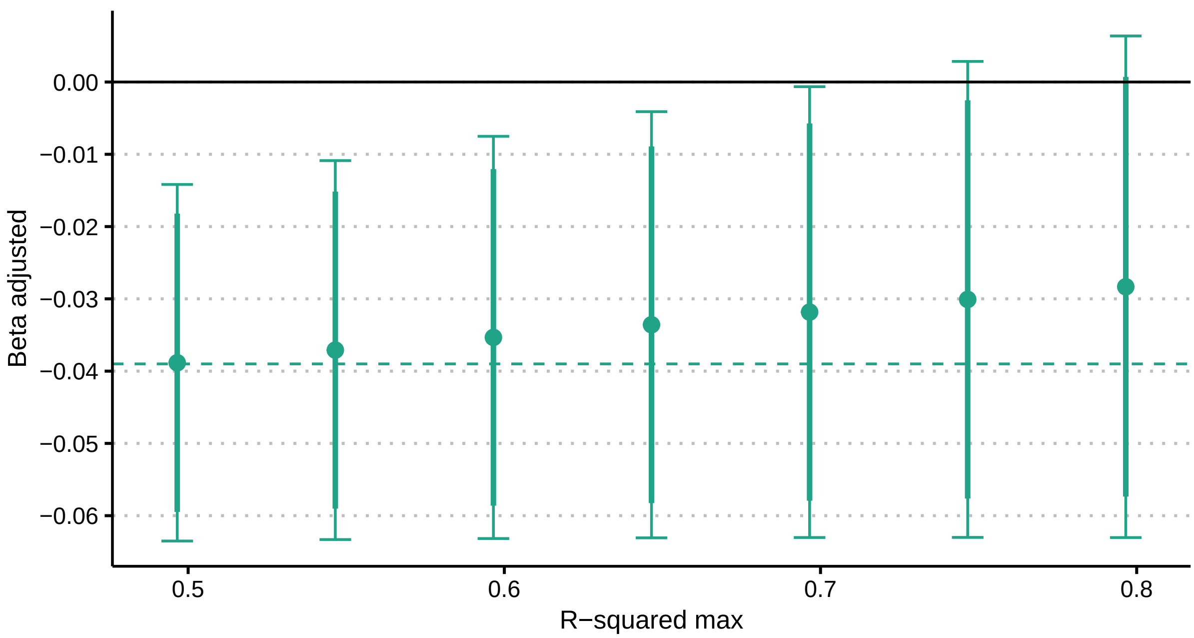


Figure A.8: Aggregate data by region and year: Actual survey and interpolation values

Notes: Stratified regions constructed using LAPOP's survey methodology and technical reports. Each row represents a subnational region. Each column represents a year. Actual survey value indicates that racial inequality measures, as well as all aggregate measures constructed using LAPOP's AmericasBarometer data, are from actual year surveys and not interpolated values.



