Environmental justice beyond race: Skin tone and exposure to air pollution

Sandra Aguilar-Gomez[a,1], **Juan Camilo Cardenas**[a,b], **and Ricardo Salas Diaz**[c]

This manuscript was compiled on November 22, 2024

Driven by environmental justice activism and policy reforms, recent social science research, mostly focused on the United States and Western Europe, shows that marginalized communities face greater environmental degradation. However, the ethnoracial categories used in these studies may not fully capture environmental inequality in the Global South. This study presents novel findings that quantify and examine potential mechanisms behind the link between skin tone and ambient air pollution exposure in Colombia, moving beyond conventional ethnoracial variables. By matching household geolocations from a large-scale longitudinal survey with satellite-based pollution estimates, we find that skin tone predicts both initial pollution levels and their changes over time. Although average pollution levels remained relatively stable during our study period, the environmental justice landscape in Colombia underwent a complete transformation. In 2010, lighter-skinned individuals faced higher particulate pollution exposure, but darker-skinned individuals experienced much steeper increases in the following years. By 2016, the environmental justice gap had reversed, with those with the darkest skin tones exposed to nearly one standard deviation higher pollution levels. Regression analyses reveal that these patterns are robust to the inclusion of a comprehensive set of theoretically relevant covariates, including ethnoracial self-identification and income. Decomposition analyses suggest that fire exposure from biomass burning and other human activities, as well as local collective action, account for a portion of observed disparities. With climate change expected to intensify fire incidence, the disproportionate disease burdens that vulnerable groups face might deepen unless policy measures are taken to reverse this trend.

Environmental justice | Air pollution | Skin tone | Colombia

Significance Statement

Each year, outdoor air pollution claims three million lives worldwide. In Colombia, particulate matter smaller than 2.5 μm ($PM_{2.5}$) levels consistently exceed WHO guidelines, with severe health consequences. Environmental justice (EJ) studies typically focus on broad ethnoracial categories, but recent research suggests skin tone better captures racialized inequality in the Global South. This study is the first to use skin tone to measure environmental disparities. Using satellite-based pollution estimates and a longitudinal household survey, it uncovers Colombia's EJ landscape transformation. While individuals with darker skin tones initially experienced lower pollution levels, by 2016, they faced significantly worse air quality. Decomposition analysis reveals that urbanization, fire pollution, income, and local collective action account for a substantial portion of observed disparities.

Air pollution exposure has strong and enduring negative health and human capital impacts (1–3). Exposure to particulate matter smaller than 2.5 μm ($PM_{2.5}$) is robustly associated with numerous health issues, including respiratory and cardiovascular disease and premature death (4–6). Pollution exposure negatively impacts labor productivity, cognitive performance and academic achievement, ultimately reducing economic well-being (7). For these reasons, the World Health Organization (WHO) has set guidelines for safe levels of $PM_{2.5}$, recommending that the mean annual concentration of $PM_{2.5}$ stay below 5 $\mu g/m^3$ (6). In Colombia, the setting of this study, the economic costs of particulate pollution are estimated to account for 1.5% of GDP (8). In Latin America and other areas in the world, many chronic conditions and mortality are often distributed along inequitable societal lines, with populations of lower socioeconomic status bearing the brunt of the disease burden (9–11), but little is known about the contribution of environmental inequalities to such disparities.

Research and activism have brought attention to the disproportionate pollution exposure suffered by marginalized communities (12–14), with growing concern for particulate matter. In recent years, social science research has made significant strides in understanding the drivers and consequences of the unequal exposure of racialized and economically disadvantaged groups to air pollution, an inequality sometimes referred to as the environmental justice (EJ) gap (15–17). Disproportionate exposure $PM_{2.5}$, has been hypothesized to explain a significant portion of the racial income gap (18, 19). However, much of this research has focused on ethnic or racial categories specific to the historical and institutional context of the United States and thus may not be applicable to most of the world.

In many regions of the Global South, institutionalized racism is deeply entrenched in urban planning and environmental policies (20–22). Social stratification in these regions, however, does not necessarily conform to the census-style ethnoracial categories commonly used in surveys and the environmental inequality literature. Scholars specializing in the Global South argue that evaluating phenotype, with

Author affiliations: [a]Economics Department, Universidad de los Andes; [b]Economics Department, University of Massachusetts, Amherst; [c]Economics Department, Dartmouth College

The authors declare no conflict of interest

[1]To whom correspondence should be addressed. E-mail: s.aguilargomez@uniandes.edu.co

an emphasis on skin tone, might be a more apt metric of racialized stratification (23–25). Currently, many surveys in Latin America employ the PERLA (Project on Ethnicity and Race in Latin America) scale, which categorizes skin tones into a set of discrete shades, ranging from light to dark. Nonetheless, skin tone has not yet been explored as a significant variable in environmental stratification studies.

In this paper, we employ a unique geo-referenced longitudinal household survey and satellite-based air pollution estimates to produce novel findings quantifying the link between skin tone, measured by the PERLA scale, and exposure to ambient air pollution in Colombia. Previous research in Latin America and elsewhere has focused on poverty and ethnic categories, making this the first study of its kind. We document a transformation of the environmental justice landscape in Colombia. The analysis indicates that while, on average, particulate matter exposure in our sample decreased insignificantly between 2010 and 2016 (-0.01 $\mu g/m^3$), this pattern masks substantial disparities. Skin tone and ethnoracial self-identification robustly predict different trajectories of environmental quality, with a larger proportion of people with darker skin tones exposed to air quality deterioration. Notably, while people with the darkest skin started at a lower level of pollution exposure, this group experienced a 1.6 $\mu g/m^3$ increase in $PM_{2.5}$ exposure, an increase amounting to 8% from the 2010 sample mean. Thus, by the end of our study period, the association between skin tone and particulate exposure completely reverts. This phenomenon is present in both urban and rural settings, but stronger for the latter. It is consistent for households who changed residences and those who did not move, ruling out residential sorting as a major driver. By contrasting the results obtained when we use a continuous measure of skin tone and a categorical race variable in ELCA, our study highlights the limitations of traditional racial categories for analyses of environmental injustice in the Global South.

Our findings provide descriptive evidence of potential mechanisms behind these results. In 2010, fire exposure accounted for just 2% of the variation in pollution exposure in our sample, as other urban factors were more dominant predictors of pollution during that time. In that year, regulation was passed establishing air quality guidelines—including for $PM_{2.5}$—and mandating monitoring of this pollutant in a network that, even today, remains concentrated in urban areas.*. Over the following decade, pollution abatement efforts concurred with two relevant processes. First, the growing influence of climate change on fire activity, which has been identified as relevant for Colombia (26–28). Second, the peace process consolidated in 2014 and has since reshaped the geography of coca cultivation and cattle ranching—both activities frequently involving deforestation through fire (29). Consistent with these large-scale phenomena, fire exposure in 2016 explained nearly 20% of particulate pollution in the sample, highlighting the increasing influence of rural factors on overall pollution levels.

Using satellite-based estimations of fire intensity (30, 31) and wind direction (32), our findings provide suggestive evidence that fire exposure is a significant contributor to the EJ gap. We observe a positive gradient between skin pigmentation and fire pollution exposure in both years. Furthermore, employing standard decomposition techniques reveals that upwind fire exposure plays a role in both the composition and returns components of the EJ gap across both years. The decomposition analysis also highlights that local collective action, urbanization, and, to a lesser extent, socioeconomic status, contribute to the gap. Together with fires and migration, these variables explain two-thirds of the 2016 EJ gap.

Our findings are consistent with other works documenting that analyses considering ethnoracial categories and skin tone can play a complementary role to advance our understanding of the complexities of race and discrimination in Latin America, with skin tone capturing more significant variance (33, 34).† Also congruently with other works looking at nonenvironmental stratification, class is a confounding factor that holds significance in predicting outcomes, but it does not fully account for skin tone differences. The findings presented in this study indicate that further research investigating environmental justice in Colombia should closely examine whether the abatement policies and sociopolitical change has disproportionately benefited certain segments of the population while neglecting others.

Data and Context

Data Sources. The main analyses in this study utilize individual-level georeferenced data from the ELCA, a nationally representative longitudinal household survey that tracked 10,000 Colombian households every three years between 2010 and 2016, recording their exact coordinates in each round. In the 2013 survey round, a PERLA scale module recording household members' skin tone was included. We impute the average skin tone of household adults to nonreported and unobserved children and group PERLA skin tones into seven categories. Detailed descriptions of such imputation are available in SI Appendix Figures S1 and S2.

The panel structure of the survey allows us to examine migration patterns and the evolution of pollution exposure levels. Pollution estimates are obtained from the Atmospheric Composition Analysis Group (ACAG) at Washington University in St. Louis (36). ACAG estimates are produced by a combination of Aerosol Optical Depth (AOD) retrievals from various NASA satellite instruments with a chemical transport model, which is then calibrated using ground-based observations. This publicly available resource provides raster data with global coverage at a spatial resolution of 0.01 degrees. We utilize annual estimates of average ground-level fine particulate matter ($PM_{2.5}$) for 2010, 2013 and 2016.

Our analysis calculates pollution exposure by averaging the cells of the $PM_{2.5}$ raster within 1 km, 5 km, and 10 km buffers around households' coordinates. The variation in buffer radius helps ensure the robustness of our results. This method has been widely used in recent environmental justice literature (17, 37). An implicit assumption is that the place of residence is a relevant location to calculate pollution exposure. While it would be ideal to also capture workplace-based exposure, ELCA does not contain such information.

*Resolution 610 from 2010

†While there are significant variations in skin color among the populations of the Americas, countries such as Mexico, Brazil, Bolivia, Ecuador, and Peru exhibit patterns in skin tones similar to those observed in Colombia. Ethnoracial categories are more related to skin color in some countries, such as Panama, than in others. However, both variables generally capture distinct information (35).

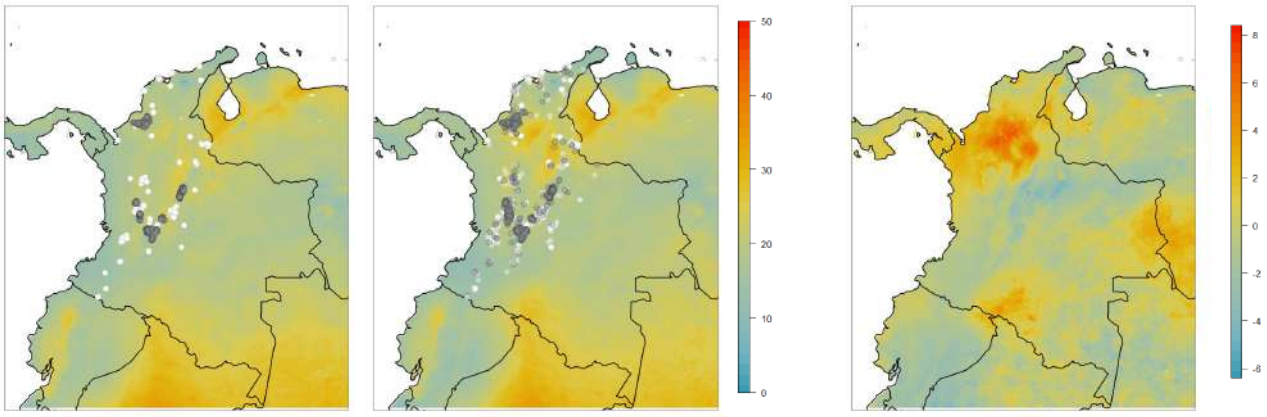


Fig. 1. $PM_{2.5}$ and geographic location of ELCA households

Notes: Variations in Colombia's $PM_{2.5}$ ($\mu g/m^3$) and ELCA households' location. Panels A and B show annual mean $PM_{2.5}$ concentrations in 2010 and 2016. Highlighted circles present the location of surveyed urban households in white and surveyed rural households in gray. Panel C shows the variations in annual mean $PM_{2.5}$ concentrations between 2010 and 2016.

To assess the role of fire pollution exposure on EJ patterns, fire exposure at the household level is measured as the intensity- and probability-weighted yearly upwind fire counts within a 50 km radius. Fire intensity and probability are obtained from the Fire Information for Resource Management System (FIRMS), based on VIIRS-MODIS products (30, 31). Wind patterns are analyzed using the ERA5 and ERA5-Land datasets (38). Yearly fire counts are weighted by intensity and probability, with fire intensity measured in radiance temperature (K), and only high-confidence fires (over 60% probability) are included. Fires are classified as upwind based on prevailing wind direction within 90-degree quadrants, following previous works (39, 40).

Urban and rural fire exposure is calculated separately to distinguish between biomass burning in agriculture and forest fires (rural) and fire activity more likely involving houses or infrastructure (urban). Urban fires are defined broadly, including peri-urban and rural areas as long as the fire occurred within the boundaries of a populated center, defined as any conglomerate of at least 20 contiguous residences. Since we are employing a 50 km radius as our boundary for area of impact of fires, following previous work (39, 40), we allow for rural households to be impacted by an urban fire and vice-versa if they are located downwind from the fire and fall within this perimeter.

Technical details regarding the PERLA scale and ELCA are available in the SI Appendix, Section Data Sources, along with a detailed explanation of the ACAG data and fire exposure assessment. Descriptive statistics are provided in SI Appendix Tables S1, S2, S3, and S4.

Context. The average individual in the sample experienced a slight, statistically insignificant decrease in exposure to $PM_{2.5}$ between 2010 and 2016. Specifically, urban households experienced an average increase of -0.1 $\mu g/m^3$. Meanwhile, the reduction for rural households was 0.11 $\mu g/m^3$ (refer to SI Appendix Table S3).

Figure 1 provides a visual representation of the change in annual particle concentrations ($\mu g/m^3$) during this period. Additionally, the figure displays the geographic distribution of observations from households participating in ELCA. The concentration levels 5km around the ELCA households each year range from 12.4 to 28.5 $\mu g/m^3$, and the intertemporal change ranges from -6.8 to 9.3 $\mu g/m^3$. It is notable that in 2010 and 2016, over 98% of these households were exposed to annual average concentrations greater than 15 $\mu g/m^3$, three times the current WHO guideline (41). SI Appendix Table S4 classifies ELCA households by their $PM_{2.5}$ exposure levels in 2010 and 2016, following WHO guidelines and interim targets.

