World Inequality Lab – Working Paper N° 2022/02

Unveiling the Cosmic Race: Racial Inequalities in Latin America

L. Guillermo Woo-Mora

February 2022

Unveiling the Cosmic Race: Racial Inequalities in Latin America*

L. Guillermo Woo-Mora†

For most recent version click here

February 13, 2022

Abstract

This paper uses skin tone and income information for 100 thousand individuals across 31 countries during the last decade to study the skin tone premium and its welfare consequences in Latin America. Firstly, I estimate the welfare consequences of racial inequality: subnational regions with higher income inequality between racial groups correlate with worse economic development. Then, I provide evidence of a skin tone premium in the region: every darker skin tone has 3 percent less income per capita out of an eleven-color palette. Results suggest that the main mechanism is racial discrimination, with relevant heterogeneity between countries.

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

*I am grateful to Suanna Oh, Thierry Verdier, Thomas Piketty, Ignacio Flores, Ekaterina Zhuravskaya, Hillel Rapoport, Eva Arceo-Gómez, Raymundo Campos-Vázquez, David Margolis, and Denis Cogneau. Also thanks to Mauricio Tenorio-Trillo, Sandra Ley, 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, PSE Casual Friday Development Seminar, PSE Petit Seminaire Informel, and DEMOSAL Conference for their helpful comments. All remaining errors are my own.

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

1 Introduction

Inequalities lie beyond class struggle. Besides income or wealth, other social dimensions also explain the distributive conflict (Akerlof and Kranton 2000; Fleurbaey and Schokkaert 2011; Piketty 2020; Shayo 2009, 2020). Most economic literature studying inequalities has mainly relied on interpersonal comparisons of the first dimensions, focusing little on the latter. Such is the case of race (Advani et al. 2021).

This paper analyses racial inequalities in Latin America. Latin America is one of the most unequal regions globally (Lustig, Lopez-Calva, and Ortiz-Juarez 2016; De Rosa, Flores, and Morgan 2020) and one of the most diverse, both racially and ethnically. Race-related studies mainly focus on the US due to the historical experience of slavery, segregation, discrimination (Cook and Logan 2020), as well as due to the substantial racial disparities in many dimensions (Brouillette, Jones, and Klenow 2021; Chetty et al. 2020; Cook, Logan, and Parman 2016; Derenoncourt and Montialoux 2020; Derenoncourt 2022; Logan and Parman 2017) Nevertheless, racial inequalities are also salient in other countries and regions, some with shared historical experiences of slavery and segregation (Baldomero-Quintana, De la Rosa-Ramos, and Woo-Mora 2022; Fujiwara, Laudares, and Valencia Caicedo 2021).

This paper overcomes a relevant conceptual challenge to analyze racial inequalities: disentangle inequalities across cultural dimensions and physical phenotype. Sociologists argue that *race* or *racialization* are physical characteristics or phenotype that can define group membership, while *ethnicity* membership is based on cultural characteristics (Telles and Martínez Casas 2019). Thus, as the *constructivist* approach argues, race is a socially constructed identity rather than innate biological factors, but where the latter influence the formation of such social identities (Rose 2022). Most ethno- or race-related studies in economics account for differences between broad ethno-racial categories as “White,” “Black,” “Asian,” “Indigenous,” or “Latino,” which can describe both phenotype and cultural characteristics simultaneously.

To overcome such challenge empirically, I use a data set of repeated cross-sections with representative information on skin tone, self-reported broad ethnic categories, and income for nearly 100 thousand individuals in more than 31 countries for the last decade. I use the AmericasBarometer Survey from the Latin American Public Opinion Project (LAPOP), which includes the color palette of the Project on Ethnicity and Race in Latin America (PERLA). Using the racial dimension rather than the ethnic one, this paper contributes with evidence of the skin tone premium, mainly driven by discrimination, and its welfare consequences on aggregate economic development for Latin America. I exploit such data to provide evidence to two relevant questions.

Do disparities between racial groups explain disparities in comparative development? Are there any economic consequences or costs of racial disparities? Group-based inequalities matter for comparative development: they can lead to political inequality, discriminatory policies between groups, inadequate public goods provision, misallocation of talent, and losses in technological innovation

(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). I contribute to the literature by providing evidence of the welfare consequences of racial inequalities. I estimate new measures of racial inequality –inequality between racial groups– at the subnational level using skin tone rather than the conventional broadly defined ethnic categories. Using cross-regional variation for the last decade and controlling for time-invariant characteristics and common-shocks trends, I find that higher between-racial-groups income inequality correlates to a lower GDP per capita. Along with the group-based inequalities hypothesis, higher racial inequality in income hinders economic development.

Could racial inequality be, in some part, driven by the causal effect of skin tone? Is skin tone a relevant dimension shaping inequalities at the individual level? In the second part of the paper, I provide evidence of a skin tone premium in Latin America. Scholars have compiled extensive evidence for Latin America on how ethno-racial components determine disparities in the labor market and wages (Arceo-Gomez and Campos-Vazquez 2014; Arceo-Gómez and Campos-Vázquez 2019; Card et al. 2018; Campos-Vázquez 2020; Derenoncourt et al. 2021; Garavito et al. 2013; Ñopo, Saavedra, and Torero 2007; Ñopo 2012), educational attainment (Botelho, Madeira, and Rangel 2015; Telles and Steele 2012; Telles and Martínez Casas 2019), social mobility (Campos-Vázquez and Medina-Cortina 2019; Monroy-Gómez-Franco and Vélez-Grajales 2020; Solís, Güémez Graniel, and Lorenzo Holm 2019), as well as access to the financial sector (Hernández-Trillo and Martínez-Gutiérrez 2021).[1] Most of the previous literature focuses on a single country. Moreover, there is little comparability on the measures of ethnicity and race since their meaning and understanding changes country by country.

This paper contributes with evidence of the racial gap in Latin America using skin tone measures, a sizable sample of individuals across multiple countries, and the most solid available methodology for cross-sectional data.[2] 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: a darker skin tone has 3 percent less monthly income per capita, out of a color palette with eleven tones. To test whether the gap is driven by racial discrimination, I combine SFD with an Oaxaca-Blinder decomposition for continuous variables (Ñopo 2008b) and provide estimates showing that at least 80% of the racial gap is due to racial discrimination, with substantial country-specific heterogene-

[1]Racial components also determine social norms and identities (Campos-Vázquez and Medina-Cortina 2017); electoral preferences (Aguilar 2011; Campos-Vázquez and Rivas-Herrera 2021) and increase the perceptions of racial discrimination (Chong and Ñopo 2008; Ñopo, Chong, and Moro 2009; Trejo and Altamirano 2016).

[2]To my best knowledge, there are few studies on ethno-racial disparities with comparable data and methodology for more than one Latin American country. Cernat, Sakshaug, and Castillo (2019) present evidence of measurement error concerns. Ñopo (2012) studies the wage gap between broad ethnic groups for Bolivia, Peru, Brazil, Guatemala, Paraguay, Chile, and Ecuador, using an Oaxaca-Blinder decomposition with matching techniques (Ñopo 2008a).

ity. Consistent with the race discrimination hypothesis, individuals with darker skin tones report higher discrimination against them. Thus, race shapes differences in income mainly through racial discrimination.

The rest of the paper is structured as follows. Section 2 presents a brief historical overview of 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 review: The racial question in Latin America

Race and ethnicity have been central dimensions for Latin America (Tenorio-Trillo 2017). The encounter of the original Indigenous population, European conquerors, and African populations, mostly brought as enslaved people for forced labor, produced an early miscegenation process throughout the continent. The result was a complex administrative and social caste system dependent on race and ethnicity. Even when the colonial caste system was more *flexible* than the racial hierarchies in the US (Graham 2013), overall, whiter meant closer to European and thus with higher social status. 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 as segregation and anti-miscegenation in the US or racial hate from 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. 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* (*La Raza Cósmica: Misión de la Raza Iberoamericana*), 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 *third way* to racial politics. Nowadays, most Latin Americans define themselves as *mestizos* or *mulatos*, even when there is substantial racial diversity within such ethnic category. Thus, Latin American countries could veil racial disparities and inequalities within them, given that everyone became *mestizo* or *mulato*.

Two hundred years 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. In the extreme, some countries are characterized as *Pigmentocracies* (Telles and

Martínez Casas 2019). As this work will show, the use of color palettes represents an improvement in studying racial disparities in Latin America given the extended *mestizo/mulato* identity, resulting from the historical experience previously described. This methodology, or the 'colorism' agenda, might significantly benefit social science research (Dixon and Telles 2017). More importantly, learning from the Latin American experience, both historically and with contemporary evidence, can shed light on tackling racial inequalities in other latitudes.

3 Data

Measuring racial inequalities is a challenging task due to data restrictions. Few surveys and censuses register individual racial characteristics besides ethnic self-recognition –whether the respondent sees herself as part of a specific ethnic group: White, Afro, Latino, Asian, among others–. In other cases, the interviewer infers the ethnicity of the respondent. In some extreme cases, there are no available measures of racial categories (The Economist 2020).[3]

I exploit a rich data set that disentangles the ethnic and racial 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 samples of voting-age adults, using a common questionnaire score and country-specific modules.[4] 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.[5] The scale ranges from 1 to 11, where one is the lightest skin tone and eleven is the darkest.[6] Figure 4 in Appendix 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 constructed sample includes more than 100 thousand individual observations from 31 countries across four waves (2012, 2014, 2016/2017, 2018/2019). Table 4 in Appendix B shows the sample size by country and wave.

Figure 6 in Appendix A shows the distribution of skin tones by country using the PERLA color palette. Racial characteristics measured by skin tone vary substantially across and within countries. For instance, the darkest-skin tones are a majority in the Caribbean. Nevertheless, every country has a black-skinned population, usually omitted since they are a minority. Medium-dark tones are the majority in Central American countries and countries with a high miscegenation historical

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

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

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

[6]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).

experience, such as Mexico, Brazil, Bolivia, Colombia, Ecuador, or Peru. Lastly, the whitest-skin tones are more usual in countries with little miscegenation that experienced high shares of 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.[7]

Since ethnicity refers to cultural characteristics, it might differ from the racial phenotype. LAPOP works with six broad ethnic categories: Afro, Indigenous, Mestiza, Mulata, White, and other ethnic groups (i.e., Asian, Jew, among others). Figure 7 in Appendix A shows the ethnic distribution for each country. The majority in Caribbean countries define themselves as Afro origin, even when there is a high diversity of racial phenotype measured by skin tone. People who define themselves as White are the majority in Argentina, Chile, Costa Rica, and Uruguay. Interesting patterns of ethnicity and racialization arise when analyzing countries with a high percentage of white and medium-dark-skinned populations. For instance, in every other country besides the countries where there is an Afro or White majority, most of the population defines themselves as Mestiza or Mulata. Consistent with the historical and anecdotal evidence presented in Section 2, such countries also happen to have a high historical miscegenation experience and a strong presence of *mestizaje* ideology.