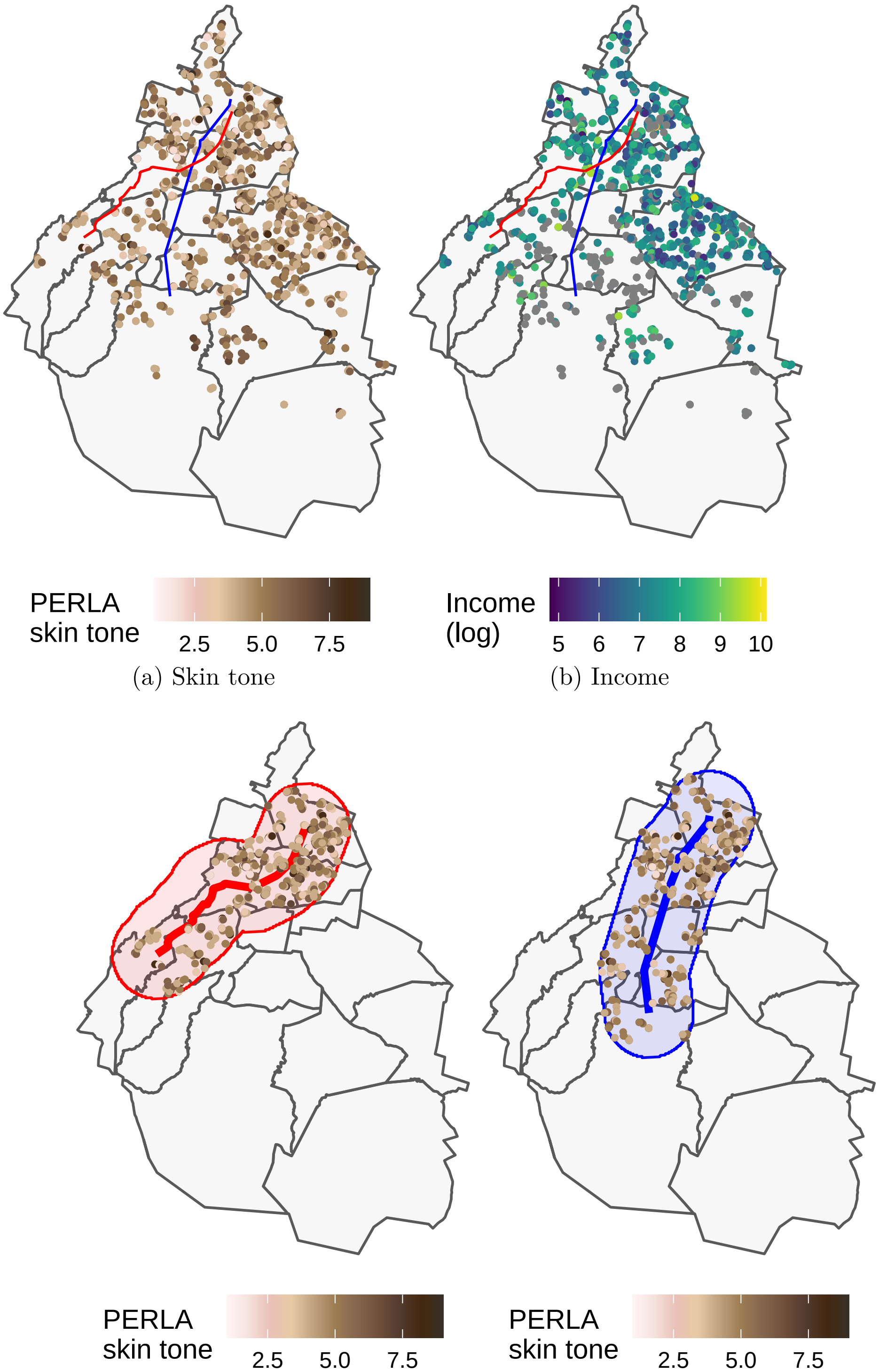
(a) Complete sample



(b) Workers subsample (wave 2018/19)

Figure A.9: Unobservable selection and coefficient stability

Notes: Beta adjusted coefficients using Oster (2019) coefficient stability and unobservable selection approach, assuming different values of a maximum R-squared. Confidence intervals computed using 1000 bootstrapped replications. Error bars represent confidence intervals at 95% statistical significance. Solid bars represent confidence intervals at 90% statistical significance. Panel (a) uses the complete sample. Pabel (b) uses the subsample in wave 2018/2019 with information on labor occupation. Dotted lines represent the coefficients from the most robust specification for each sample.