Despite government efforts to regulate and enforce air quality standards, average air quality in Colombia has not improved over the past decade. The causes of sub-optimal air quality vary between urban and rural areas, although biomass burning is a dominant factor nationwide, accounting for 75% of particulate pollution. Emissions inventories for the country's urban centers indicate that particulate matter primarily arises from fossil fuel combustion, motor vehicle transportation, and industrial activities. In Colombian cities, 80% of $PM_{2.5}$ emissions are attributed to mobile sources, with the remaining 20% coming from stationary sources (8). Recent studies (42) have highlighted the increasing contribution of motorbikes to urban pollution levels. Additionally, unfavorable topography and meteorological conditions hinder the proper dispersion of pollutants in the atmosphere (8). Improper waste disposal practices, especially uncontrolled waste burning, remain a significant cause of air pollution, particularly in urban peripheries and rural areas where settlements lack adequate waste disposal services. Recent migration and displacement have further exacerbated this problem (43, 44).

In rural Colombia and neighboring countries, several factors contribute to high levels of air pollution. Open biomass burning, including agricultural and forest fires, is a major source of pollution in South America, making the region one of the largest global contributors to such emissions (45, 46). In Colombia, deforestation—often linked to illicit activities such as coca cultivation and trafficking—aggravates

the problem, as land clearing for both legal and illegal crops often involves intentional fires (47, 48). The peace negotiations and FARC's ceasefire in 2014 further accelerated deforestation and altered conservation dynamics (49, 50). Moreover, extractive industries such as open-pit coal mining and quarrying have significantly deteriorated air quality across numerous municipalities (51). Pollution from nearby urban centers also affects rural populations, compounding the issue. Climate change intensifies these challenges by increasing the frequency of forest fires, further raising particulate matter ($PM_{2.5}$) concentrations (27, 28).

Methodology

To measure the transformation of environmental inequality in Colombia, we quantify and decompose the skin-tone differences in pollution exposure observed in 2010, 2016 and the inter-temporal change during this period. Without any additional covariates, these relationships are of academic and policy interest because they reveal that skin tone is a relevant axis of environmental stratification in Colombia. In the environmental justice literature, researchers are interested in measuring the robustness of the inequitable exposure finding to the inclusion of statistical controls. In particular, a number of studies have questioned whether race is still significant after conditioning on income and proxies for wealth, or other local characteristics that might drive firms' production decisions or governments' environmental policies, even if some of these characteristics are themselves a consequence of race (17, 52, 53).

Quantification of the environmental justice gap. Our econometric specifications estimate an ordinary least squares regression using as our dependent variable estimated $PM_{2.5}$ exposure, which may be 2010 levels, 2016 levels, or the change between 2010 and 2016. Our independent variables of interest are a series of skin tone indicators. Specifically, we estimate the following equation:

$$Y_{it} = \alpha_0 + \sum_{j \in [3,7]} \alpha_{jt} \mathbb{1}[PERLA_i = j] + \chi_{it} + \epsilon_{it} \qquad [1]$$

where Y_{it} is the average pollution measurement of interest in a 5-km buffer around the residence of individual i. These regressions are estimated separately for $t = 2010, 2016, 2016 - 2010$. $\mathbb{1}[PERLA_i = j]$ represents an indicator variable for the PERLA skin tone categories as defined in this study, $j \in [3, 7]$. The omitted category corresponds to individuals with lighter skin (tones 1 and 2 on the PERLA scale). The coefficients α_{jt} measure the relationship between $PM_{2.5}$ exposure at time t.

Theoretically relevant covariates from the EJ literature are then progressively added to χ_i. First, to insert our analyses into the "race versus class" debate (53), we include monthly household income in t as well as an indicator of at least one unmet need. Second, as we are interested in showing that skin tone is a robust measure of environmental stratification regardless of ethnoracial self-identification, we include enthoracial indicators. Degree of urbanization in municipality of residence is also accounted for.

We also include migratory status, indicating if the have changed municipalities in the past five years. Internal migration in Colombia is driven by the availability of work, the need to mitigate the impact of climatic events, and the influence of the armed conflict. Indigenous and Afro-descendant populations are overrepresented among the displaced population (54), and are more likely to be displaced by the Colombian conflict than to migrate for economic reasons (55), potentially constraining there relocation choices with implications for the environmental quality they have access to. Urban and rural upwind fire pollution exposure in a 50 km radius is included. Most recent election turnout is also included as a control, as it has been conceptualized as a proxy for local collective action and found to be a predictor of pollution exposure and location of hazardous waste facilities (56, 57). The SI Appendix describes the covariate construction in greater detail.

Decomposition of the environmental justice gap. To further dissect the observed disparities, we employ the Kitagawa–Oaxaca–Blinder decomposition (KOBD) framework (58–60). This approach involves a series of counterfactual exercises, such as estimating how the EJ gap would look if people with darker skin tones had the same composition of observable characteristics—such as income or fire pollution exposure—as those with lighter tones, or how pollution exposure for individuals with darker skin tones would change if they experienced the same returns to standard predictors of environmental quality as those with lighter skin tones. This methodology allows us to quantify the extent to which various factors contribute to the observed skin tone gap in pollution exposure. Specifically, it helps identify what portion of the EJ gap is attributable to differences in the distribution of observable characteristics across skin tone categories. The unexplained portion of the EJ gap arises from differences in returns to both observable and unobservable characteristics between skin tones. This difference in returns is particularly significant because, when group membership is linked to immutable characteristics, it is often interpreted as a measure potentially associated with discrimination (61). Still, it is important to acknowledge that some of the variables included in the explained portion of the decomposition may themselves be influenced by systemic discrimination.

PERLA skin tones are categorized into light L ($\text{PERLA}_i \leq 2$) and dark D ($\text{PERLA} \geq 3$), where we treat categories 1 and 2 as the base category to align with the structure used throughout the paper. Letting $D_D = 1$ indicate a darker skin tone, then under certain assumptions, the overall mean EJ gap $\Delta_O = \mathbb{E}[Y_D \mid D_D = 1] - \mathbb{E}[Y_L \mid D_D = 0]$ can be separated into:

$$\begin{aligned} \hat{\Delta}_O &= \bar{X}_D(\hat{\beta}_D - \hat{\beta}_L) + (\bar{X}_D - \bar{X}_L)\hat{\beta}_L \\ &= \hat{\Delta}_S + \hat{\Delta}_X \end{aligned} \qquad [2]$$

The first term in Eq. 2 is the environmental stratification component, $\hat{\Delta}_S$, which captures differences in coefficients of included controls, while the second term is the composition component, $\hat{\Delta}_X$, capturing the differences in characteristics between groups L and D. The decomposition is performed incorporating in X the covariates employed in Eq. 1. The required assumptions, simple counterfactual treatment, overlapping support between groups and ignorability of unobserved covariates, are discussed in detail in the SI

Methodological Appendix Section, along with the limitations they entail.

Under the same assumptions required to estimate Eq. 2, $\hat{\Delta}_S$ and $\hat{\Delta}_X$ can be further decomposed and expressed as the sum of the contributions of the covariates. The detailed decomposition then allows us to examine the evolution of the relative relevance of covariates and their returns to the EJ gap. Each element of the sum $\hat{\Delta}_S$ can be interpreted as the contribution of the difference in the returns to the kth covariate to the total environmental stratification portion, evaluated at the mean value of $\bar{X}_{Bk}$. A comprehensive discussion on the estimation techniques, underlying assumptions, and potential limitations of the overall and detail KOBD is available in the SI Methodological Appendix Section. In Section , we discuss the implications of the main assumptions for this KOBD application and corroborate that our results are robust to key modeling choices.

Results

Skin tone as consistent axis of environmental stratification. Previous EJ research typically relies on ethnoracial categories (16, 53) or foreign-born status (62–64) to assess environmental inequality. However, skin tone has been identified as "a central axis of social stratification in at least several Latin American countries, though it is often ignored" (24, p.3). The reluctance of survey respondents to self-identify with one of the available categories in ELCA ("White," "Black/Afrodescendant," "Indigenous," "Palenquero," "Raizal")—where a striking 30% of respondents select none of these options—reflects the methodological challenges in conducting environmental justice assessments with these categories. Along the same lines, Figure 2 illustrates the significant overlap of observed skin tone that exists even among individuals who self-report as Black and white. Indigenous people, who are a diverse group in terms of phenotype, also potentially face different manifestations and levels of institutionalized colorism. The disconnect between ethnoracial categories as traditionally defined and the local realities of Latin America may be behind the lack of findings of statistically significant links between pollution and ethnoracial identity in previous research in some of these contexts (e.g., (65–67)).

Before proceeding to the regression analyses, average pollution exposure is computed for skin tone categories using different buffer radii. Figure 3 (a) illustrates a linear gradient in 2010, where individuals in the darkest category benefited from lower pollution exposure—7.9% less than those in categories 1 and 2, a difference of 1.52 $\mu g/m^3$. However, by 2016, this relationship had reversed. Individuals with lighter skin tones ($\leq$4 on the PERLA scale) disproportionately benefited from air quality improvements, while those with darker skin tones ($\geq$5 on the PERLA scale) were increasingly exposed to air quality degradation throughout the study period. Specifically, the darker skin tone group (PERLA$\geq$7) saw an increase of 1.55 $\mu g/m^3$ (8.1%) by 2016, while the lightest skin tone group experienced an almost equal improvement of -1.21 $\mu g/m^3$ (-5.9%). As a result, by 2016, individuals in the darkest skin tone categories had an average pollution exposure that was 6% higher than those with the lightest skin tones. These patterns are remarkably robust to defining exposure with different buffer radii (1 km, 5 km, and 10 km).

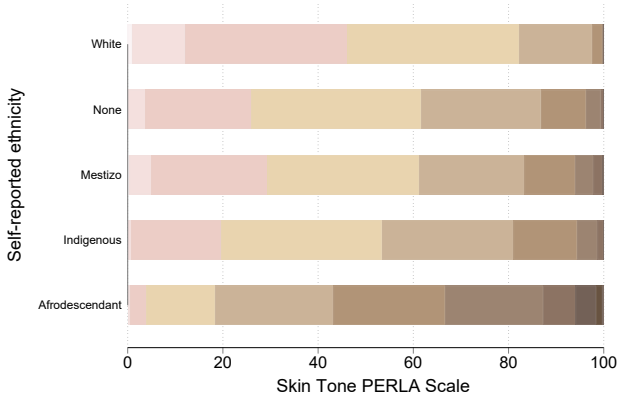


Fig. 2. Ethnoracial categories and skin tone in Colombia

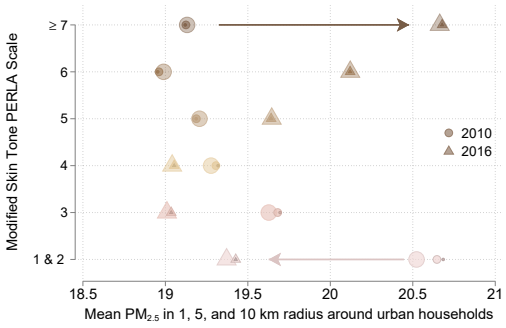
Notes: Figure shows the distribution of skin tone using the PERLA scale within each one of the self-reported ethnoracial categories.

In addition to aggregate estimates, we conducted a separate analysis by dividing the sample based on the type of residence (urban or rural) at the time of the first interview and whether households moved more than five kilometers from their initial residence during the study period (Panels (b) and (c)). The direction of the changes is robust to these varying sample dissections. While the magnitude of the change was considerably greater among those living in a rural area in 2010, in the next section we show that that the pattern of worsening air pollution with darker skin tones is statistically significant and robust for urban respondents as well.

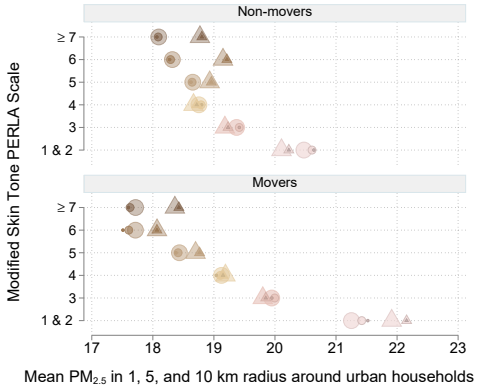
Previous EJ research has debated whether systematic inequalities in exposure to air pollution are driven by disproportionate siting of emitting facilities in minority neighborhoods (or selective implementation of environmental policies) or from sorting[‡]. The panel structure of the ELCA allows us to examine the potential role of sorting and siting as two relevant mechanisms contributing to the emergence of EJ gaps in Colombia. Panels (b) and (c) of Figure 3 situate our findings within the sorting versus siting debate by comparing movers with non-movers —i.e. those who relocated more than five kilometers between ELCA rounds from those who did not. In our sample 9.8% of people moved under this criterion, while 4.8% changed their municipality of residence between 2010 and 2016. Residential sorting —moving looking for lower prices or towards specific amenities— can occur within the same neighborhood and has been found to explain some environmental inequalities (17, 53).

If sorting were the primary driver of the observed environmental fortune reversal, we would expect the aggregate pattern to dissipate among non-movers. For respondents who began in urban areas, the observed reversal is primarily driven by non-movers, as movers in some skin tone categories relocated to areas with slightly better air quality. This result is even more pronounced for rural respondents. Regardless of whether they moved, individuals with the lightest skin tones experienced a roughly 2 $\mu g/m^3$ (10%) improvement in air quality over the study period, the opposite of what we

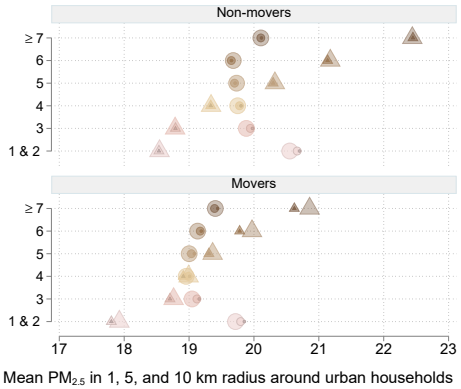
[‡] Other policy-relevant drivers are differential access to information and discriminatory housing markets, which yield a phenomenon known as steering (68, 69).