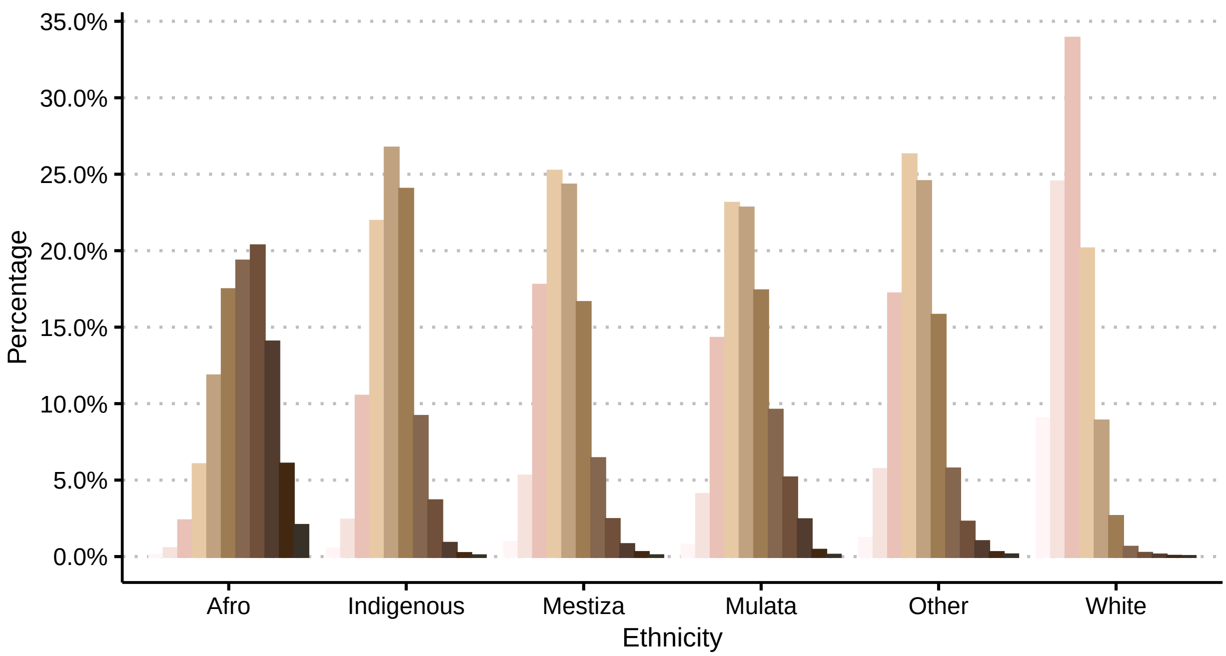

Figure 1: Ethniciy and Race

Figure 1 shows the distribution of skin tones by each of the ethnic categories for the whole sample.

[7] As Figure 5 in Appendix A shows, either the whitest and the darkest PERLA colors are outliers dependent on the country. Then, I use the minimum (maximum) skin tone within each country's boxplot's whiskers and replace all skin tones below (above) with such value. As shown later, results are robust to both the standard PERLA color scale and the *collapsed* PERLA color scale.

The patterns previously described persist: people who define themselves as Afro have darker skin tones, while those who define themselves as White have whiter skin tones. People who define themselves as Indigenous, Mestiza, Mulata, or 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 phenotypes within an ethnic group. Consistent with the *constructivist* approach, race and ethnicity are fluid and contextually dependent (Rose 2022).

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

Besides ethnicity- and race-related questions, LAPOP also includes information on socio-demographics, such as age, gender, region, urban or rural household, years of schooling, occupational status, marital status, and household size. The survey asks about self-reported monthly household income by brackets, but the brackets' values change for each country and wave. 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 5 in Appendix B shows the sample descriptive statistics.

4 Racial Inequality and Economic Development

What are the aggregate welfare implications of racial disparities? Alesina, Michalopoulos, and Papaioannou (2016) show that, rather than ethnic diversity (Alesina and La Ferrara 2005), economic inequalities between ethnic groups are correlated with lower economic development at the national level. When analyzing specific world regions, the authors find that the negative correlation between ethnic inequality and economic development is not statistically significant for Western Europe and the Americas. A plausible hypothesis is that ethnic inequality is not as salient as racial inequality in Latin America, given the blurry border between ethnicity and race resulting from the *mestizaje* ideology. This section exploits LAPOP data to compute aggregated measures of racial inequality, or inequality between racial groups, at the subnational level. The overreaching goal is to test whether racial inequalities correlate with lower economic development.

LAPOP data is representative at the subnational level. Each country is stratified between three and eight regions, where some small countries are not stratified. Figure 8 in Appendix A plots the stratified regions. Figure's 8 panel (a) shows the mean skin tone by each of LAPOP's stratified sampling regions. There is substantial racial diversity between country regions. Thus, racial statistics at the country-level mask relevant within-country heterogeneity.

To obtain measures of regional economic development, I exploit multiple sources. Firstly, I use Kummu, Taka, and Guillaume (2018) gridded subnational data on Gross Domestic Product and Human Development Index.[8] Secondly, I also use NASA/NOAA Visible Infrared Imaging Radiometer Suite (VIIRS) nightlight yearly data (Elvidge et al. 2021) to proxy for economic development. To account for differences in population between regions, I use Gridded Population of the World data (NASA Socioeconomic Data and Applications Center 2018). Figure's 8 panel (b) shows GDP per capita (log) by each country-region.

Using LAPOP data, I compute inequality measures without any interpersonal comparison besides income. Using the computed monthly income per capita, I construct Gini indexes for each country region and year in the LAPOP sample. Figure's 8 panel (c) shows the mean Gini index for each region in the analysis. LAPOP data allows decomposing income inequality measures by racial group. Using the PERLA scale color categories, I use the mean log deviation (MLD) index (Foster and Shneyerov 2000) to obtain the between and within components for income and educational inequality for each subnational region and year of LAPOP's data.[9] Besides the racial inequality measures, I also use the broad ethnic categories used by LAPOP data to compute alternative ethnic inequality measures. Figure's 8 panel (d) shows the mean Income MLD between-group racial component.

Figure 2 panel (a) shows the unconditional correlation of the mean income MLD between racial group component and mean (log) GDP per capita. Panel (b) shows the same correlation but uses the mean years of schooling MLD between racial group component and mean (log). There is a negative correlation between racial inequality and lower economic development. To test more robustly the relation between racial inequality and economic development I use the following econometric specification:

$$y_{rct} = \beta Racial\ Inequality_{rct} + \gamma X_{rc} + \theta_c \times \eta_t + \varepsilon_{rct} \qquad (1)$$

Where y_{crt} represents (log) GDP per capita of subnational region r in country c at time t; $Racial\ Inequality_{rct}$ is the (log) MLD between racial groups component; X_{rc} represents a set of time-invariant controls at the country-region level (geographical variables and economic activity proxies). One improvement of this analysis with respect to Alesina, Michalopoulos, and Papaioannou (2016) is the availability of multiple observations for each subnational region. I include country

[8] 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.

[9] The MLD or the Generalized Entropy index GE(0) fulfills the properties of the axiomatic approach –transfer principle, population principle, decomposability principle, and scale-invariant– (Shorrocks 1984), as well as the path independent decomposability .

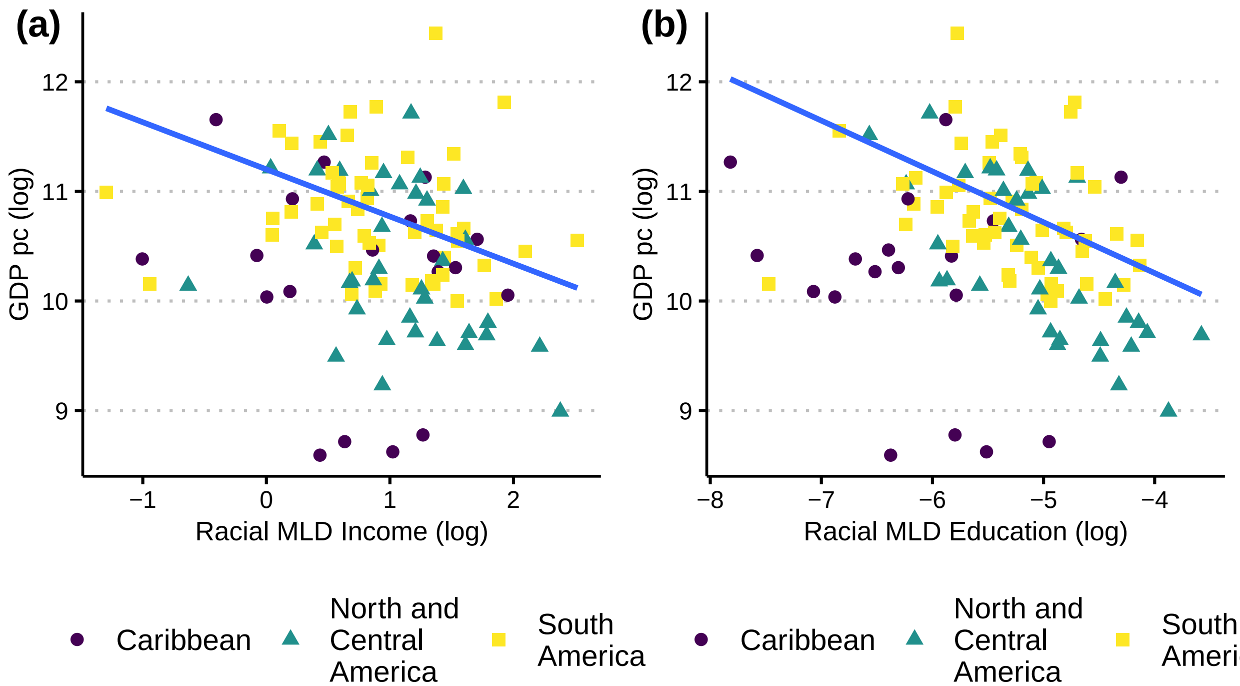

Figure 2: Racial-inequality and Economic Development

fixed effects θ_c interacted with year fixed effects η_t 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 the country for a given year. To give more weight to more populated regions, I weight the specifications by population.