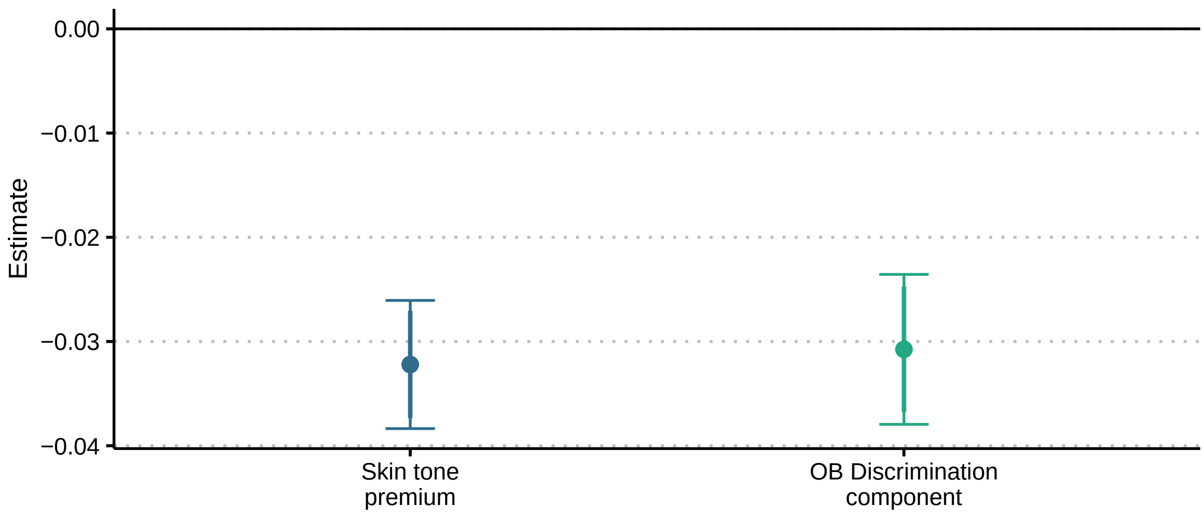


(a) Skin tone (b) Income

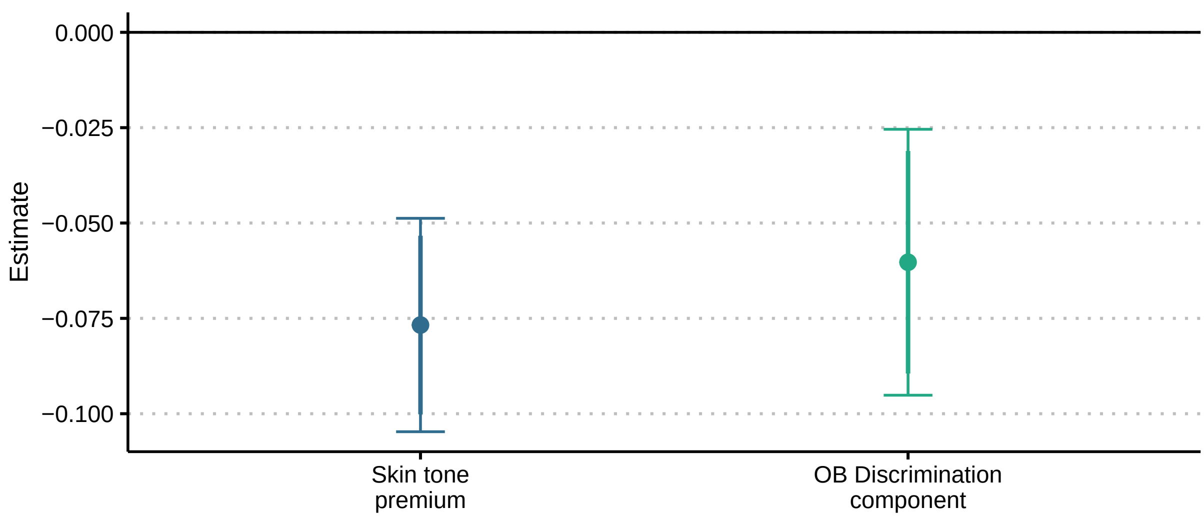
(c) Reforma Av. (5km) (d) Insurgentes Av. (5km)

Figure A.10: Skin tone premium for Mexico City

Notes: Data from ESRU-EMOVI 2017 survey. Map represents Mexico City and the municipality limits within. Each dot is an individual observation. Red line represents Reforma Avenue. Blue line represents Insurgentes Avenue. Panel (a) shows the skin tone of each individual. Panel (b) shows the income (log) of each individual. Panel (c) shows the skin tone for the subsample of observations within a 5km buffer around Reforma Avenue. Panel (b) shows the skin tone for the subsample of observations within a 5km buffer around Insurgentes Avenue.



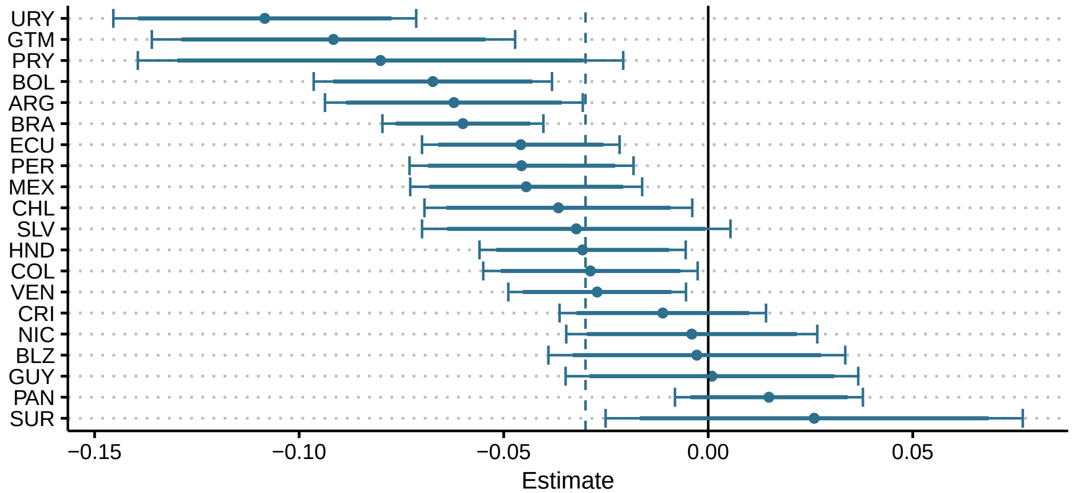
(a) Complete sample



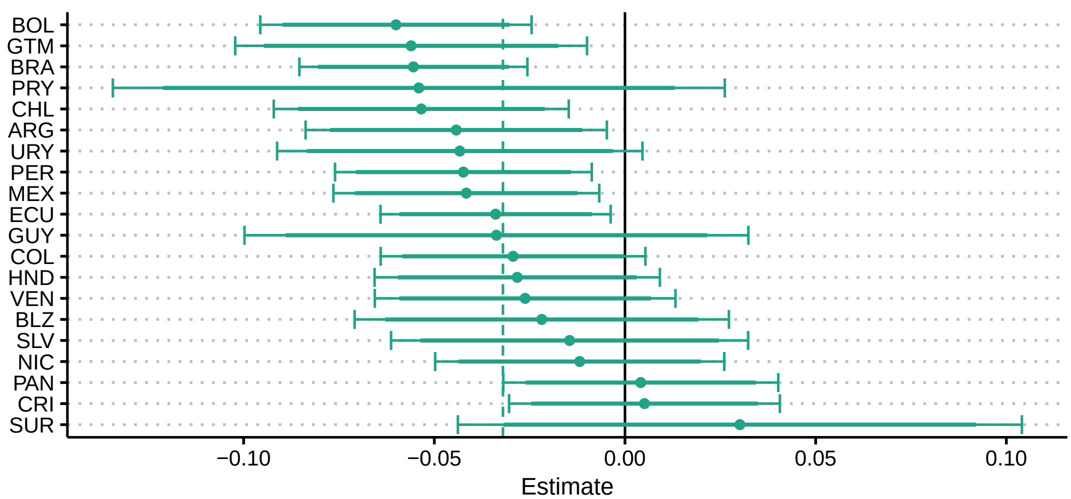
(b) Workers subsample (wave 2018/19)

Figure A.11: Oaxaca-Blinder decomposition with SFD

Notes: Components using Oaxaca-Blinder decomposition for continous variables (Ñopo 2008) and SFD (Druckenmiller and Hsiang, 2018). Confidence intervals computed using 1000 bootstrapped replications. Error bars represent confidence intervals at 95% statistical significance. Solid bars represent confidence intervals at 90% statistical significance. Panel (a) uses the complete sample. Pabel (b) uses the subsample in wave 2018/2019 with information on labor occupation.