(a) Full sample



(b) Urban sample



(c) Rural sample

Fig. 3. Average $PM_{2.5}$ exposure by skin tone

Notes: Estimates using 2010 and 2016 ELCA. **(b)** includes households residing in urban areas when first interviewed and **(c)** households residing in rural areas. Movers are defined as those who relocated more than five kilometers between ELCA rounds.

observe for the darkest-skinned respondents. Both movers and non-movers experienced increased particulate pollution exposure, (6.3% and 11.7%, respectively) resulting in a 2016 gap between the lightest and darkest skin of 21.1% for non-movers and 15.8% for movers.

Using predetermined racial categories, we observe some disparities in exposure levels. To elaborate, in 2010, individuals identifying as white had exposure levels of 19.78 $\mu g/m^3$, 0.44 $\mu g/m^3$ higher than mestizos, 0.74 higher than indigenous and more than a unit higher than the pollution exposure of Afro-descendants (18.69 $\mu g/m^3$). Over time, both white and mestizo populations experienced notable improvements in air quality. On the other hand, the air quality for indigenous and Afro-descendant individuals worsened, with $PM_{2.5}$ increments of 1.02 $\mu g/m^3$ and 0.54 $\mu g/m^3$, respectively. This relationship is even more pronounced in rural samples. Detailed results are presented in SI Appendix Figure S5 in the appendix.

Figure S6 in the SI Appendix confirms that the 2013 gradient was an intermediate step between 2010 and 2016, indicating a gradual transformation of the environmental justice landscape in Colombia during the study period.

The environmental justice gap. The results from estimating Eq.1 for 2010, 2016, and their difference, as shown in Figure 4, once again reveal a reversal of the skin tone-pollution gradient in Colombia. Panel (a) illustrates that in 2010, individuals with darker skin tones were exposed to lower pollution levels. However, this difference becomes less pronounced as more statistical controls are applied, with local election turnout contributing to the most significant reduction. Specifically, the association between having a skin tone rating of ≥ 7 (the darkest category) and pollution exposure in 2010 shifts from -1.5 in the raw correlations to -0.6 $\mu g/m^3$ in the specification with all the controls.

In 2016, the shape of this gradient shifts notably. While skin tones 3 and 4 are associated with pollution levels approximately half a unit lower than the lightest skin tone, an increasing pattern emerges for tones 4-7, as shown in Panel (b). The estimate for $\hat{\alpha}_{7,2016}$ is approximately 1 $\mu g/m^3$ across all specifications. This pattern holds even after controlling for a broad set of theoretically relevant covariates. However, the inclusion of urban and rural fire exposure variables (the final set of coefficients) reduces the strength of this pattern, indicating that differential exposure to fire pollution may play a role in explaining the environmental justice gap in Colombia.

It is worth examining the intertemporal changes in pollution, which allows us to fix a series of individual characteristics and by following them over time, explore how skin tone is linked to their environmental destinies. The results indicate a strikingly strong association between skin tone and air quality deterioration. The group with darkest skin tones experienced an increase in pollution exposure ranging between 2.1 and 2.8 $\mu g/m^3$, depending on the covariates adjusting this relationship, a magnitude corresponding to equivalent 1-1.4 SD of of exposure changes in our dataset. Findings from comparable contexts suggest that this effect is not only statistically significant but also economically relevant (70). For example, recent studies estimate that a 10 $\mu g/m^3$ increase in $PM_{2.5}$ during the first year of life lead to a 9.2% increase in the infant mortality rate in Sub-Saharan Africa (71), a region baseline pollution averages 25.2 $\mu g/m^3$, on par with to the 19.01 $\mu g/m^3$ average in our sample. Furthermore, estimates from Chile (26.32 $\mu g/m^3$) indicate that an increase as small as 1 $\mu g/m^3$ in $PM_{2.5}$ exposure for a single day increases ER visits for respiratory illness by 0.36% (72).

As skin tone lightens, this coefficient linearly declines, culminating in an estimate ranging between 0.3 and 0.3 $\mu g/m^3$ for individuals categorized with a skin tone of 3—the second-to-lightest shade on our adapted PERLA scale.

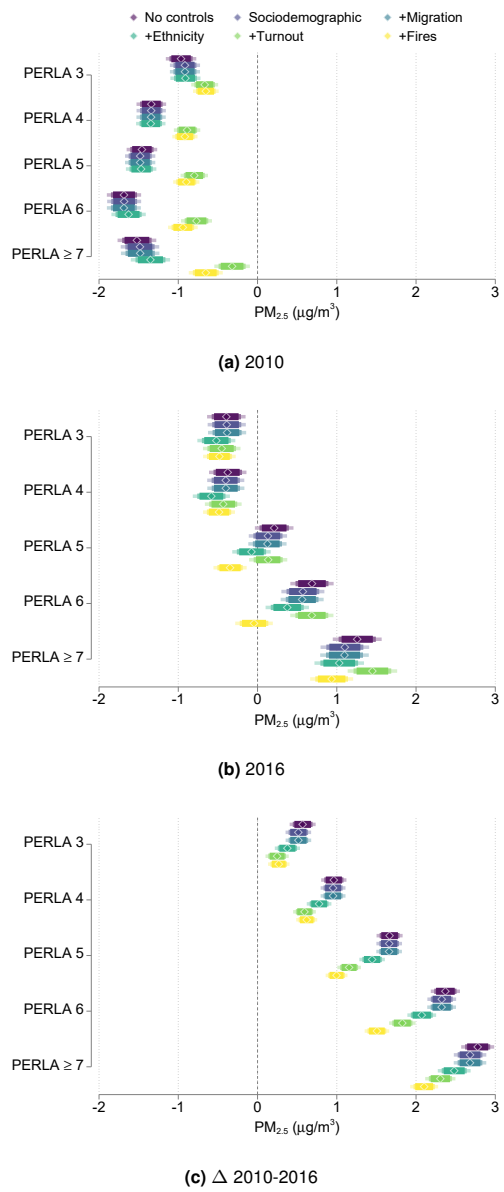


Fig. 4. Average $PM_{2.5}$ exposure by skin tone

Notes: Associations between skin tone and $PM_{2.5}$ in 2010 (a), 2016 (b), and the difference in $PM_{2.5}$ between 2016 and 2010 (c). The figure includes 99%, 95%, and 90% confidence intervals (fading color bars). Covariates are progressively added, starting with no controls (purple) and ending with the most complete specification (yellow), which includes sociodemographic controls (income, an unmet needs indicator, urbanization of the county of residence), migrant status, ethnicity indicators (none, mestizo, indigenous and afrodescendant), voter turnout, and urban and rural upwind probability-weighted fire exposure.

Figure 5 indicates that EJ issues are also somewhat evident along traditional ethnoracial bounds, with indigenous groups experiencing the best air quality initially but the most substantial $PM_{2.5}$ increments subsequently. The omitted category consists of respondents identifying as white (20% of the sample). Controlling for fire exposure and local collective action mutes the 2010 relationships, but not the 2016 or the intertemporal pattern. Overall, all groups experienced similar pollution increases compared to white respondents, with indigenous individuals being the most affected, showing estimated PM2.5 increases ranging from 1.8 to 2.2 $\mu g/m^3$.

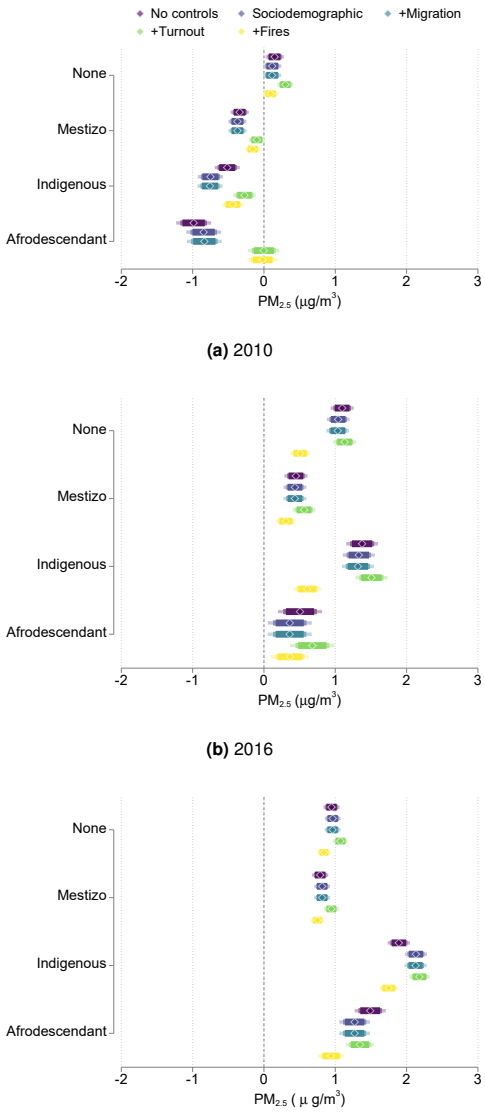


Fig. 5. Average $PM_{2.5}$ exposure by Ethnicity

Notes: Associations between ethnicity and $PM_{2.5}$ in 2010 (a), 2016 (b), and the difference in $PM_{2.5}$ between 2016 and 2010 (c). The figure includes 99%, 95%, and 90% confidence intervals (fading color bars). Covariates are progressively added, starting with no controls (purple) and ending with the most complete specification (yellow), which includes sociodemographic controls (income, an unmet needs indicator, urbanization of the county of residence), migrant status, voter turnout, and urban and rural upwind probability-weighted fire exposure.

Figure S6 in the SI Appendix shows that findings from Figure 4 remain consistent when urban and rural households are analyzed separately. Among urban respondents, those with the darkest skin tones experienced an increase in pollution exposure ranging from 0.7 to 1 $\mu g/m^3$, equivalent to 0.35-0.5 SDs of the exposure changes among this sample.

The pattern of increasing exposure with skin pigmentation, and remains stable even when controls are added. For rural households, the pattern is more pronounced than in the general sample, with relative increases for the darkest skin tones reaching up to 4.3 $\mu g/m^3$ in the raw specification and 2.7 $\mu g/m^3$ after including all controls.

Mechanisms. In Colombia, residential firewood use, agricultural burns, and natural sources, such as forest fires, contribute approximately 75% of $PM_{2.5}$ emissions, with the remaining 25% originating from fixed and mobile emission sources (8). Forest and agricultural fires alone account for half of the particulate emissions in the country. Our analyses in the previous sections suggest that exposure to fire pollution may account for part of the observed disparities. In this section, we formalize this argument by examining the evolution of the relationship between skin tone and fire pollution exposure, and then proceed with decomposition analyses to further investigate these dynamics.

Figure 6 depicts the relationship between skin tone and potential fire smoke exposure in Colombia for the years 2010 and 2016. It is important to remind the reader that urban fires are broadly defined to include peri-urban and rural areas, as long as the fire took place within the boundaries of a populated area, which is characterized as any cluster of at least 20 adjacent or neighboring residences. Under this definition, the frequency of rural fires in our sample is an order of magnitude higher than the frequency of urban fires. In both years, individuals with darker skin tones were exposed to more upwind rural and urban fires. This pattern remains constant across time, which indicates that fire pollution is a persistent determinant of air quality inequities but is less likely to fully explain the observed changes. Additionally, Figure S9 in the SI Appendix shows that fire exposure explained 2% of the variation in $PM_{2.5}$ in 2010 but accounted for nearly 20% of the variation in particulate pollution in the sample by 2016, underscoring the increasing significance of this factor. This raises the question of what other components of the EJ gap have gained or lost explanatory power over time.

The results of the specifications presented in Figure 4 reveal an influence of other included covariates on the magnitude of the skin color coefficients. Intuitively, a portion of pollution disparities stems from the correlation of standard predictors of environmental quality, such as socioeconomic status and local collective action (quantified through voter turnout in recent presidential elections), with skin tone. Figure 7 displays the results of implementing the KOBD described in Equation 2, showing what fraction of the EJ disparities can be attributed to differences in the included covariates. The explained portion represents the amount by which the EJ gap would be reduced in the hypothetical world where, other things equal, dark-skinned individuals had the same observable characteristics as light-skinned individuals, measured using the included covariates. This decomposition further illuminates whether specific explanatory variables disproportionately benefit white individuals, leading to an *unexplained* segment of the EJ gap.

Panels (a) and (c) show that a third portion of the racial difference in pollution exposure remains unaccounted for by the conventionally relevant variables in the domain. Panels (b) and (d) present the contributions of specific covariates to the explained (left) and unexplained (right) portions of the gap. They offer two additional insights. In terms of the explained portion, differences in voter turnout and urbanization had a significant role in 2010, with higher local collective action levels among darker-skinned individuals

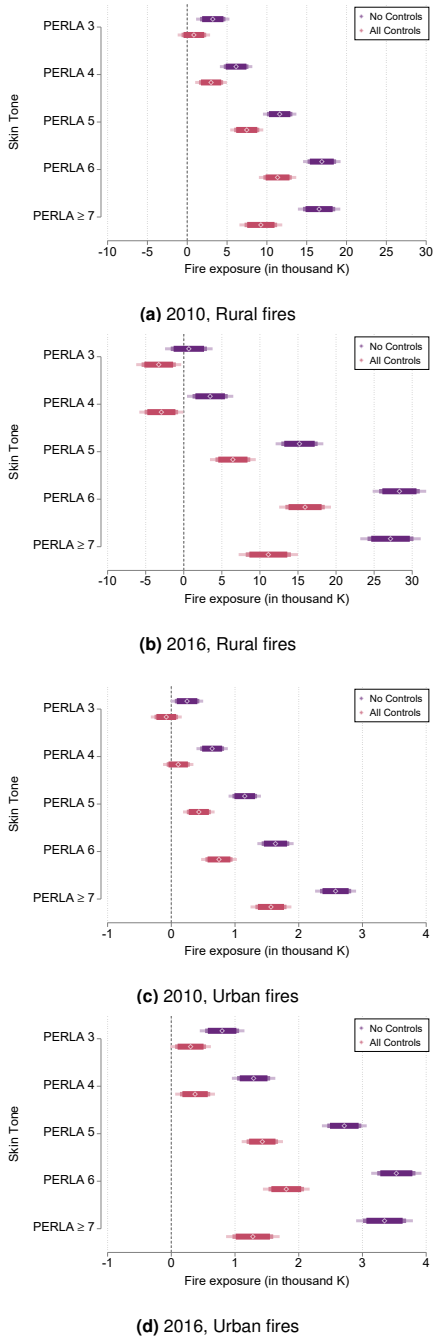


Fig. 6. Skin tone and potential fire smoke exposure in 2010 and 2016 across rural and urban areas

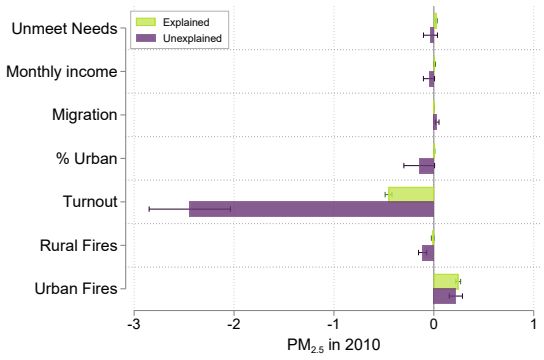
Notes: Each subfigure shows the estimated exposure to fire smoke based on skin tone for the years 2010 and 2016, separated by rural and urban fires. Exposure is calculated as the probability- and intensity- weighted upwind fire incidence in a 50km radius around each household's coordinates. The figure includes 99%, 95%, and 90% confidence intervals (fading color bars). All controls include income, an unmet needs indicator, ethnicity, urbanization of the county of residence, migrant status and voter turnout.