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 and inequality in years of schooling between racial groups, an increase in one percent on racial income inequality correlates with a decrease in regional GDP per capita of almost 14%. Column 2 includes total income inequality and income inequality measures between broad ethnic groups (Afro, White, Mestizo, Mulato, Other) computed with LAPOP data. Column 3 includes racial and ethnic fractionalization measures and the number of ethnic groups by country region, and the median skin tone. Columns 4 and 5 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 correlation between racial inequality and GDP per capita remains statistically significant.

The most robust specification in Column 6 uses sub-national region and year fixed effects. Accounting for all invariant characteristics at the regional level and common-shock trends to all regions, an increase in one percent in racial inequality correlates with a decrease of one percent on regional GDP per capita. In contrast, racial inequalities in education have no statistically significant effect on economic development. To test the robustness of the results, Columns 7 and 8 use HDI and total nightlights as an alternative measure of economic development at the subnational region

Table 1: Racial Inequality and Economic Development

	Dependent variable:							
	GDP per capita (log)						HDI (log)	Nightlights (log)
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Racial Income MLD (log)	-0.138***	-0.171**	-0.109*	-0.101**	-0.088***	-0.011**	-0.007***	-0.087**
	(0.047)	(0.069)	(0.061)	(0.042)	(0.027)	(0.004)	(0.001)	(0.038)
Racial Education MLD (log)	-0.246*	-0.248*	-0.209*	-0.121**	-0.016	-0.006	-0.003	-0.016
	(0.136)	(0.141)	(0.108)	(0.050)	(0.028)	(0.007)	(0.002)	(0.032)
Income MLD (log)		0.070	-0.225	-0.128	0.021		0.001	0.055
		(0.235)	(0.254)	(0.128)	(0.100)		(0.006)	(0.115)
Ethnic Income MLD (log)		0.129	0.140	0.134**	0.088**		0.010***	0.010
		(0.178)	(0.144)	(0.059)	(0.039)		(0.002)	(0.045)
Racial Fractionalization index (log)			0.876	0.210	-0.130		0.053	0.499
			(1.056)	(0.479)	(0.361)		(0.035)	(0.473)
Ethnic Fractionalization index (log)			-0.210	-0.174*	-0.297***		-0.010	0.120
			(0.185)	(0.097)	(0.090)		(0.008)	(0.147)
Mestizo share (log)			-0.148	0.216	0.431***		-0.001	0.023
			(0.270)	(0.174)	(0.134)		(0.010)	(0.142)
No. GREG Groups			-0.005	-0.003	-0.001		-0.001***	-0.001
			(0.003)	(0.003)	(0.002)		(0.000)	(0.003)
Median skin tone			-0.145	0.051	-0.032		-0.001	0.060
			(0.119)	(0.068)	(0.048)		(0.003)	(0.040)
Population density (log)				0.101**	0.117*		0.005	0.270**
				(0.039)	(0.061)		(0.004)	(0.110)
Years of schooling					0.094*		0.008***	0.128**
					(0.052)		(0.003)	(0.061)
Num.Obs.	391	391	391	391	391	391	391	391
R2	0.565	0.571	0.635	0.796	0.874	0.996	0.970	0.980
R2 Adj.	0.431	0.435	0.511	0.719	0.823	0.995	0.958	0.971
R2 Within	0.153	0.163	0.288	0.602	0.754	0.031	0.845	0.768
FE: Country × Year	X	X	X	X	X		X	X
FE: Sub-national region						X		
FE: Year						X		
Geographic controls				X	X		X	X
Economic controls					X		X	X

Notes: Standard errors clustering by Sub-national 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, include border length, coastline length, an indicator whether the country-region host the country's capital, roads length, and number of airports. *p<0.1; **p<0.05; ***p<0.01

level. In both specifications, the racial inequality estimate is negative and statistically significant. Consistent with the group-based inequalities hypothesis, and rather than ethnic-based inequalities, racial-based inequalities hinder economic development in Latin America.

5 Skin tone premium

In this section, I use the LAPOP data at the individual level to analyze the skin tone premium on a set of socio-economic outcomes following the disparities approach (Fleurbaey and Schokkaert 2011). To estimate the skin tone premium, I use the following econometric specification:

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

Where y_i is the (log) monthly income per capita 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 cluster fixed-effects, and η is a time fixed effect. Note I use the interaction of intra-municipality cluster fixed-effects with year fixed effects.[10] I use LAPOP sample weights to make the results comparable across countries and waves representative at the national level (Castorena 2021).

Table 2 takes as a benchmark the specification of skin tone on income with intra-municipal cluster times year fixed-effects and the polynomial of degree two and includes several observable characteristics as controls in each column.[11] The first specification considers differences in income and skin tone between intra-municipal clusters for each country and year and 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 percent, statistically significant at one percent. 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, Table's 2 Column 6 shows that an increase on one skin tone in the collapsed PERLA scale correlates with a decrease of 7.7 percent in income, statistically significant at one percent. To account for measurement error concerns given the survey's methodology to register respondent skin tone, Column 7 uses a subsample with information on the interviewer information and includes interviewer's skin tone as control and interviewer fixed effects: the semi-elasticity of

[10] Note that I do not observe the same unit i across multiple periods.

[11] Table 6 shows in Appendix B the gradient between skin tone and income accounting for non-linearities and between-geographic heterogeneity. Panel A uses the original PERLA skin tone palette, while Panel B uses the *collapsed* PERLA palette described in section 3. Both panels show that, for each specification, 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. As the R-squared statistics in Table 6 show, a substantial part of the variation can be explained by geographical heterogeneity.

Table 2: Skin tone premium on income: OLS

	Dependent variable:									
	Monthly income per capita (log)									
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
PERLA scale	-0.100***	-0.109***	-0.086***	-0.074***	-0.077***	-0.076***	-0.111***	-0.129***	-0.075***	-0.084***
	(0.008)	(0.008)	(0.008)	(0.007)	(0.008)	(0.008)	(0.014)	(0.029)	(0.029)	(0.031)
PERLA scale (squared)	0.006***	0.006***	0.005***	0.004***	0.005***	0.005***	0.008***	0.009***	0.005*	0.006*
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.003)	(0.003)	(0.003)
1[Female]		-0.292***	-0.279***	-0.268***	-0.268***	-0.185***	-0.226***		-0.269***	-0.270***
		(0.006)	(0.006)	(0.006)	(0.006)	(0.006)	(0.010)		(0.020)	(0.022)
Years of Schooling			0.054***	0.068***	0.067***	0.062***	0.070***		0.070***	0.069***
			(0.001)	(0.001)	(0.001)	(0.001)	(0.001)		(0.003)	(0.004)
Age				0.012***	0.012***	0.005***	0.011***		0.021***	0.022***
				(0.001)	(0.001)	(0.001)	(0.002)		(0.004)	(0.005)
Age (squared)				0.000***	0.000***	0.000	0.000*		0.000***	0.000***
				(0.000)	(0.000)	(0.000)	(0.000)		(0.000)	(0.000)
1[Afro]					-0.041***	-0.037**	-0.061**		-0.068	-0.038
					(0.015)	(0.015)	(0.025)		(0.054)	(0.058)
1[Indigenous]					-0.122***	-0.112***	-0.134***		-0.152***	-0.168***
					(0.014)	(0.014)	(0.022)		(0.047)	(0.048)
1[Mulata]					-0.015	-0.018	-0.020		-0.044	-0.032
					(0.016)	(0.016)	(0.022)		(0.050)	(0.053)
1[Other]					-0.096***	-0.096***	-0.092***		-0.170***	-0.161***
					(0.014)	(0.014)	(0.018)		(0.041)	(0.044)
1[White]					-0.032***	-0.034***	-0.061***		-0.042	-0.042
					(0.009)	(0.009)	(0.014)		(0.029)	(0.031)
1[Agriculture]									-0.198***	-0.205***
									(0.045)	(0.048)
1[Directors and managers]									0.432***	0.405***
									(0.078)	(0.087)
1[Elemental occupations]									-0.102***	-0.140***
									(0.035)	(0.039)
1[Military]									0.214**	0.243**
									(0.102)	(0.106)
1[Artisans]									0.031	0.003
									(0.055)	(0.059)
1[Mechanics]									0.031	0.003
									(0.055)	(0.059)
1[Administrative personnel]									0.190***	0.184***
									(0.037)	(0.040)
1[Scientists and intellectuals]									0.280***	0.266***
									(0.042)	(0.046)
1[Technicians and professionals]									0.212***	0.213***
									(0.036)	(0.039)
Num.Obs.	108,024	108,024	106,936	106,936	106,936	102,092	44,479	12,177	11,563	11,258
R2 Adj.	0.427	0.445	0.477	0.492	0.493	0.511	0.484	0.378	0.480	0.461
R2 Within	0.005	0.037	0.092	0.118	0.119	0.150	0.179	0.007	0.171	0.168
FE: Intra-Mun. Cluster × Year	X	X	X	X	X	X	X	X	X	X
Socio-demographic Controls						X	X		X	X
FE: Interviewer							X			X

Notes: Standard errors clustering by intra-municipality geographic cluster and year in parenthesis. Besides the coefficients in table, socio-demographic 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. *p<0.1; **p<0.05; ***p<0.01

skin tone on income increases to minus 11. If anything, the measurement error seems to underestimate the coefficient of skin tone on income in absolute terms. Columns 8 to 10 use a subsample with information on the respondent's type of occupation: whether she works on services or retail, agriculture, artisan, mechanic, technician, administrative, science, manager or director, military occupations. As shown by the last columns, even after accounting for the type of occupation, the negative gradient between skin tone and income persist: the most robust specification shows that one darker skin tone correlates with a decrease of almost 8.5 percent on income, statistically significant at one percent.[12]

5.1 Purging unobserved heterogeneity

The previous estimates can be biased due to unobserved characteristics that affect monthly household income per capita and correlate with skin tone –parental background, early childhood, social networks–. I use the Spatial First Difference (Druckenmiller and Hsiang 2018) research design to purge 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 can be purged using a differencing approach where the spatial position of observations can be located and organized. Thus, one uses first-differencing between adjacent units exploiting the spatial dimension, 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,[13] the adjacent identification number units' are adjacent in space.[14] Thus, by using SFD, I can purge common unobserved factors between neighboring observations. For each intra-municipal cluster and year, I arrange the observations in the survey by their identification number assigned. Afterwards, I use first-differencing and estimate the following SFD specification:

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

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 cluster fixed-effects, θ, interacted with year fixed

[12] The previous specifications only include observations with strictly positive income. To test the robustness of the results, I use a Poisson regression, including all observations that reported zero income and use. As shown in Table 7 in Appendix B, the negative gradient is still statistically significant to the inclusion of observations that report zero income by using Poisson regression for the complete sample, whereas it is not statistically significant for the workers' subsample.