(a) Skin tone premium



(b) OB Discrimination component

Figure A.12: OB decomposition by country

Notes: Components using Oaxaca-Blinder decomposition for continous variables (Ñopo 2008) and SFD (Druckenmiller and Hsiang, 2018). Confidence intervals computed using 1000 bootstrapped replications. Error bars represent confidence intervals at 95% statistical significance. Solid bars represent confidence intervals at 90% statistical significance. Panel (a) uses the complete sample. Pabel (b) uses the subsample in wave 2018/2019 with information on labor occupation.

B Tables

Table B.1: LAPOP AmericasBarometer: Sample size by country and wave.

Country	ISO Code	2012	2014	2016/17	2018/19	Total
Mexico	MEX	1,484	1,493	1,547	1,563	6,087
Argentina	ARG	1,482	1,462	1,519	1,514	5,977
Bolivia	BOL	2,919	3,047	1,686	1,676	9,328
Brazil	BRA	1,496	1,500	1,530	1,495	6,021
Chile	CHL	1,552	1,560	1,618	1,627	6,357
Colombia	COL	1,433	1,491	1,540	1,661	6,125
Costa Rica	CRI	1,444	1,525	1,509	1,493	5,971
Dominican Republic	DOM	1,508	1,513	1,512	1,514	6,047
Ecuador	ECU	1,469	1,483	1,539	1,522	6,013
El Salvador	SLV		1,509	1,547	1,505	4,561
Guatemala	GTM	1,489	1,505	1,539	1,590	6,123
Honduras	HND	1,688	1,550	1,550	1,560	6,348
Jamaica	JAM	1,487	1,496	1,502	1,501	5,986
Nicaragua	NIC	1,679	1,544	1,558	1,541	6,322
Panama	PAN	1,616	1,470	1,520	1,549	6,155
Paraguay	PRY	1,500	1,490	1,524	1,512	6,026
Peru	PER	1,494	1,485	2,643	1,515	7,137
Uruguay	URY	1,502	1,511	1,510	1,580	6,103
Haiti	HTI	1,147	875			2,022
Belize	BLZ	851	995			1,846
Guyana	GUY	1,121	1,105			2,226
Suriname	SUR	865	1,819			2,684
Trinidad and Tobago	TTO	801	1,651			2,452
Venezuela	VEN	948	1,183			2,131
Bahamas	BHS		1,739			1,739
Barbados	BRB		1,746			1,746
Dominica	DMA		584			584
Grenada	GRD		422			422
St. Kitts and Nevis	KNA		634			634
St. Lucia	LCA		522			522
St. Vincent	VCT		380			380
Total by wave		32,975	42,289	28,893	27,918	132,075

Notes: Based on LAPOP's AmericasBarometer data.

Table B.2: LAPOP AmericasBarometer: Descriptive Statistics

Statistic	Mean	St. Dev.	Min	Pctl(25)	Pctl(75)	Max
Age	39.75	15.86	16	27	50	89
Skin tone (PERLA scale)	4.59	1.90	1	3	6	11
Years of schooling	9.83	4.23	0	6	12	18
Income (PPP 2019)	251.97	297.74	1.22	75.59	314.40	2,955.34
Household size	4.39	2.14	1	3	5	17
Interpersonal trust	2.77	0.91	1	2	3	4
Gender						
Female	0.50	0.50	0	0	1	1
Male	0.0.50	0.50	0	0	1	1
Ethnicity						
Afro	0.12	0.33	0	0	0	1
Indigenous	0.08	0.27	0	0	0	1
Mestiza	0.45	0.50	0	0	1	1
Mulata	0.05	0.22	0	0	0	1
Other	0.07	0.25	0	0	0	1
White	0.23	0.42	0	0	0	1
Occupation						
Actively looking for a job	0.09	0.28	0	0	0	1
Not working and not looking for a job	0.03	0.17	0	0	0	1
Not Working but have job	0.04	0.20	0	0	0	1
Retired	0.08	0.28	0	0	0	1
Studying	0.07	0.26	0	0	0	1
Taking care of the home	0.18	0.39	0	0	0	1
Working	0.50	0.50	0	0	1	1
Marital Status						
Divorced or Separated	0.06	0.24	0	0	0	1
Living together	0.25	0.43	0	0	0	1
Married	0.33	0.47	0	0	1	1
Single	0.32	0.47	0	0	1	1
Widowed	0.04	0.20	0	0	0	1
Urbanization						
Big City	0.17	0.38	0	0	0	1
Medium City	0.16	0.36	0	0	0	1
Small City	0.15	0.36	0	0	0	1
Metropolitan area	0.22	0.41	0	0	0	1
Rural area	0.30	0.46	0	0	1	1
Religion						
Agnostic Atheist	0.02	0.13	0	0	0	1
Catholic	0.56	0.50	0	0	1	1
Evangelical	0.19	0.39	0	0	0	1
Hindu	0.003	0.06	0	0	0	1
Jehovah Witness	0.01	0.09	0	0	0	1
Jewish	0.001	0.02	0	0	0	1
Mormon	0.004	0.06	0	0	0	1
Muslim	0.002	0.04	0	0	0	1
Non-Christian Eastern Religion	0.02	0.14	0	0	0	1
None	0.10	0.29	0	0	0	1
Other	0.01	0.12	0	0	0	1
Protestant	0.08	0.27	0	0	0	1
Traditional or Native Religion	0.01	0.08	0	0	0	1

Notes: Unit of analysis is the individual. Based on LAPOP's AmericasBarometer data.