	$PM_{2.5}$ in 2010
Darker Skin Tones	19.12***
	(0.0222)
Lighter Skin Tones	19.55***
	(0.0201)
Difference	-0.427***
	(0.0299)
Explained	-0.184***
	(0.0190)
Unexplained	-0.243***
	(0.0309)
Observations	23,406

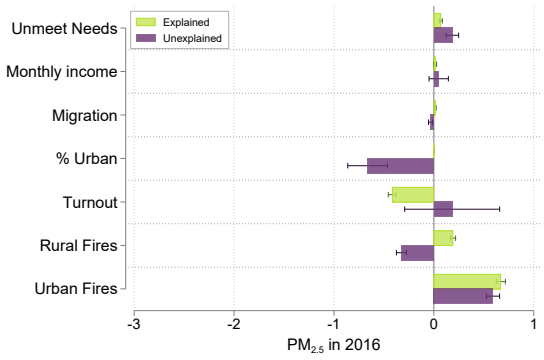
(a) 2010, Summary

	$PM_{2.5}$ in 2016
Darker Skin Tones	19.90***
	(0.0337)
Lighter Skin Tones	19.07***
	(0.0226)
Difference	0.829***
	(0.0406)
Explained	0.545***
	(0.0295)
Unexplained	0.284***
	(0.0375)
Observations	23,408

(c) 2016, Summary



(b) 2010, Detail



(d) 2016, Detail

Fig. 7. Decomposition of EJ gap

Notes: Kitagawa–Oaxaca–Blinder decomposition of the association between skin tone and $PM_{2.5}$ pollution exposure. Panels (a) and c) present the overall results of the decomposition, using as cutoff PERLA$\leq$2. Panels (b) and d) present the detailed decomposition. Standard errors in parentheses *** p<0.01, ** p<0.05, * p<0.1.

having high explanatory power for their relatively lower pollution exposure. However, by 2016, income differences between skin tones and urban/peri-urban and rural fire exposure began to play a detrimental role for darker-skinned individuals.

On the unexplained side, the environmental stratification function—which reflects differential returns to these covariates—also underwent a marked shift. In 2010, the unexplained portion suggested that the lower pollution exposure for darker-skinned individuals was partly due to favorable returns on factors such as voter turnout and urbanization. However, by 2016, these returns had reversed, with lighter-skinned individuals gaining more from local political participation. Notably, the (negative) returns to fire exposure increased, exacerbating the EJ. Importantly, a substantial proportion of the disparity remains unaccounted for by our study's included covariates' composition and returns.

Robustness. SI Appendix explores heterogeneity in the results across different modeling choices as well as the results' sensitivity to a rich set of robustness checks. First of all, all our main analysis calculate pollution exposure employing a 5 km radius from the household's coordinates. The relationships displayed in Figure 3 prove consistent regardless of buffer size, as shown in SI Appendix Figures S5 and S6, where we employ 1km and 10 km radii instead.

Our results remain robust across different imputation and clustering techniques. SI Appendix Tables S7 and S8 demonstrate that the sign, significance, and magnitude of our main findings hold under the following specifications: 1) excluding individuals with imputed skin tones, 2) clustering standard errors at the household level, and 3) restricting the sample to only household heads. It is important to note that the ELCA tracks individuals rather than households, allowing us to follow household members as they relocate and form new households by 2016, making the individual-level regression our preferred specification.

One could relax the assumption that inter-group differences only come from Δ_S and Δ_X. A threefold decomposition allows the gap may vary depending on the interaction between the characteristics and their respective returns. SI Appendix Figure S8 shows that in a threefold decomposition, the interaction effect is relatively small in contrast with Δ_S and Δ_X. One could relax the assumption of no general equilibrium effects implied in standard KOBD techniques. An option to relax this assumption is proposed by (61). A pooled decomposition employs as a counterfactual a weighted average $\beta^* = \Omega\beta_L + (I - \Omega)\beta_D$ where $\Omega = w$ reflects a weighting corresponding to the share of the two groups in the

population. SI Appendix Figure S8 shows that the results are virtually identical to figure 7.

A limitation of decomposition analysis, as with any counterfactual exercise, is that a reference group has to be chosen by the researcher, and this choice mechanically impacts the estimates resulting from the decomposition (61). As a robustness test, we conduct five decompositions based on different cutoffs between darker and lighter skin tones. We represent each of these coefficients using a marker with the color of each skin tone cutoff in Appendix Figure S9. This figure shows that the qualitative findings of the general and detailed decompositions do not change drastically when progressively expanding the reference group to include additional skin tones beyond 1 and 2. Still, the role of some covariates can be sensitive to this coding decision, suggesting caution is warranted when interpreting the detailed KOBD results.

Discussion

We are the first to show that skin tone is a relevant dimension of environmental stratification, measured as exposure to $PM_{2.5}$. While people with darker skin tones were initially exposed to better air quality, our findings reveal that improvements in environmental quality have been unequally distributed along racial and ethnic lines. Lighter-skinned individuals disproportionately benefited from reductions in air pollution over the time period that we examine, while people with darker skin experienced increases in pollution. By 2016, we document an environmental fortune reversal, with people with darker skin tones and non-white ethnoracial self identification experiencing worse particulate pollution levels.

We explore the mechanisms driving these patterns. Two thirds of the observed EJ gaps in both years can be explained with theoretically relevant covariates, measuring socioeconomic status, migration status, local collective action, urban residence, and exposure to pollution from fires. This last factor presents an increasing gradient with skin tone pigmentation and is revealed to be an important covariate in decomposition analyses, especially in 2016. A third of the gap remains unexplained. In the literature, this phenomenon is considered consistent, among other things, with potential systemic discrimination.

Discrimination in this context could manifest, for instance, through selective cleanup or enforcement of environmental policies, as seen in other cases of environmental injustice (69, 73). Government inaction and varying levels of regulatory stringency have been identified as drivers of environmental disparities in some settings, with local poverty, collective action potential, and, in some cases, race emerging as relevant predictors in the U.S. context (37). In Colombia, open agricultural fires, forest fires, and waste burning are prohibited. While controlled cropland burning is legal, it must adhere to strict requirements regarding the distance from dwellings, infrastructure, and priority ecosystems. Moreover, monitoring reports must be submitted to the local environmental authority for a burn to be considered controlled (Decree 948 of 1995 and Decree 4296 of 2004). However, regulators with resource and time constraints must choose to prioritize regulation and remediation across various sites. This raises the possibility that uneven enforcement could lead to inequitable exposure to biomass burning pollution in Colombia, particularly affecting marginalized communities. An example could be the case of sugarcane cultivation. Controlled burning is explicitly forbidden in many parts of Valle del Cauca and Cauca, two of the main sugarcane-producing departments, but research has found high emissions from this sector (74). In these regions Afro-descendants make up almost one-third of the population, compared to 10% nationally. Future research should explore whether disparities exist in enforcement practices in this and other regulations in Colombia and abroad.

Our study has a number of limitations that future work can address. First, we lack exogenous variation that would allow us to confidently establish a causal link between fire exposure —and in general, the potential mechanisms assessed— and pollution disparities. Second, even after including a wide range of theoretically relevant controls, a substantial portion of the environmental justice gap remains unexplained, revealing data limitations in measuring additional relevant contributing factors. Third, while the overall robustness of the widely-used ACAG dataset—derived from a combination of AOD measurements and a chemical transport model calibrated with ground-based data—makes it a suitable source for our analysis, the pollution estimates still carry some degree of uncertainty. This uncertainty stems, among other factors, from local variations in aerosol properties and the conversion of AOD to $PM_{2.5}$. The high resolution we use in this project poses additional challenges, as finer-scale PM2.5 gradients may not be fully captured due to the influence of coarser resolution data sources, which might introduce measurement error.

Previous research has underscored the considerable economic and educational repercussions of even marginal shifts in $PM_{2.5}$ levels. Given these insights, the observed fluctuations in exposure levels among diverse groups can substantially influence racial equality of opportunity and economic disparities in both the immediate future and the long run. The findings presented in this study indicate that further research investigating environmental justice in Colombia should closely examine whether the mitigation policies implemented in the last decade have, by design or by accident, disproportionately benefited certain segments of the population while neglecting others. Moreover, without further intervention, some of the observed dynamics might be exacerbated by climate change and current internal conflict dynamics, and thus likely to increase the disproportionate disease burden on vulnerable groups unless policy measures are implemented to reverse this trend.

ACKNOWLEDGMENTS. We thank Christopher Timmins, Danae Hernandez-Cortes, and Manuel Pastor for their helpful comments on this paper. We also thank participants of the International Economic Association 2024 Conference and the Universidad Iberoamericana Economics Seminar for their comments and suggestions. Diana Millan-Orduz and David Chaparro-Alvarez provided excellent research assistance. We thank Samantha Eyler-Driscoll for her support in editing the manuscript.