[13] See LAPOP's sample design: https://www.vanderbilt.edu/lapop/insights/IMN004en.pdf.

[14] See Figure 13 in Appendix C.

effects, η. Note that the SFD specification compares adjacent units within an intra-municipal cluster. Then, the intra-municipal cluster times year fixed-effect controls for between-cluster differences, and SFD purges common unobserved heterogeneity within an intra-municipal cluster.

Table 3 shows the coefficients for the SFD specification in equation (3) by adding multiple observable covariates in each column.[15] The SFD coefficients show that OLS estimates are biased from zero in absolute terms. The most robust specification for the complete sample in Column 7, using all available observable covariates and interviewer fixed-effects, shows that an increase on one darker skin tone correlates with a decrease of 3 percent. Using the subsample of workers, the most robust specification in Column 10 shows a semi-elasticity for skin tone and income is 3.9 percent.

To test the relative skin tome premium on income, I reproduce the most robust specification in Table 2, Column 7, but using dummy variables for each of the eleven skin tones using as reference the whitest skin tone. Figure 3 plots the SFD coefficients for each of the collapsed PERLA scale skin tones with respect to the whitest skin tone. Panel (a) shows the estimates for the complete sample, while panel (b) shows the results for the subsample of workers. Panel (a) shows a non-linear skin tone premium, where all skin tones but the second whitest and the darkest one have a statistically significant lower income than the whitest skin tone. Panel (b) shows that once accounting for labor market occupation, the only skin tones that have less income than the whitest skin tone is the eighth skin tone, with a semi-elasticity of minus 33, statistically significant at five percent, and the seventh skin tone, with semi-elasticities of minus 23, statistically significant at ten percent.

5.2 Robustness

To test the robustness of the results, I firstly use the SFD specification and three alternative measures of individual economic welfare: predicted income using household assets, a household asset index, and years of schooling. Figure 9 in Appendix A 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.

Secondly, I use the unobservable selection and coefficient stability approach (Altonji, Elder, and Taber 2005; Oster 2019) to test whether, after using the SFD research design, other unobservable factors are biasing my estimates. González and Miguel (2015) argue that the maximum R-squared is way below one in contexts where there is measurement error. I use the SFD specifications, equation (3), to compute the adjusted coefficients for different values of the maximum R-squared using the complete sample and the subsample with info on labor occupations and construct confidence intervals using bootstrapping.

[15]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.

Table 3: Skin tone premium on income: SFD

	Dependent variable:									
	Monthly income per capita (log)									
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
PERLA scale	-0.034***	-0.045***	-0.032***	-0.031***	-0.030***	-0.028***	-0.031***	-0.070***	-0.046***	-0.039***
	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.005)	(0.011)	(0.012)	(0.014)
PERLA scale (squared)	-0.001	0.000	0.000	-0.001	-0.001	-0.001	0.001	-0.003	-0.001	0.000
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.002)	(0.004)	(0.004)	(0.005)
Female		-0.294***	-0.285***	-0.270***	-0.270***	-0.189***	-0.238***		-0.273***	-0.266***
		(0.007)	(0.007)	(0.007)	(0.007)	(0.008)	(0.013)		(0.029)	(0.032)
Years of Schooling			0.054***	0.069***	0.068***	0.063***	0.072***		0.070***	0.070***
			(0.001)	(0.001)	(0.001)	(0.001)	(0.002)		(0.005)	(0.005)
Age				0.012***	0.012***	0.005***	0.010***		0.025***	0.025***
				(0.001)	(0.001)	(0.001)	(0.002)		(0.006)	(0.007)
Age (squared)				0.000**	0.000**	0.000	0.000		0.000***	0.000***
				(0.000)	(0.000)	(0.000)	(0.000)		(0.000)	(0.000)
1[Afro]					-0.034*	-0.035*	-0.073**		-0.010	-0.014
					(0.018)	(0.019)	(0.031)		(0.076)	(0.085)
1[Indigenous]					-0.110***	-0.102***	-0.121***		-0.243***	-0.198***
					(0.017)	(0.018)	(0.027)		(0.070)	(0.075)
1[Mulata]					-0.016	-0.021	-0.045		-0.008	-0.028
					(0.019)	(0.020)	(0.029)		(0.070)	(0.078)
1[Other]					-0.097***	-0.091***	-0.082***		-0.154***	-0.193***
					(0.017)	(0.018)	(0.024)		(0.060)	(0.067)
1[White]					-0.019*	-0.026**	-0.051***		-0.014	0.015
					(0.011)	(0.011)	(0.017)		(0.041)	(0.046)
Num.Obs.	85,362	85,362	83,783	83,783	83,783	76,747	33,587	7,927	7,174	6,834
R2	0.083	0.113	0.169	0.194	0.195	0.238	0.294	0.301	0.445	0.508
R2 Adj.	-0.227	-0.187	-0.117	-0.084	-0.082	-0.049	-0.019	-0.276	-0.057	-0.173
R2 Within	0.003	0.036	0.093	0.120	0.121	0.151	0.185	0.010	0.191	0.188
FE: Intra-Mun. Cluster × Year	X	X	X	X	X	X	X	X	X	X
Socio-demographic controls						X	X		X	X
Occupational controls								X	X	X
FE: Interviewer							X			X

Notes: Standard errors clustering by intra-municipality geographic cluster and year in parenthesis. Besides the coefficients in table, socio-demographic 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–. *p<0.1; **p<0.05; ***p<0.01

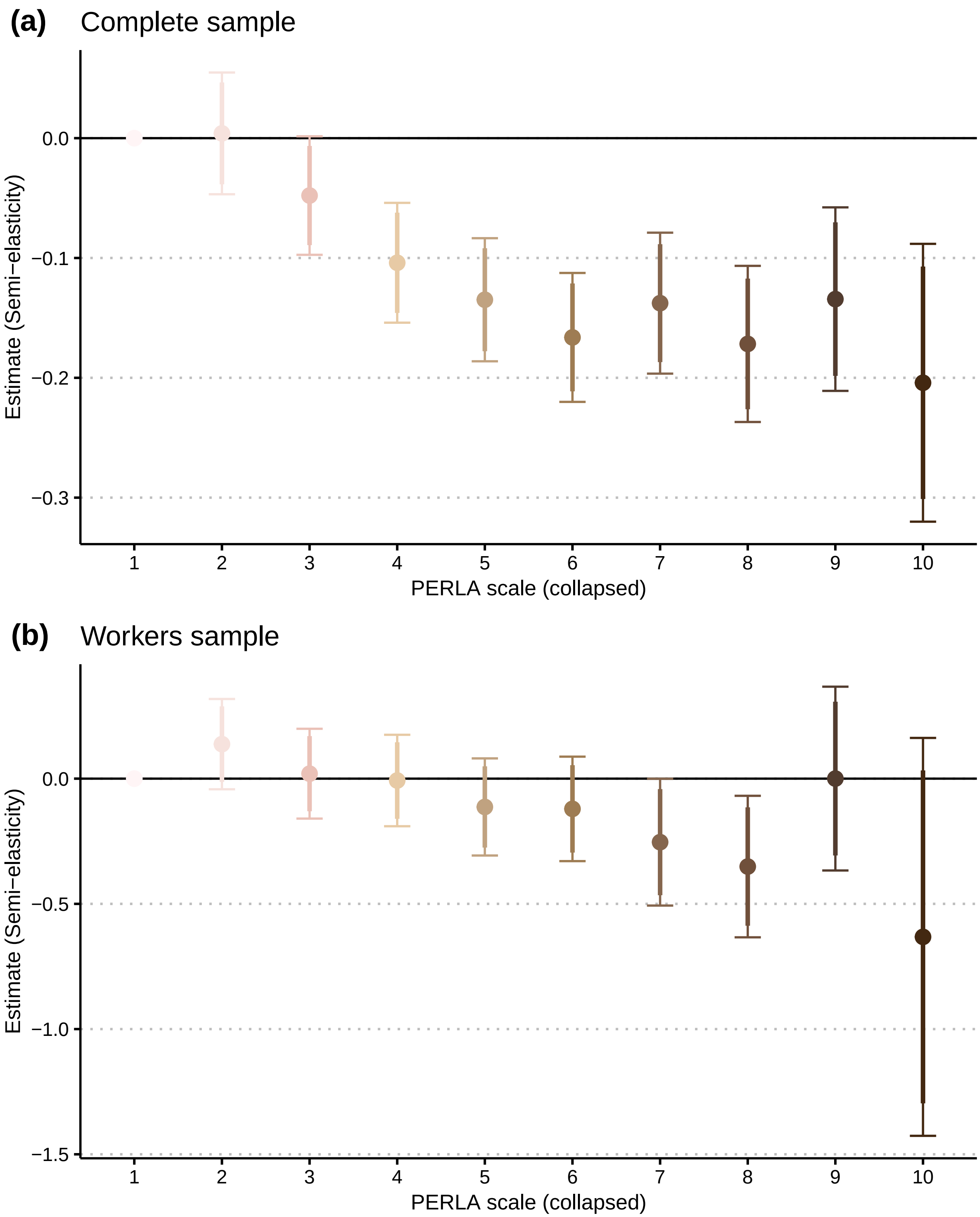

Figure 3: SFD Semi-elasticities by skin tone

Figure 10 in Appendix A 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 of workers. Thus, the latter is evidence to attend concerns for omitted variable bias on the skin tone effect on income.

Finally, I use Spatial Second Differences (SSD) (Druckenmiller and Hsiang 2018). SSD extends SFD: instead of using a one-spatial lag difference between neighboring units, it uses two-spatial lag differences. Druckenmiller and Hsiang (2018) argue that if there are remaining unobserved spatial heterogeneity biasing the results, one can further use spatial differences to filter out such variation. Table 8 in Appendix B replicates the specifications in each column in Table 3, but using SDD. In all specifications, the coefficient of skin tone on income is negative, and only in the specification for the subsample of workers with complete covariates, geographical fixed-effects, and interviewer fixed-effects, the coefficient is not statistically significant. The latter is most likely due to loss in statistical power given SDD drops even more observations for the second spatial lag.