Table B.3: Racial Inequality and Economic Development

	HDI (log)	Nightlights (log)	GDP pc (log)	
	(1)	(2)	(3)	(4)
Racial Income MLD (log)	-0.007***	-0.043*	-0.070**	-0.157***
	(0.003)	(0.025)	(0.030)	(0.046)
Racial Education MLD (log)	-0.004	0.022	-0.014	0.031
	(0.003)	(0.035)	(0.025)	(0.041)
No. Obs.	824	824	385	202
No. Country-regions	103	103	103	103
No. Years	8	8	8	8
R2	0.973	0.987	0.877	0.888
R2 Within	0.848	0.983	0.759	0.766
FE: Country $\times$ Year	X	X	X	X
Actual survey			X	X
Year $\leq$ 2015				X
Inequality controls	X	X	X	X
Racial-ethnic controls	X	X	X	X
Geographic controls	X	X	X	X
Economic controls	X	X	X	X

Notes: Unit of analysis is the country-region level. Standard errors clustering by country-region in parenthesis. All regressions are weighted by population. HDI (log) and GDP per capita (log) by country-region based on Kummu, Taka, and Guillaume (2018) data grids. Total nightlights (log) based on Elvidge et al. (2021) data. Racial inequality measures constructed using LAPOP's AmericasBarometer data. Inequality controls include total income inequality and income inequality between broad-ethnic groups. Racial-ethnic controls include racial and ethnic fractionalization indexes, the share of mestizo, indigenous, afro, and other population, the number of ethnic groups in GREG, and the median skin tone. Geographic controls include area, longitude, latitude, altitude, ruggedness, mean temperature, mean precipitation, and mean solar radiation. Economic controls, an indicator whether the country-region host the country's capital, an indicator whether the region has coastline, an indicator whether the region shares an international border, total roads' length, the number of airports, and mean total nightlights (log). *p<0.1; **p<0.05; ***p<0.01

Table B.4: Skin tone premium on income: OLS

	Income (log)						
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Panel A. Original PERLA color scale							
PERLA scale	-0.053***	-0.388***	-0.163***	-0.143***	-0.140***	-0.124***	-0.101***
	(0.003)	(0.010)	(0.008)	(0.008)	(0.008)	(0.008)	(0.008)
PERLA scale (squared)		0.034***	0.010***	0.008***	0.008***	0.007***	0.006***
		(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
R2 Adj.	0.008	0.025	0.296	0.327	0.346	0.384	0.427
R2 Within			0.014	0.012	0.011	0.008	0.005
Panel B. Collapsed PERLA color scale							
PERLA scale	-0.050***	-0.368***	-0.159***	-0.139***	-0.137***	-0.123***	-0.100***
	(0.003)	(0.009)	(0.008)	(0.007)	(0.007)	(0.007)	(0.008)
PERLA scale (squared)		0.031***	0.009***	0.008***	0.008***	0.007***	0.006***
		(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
R2 Adj.	0.007	0.025	0.296	0.327	0.346	0.384	0.427
R2 Within			0.014	0.012	0.010	0.008	0.004
No. Obs.	108,024	108,024	108,024	108,024	108,024	108,024	108,024
FE: Country × Year			X				
FE: Country-Region × Year				X			
FE: Province × Year					X		
FE: Municipality × Year						X	
FE: Intra-Mun. × Year							X

Notes: Unit of analysis is the individual. Standard errors clustering by intra-municipality geographic unit and year in parenthesis. Based on LAPOP's AmericasBarometer data. *p<0.1; **p<0.05; ***p<0.01