1. R Burnett, et al., Global estimates of mortality associated with long-term exposure to outdoor fine particulate matter. *Proc. Natl. Acad. Sci.* **115**, 9592–9597 (2018).
2. J Currie, JG Zivin, J Mullins, M Neidell, What do we know about short-and long-term effects of early-life exposure to pollution? *Annu. Rev. Resour. Econ.* **6**, 217–247 (2014).
3. J Graff Zivin, M Neidell, Environment, health, and human capital. *J. economic literature* **51**, 689–730 (2013).
4. CA Pope Iii, et al., Lung cancer, cardiopulmonary mortality, and long-term exposure to fine particulate air pollution. *Jama* **287**, 1132–1141 (2002).
5. DW Dockery, et al., An association between air pollution and mortality in six US cities. *New Engl. journal medicine* **329**, 1753–1759 (1993).
6. WHO, *Air pollution and child health: prescribing clean air.* (World Health Organization), (2016).
7. S Aguilar-Gomez, H Dwyer, J Graff Zivin, M Neidell, This is air: The "nonhealth" effects of air pollution. *Annu. Rev. Resour. Econ.* **14**, 403–425 (2022).
8. M de Ambiente, Estrategia nacional de calidad del aire, (DIRECCIÓN DE ASUNTOS AMBIENTALES, SECTORIAL Y URBANA), Technical report (2019).
9. N Fleisher, AD Roux, M Alazraqui, H Spinelli, FD Maio, Socioeconomic Gradients in Chronic Disease Risk Factors in Middle-Income Countries: Evidence of Effect Modification by Urbanicity in Argentina. *Am. J. Public Heal.* **101**, 294–301 (2011).
10. P Hessel, P Rodriguez Lesmes, D Torres, Socio-economic inequalities in high blood pressure and additional risk factors for cardiovascular disease among older individuals in Colombia: Results from a nationally representative study. *PLOS ONE* **15**, e0234326 (2020).
11. RA Pérez, CAO Tejada, LM Triaca, AD Bertoldi, AMA dos Santos, Socioeconomic inequality in health in older adults in Brazil. *Dialogues Heal.* **1**, 100009 (2022).
12. L Pulido, Geographies of race and ethnicity II: Environmental racism, racial capitalism and state-sanctioned violence. *Prog. human geography* **41**, 524–533 (2017).
13. N Brazil, Environmental inequality in the neighborhood networks of urban mobility in US cities. *Proc. Natl. Acad. Sci.* **119**, e2117776119 (2022).
14. SE Chambliss, et al., Local- and regional-scale racial and ethnic disparities in air pollution determined by long-term mobile monitoring. *Proc. Natl. Acad. Sci.* **118**, e2109249118 (2021).
15. D Hernandez-Cortes, KC Meng, Do environmental markets cause environmental injustice? Evidence from California's carbon market. *J. Public Econ.* **217**, 104786 (2023).
16. P Mohai, D Pellow, JT Roberts, Environmental justice. *Annu. review environment resources* **34**, 405–430 (2009).
17. S Banzhaf, L Ma, C Timmins, Environmental justice: The economics of race, place, and pollution. *J. Econ. Perspectives* **33**, 185–208 (2019).
18. D Alexander, J Currie, Is it who you are or where you live? Residential segregation and racial gaps in childhood asthma. *J. health economics* **55**, 186–200 (2017).
19. A Aizer, J Currie, P Simon, P Vivier, Do low levels of blood lead reduce children's future test scores? *Am. Econ. Journal: Appl. Econ.* **10**, 307–341 (2018).
20. JM Kohima, UE Chigbu, ML Mazambani, MR Mabakeng, (Neo-)segregation, (neo-)racism, and one-city two-system planning in Windhoek, Namibia: What can a new national urban policy do? *Land Use Policy* **125**, 106480 (2023).
21. L Melgaço, LXP Coelho, Race and Space in the Postcolony: A Relational Study on Urban Planning Under Racial Capitalism in Brazil and South Africa. *City & Community* **21**, 214–237 (2022).
22. J Sundberg, Placing Race in Environmental Justice Research in Latin America. *Soc. & Nat. Resour.* **21**, 569–582 (2008).
23. L Monroy-Gómez-Franco, Shades of social mobility: Colorism, ethnic origin and intergenerational social mobility. *The Q. Rev. Econ. Finance* **90**, 247–266 (2023).
24. E Telles, *Pigmentocracies: Ethnicity, race, and color in Latin America.* (UNC Press Books), (2014).
25. AR Dixon, EE Telles, Skin color and colorism: Global research, concepts, and measurement. *Annu. Rev. Sociol.* **43**, 405–424 (2017).
26. JG Canadell, et al., Multi-decadal increase of forest burned area in Australia is linked to climate change. *Nat. communications* **12**, 6921 (2021).
27. MW Jones, et al., Global and regional trends and drivers of fire under climate change. *Rev. Geophys.* **60**, e2020RG000726 (2022).
28. EW Butt, et al., Achieving brazil's deforestation target will reduce fire and deliver air quality and public health benefits. *Earth's Futur.* **10**, e2022EF003048 (2022).
29. D Armenteras, L Schneider, LM Dávalos, Fires in protected areas reveal unforeseen costs of colombian peace. *Nat. ecology & evolution* **3**, 20–23 (2019).
30. NASA, Viirs-modis data (https://www.earthdata.nasa.gov/learn/find-data/near-real-time/viirs) (year?) Accessed: 2024-07-18.
31. NASA, Firms data (https://www.earthdata.nasa.gov/learn/find-data/near-real-time/firms/about-firms) (year?) Accessed: 2024-07-18.
32. European Centre for Medium-Range Weather Forecasts, Era5-land: Ecmwf's global reanalysis for land applications (2024) Accessed: 2024-08-21.
33. E Telles, RD Flores, F Urrea-Giraldo, Pigmentocracies: Educational inequality, skin color and census ethnoracial identification in eight Latin American countries. *Res. Soc. Stratif. Mobil.* **40**, 39–58 (2015).
34. AR Dixon, Colorism and classism confounded: Perceptions of discrimination in Latin America. *Soc. Sci. Res.* **79**, 32–55 (2019).
35. E Telles, T Paschel, Who Is Black, White, or Mixed Race? How Skin Color, Status, and Nation Shape Racial Classification in Latin America. *Am. J. Sociol.* **120**, 864–907 (2014).
36. A Van Donkelaar, et al., Monthly Global Estimates of Fine Particulate Matter and Their Uncertainty. *Environ. Sci. & Technol.* **55**, 15287–15300 (2021).
37. HS Banzhaf, L Ma, C Timmins, Environmental justice: Establishing causal relationships. *Annu. Rev. Resour. Econ.* **11**, 377–398 (2019).
38. European Centre for Medium-Range Weather Forecasts, How to calculate wind speed and wind direction from u and v components of the wind (2024) Accessed: 2024-07-22.
39. HK Pullabhotla, M Souza, Air pollution from agricultural fires increases hypertension risk. *J. Environ. Econ. Manag.* **115**, 102723 (2022).
40. MA Rangel, TS Vogl, Agricultural fires and health at birth. *Rev. Econ. Stat.* **101**, 616–630 (2019).
41. WH Organization, *WHO global air quality guidelines: particulate matter (PM2.5 and PM10), ozone, nitrogen dioxide, sulfur dioxide and carbon monoxide.* (World Health Organization), (2021).
42. NY Rojas, et al., Road transport exhaust emissions in Colombia. 1990–2020 trends and spatial disaggregation. *Transp. Res. Part D: Transp. Environ.* **121**, 103780 (2023).
43. DF Florez Trujillo, A Valencia, B Avendano-Uribe, Informal settlement fires in colombia. *Fire Technol.* pp. 1–24 (2023).
44. E Giraldo Gómez, C García, C del Rosario, *Manejo integrado de residuos sólidos municipales.* (Universidad de los Andes and Ministerio de Ambiente), (1997).
45. K Ballesteros-González, AP Sullivan, R Morales-Betancourt, Estimating the air quality and health impacts of biomass burning in northern South America using a chemical transport model. *Sci. The Total. Environ.* **739**, 139755 (2020).
46. GR Van der Werf, et al., Global fire emissions and the contribution of deforestation, savanna, forest, agricultural, and peat fires (1997–2009). *Atmospheric chemistry physics* **10**, 11707–11735 (2010).
47. B Baptiste, et al., Greening peace in Colombia. *Nat. Ecol. & Evol.* **1**, 0102 (2017).
48. R Rocha, AA Sant'Anna, Winds of fire and smoke: Air pollution and health in the Brazilian Amazon. *World Dev.* **151**, 105722 (2022).
49. M Prem, S Saavedra, JF Vargas, End-of-conflict deforestation: Evidence from colombia's peace agreement. *World Dev.* **129**, 104852 (2020).
50. M Prem, JF Vargas, D Mejía, The Rise and Persistence of Illegal Crops: Evidence from a Naive Policy Announcement. *The Rev. Econ. Stat.* **105**, 344–358 (2023).
51. JI Huertas, ME Huertas, S Izquierdo, ED González, Air quality impact assessment of multiple open pit coal mines in northern Colombia. *J. Environ. Manag.* **93**, 121–129 (2012).
52. DL Anderton, AB Anderson, JM Oakes, MR Fraser, Environmental equity: the demographics of dumping. *Demography* **31**, 229–248 (1994).
53. L Cain, D Hernandez-Cortes, C Timmins, P Weber, Recent findings and methodologies in economics research in environmental justice. *Rev. Environ. Econ. Policy* **18**, 116–142 (2024).
54. AM Ibáñez, A Velásquez, *El impacto del desplazamiento forzoso en Colombia: condiciones socioeconómicas de la población desplazada, vinculación a los mercados laborales y políticas públicas.* (Cepal), (2008).
55. J Navarrete-Suárez, C Masferrer, , et al., Heterogeneidad de la integración laboral en colombia: diferencias según el sexo y la pertenencia étnica de desplazados forzados y otros migrantes internos. *Revista Latinoamericana de Población* **14**, 89–123 (2020).
56. JT Hamilton, Politics and social costs: estimating the impact of collective action on hazardous waste facilities. *The RAND journal economics* pp. 101–125 (1993).
57. N Brooks, R Sethi, The distribution of pollution: community characteristics and exposure to air toxics in *Distributional Effects of Environmental and Energy Policy.* (Routledge), pp. 301–318 (2017).
58. EM Kitagawa, Components of a difference between two rates. *J. american statistical association* **50**, 1168–1194 (1955).
59. R Oaxaca, Male-female wage differentials in urban labor markets. *Int. economic review* **14**, 693–709 (1973).
60. AS Blinder, Wage discrimination: reduced form and structural estimates. *The J. Hum. Resour.* pp. 436–455 (1973).
61. N Fortin, T Lemieux, S Firpo, Decomposition methods in economics in *Handbook of labor economics.* (Elsevier) Vol. 4, pp. 1–102 (2011).
62. O Raaschou-Nielsen, et al., Air pollution at the residence of Danish adults, by socio-demographic characteristics, morbidity, and address level characteristics. *Environ. research* **208**, 112714 (2022).
63. C Kõnig, Neighbourhood structure and environmental quality: A fine-grained analysis of spatial inequalities in urban Germany. *Urban Stud.* p. 00420980231224224 (2024).
64. T Rüttenauer, Neighbours matter: A nation-wide small-area assessment of environmental inequality in Germany. *Soc. Sci. Res.* **70**, 198–211 (2018).
65. J Pearce, S Kingham, P Zawar-Reza, Every breath you take? Environmental justice and air pollution in Christchurch, New Zealand. *Environ. Plan. A* **38**, 919–938 (2006).
66. AR Germani, P Morone, G Testa, Environmental justice and air pollution: A case study on Italian provinces. *Ecol. Econ.* **106**, 69–82 (2014).
67. L Chakraborti, J Voorheis, Is air pollution increasing in poorer localities of mexico? evidence from pm 2.5 satellite data. *Environ. Dev. Econ.* pp. 1–18 (2024).
68. S Aguilar-Gomez, Addressing environmental quality inequities: The role of information and collective action in *Equity in the Urban Built Environment*, ed. B Bereitschaft. (Routledge), (2024).
69. P Christensen, C Timmins, Sorting or steering: The effects of housing discrimination on neighborhood choice. *J. Polit. Econ.* **130**, 2110–2163 (2022).
70. S Aguilar-Gomez, NM Rivera, Pollution in the global south: An overview of its sources and impacts. *Univ. Chile, Dep. Econ. Work. Pap.* (2024).
71. S Heft-Neal, J Burney, E Bendavid, M Burke, Robust relationship between air quality and infant mortality in Africa. *Nature* **559**, 254–258 (2018).
72. E Dardati, R de Elejalde, E Giolito, On the short-term impact of pollution: The effect of PM 2.5 on emergency room visits. *Heal. Econ.* **33**, 482–508 (2024).
73. D Taylor, Toxic communities in *Toxic Communities.* (New York University Press), (2014).
74. L Mateus-Fontecha, et al., Understanding aerosol composition in an inter-andean valley impacted by sugarcane intensive agriculture and urban emissions. *Atmospheric Chem. Phys. Discuss.* **2021**, 1–36 (2021).

Supporting Information for

Environmental justice beyond race: Skin tone and exposure to air pollution

Sandra Aguilar-Gomez, Juan Camilo Cardenas and Ricardo Salas Diaz

Corresponding Author: Sandra Aguilar-Gomez
E-mail: s.aguilargomez@uniandes.edu.co

This PDF file includes:

Description of data and summary statistics

ELCA. The Encuesta Longitudinal Colombiana (ELCA) tracked approximately 10,000 households at three-year intervals for six years from 2010 to 2016. The survey incorporates ethnic self-identification questions and interviewer-based skin tone classification, among many other sociodemographic, economic, and cultural questions. We restrict our sample to ELCA households interviewed in 2010, 2013, and 2016. Skin tone information was collected in the 2013 round. The baseline included 10,164 households, 5,446 urban and 4,718 rural, of which 9,853 remained in the second follow-up, 5,275 urban and 4,578 rural. Attrition was concentrated among households of senior citizens over 64 in the baseline with no children (1). Importantly, as some household members relocated during these six years, which accounts for some of the pollution variability, we use all the surveys at the individual level.

ELCA data collection followed a probabilistic, stratified, multistage, and cluster sampling design. The urban sample is nationally representative of the households between socioeconomic strata 1 and 4, which cover 97% of the population, excluding the top 3% in terms of socioeconomic status.*. This sample encompasses diverse regions, income levels, and degrees of urbanization. The rural sample covers four large subregions: Mid-Atlantic (north), Cundi-Boyacense (center), coffee axis (center-west), and center-east. Noncovered regions with significant rural populations include Amazonas (south), Orinoquia (east), and Pacífico (west). While their omission from ELCA sampling means our study is not representative of the entire country, the noncovered rural regions contain only approximately 3% of the Colombian population and hence are unlikely to change our results.

The 2013 data collection round of ELCA included various ethnoracial categories, including "Indigenous," "ROM or Romani," "Raizal," "Palenquero," "Black or Mulatto (Afro-descendant)", "White," and "Mestizo." The Raizal population pertains to those born in San Andrés and Providencia (Caribbean islands) with distinctive cultural traditions. Palenqueros are individuals born in the palisade towns established by slaves during colonial times. Following the National Statistical Department, we group the "Raizal" and "Palenquero" categories along with the "Black and Mulatto" into the "Afro-descendant" group. Table S1 presents the population distribution.

In 1999, the Romani people were officially recognized as an ethnic group in Colombia. However, due to the very small sample size—five urban and six rural observations, respectively—their data was excluded. Additionally, the Venezuelan population in Colombia surged from 40,000 in 2015 to 1.1 million in 2018 (2), but our estimates, based on the 2010 ELCA design, do not account for this group.

	Indigenous	Afro-descendant	None	Mestizo	White	Total
Rural	1382 (7.45%)	132 (0.71%)	3279 (17.67%)	3026 (16.30%)	1741 (9.38%)	9560 (51.51%)
Urban	296 (1.59%)	590 (3.18%)	2962 (15.96%)	3174 (17.10%)	1977 (10.65%)	8999 (48.49%)
Total	1678 (9.04%)	722 (3.89%)	6241 (33.63%)	6200 (33.41%)	3718 (20.03%)	18559 (100%)

Table S1. Number of observations in urban and rural samples, by ethnic self-identification.

Notes: The table displays the count of individuals for each self-identification category. We do not include the 3,070 individuals from the rural sample and the 3,505 individuals from the urban sample who did not respond to this question. The percent of total observations appears in parentheses.

Skin tone assessment and imputation. In the 2013 round, interviewers employed a standardized tone palette to assess respondents' facial skin tones. This palette, designed by the Project of Ethnicity and Race in Latin America (PERLA), was devised to complement the ethnic self-identification categories. The tone spectrum has a lightest shade of 1 and a darkest shade of 11. To assess the skin tone of household members, ELCA enumerators are provided with a physical color palette and record the number (1-11) that most closely matches the interviewee's skin tone. The specific instructions given to the interviewers are as follows:

"At the end of this chapter, the interviewer will answer question 1000, assigning the code that, in their perception, best represents the interviewee's skin color. This question must be handled with great discretion so that the interviewee does not perceive that their skin color is being coded."

P1000: Use the COLOR PALETTE (card 5) and indicate the number that most closely matches the skin color of the interviewee's face. Remember that this question is to be completed by observation and should be handled discreetly by the interviewer."

There are instances when we impute skin tones for certain individuals. Imputation was performed for individuals coded as "12", which indicates they were not directly observed by the interviewer, and to those individuals who participated in both the 2010 and the 2016 waves but did not have an answer for question 1000 in 2013. To address this, we calculate an average skin tone for all observed household members and assign the closest integer value of this average to those lacking direct observations.

For visual insight into this imputation method across urban and rural households, refer to Figures S1 and S2.

*The socioeconomic strata is a scaled employed for targeting social policy. Every urban household is classified from 1 to 6.