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 has argued 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 innocuous, both for experimental and observational data (Rose 2022). To overcome such a problem, I combine an Oaxaca-Blinder (OB) decomposition for continuous variables proposed by Ñopo (2008b) with the SFD design to test whether racial discrimination explains the racial gap. See Appendix D for a detailed exposition of the methodology.

Figure 11 in Appendix A 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 through SFD, an increase in one darker skin tone decreases monthly income per capita by 3.2 percent. The OB discrimination estimate implies that 3.08 percent is due to discriminatory traits. Both coefficients are statistically significant at one percent. The latter results imply that differences in observed average characteristics cannot explain more than 90% of the gap ($\frac{\Delta^0_{OBD}}{\alpha_1} = \frac{-0.030}{-0.032} = 93.75\%$). I also compute the OBD-SFD for the subsample containing information on labor occupation. The racial gap semi-elasticity is minus 7.6 percent, where discrimination can explain almost 80% of the latter ($\frac{\Delta^0_{OBD}}{\alpha_1} = \frac{-0.060}{-0.076} = 78.94\%$). Thus, racial discrimination would explain between 80% and 90% of the skin tone premium, accounting for unobserved heterogeneity. To provide another set of evidence that confirms the racial discrimination hypothesis, I use LAPOP questions self-reported experiences of discrimination. Using SFD research design, Table

9 in Appendix B 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.

Figure 12 in Appendix A 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 –minus 0.030– for countries such as Uruguay, Guatemala, Paraguay, Bolivia, Argentina, and Brazil. Countries such as Ecuador, Peru, Mexico, Chile, El Salvador, Honduras, Colombia, and Venezuela, have a skin tone premium alongside the mean in the region. Finally, 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 Bolivia, Guatemala, Brazil, Chile, Argentina, Uruguay, Peru, Mexico, and Ecuador, but not different for the regional mean. If racial inequalities and discrimination are global phenomena, the heterogeneous effects show that their expressions are local.

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. Consistent with the historical and anecdotal evidence, racial discrimination drives the significant income 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. Furthermore, besides notions of justice, racial inequalities at the individual level have a critical role in aggregate economic development.

An essential shortcoming of the paper is the cross-sectional nature of the data. While the data disentangles inequalities from a cultural to a phenotype dimension, it fails to provide a more solid identification strategy. Moreover, I 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 social characteristics. 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.

There is a valuable room for public policies. 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. However, consistent with recent literature (Derenoncourt and Montialoux 2020; Derenoncourt et al. 2021), the results suggest that progressive income and wealth taxation is also progressive in racial disparities. 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.

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

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

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

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

Alesina, Alberto, 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-Vazquez. 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.

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/&%7Dcontext=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. 2020. "The higher price of whiter skin: an analysis of escort services." *Applied Economics Letters*. https://doi.org/10.1080/13504851.2020.1725229.

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

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

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.

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.

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

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

Cook, Lisa D. 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.

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

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.

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."

Garavito, César Rodríguez, Juan Camilo Cárdenas C., Juan David Oviedo M., and Sebastián Villamizar S. 2013. "La discriminación racial en el trabajo. Un estudio experimental en Bogotá."

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.

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.

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.

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.*

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.

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

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

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

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.

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

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

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

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

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

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.

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

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

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

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

———. 2020. “Social identity and economic policy.” Annual Reviews Inc. https://doi.org/10.1146/annurev-economics-082019-110313.

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

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 Liza Steele. 2012. “Pigmentocracia en las Américas: ¿cómo se relaciona el logro educativo con el color de piel?” Perspectivas Desde El Barómetro de Las AMéricas.

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

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.

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.