Table B.5: Skin tone gradient on income: Robustness

	Income (log)			
	Poisson	SFD	SFD	SSD
	(1)	(2)	(3)	(4)
PERLA scale	-0.058***	-0.031***	-0.039***	-0.029***
	(0.010)	(0.005)	(0.014)	(0.007)
PERLA scale (squared)	0.003***	0.001	0.000	-0.001
	(0.001)	(0.002)	(0.005)	(0.001)
Female	-0.178***	-0.238***	-0.266***	-0.246***
	(0.008)	(0.013)	(0.032)	(0.017)
Years of Schooling	0.064***	0.072***	0.070***	0.071***
	(0.001)	(0.002)	(0.005)	(0.002)
Age	0.011***	0.010***	0.025***	0.011***
	(0.001)	(0.002)	(0.007)	(0.003)
Age (squared)	0.000***	0.000	0.000***	0.000
	(0.000)	(0.000)	(0.000)	(0.000)
1[Afro]	-0.001	-0.073**	-0.014	-0.089**
	(0.020)	(0.031)	(0.085)	(0.040)
1[Indigenous]	-0.082***	-0.121***	-0.198***	-0.111***
	(0.021)	(0.027)	(0.075)	(0.036)
1[Mulata]	0.006	-0.045	-0.028	-0.064*
	(0.019)	(0.029)	(0.078)	(0.036)
1[Other]	-0.117***	-0.082***	-0.193***	-0.067**
	(0.019)	(0.024)	(0.067)	(0.030)
1[White]	-0.017	-0.051***	0.015	-0.059***
	(0.011)	(0.017)	(0.046)	(0.023)
No. Obs.	130,189	33,587	6,834	24,134
R2 Pseudo	0.722			
R2		0.294	0.508	0.323
R2 Adj.		-0.019	-0.173	-0.119
R2 Within		0.185	0.188	0.188
FE: Intra-Mun. × Year	X	X	X	X
FE: Interviewer		X	X	X
Sociodemographic controls	X	X	X	X
Occupational controls			X	

Notes: Unit of analysis is the individual. Standard errors clustering by intra-municipality geographic unit and year in parenthesis. Besides the coefficients in table, sociodemographic controls include: marital status, occupation status –working, taking care of the home, actively looking for a job, not working and not looking for a job, not working but have a job, studying, retired–, locality size –metro area, big city, medium city, small city, or rural area–, religion, and interpersonal trust. The ethnicity group of reference is Mestizo category. The occupation group of reference is working on services or retail. Occupational controls include indicator variables of types of labor –agriculture, artisans and mechanics, laborers, technician, administrative staff, scientist, managers, military–. Based on LAPOP's AmericasBarometer data. *p<0.1; **p<0.05; ***p<0.01

Table B.6: Skin tone gradient on income: Mexico

	Income (log)						
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	OLS	SFD					
PERLA scale	-0.263***	-0.042***	-0.037***	-0.039***	-0.018***	-0.091***	-0.073***
	(0.094)	(0.007)	(0.007)	(0.009)	(0.006)	(0.028)	(0.028)
PERLA scale (squared)	0.022**	0.012***	0.012***	0.011*	0.006***	0.004	-0.007
	(0.010)	(0.004)	(0.004)	(0.006)	(0.002)	(0.019)	(0.024)
Female	-0.192***	-0.099***	-0.093***	-0.078***	-0.032***	-0.174***	-0.107**
	(0.019)	(0.013)	(0.013)	(0.017)	(0.011)	(0.048)	(0.052)
Years of schooling	0.043***	0.016***	0.012***	0.010***	0.004***	0.020***	0.017**
	(0.002)	(0.001)	(0.002)	(0.002)	(0.001)	(0.006)	(0.007)
Age	-0.025***	0.003***	0.005***	0.004***	0.001**	0.009***	0.007***
Num.Obs.	15,228	14,842	14,842	9,139	8,717	619	541
R2	0.402	0.349	0.355	0.382	0.364	0.287	0.259
R2 Within	0.120	0.044	0.053	0.044	0.021	0.147	0.106
Sample	Mexico	Mexico	Mexico	Mexico	Mexico City	Reforma Av.	Insurgentes Av.
FE: Municipality	X	X	X	X	X	X	X
Socio-demographic Controls	X	X	X	X	X	X	X
Parental controls			X	X	X	X	X
Occupational controls				X	X	X	X

Notes: Unit of analysis is the individual. Conley standard errors in parenthesis. Sociodemographic controls include: age squared, indigenous dummy, and locality size –urban or rural area–. Parental controls include parents mean years of schooling, household wealth index in early childhood. Occupational controls include indicator variables of occupation status –working, taking care of the home, actively looking for a job, not working and not looking for a job, not working but have a job, studying, retired– and types of labor –agriculture, artisans and mechanics, laborers, technician, administrative staff, scientist, managers, military–. Based on ESRU-EMOVI 2017 data. *p<0.1; **p<0.05; ***p<0.01

Table B.7: Self-reported measures of Racial discrimination: SFD

	1[Self-reported discriminatory experience]		
	on work or school	on public places	by governmental institutions
	(1)	(2)	(3)
PERLA scale	0.005**	0.005**	0.006***
	(0.002)	(0.002)	(0.002)
PERLA scale (squared)	0.000	0.001**	0.000
	(0.001)	(0.001)	(0.001)
Number of Observations	19,728	19,728	19,728
R2	0.055	0.055	0.058
R2 Adj.	-0.190	-0.190	-0.187
R2 Within	0.000	0.001	0.001
FE: Intra-Mun. Cluster × Year	X	X	X

Notes: Unit of analysis is the individual. Standard errors clustering by intra-municipality geographic cluster and year in parenthesis. Based on LAPOP's AmericasBarometer wave 2012. *p<0.1; **p<0.05; ***p<0.01

C Spatial First Differences (SFD)

Researches face several restrictions for causal inference with cross-sectional data. When one wants to estimate a treatment unbiased effect on an outcome of interest, and there is no panel-data structure, a discontinuity on a score, or available instruments, relevant endogeneity concerns arise from omitted variables bias. Druckenmiller and Hsiang (2018) propose a new research design to deal with unobserved heterogeneity using cross-sectional data to identify unbiased effects: Spatial First Differences (SFD). Druckenmiller and Hsiang (2018) argue that "unobserved heterogeneity in cross-sectional context is captured by non-stationary trends in outcomes across space." Alternatively, as Tobler's First Law of Geography points out: *"everything is related to everything else, but near things are more related than distant things"* (Kelly 2020). Thus, if units are *densely packed across physical space*, omitted variables bias can be purged using a differencing approach where the spatial position of observations can be located and organized, such as it is commonly used in time-series contexts.