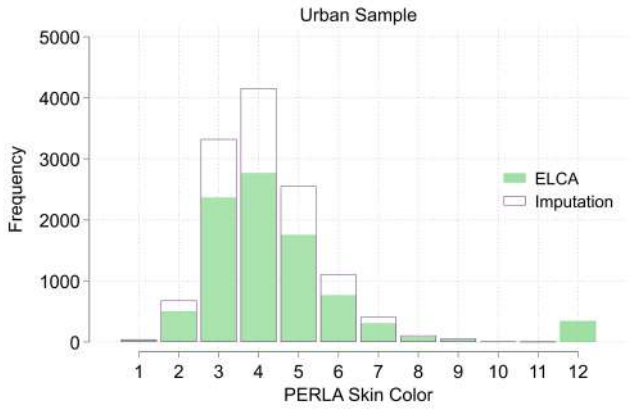


Fig. S1. Observations per PERLA skin tone scale, urban sample

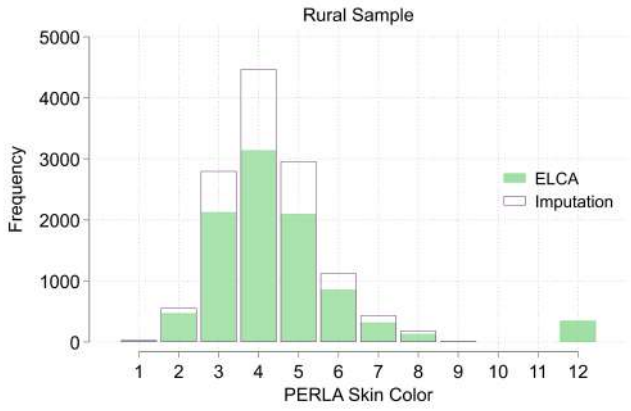


Fig. S2. Observations per PERLA skin tone scale, rural sample

Notes: The green bars represent the original skin coding, as measured by the surveyor in the 2013 wave of the ELCA. The white bars represent the imputed values for the children reported as unseen (category 12) and the adults in the urban sample who were categorized as unseen or who did not take part in the 2013 survey.

We employ a grouping approach to consolidate the PERLA skin tone distribution's tails to address the potential limitations arising from limited observations in the extreme categories. Specifically, we combine the first two categories into one category and the final four into another. All of the individuals in the five of the PERLA skin tone categories grouped (1, 8, 9, 10, and 11) sum up less than 2% of the entire sample. The grouping procedure is illustrated in Figure S3. The number of individuals in each of the samples by our grouping of PERLA skin tones after imputation is presented in Table S2.

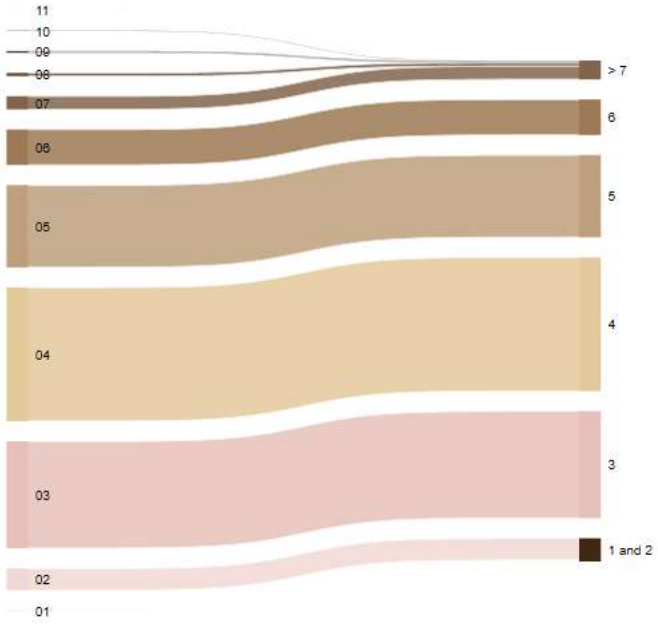


Fig. S3. Visual overview of grouping of PERLA categories

Notes: The Sankey plot presents on the left the original skin tones according to the PERLA Palette and on the right the grouping employed in this paper.

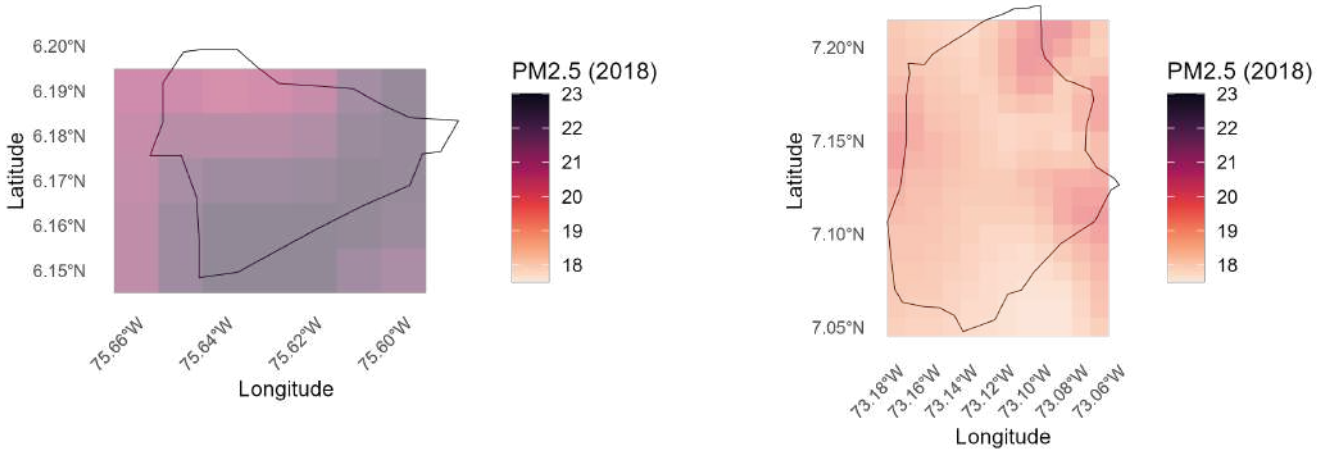
	Sample		
	Rural	Urban	Total
1 and 2	585	723	1308
%	2.34	2.89	5.23
3	2800	3324	6124
%	11.19	13.28	24.47
4	4469	4158	8627
%	17.86	16.61	34.47
5	2961	2553	5514
%	11.83	10.20	22.03
6	1126	1106	2232
%	4.50	4.42	8.92
≥ 7	633	589	1222
%	2.53	2.35	4.88
Total	12574	12453	25027
%	50.24	49.76	100.00

Table S2. Descriptive statistics: Number of observations in ELCA urban and rural samples, by skin tone

Notes: Sample is restricted to ELCA respondents interviewed in 2010 and 2016. PERLA skin tone scale is regrouped into seven categories.

Pollution. We sourced our pollution variables from a reanalysis of satellite-based estimates by the Atmospheric Composition Analysis Group (ACAG) at Washington University in St. Louis. These data encompass annual ground-level fine particulate matter modeled concentrations ($PM_{2.5}$) spanning from 1998 to 2020 (3). These estimates are achieved by merging aerosol optical depth (AOD) retrievals from various NASA satellite instruments with the GEOS-Chem chemical transport model, and subsequently calibrating to global ground-based observations using a geographically weighted regression. ACAG's ground-level raster offers a resolution of $0.01 \times 0.01^{\circ}$, which equates to approximately $1.1 \times 1.1km$. To help readers visualize this, Figure S4 illustrates two Colombian counties of distinct sizes.

Limitations of $PM_{2.5}$ estimates. As described above, he ACAC's hybrid approach employ a GWR and a chemical transport model to produce PM estimates. The GWR models provide high predictive power, but one limitation is that they do not explicitly



(a) Itagüí, Antioquia **(b)** Bucaramanga, Santander

Fig. S4. Pollution levels from raster and county boundaries for **(a)** Itagüí, one of the smallest counties in Colombia, and **(b)** Bucaramanga, a medium-sized county. County boundaries are obtained from the Dirección Nacional de Planeación (DNP).

rely on scientific relationships, which reduces their interpretability (4). Moreover, local variations in aerosol optical properties and vertical profiles of aerosols complicate estimation, as the conversion factor between AOD and $PM_{2.5}$ can vary temporally and spatially, introducing potential errors in PM2.5 estimations (5).

In contrast, process-driven methods, such as the GEOS-Chem chemical transport model, simulate the physical and chemical processes that govern the AOD-PM2.5 relationship (4). These models provide estimates that are largely independent of some error sources, such as surface reflectivity, which can affect satellite retrievals (5). However, their accuracy depends on the quality of input parameters like emissions data and meteorological conditions, which determine the conditions of the simulation and influence the results (5). While process-driven models offer a scientifically grounded approach to estimating PM2.5, they are susceptible to inaccuracies when the input data are imprecise. By combining the two components described and then calibrating with ground-based measurement, ACAG mitigates the drawbacks of each of them, generating a robust and widely used data product, but it is still susceptible to some classical measurement error, which can produce downward bias in our estimates (6).

ACAG data products that are provided at the finest resolution may, in some regions, not fully resolve PM2.5 gradients at the highest resolution due to influence by information sources at coarser resolution (3), which represents an important limitation to our analyses. To be clear, classical measurement error in which the measured exposure is expected to have more variation than the truth, tends to bias the effect estimate toward the null; whereas under Berkson error, in contrast, occurs as a result of using aggregated instead of individual exposure data, as the case with the coarser resolution of some information sources used in the ACAG $0.01 \times 0.01^{\circ}$ data product. In this case, the measurement is less variable than the truth, which results in an unbiased effect estimate but greater variance, provided that the true relationship between exposure and outcome is linear. Systematic error may also occur owing to the spatial correlation between neighboring areas. In the presence of confounders, the direction of bias becomes unclear in principle and depends on the correlation of exposure with the confounders (7).

Fire exposure. Fire counts are obtained from the Fire Information for Resource Management System (FIRMS), which provides near real-time active fire locations to natural resource managers based on VIIRS-MODIS products (8, 9). Data from the ERA5 and ERA5-Land Year Aggregated datasets are used for the analysis of wind patterns. ERA5 provides the u and v components of wind, representing east-west and north-south velocities, respectively. According to ERA5 documentation, these data are calculated using the mode of hourly air temperature observations at 2 meters. This methodology is used because it has been found that air temperature at this height accurately and reliably reflects annual wind conditions, providing a dependable representation of average wind conditions throughout the year (10). Using the u and v components, both wind speed and direction are calculated as per ECMWF guidelines (11).

To calculate fire exposure at the household level, we determine the intensity- and probability-weighted yearly upwind fire counts within a 50 km radius of the household's coordinates. First, following (12), we sum the yearly number of 1 km grid cells containing a fire within this radius, multiplied by the intensity of each fire. Intensity is measured in brightness temperature or radiance temperature (K), which is a standard measure of the intensity of electromagnetic energy coming from a source. All fires are assigned a confidence or likelihood index ranging from 0% to 100%. Following standard practice, we retain only high-confidence fires with a probability of 60% or higher (13). To classify fires as upwind, we use the method from (14), which identifies the prevailing wind direction within 90-degree quadrants, starting from 0 degrees North. We then divide the 50 km radius around each household into 90-degree sectors corresponding to these wind quadrants. By matching these sector buffers with the yearly fire intensity- data, we estimate the yearly upwind fire exposure at the household level.

To differentiate between urban and rural fires, the populated centers layer was used, which identifies all concentrations of at

least twenty contiguous, neighboring, or adjoining houses as identified by the National Administrative Department of Statistics (DANE) (15). The objective is to separate agricultural or forest-related fires, which will be defined as rural fires, from those involving houses or infrastructure, which have different causes.

Covariate construction. All rounds of ELCA include a question collecting information the household monthly income in a hundred thousand Colombian pesos. It includes the sum of all the household monthly salary income, retirement income, rent income, interest income, family or government transfers, and other sources. In 2010, we imputed income information for 1630 individuals, and in 2016, we imputed income information for 542 individuals. The imputation process involved the use of an OLS model that included the age, educational attainment, sex, and region of individuals over 16 years old. Any negative estimations were excluded.

The marginalization index (MI), also referred to as the basic unmet needs index, quantifies the presence of at least one unmet need and is calculated from the following components: 1) housing needs, which considers the maximum value between inadequate walls and flooring; 2) household critical overcrowding, calculated as the presence of more than three individuals per bedroom; 3) household utilities access, determined by the maximum between inadequate sewer waste elimination and access to improved drinking water; 4) economic dependency, calculated for households with more than three individuals per occupied worker; and 5) schooling, measured as the number of children aged 6 to 12 who are not attending school.

In the regression analysis presented in figures 4-7, S7-S9, and tables S5, S6, and S8, migration refers to the people who reported living in a different municipality in 2016 than that reported in 2010. In Figures 3B and 3C, as well as Figure S5, we compare movers with non-movers, defined as people who moved more than 5 km from their original place of residency, independently of whether they still reside in the same municipality.

Urbanization measures the percentage of each municipality's urban population in 2010. We assigned this value to each household by the municipality in which they were located in 2010 and 2016. We use the CEDE electoral dataset to calculate the closest presidential municipality's election turnout. The database contains the results of every election for the congress and the executive branch at the country, departmental, and municipal levels from 1958 up to the present. The values employed in our regression analyses correspond to the 2010 election turnout of the municipality in which the household was located in 2010 and the 2014 election turnout of the municipality in which they were located in 2016.

Additional descriptive statistics and context. Table S3 presents our dependent variable, average annual $PM_{2.5}$ calculated for different radii around the ELCA households for 2010 and 2016, and the differences between both years for the rural, urban, and full samples. The World Health Organization's (WHO's) air quality guidelines specify that annual average concentrations of $PM_{2.5}$ should not exceed 5 µg/m^3 (16). This guideline is articulated as the air quality guideline (AQG) category in Table S4. To facilitate incremental progress towards cleaner air, especially in areas with heightened air pollution levels, interim targets have been proposed. Specifically, the WHO has set the following interim targets for annual exposure: 35 µg/m^3 (IT1), 25 µg/m^3 (IT2), 15 µg/m^3 (IT3), and 10 µg/m^3 (IT4). Our investigation using the ELCA data revealed that the yearly exposure to PM2.5 within a 5 km radius surrounding most households falls within the IT2 range at any given year.