Appendix

A Figures

Figure 4: PERLA Palette

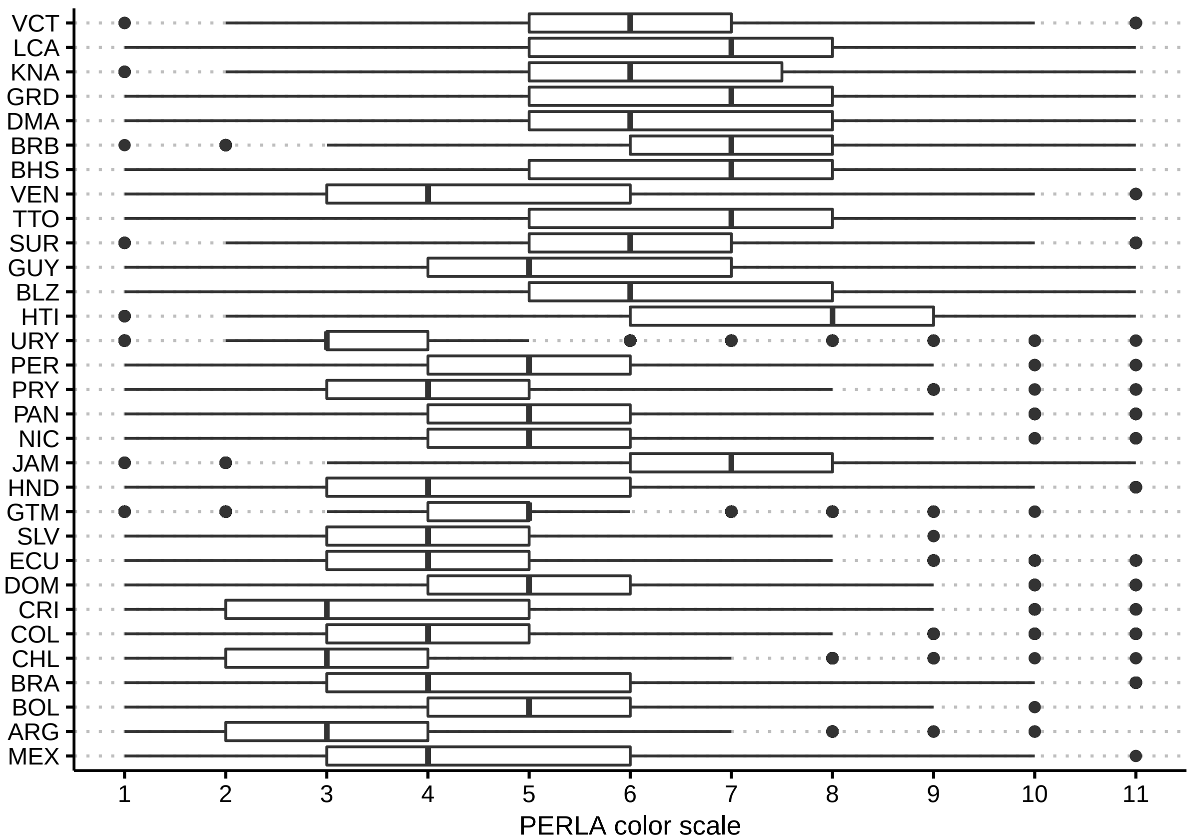

Figure 5: PERLA skin tone boxplot by country

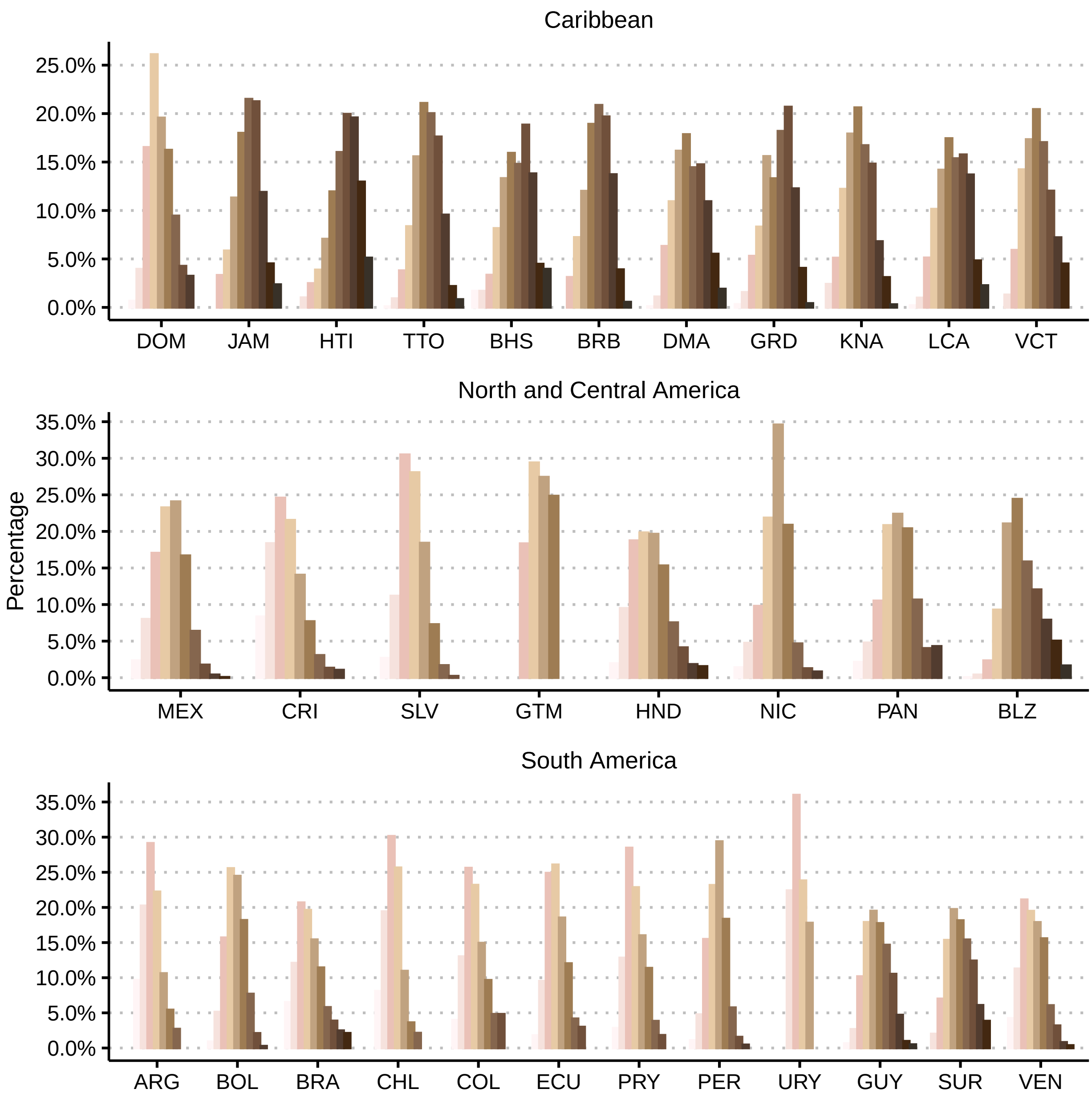

Figure 6: Skin tone distribution by country

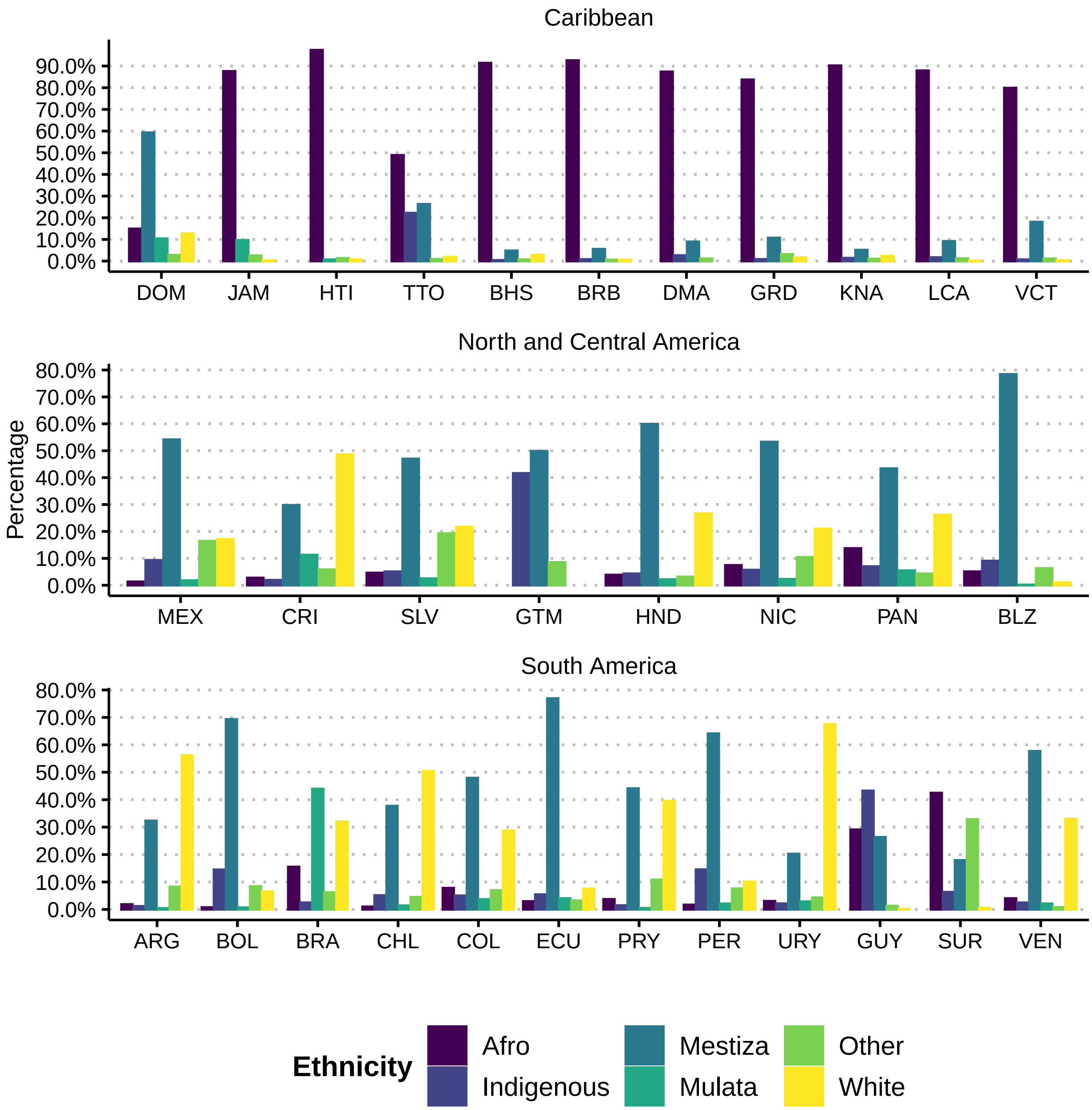

Figure 7: Ethnicity distribution by country

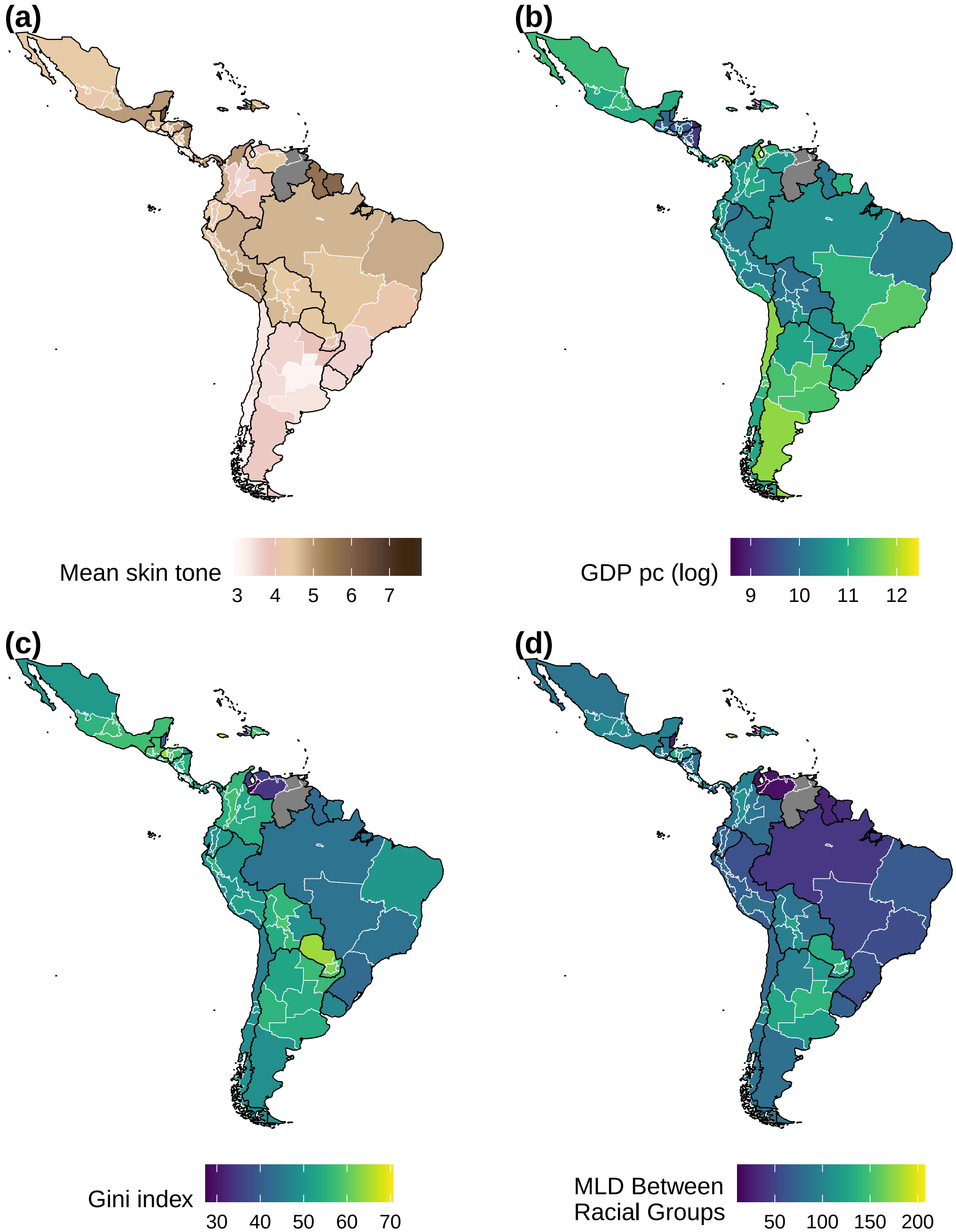

Figure 8: LAPOP Stratified Regions

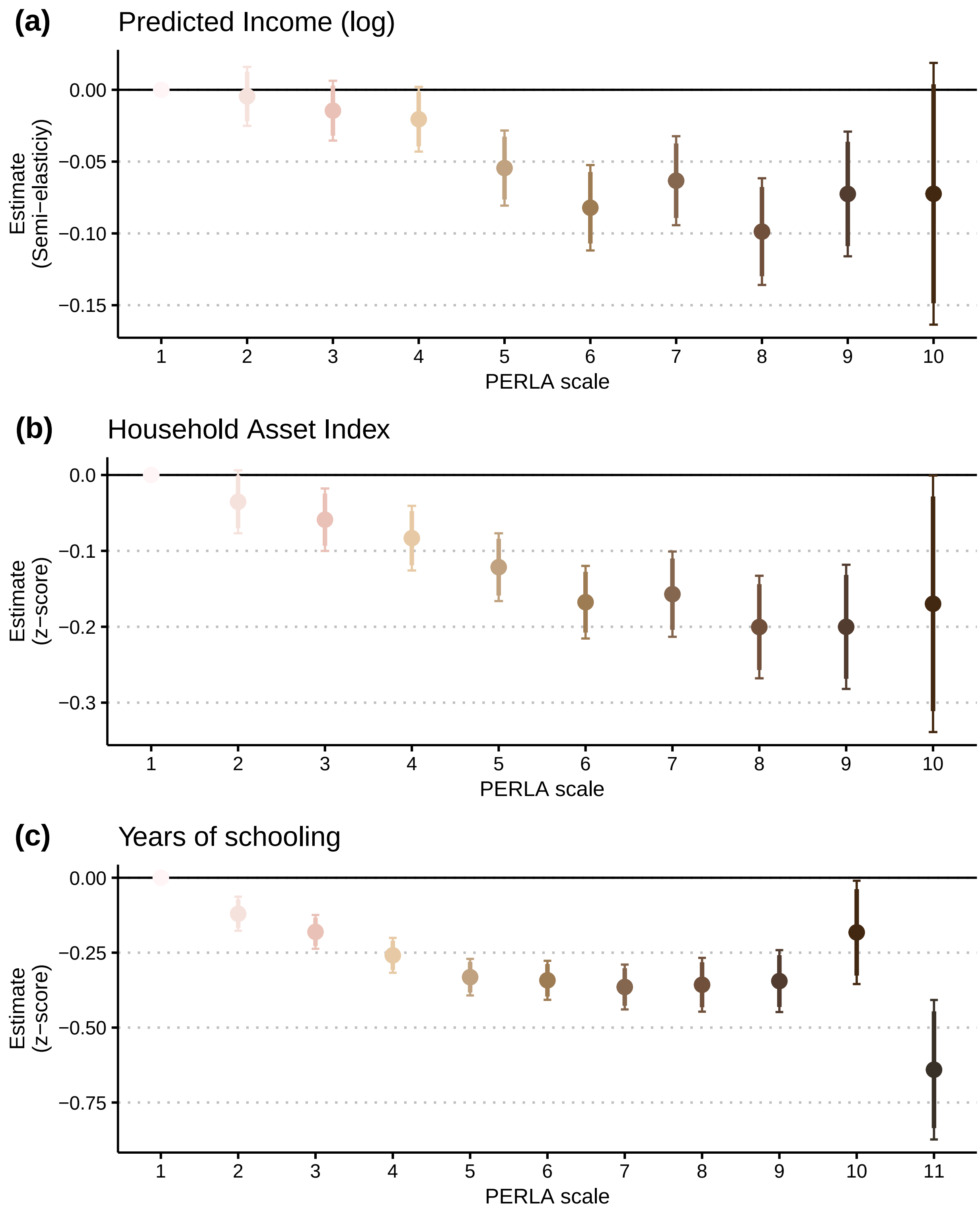

Figure 9: SFD Semi-elasticities by skin tone: Alternative measures

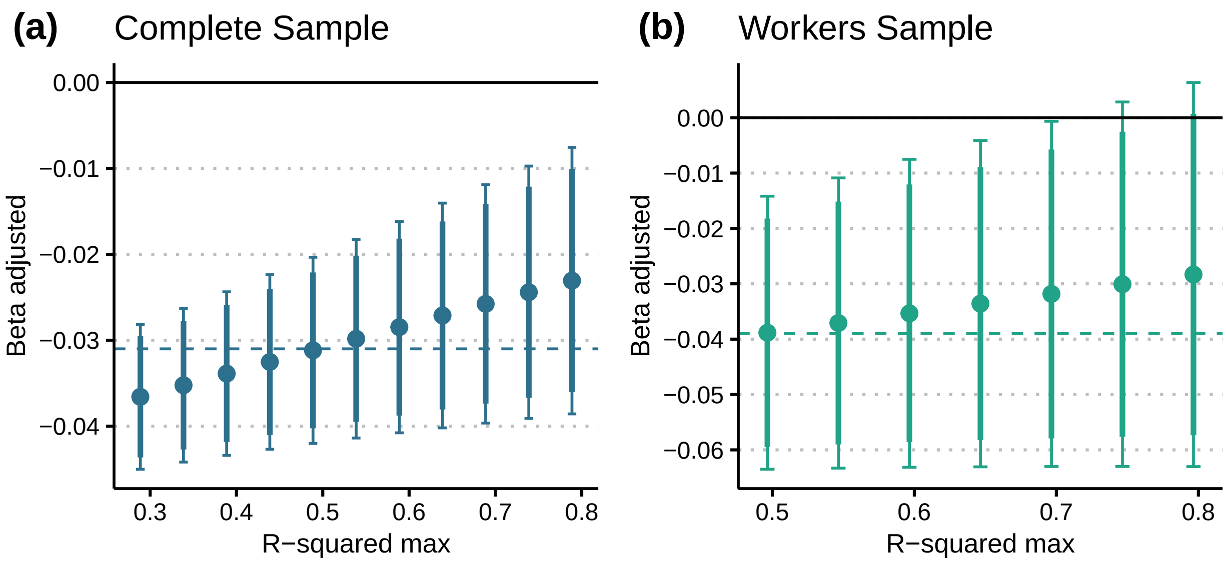

Figure 10: Unobservable selection and coefficient stability

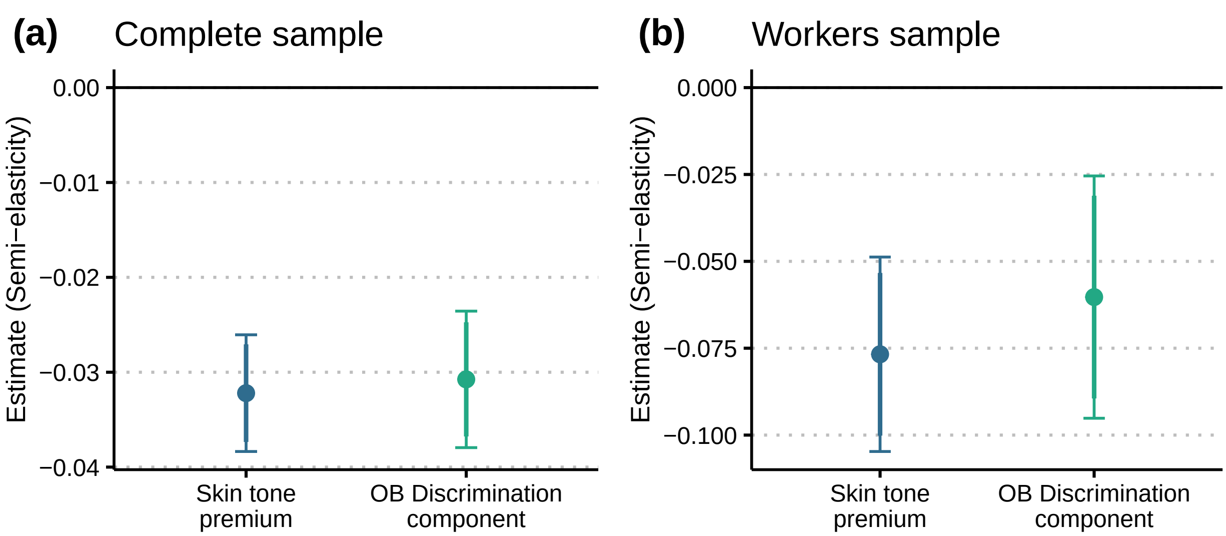

Figure 11: Oaxaca-Blinder decomposition with SFD

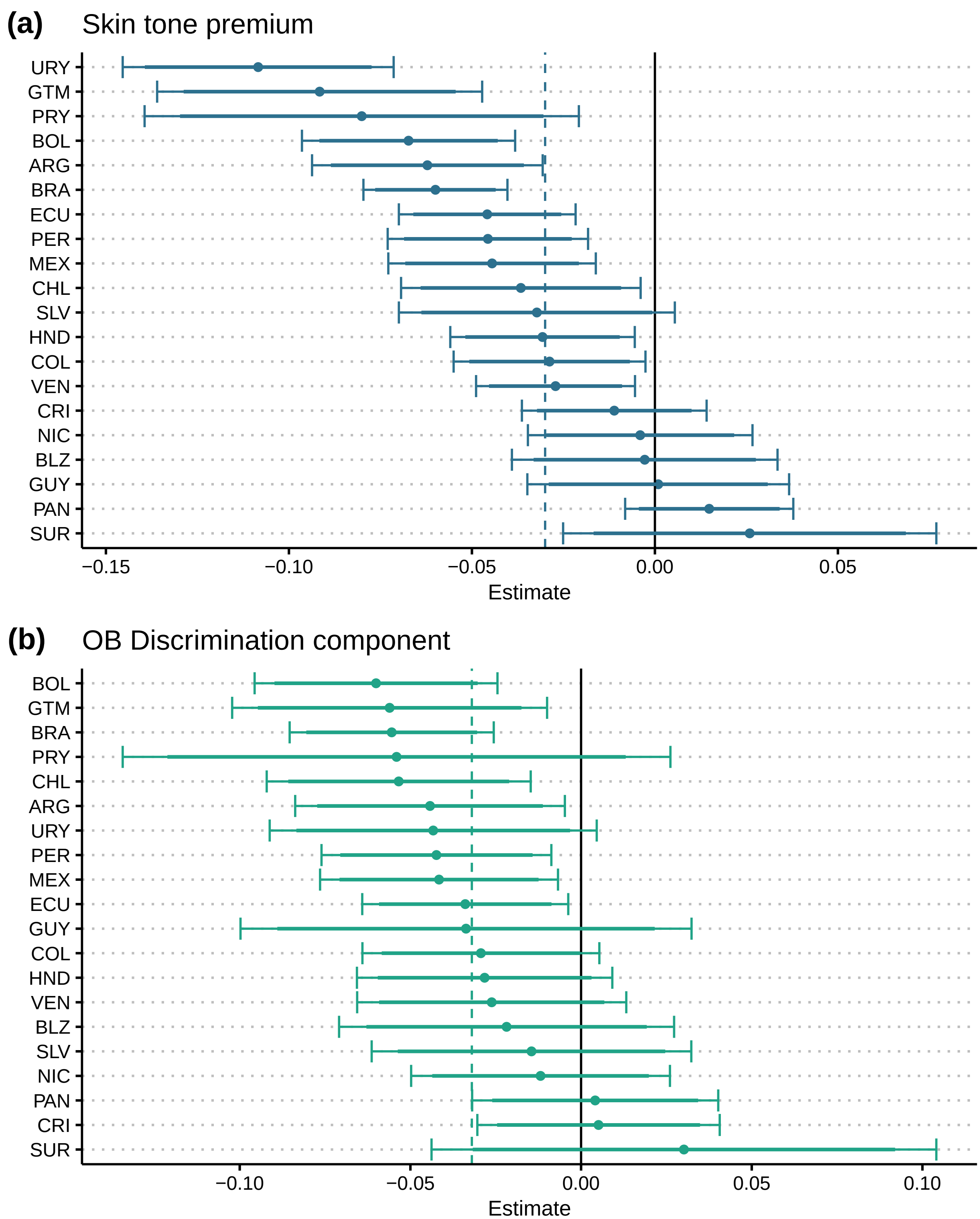

Figure 12: OB decomposition by country

B Tables

Table 4: 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

Table 5: 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
Monthly Income per capita (PPP 2019)	251.97	297.74	1.22	75.59	314.40	2,955.34
Predicted Income per capita (PPP 2019)	219.15	139.94	0.09	119.02	289.85	808.12
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

Table 6: Skin tone premium on income: OLS

	Dependent variable:						
	Monthly income per capita (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
Num.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. Cluster × Year							X

Notes: Standard errors clustering by intra-municipality geographic cluster and year in parenthesis. *p<0.1; **p<0.05; ***p<0.01