In terms of identification, the authors argue that the Conditional Independence Assumption:

$$E[y_i|t_j] = E[y_j|t_j] \quad \forall i \neq j \tag{4}$$

where the potential outcome of unit i under treatment t_j, is equal to potential outcome of unit j under t_j, is a demanding assumption since it assumes all units are comparable between each other. However, if one assumes a Local Conditional Independence Assumption, where:

$$E[y_i|t_{i-1}] = E[y_{i-1}|t_{i-1}] \quad \forall \{i, i-1\} \tag{5}$$

one accepts that the potential outcome of two adjacent units, i and $i-1$ is equal under treatment t_{i-1}. Thus, neighboring units are better countrerfactuals of each other. Note that assumption in equation (5) is less demanding than the same in equation (4). More importantly, it is fairly similar as the Continuity Assumption used in RD designs. Or as Druckenmiller and Hsiang (2018) points out, Local Conditional Independence assumes a discontinuity for every pair of adjacent units. Druckenmiller and Hsiang (2018) use simulations, two empirical applications, and a specification experiment omitting all possible observable covariates, to show how SFD successfully deals with omitted variable bias.

To estimate SFD, Druckenmiller and Hsiang (2018) propose a simple estimator that compares each units to its immediately adjacent neighbor:

$$\begin{aligned} y_i - y_{i-1} &= (t_i - t_{i-1})\beta + (x_i - x_{i-1})\alpha + (c_i - c_{i-1})\gamma + (\varepsilon_i - \varepsilon_{i-1}) \\ \Delta y_i &= \beta_{SFD}\Delta t_i + \alpha_{SFD}\Delta x_i + \Delta c_i\gamma + \Delta\varepsilon_i \\ \Delta y_i &= \beta_{SFD}\Delta t_i + \alpha_{SFD}\Delta x_i + \Delta\varepsilon_i \end{aligned} \tag{6}$$

where $\Delta c_i \gamma = 0$ under Local Conditional Independence Assumption (equation (5)). Such estimator filters the influence of variables that vary little across space. SFD can be implemented in single- or two-dimensional space by organizing observations and pairing units to its adjacent unit.