	2010			2016			Diff 2016 - 2010		
	Rural	Urban	Full	Rural	Urban	Full	Rural	Urban	Full
$PM_{2.5}10km$	19.75	18.92	19.37	19.68	19.01	19.37	-0.07	0.09	0.00
	1.60	2.67	2.20	2.50	3.20	2.87	2.30	1.63	2.02
$PM_{2.5}5km$	19.77	18.94	19.39	19.67	19.06	19.39	-0.11	0.10	-0.01
	1.70	2.80	2.31	2.53	3.30	2.93	2.36	1.66	2.07
$PM_{2.5}1km$	19.80	18.94	19.40	19.67	19.06	19.39	-0.13	0.12	-0.01
	1.76	2.86	2.37	2.55	3.37	2.97	2.40	1.71	2.11
N	23410	23410	23410	23408	23408	23408	23408	23408	23408

Table S3. $PM_{2.5}$ in 2010, 2016 and difference with different-size buffers around households

Notes: Sample corresponds to ELCA respondents interviewed in 2010 and 2016. Pollution exposure corresponds to exposure within the specified radii around the coordinates of the household. Urban and rural categories are assigned based on the place of residence when the respondent was first interviewed in 2010. Standard deviation in parentheses.

	Exposure Range	ELCA Households					
	(µg/m³)	2010			2016		
		Rural	Urban	Total	Rural	Urban	Total
Interim Target 1	25–35	0	272	272	1	494	495
Interim Target 2	15–25	12630	11887	24517	12604	11177	23781
Interim Target 3	10–15	0	345	345	25	833	858
Interim Target 4	5–10	0	0	0	0	0	0
Air Quality Guideline	<5	0	0	0	0	0	0
N		12630	12504	25134	12630	12504	25134

Table S4. Air quality guidelines for $PM_{2.5}$ and interim targets

Notes: The table shows the number of individuals in the range of each interim target of the WHO guidelines for $PM_{2.5}$. Exposure is calculated from the household's exact coordinates. No households had an annual exposure higher than 35 µg/m^3.

Methodological Appendix

Kitagawa–Oaxaca–Blinder decomposition: Assumptions and limitations. While in practice, the KOBD is a simple and widely used counterfactual analysis, it relies on a number of assumptions that can have implications for interpretation. Fortin et al (2011) (17) offer a meticulous description of these assumptions. We rely on their notation and structure to discuss the implications of such assumptions for our analyses.

First of all, the environmental stratification function estimated for each group is assumed to be linear and additively separable in observable (X_i) and unobservable (ϵ_i) characteristics:

$$Y_{ig} = X_i\beta_g + v_{ig}, \quad \text{for} \quad g = L, D$$

where v_g is a function of unobservable characteristics $v_{ig} = h_g\epsilon_i$ with zero conditional mean.

To be able to write the overall gap Δ_O as the sum of Δ_S and Δ_X, three additional assumptions are required. The first assumption is that the environmental stratification function $m_L(X, \epsilon)$ can serve as a counterfactual for individuals in group D and $m_D(X, \epsilon)$ as a counterfactual for individuals in group L. This assumption implies that there would be no general equilibrium effects if dark-skinned people had access to the environmental returns experienced by light-skinned people, and vice versa. The validity of this assumption hinges on whether changes in pollution exposure stem from overall improvements or deteriorations in pollution levels, or from the displacement of pollution sources from one group's predominant region to the other's. An option to relax this assumption proposed by (17) is implemented in the robustness section. A pooled decomposition employs as a counterfactual a weighted average $\beta^* = \Omega\beta_L + (I - \Omega)\beta_D$, where $\Omega = w$ reflects a weighting corresponding to the share of the two groups in the population.

The second assumption, overlapping support, rules out cases where different predictors might exist for environmental quality between groups D and L. The third assumption, ignorability, implies that manipulations of the distribution of observables X are not confounded by changes in the distribution of the error term. This assumption requires that the conditional distribution of unobservables given X is the same across groups L and D. In essence, the "selection based on observables" assumption permits selection biases as long as they are consistent between the two groups. In our context, ignorability implies that individuals with the same observable characteristics select into environmental quality similarly, regardless of skin tone. This assumption also implies a similar willingness to pay function between individuals from groups D and L with similar observables. However, this assumption represents a limitation of the analysis, as it may not hold true in reality, and the direction of the bias is unclear. For a more detailed discussion on recent findings linking willingness to pay with EJ outcomes, see (18). While we do not directly address the magnitude of selection based on unobservables in our decomposition results, we employ bounding exercises as robustness tests to our main results.

Under the assumptions above:

$$\begin{aligned}\Delta_O &= \mathbb{E}[Y_D \mid D_D = 1] - \mathbb{E}[Y_L \mid D_D = 0] \\ &= (\mathbb{E}[X|D_D = 1]\beta_D + \mathbb{E}[v_D|D_D = 1]) - \\ &\quad (\mathbb{E}[X|D_D = 0]\beta_L + \mathbb{E}[v_L|D_D = 0])\end{aligned} \tag{1}$$

Where where v_g is a function of unobservable characteristics with zero conditional mean. Replacing the expected value of the covariates $\mathbb{E}[X \mid D_D = d]$, for $d = 0, 1$, by the sample averages $\bar{X}_g$, and adding and subtracting the average contrafactual exposure $\bar{X}_D\hat{\beta}_L$, the decomposition can be expressed and estimated as:

$$\begin{aligned}\hat{\Delta}_O &= \bar{X}_D\hat{\beta}_D - \bar{X}_D\hat{\beta}_L + \bar{X}_D\hat{\beta}_L - \bar{X}_L\hat{\beta}_L \\ &= \bar{X}_D(\hat{\beta}_D - \hat{\beta}_L) + (\bar{X}_D - \bar{X}_L)\hat{\beta}_L \\ &= \hat{\Delta}_S + \hat{\Delta}_X\end{aligned} \quad [2]$$

Under the same set of assumptions, the decomposition can be further disaggregated into a detailed decomposition that quantifies the contribution of each covariate to Δ_S and Δ_X. Specifically;

$$\hat{\Delta}_S = (\hat{\beta}_{D0} - \hat{\beta}_{L0}) + \sum_{k=1}^{M} \bar{X}_{Dk}(\hat{\beta}_{Dk} - \hat{\beta}_{Lk}), \quad [3]$$

$$\hat{\Delta}_X = \sum_{k=1}^{M} (\bar{X}_{Dk} - \bar{X}_{Lk})\hat{\beta}_{Lk}, \quad [4]$$

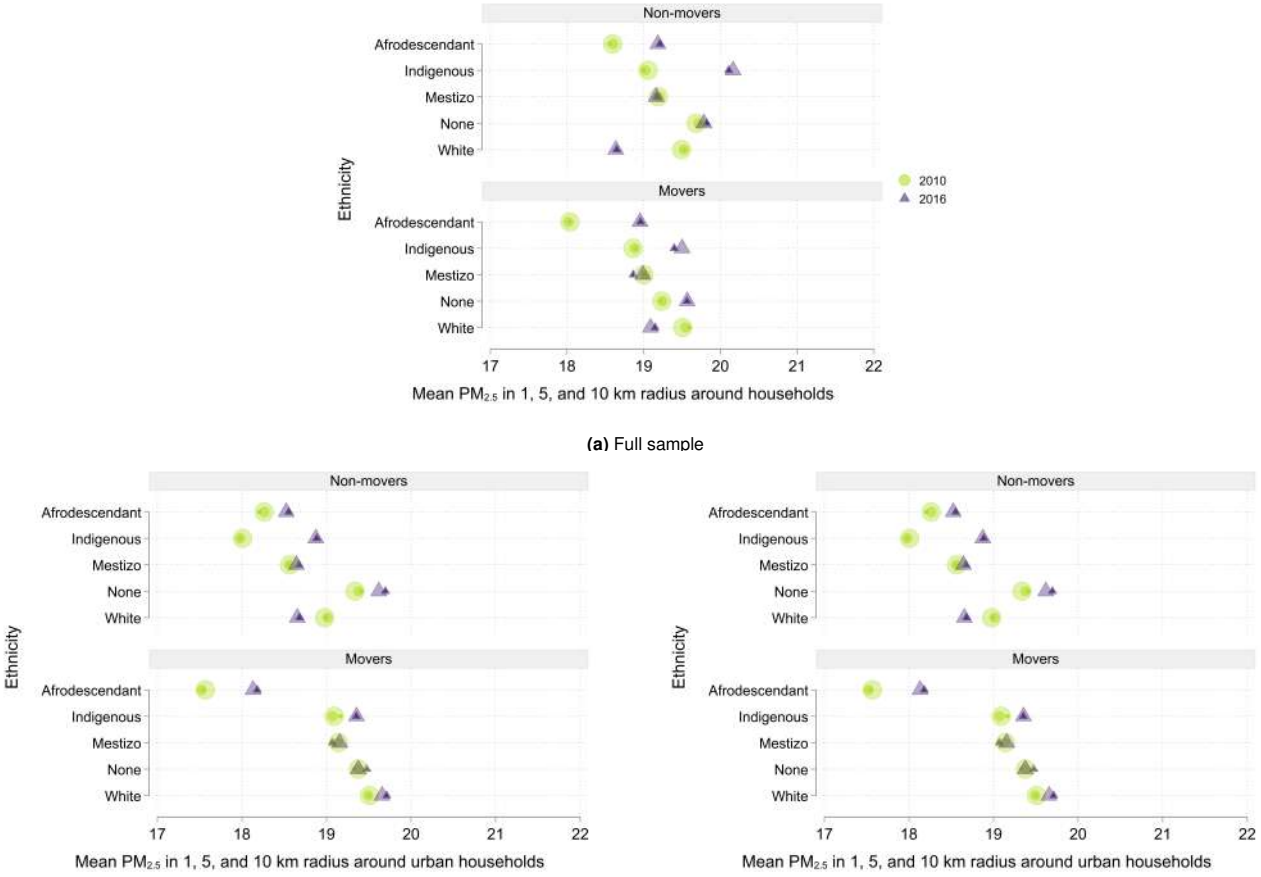
where $(\hat{\beta}_{D0} - \hat{\beta}_{L0})$ represents the omitted group differences, and where $\bar{X}_{gk}$ and $\hat{\beta}_{gk}$ represent the kth element of $\bar{X}_g$ and $\hat{\beta}_g$, respectively. $\left(\bar{X}_{Dk} - \bar{X}_{Ak}\right)\hat{\beta}_{Lk}$ and $\bar{X}_{Dk}(\hat{\beta}_{Dk} - \hat{\beta}_{Lk})$ are the respective contributions of the kth covariate to the composition and stratification effects.

While transitioning from a general to a detailed decomposition is computationally straightforward, some of the coding decisions made during the general decomposition carry greater implications for the detailed decomposition. The omitted group problem is perhaps the most significant concern. Various elements in Δ_S are influenced by the choice of the omitted group, as the interpretation of covariate coefficients hinges on the counterfactual difference in Y when each covariate k is shifted from its value in the omitted group to its average value $\bar{X}_{Dk}$. The most important implication of this limitation is that the categorical variables must be coded such that the omitted group has economic meaning. In this particular case, since our categorical variables are ethnicity (included categories are mestizo, indigenous and afro-descendant), unmet needs (=1 if ≥ 1 unmet need) and migration status (=1 if the person migrated in the past five years), our omitted group comprises white non-migrants with no unmet needs. This choice is reasonable, as it aligns with the common practice in the literature of using a less marginalized population as the reference group.

Robustness tests and additional results

Additional results.

Self-reported ethnoracial categories and pollution exposure. Figure S5 shows descriptive results comparable to Figure 3, but correlating pollution exposure with ethnicity instead of skin tone. The findings reveal that most non-white respondents experienced air quality deterioration.



(a) Full sample

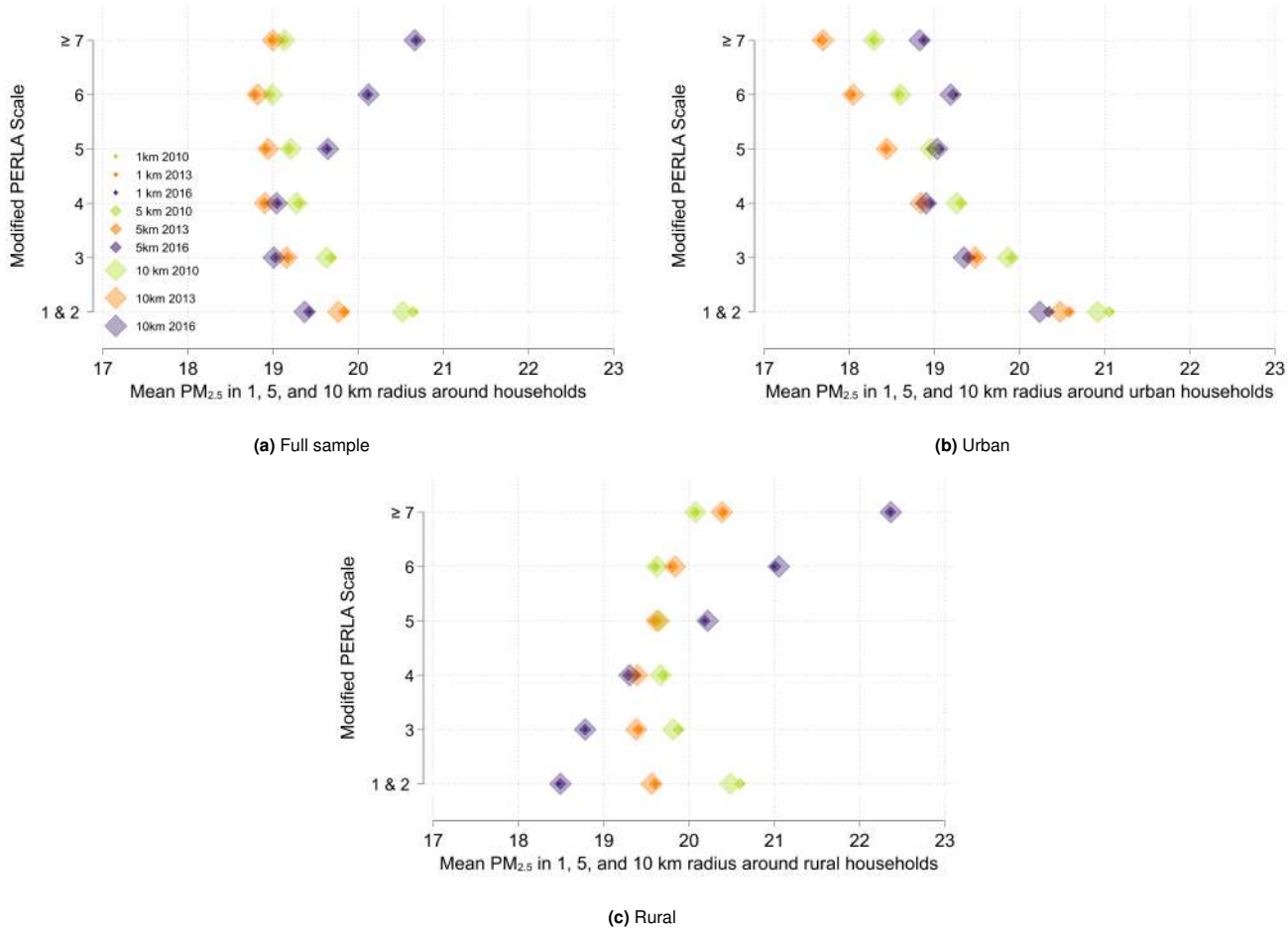
(b) Urban sample, by migration status

(c) Rural sample, by migration status

Fig. S5. Average $PM_{2.5}$ exposure by skin tone and self-reported ethnicity

Notes: Estimates from 2010 and 2016 ELCA. **(b)** includes households residing in urban areas and **(c)** households residing in rural areas when first interviewed. Annual mean $PM_{2.5}$ exposure is calculated by spatially averaging pollution values in a buffer of 1 km, 5 km, or 10 km around each household.