Table 7: Skin tone premium on income: Poisson

	Dependent variable:									
	Monthly income per capita									
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
PERLA scale	-0.077***	-0.085***	-0.063***	-0.052***	-0.050***	-0.056***	-0.080***	-0.085**	-0.036	-0.031
	(0.010)	(0.010)	(0.010)	(0.010)	(0.011)	(0.011)	(0.017)	(0.033)	(0.031)	(0.034)
PERLA scale (squared)	0.004***	0.004***	0.004***	0.002**	0.002**	0.003**	0.006***	0.005	0.002	0.001
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.002)	(0.004)	(0.003)	(0.004)
Female		-0.280***	-0.272***	-0.269***	-0.268***	-0.178***	-0.233***		-0.257***	-0.267***
		(0.007)	(0.007)	(0.007)	(0.007)	(0.008)	(0.012)		(0.022)	(0.023)
Years of Schooling			0.061***	0.072***	0.072***	0.064***	0.071***		0.055***	0.054***
			(0.001)	(0.001)	(0.001)	(0.001)	(0.002)		(0.004)	(0.004)
Age				0.022***	0.022***	0.011***	0.012***		0.022***	0.020***
				(0.001)	(0.001)	(0.001)	(0.002)		(0.004)	(0.005)
Age (squared)				0.000***	0.000***	0.000***	0.000**		0.000***	0.000***
				(0.000)	(0.000)	(0.000)	(0.000)		(0.000)	(0.000)
1[Afro]					-0.009	0.002	-0.043		-0.131**	-0.114*
					(0.020)	(0.020)	(0.029)		(0.057)	(0.061)
1[Indigenous]					-0.088***	-0.081***	-0.065**		-0.111**	-0.099*
					(0.021)	(0.021)	(0.028)		(0.050)	(0.051)
1[Mulata]					0.008	0.007	0.020		-0.053	-0.025
					(0.020)	(0.020)	(0.027)		(0.060)	(0.060)
1[Other]					-0.125***	-0.117***	-0.115***		-0.136***	-0.108**
					(0.018)	(0.018)	(0.024)		(0.047)	(0.050)
1[White]					-0.014	-0.017	-0.015		-0.037	-0.034
					(0.011)	(0.011)	(0.016)		(0.029)	(0.031)
1[Agriculture]									-0.145***	-0.164***
									(0.051)	(0.052)
1[Directors and managers]									0.346***	0.364***
									(0.080)	(0.081)
1[Elemental occupations]									-0.058	-0.110***
									(0.037)	(0.040)
1[Military]									0.320***	0.319***
									(0.118)	(0.115)
1[Artisans]									0.025	0.013
									(0.037)	(0.040)
1[Mechanics]									0.037	-0.010
									(0.061)	(0.058)
1[Administrative personnel]									0.178***	0.166***
									(0.039)	(0.041)
1[Scientists and intellectuals]									0.285***	0.285***
									(0.046)	(0.049)
1[Technicians and professionals]									0.213***	0.200***
									(0.039)	(0.042)
Num.Obs.	139,823	139,823	137,937	137,937	137,937	130,189	52,018	13,495	12,671	12,320
R2 Pseudo	0.674	0.680	0.695	0.703	0.703	0.722	0.659	0.650	0.708	0.731
FE: Intra-Mun. Cluster × Year	X	X	X	X	X	X	X	X	X	X
Socio-demographic Controls						X	X		X	X
FE: Interviewer							X			X

Notes: Standard errors clustering by intra-municipality geographic cluster and year in parenthesis. Besides the coefficients in table, socio-demographic 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. *p<0.1; **p<0.05; ***p<0.01

Table 8: Skin tone premium on income: SSD

	Dependent variable:									
	Monthly income per capita (log)									
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
PERLA scale	-0.033***	-0.044***	-0.030***	-0.029***	-0.029***	-0.028***	-0.029***	-0.071***	-0.037**	-0.014
	(0.003)	(0.003)	(0.003)	(0.003)	(0.004)	(0.004)	(0.007)	(0.017)	(0.019)	(0.023)
PERLA scale (squared)	-0.002***	-0.002***	-0.002**	-0.001**	-0.001**	-0.001	-0.001	-0.002	0.002	0.004
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.004)	(0.005)	(0.006)
Female		-0.292***	-0.282***	-0.268***	-0.267***	-0.191***	-0.246***		-0.285***	-0.265***
		(0.009)	(0.009)	(0.009)	(0.009)	(0.010)	(0.017)		(0.045)	(0.053)
Years of Schooling			0.055***	0.069***	0.068***	0.063***	0.071***		0.077***	0.075***
			(0.001)	(0.001)	(0.001)	(0.001)	(0.002)		(0.007)	(0.009)
Age				0.011***	0.011***	0.005***	0.011***		0.015	0.013
				(0.001)	(0.001)	(0.002)	(0.003)		(0.009)	(0.011)
Age (squared)				0.000	0.000	0.000	0.000		0.000	0.000
				(0.000)	(0.000)	(0.000)	(0.000)		(0.000)	(0.000)
1[Afro]					-0.028	-0.031	-0.089**		0.094	0.018
					(0.023)	(0.024)	(0.040)		(0.125)	(0.150)
1[Indigenous]					-0.100***	-0.096***	-0.111***		-0.205*	-0.158
					(0.021)	(0.023)	(0.036)		(0.117)	(0.139)
1[Mulata]					-0.020	-0.027	-0.064*		-0.068	-0.068
					(0.024)	(0.025)	(0.036)		(0.099)	(0.117)
1[Other]					-0.097***	-0.084***	-0.067**		-0.026	0.011
					(0.021)	(0.023)	(0.030)		(0.092)	(0.113)
1[White]					-0.021	-0.030**	-0.059***		0.041	0.085
					(0.014)	(0.014)	(0.023)		(0.060)	(0.075)
Num.Obs.	63,793	63,793	62,144	62,144	62,144	54,928	24,134	4,346	3,779	3,520
R2	0.095	0.126	0.185	0.209	0.210	0.253	0.323	0.359	0.505	0.615
R2 Adj.	-0.317	-0.272	-0.193	-0.159	-0.157	-0.126	-0.119	-0.452	-0.175	-0.577
R2 Within	0.003	0.037	0.096	0.122	0.123	0.153	0.188	0.011	0.209	0.200
FE: Intra-Mun. Cluster × Year	X	X	X	X	X	X	X	X	X	X
Socio-demographic controls						X	X		X	X
Occupational controls								X	X	X
FE: Interviewer							X			X

Notes: Standard errors clustering by intra-municipality geographic cluster and year in parenthesis. Besides the coefficients in table, socio-demographic 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–. *p<0.1; **p<0.05; ***p<0.01

Table 9: Self-reported measures of Racial discrimination: SFD

	Dependent variable:		
	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: Standard errors clustering by intra-municipality geographic cluster and year in parenthesis. *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.

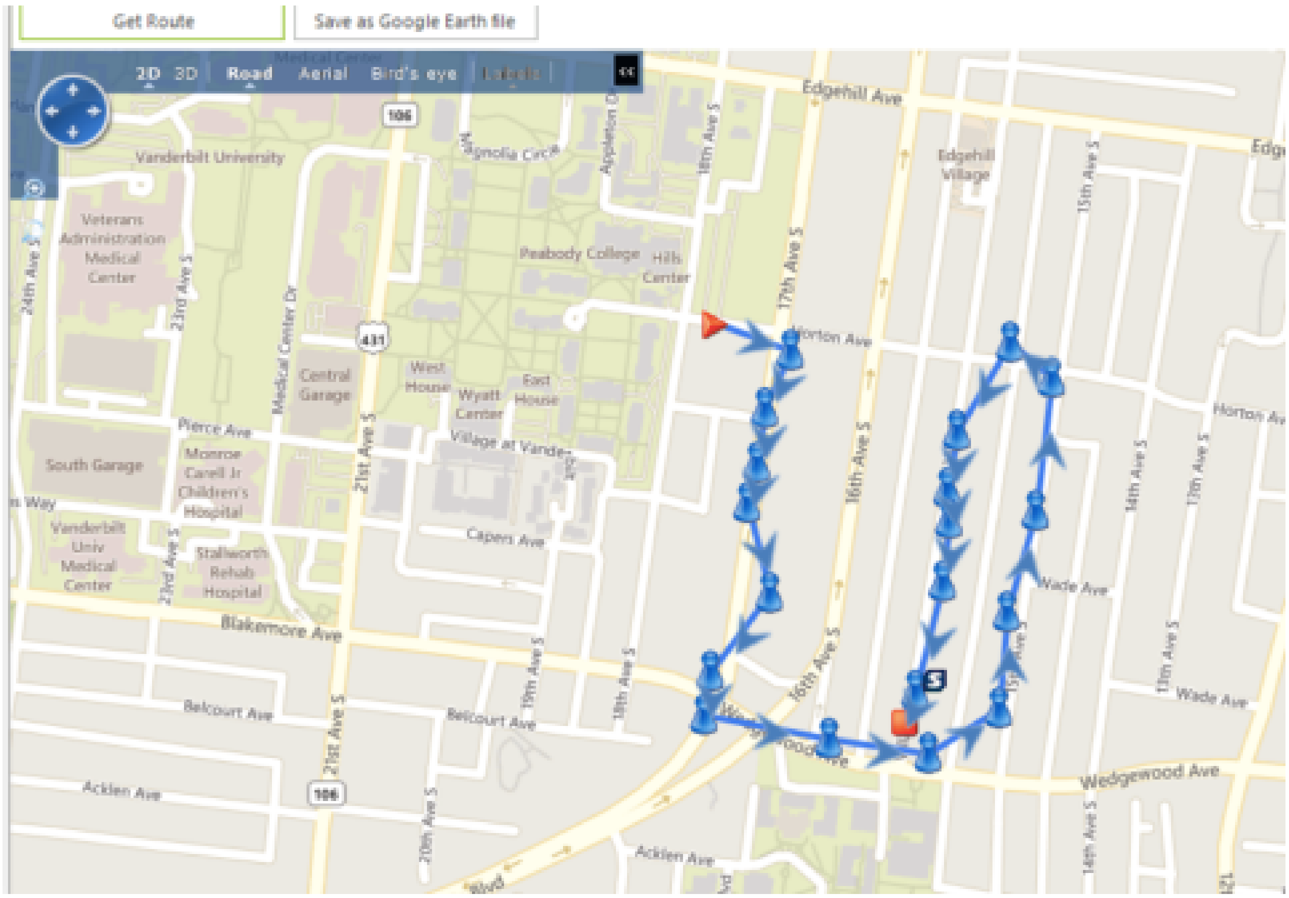

Figure 13: Spatial proximity of LAPOP interviewed units

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 2008b).

Following Ñopo (2008b), 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 (2008b) 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:[16]

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 $E[x]$.[17]

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.

[16] 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.

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