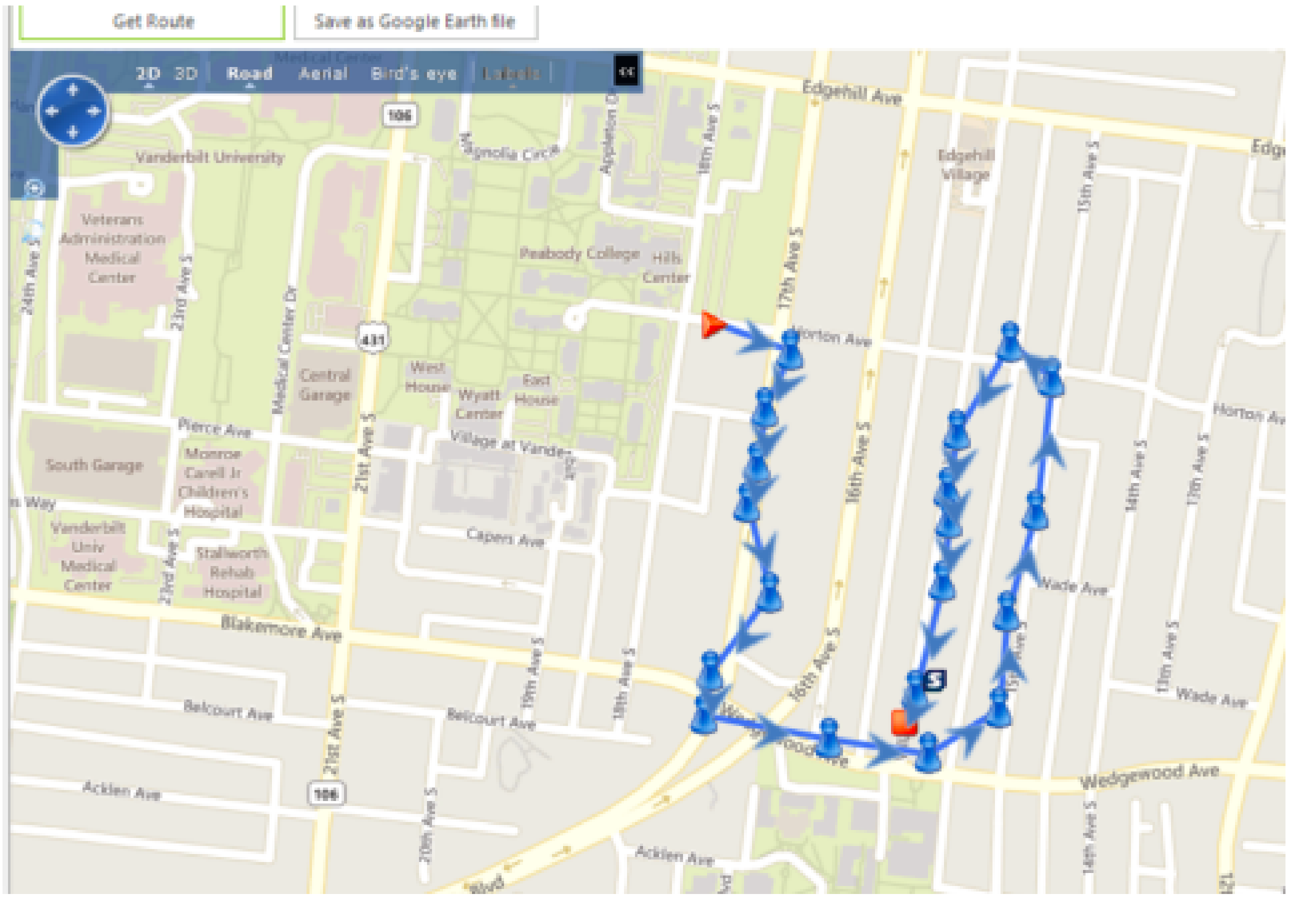


Figure C.13: Spatial proximity of LAPOP interviewed units

D Oaxaca-Blinder (OB) decomposition for continuous variables with SFD

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 one captures the differences in returns to those characteristics (Fortin, Lemieux, and Firpo 2011; Jann 2008; Ñopo 2008).

Following Ñopo (2008), the OB decomposition for a continuum of groups can be extended from a 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' equation:

$$y_i = \alpha_0 + \alpha_1 t_i + \varepsilon_i \tag{7}$$

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' equation:

$$y_i = \beta_0 + \beta_1 x_i + \beta_2 t_i + \beta_3 t_i \cdot x_i + \varepsilon_i \tag{8}$$

Where y_i and t_i represent the same variables as in Equation (7), and x_i is an 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 (7) can be expressed as:

$$\begin{aligned}
\alpha_1 &= E[(\beta_0 + \beta_2) + (\beta_1 + \beta_3)x|t = 1] - E[\beta_0 + \beta_1 x|t = 0] \\
\alpha_1 &= (\beta_0 + \beta_2) + (\beta_1 + \beta_3)E[x|t = 1] - \beta_0 - \beta_1 E[x|t = 0] \\
\alpha_1 &= \beta_1(E[x|t = 1] - E[x|t = 0]) + \beta_2 + \beta_3 E[x|t = 1] \\
\alpha_1 &= \Delta^x_{OBD} + \Delta^0_{OBD}
\end{aligned}$$

Where $\Delta^x_{OBD} \equiv \beta_1(E[x|t = 1] - E[x|t = 0])$, represents the differences in the outcome due to average observable characteristics of the individuals, or composition effect; while $\Delta^0_{OBD} \equiv \beta_2 + \beta_3 E[x|t = 1]$ represents the differences in the outcome that cannot be explained by observable characteristics. Ñopo (2008) formally shows that obtaining the two OB components is straightforward when t_i is a continuous measure.

Given the economic theory of discrimination, the component Δ^0_{OBD} is usually interpreted as a measure of discrimination. Nevertheless, when the researcher does not account for all observed and unobserved relevant factors affecting the outcome of interest, the component Δ^0_{OBD} will also capture differences due to unobserved heterogeneity between groups (Jann 2008). However, if SFD design purges the unobserved heterogeneity, when estimating OB decomposition using SFD instead of usual specifications in levels, the OB components capture differences solely due to discrimination. Thus, one can provide a causal estimate of racial discrimination.

I use both SFD and OB decomposition to obtain estimates of racial discrimination. In practice, I use the following procedure:[23]

1. Estimate the SFD 'simplified' equation:

$$\Delta y_i = \alpha_0 + \alpha_1 \Delta t_i + \theta + \eta + \varepsilon_i$$

   The coefficient of interest is α_1.

2. Estimate the SFD 'extended' equation:

$$\Delta y_i = \beta_0 + \beta_1 \Delta x_i + \beta_2 \Delta t_i + \beta_3 \Delta t_i \cdot \Delta x_i + \theta + \eta + \varepsilon_i$$

   The coefficients of interest are β_2 and β_3.

3. Estimate the average characteristics of the sample used to estimate equations in steps 1 and 2, namely $E[x]$.[24]

4. Obtain the component of the wage gap that cannot be explained by differences in average characteristics:

$$\Delta^0_{OBD} = \beta_2 + \beta_3 E[x]$$

5. Obtain the component of the wage gap that is explained by differences in average characteristics:

$$\Delta^x_{OBD} = \alpha_1 - \Delta^0_{OBD}$$

I used country-region fixed effects, θ, and year fixed effects, η, to account for geographic and time between-differences in income. For inference, I use stratified bootstrapping to obtain the empirical distributions of the aggregated and detailed decomposition components.

[23]Note Δ^x_{OBD} and Δ^0_{OBD} stand for the OB components, while Δx_i stands for the first spatial difference of unit i and $i-1$ in the variable or vector of variables x_i.

[24]The set includes: sex; age; age squared; years of schooling; marital status; occupational status; ethnicity; urbanization or locality size; religion; interpersonal trust.