Including 2013 in the analysis. Figure S6 presents the same results as in Figure 3 but includes the buffers around the households in 2013. The graphs show that the improvement in pollution exposure for lighter skin colors was gradual for both the rural and urban populations but was more significant in the rural sample.



(a) Full sample

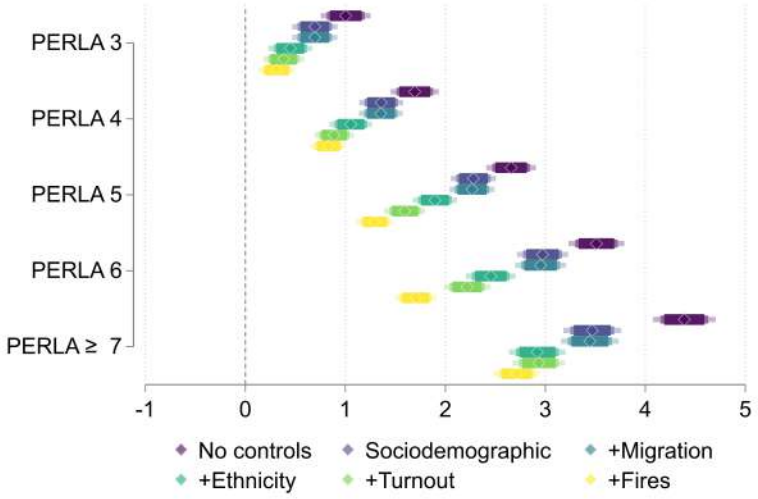
(b) Urban

(c) Rural

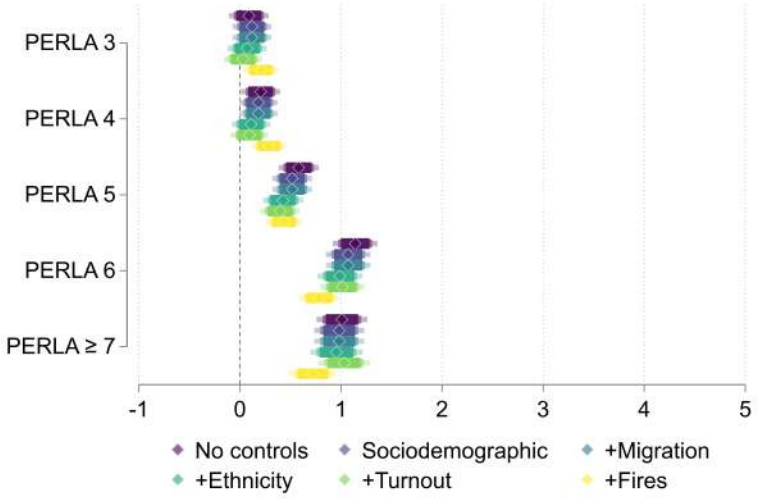
Fig. S6. Average $PM_{2.5}$ exposure by skin tone

Notes: Estimates using 2010 and 2016 ELCA. **(b)** includes households residing in urban areas when first interviewed and **(c)** households residing in rural areas. We calculate annual mean $PM_{2.5}$ exposure by spatially averaging pollution values in a buffer of 1 km, 5 km, and 10 km around each household.

Skin tone and pollution exposure, urban and rural households. Figure S7 shows the same results as Figure 4, but splitting the sample between urban and rural households (based on their 2010 place of residence).



(a) Rural



(b) Urbanl

Fig. S7. Δ 2010-2016 $PM_{2.5}$ by skin tone and urban/rural status

Notes: Associations between skin tone and the difference in $PM_{2.5}$ between 2016 and 2010. The figure includes 99%, 95%, and 90% confidence intervals (fading color bars). Covariates are progressively added, starting with no controls (purple) and ending with the most complete specification (yellow), which includes sociodemographic controls (income, an unmet needs indicator, urbanization of the county of residence), migrant status, ethnicity indicators (none, mestizo, indigenous and afrodescendant), voter turnout, and urban and rural upwind probability-weighted fire exposure.

Robustness to variations in buffer radius. This section demonstrates the robustness of the results when we vary the size of the buffer used to calculate exposure. Three dependent variables are utilized. The first dependent variable measures the average particulate matter ($PM_{2.5}$) in different-sized buffers around each household location in 2010. The second dependent variable measures the average particulate matter ($PM_{2.5}$) in different-sized buffers around each household location in 2016. The third dependent variable captures the difference between $PM_{2.5}$ in 2016 and $PM_{2.5}$ in 2010 within the respective buffer. The variables of interest in this section are the different PERLA skin tone categories, with the omitted category being the lightest shade (Perla skin tones 1 and 2). The analysis incorporates controls employed in our main specification.

10 km buffer. Table S5 presents the association between skin tone and pollution exposure, using a 10 km buffer around households' coordinates to generate the spatial average of pollution concentration.

	(1) $PM_{2.5}$ 2010	(2) $PM_{2.5}$ 2010	(3) $PM_{2.5}$ 2016	(4) $PM_{2.5}$ 2016	(5) Diff $PM_{2.5}$ 2016-2010	(6) Diff $PM_{2.5}$ 2016-2010
PERLA 3	-0.895***	-0.540***	-0.362***	-0.514***	0.532***	0.156***
	(0.0710)	(0.0586)	(0.0923)	(0.0813)	(0.0622)	(0.0522)
PERLA 4	-1.246***	-0.773***	-0.330***	-0.517***	0.914***	0.485***
	(0.0690)	(0.0574)	(0.0897)	(0.0797)	(0.0605)	(0.0511)
PERLA 5	-1.316***	-0.717***	0.274***	-0.266***	1.588***	0.890***
	(0.0712)	(0.0598)	(0.0926)	(0.0830)	(0.0624)	(0.0531)
PERLA 6	-1.534***	-0.761***	0.750***	0.123	2.281***	1.447***
	(0.0799)	(0.0676)	(0.104)	(0.0940)	(0.0700)	(0.0598)
PERLA $\geq$ 7	-1.391***	-0.627***	1.292***	0.914***	2.681***	1.882***
	(0.0902)	(0.0776)	(0.117)	(0.108)	(0.0790)	(0.0684)
Observations	23,304	23,304	23,305	23,305	23,302	23,302
R-squared	0.022	0.343	0.027	0.253	0.112	0.381
Controls		✓		✓		✓

Table S5. Skin tone and $PM_{2.5}$ exposure, ELCA - 10 km buffers

Notes: OLS estimation of the relationship between PERLA skin tone and PM2.5 in 2010 (Columns 1 and 2), in 2016 (Columns 3 and 4) and the difference in the PM2.5 from 2010 to 2016 (Columns 5 and 6). Exposure is calculated using a 10-km buffer around the individual's household coordinates. Controls include sociodemographic (income, an unmet needs indicator, urbanization of the county of residence), migrant status, ethnicity indicators (none, mestizo, indigenous and afrodescendant), voter turnout, and urban and rural upwind probability-weighted fire exposure. Standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

1 km buffer. Table S6 presents the association between skin tone and pollution exposure, using a 1 km buffer around households' coordinates to generate the spatial average of pollution concentration.

	(1) $PM_{2.5}$ 2010	(2) $PM_{2.5}$ 2010	(3) $PM_{2.5}$ 2016	(4) $PM_{2.5}$ 2016	(5) Diff $PM_{2.5}$ 2016-2010	(6) Diff $PM_{2.5}$ 2016-2010
PERLA 3	-0.988***	-0.609***	-0.391***	-0.543***	0.595***	0.208***
	(0.0765)	(0.0635)	(0.0957)	(0.0850)	(0.0650)	(0.0549)
PERLA 4	-1.365***	-0.862***	-0.373***	-0.551***	0.990***	0.554***
	(0.0744)	(0.0622)	(0.0931)	(0.0832)	(0.0632)	(0.0536)
PERLA 5	-1.498***	-0.859***	0.215**	-0.308***	1.710***	1.003***
	(0.0768)	(0.0648)	(0.0960)	(0.0867)	(0.0652)	(0.0558)
PERLA 6	-1.731***	-0.903***	0.694***	0.0947	2.423***	1.574***
	(0.0861)	(0.0732)	(0.108)	(0.0982)	(0.0731)	(0.0629)
PERLA $\geq$ 7	-1.564***	-0.733***	1.251***	0.913***	2.814***	1.989***
	(0.0973)	(0.0841)	(0.122)	(0.112)	(0.0826)	(0.0718)
Observations	23,304	23,304	23,305	23,305	23,302	23,302
R-squared	0.024	0.338	0.025	0.240	0.112	0.375
Controls		✓		✓		✓

Table S6. Skin tone and $PM_{2.5}$ exposure, ELCA - 1 km buffers

Notes: OLS estimation of the relationship between PERLA skin tone and PM2.5 in 2010 (Columns 1 and 2), in 2016 (Columns 3 and 4) and the difference in the PM2.5 from 2010 to 2016 (Columns 5 and 6). Exposure is calculated using a 1-km buffer around the individual's household coordinates. Controls include sociodemographic (income, an unmet needs indicator, urbanization of the county of residence), migrant status, ethnicity indicators (none, mestizo, indigenous and afrodescendant), voter turnout, and urban and rural upwind probability-weighted fire exposure. Standard errors in parentheses.*** p<0.01, ** p<0.05, * p<0.1.

Additional robustness tests.

Varying imputation and clustering techniques. Tables S7 and S8 demonstrate our main findings hold under the following specifications: 1) excluding individuals with imputed skin tones (Columns 1-3), 2) clustering standard errors at the household level (Columns 4-6), and 3) restricting the sample to only household heads (Columns 7-9). It is important to note that the ELCA tracks individuals rather than households, allowing us to follow household members as they relocate and form new households by 2016, making the individual-level regression our preferred specification.

	(1) $PM_{2.5}$ 2010 No Imputation	(2) $PM_{2.5}$ 2016 No Imputation	(3) Diff $PM_{2.5}$ 2016-2010 No Imputation	(4) $PM_{2.5}$ 2010 Household Cluster	(5) $PM_{2.5}$ 2016 Household Cluster	(6) Diff $PM_{2.5}$ 2016-2010 Household Cluster	(7) $PM_{2.5}$ 2010 Household Head	(8) $PM_{2.5}$ 2016 Household Head	(9) Diff $PM_{2.5}$ 2016-2010 Household Head
PERLA 3	-0.952***	-0.324***	0.625***	-0.976***	-0.358***	0.615***	-1.004***	-0.372**	0.633***
	(0.0848)	(0.106)	(0.0730)	(0.0209)	(0.0594)	(0.0446)	(0.139)	(0.172)	(0.115)
PERLA 4	-1.296***	-0.305***	0.988***	-1.345***	-0.339***	1.003***	-1.269***	-0.396**	0.873***
	(0.0827)	(0.104)	(0.0712)	(0.0414)	(0.0953)	(0.0601)	(0.135)	(0.167)	(0.111)
PERLA 5	-1.410***	0.228**	1.635***	-1.466***	0.195**	1.658***	-1.390***	0.121	1.511***
	(0.0855)	(0.107)	(0.0736)	(0.0391)	(0.0873)	(0.0521)	(0.139)	(0.172)	(0.115)
PERLA 6	-1.547***	0.853***	2.398***	-1.608***	0.802***	2.408***	-1.555***	0.732***	2.287***
	(0.0962)	(0.121)	(0.0828)	(0.0682)	(0.152)	(0.0906)	(0.154)	(0.192)	(0.128)
PERLA $\geq$ 7	-1.214***	1.878***	3.090***	-1.492***	1.302***	2.792***	-1.492***	1.290***	2.782***
	(0.160)	(0.200)	(0.137)	(0.0469)	(0.0856)	(0.0421)	(0.174)	(0.215)	(0.144)
Observations	15,976	15,977	15,974	17,280	17,281	17,278	6,642	6,642	6,642
R-squared	0.022	0.024	0.098	0.025	0.027	0.112	0.020	0.027	0.112
Controls									

Table S7. Skin tone and PM$_{2.5}$ exposure, additional robustness tests, no controls

Notes: OLS estimation of the relationship between PERLA skin tone and PM2.5 without imputating skin colors (Columns 1 to 3), cluster standard errors to the household level (Columns 4 to 6) and only using the information of the household head (Columns 5 and 6). Standard errors in parentheses.*** p<0.01, ** p<0.05, * p<0.1.

	(1) $PM_{2.5}$ 2010 No Imputation	(2) $PM_{2.5}$ 2016 No Imputation	(3) Diff $PM_{2.5}$ 2016-2010 No Imputation	(4) $PM_{2.5}$ 2010 Household Cluster	(5) $PM_{2.5}$ 2016 Household Cluster	(6) Diff $PM_{2.5}$ 2016-2010 Household Cluster	(7) $PM_{2.5}$ 2010 Household Head	(8) $PM_{2.5}$ 2016 Household Head	(9) Diff $PM_{2.5}$ 2016-2010 Household Head
PERLA 3	-0.551***	-0.509***	0.205***	-0.567***	-0.538***	0.190***	-0.584***	-0.518***	0.200**
	(0.0701)	(0.0933)	(0.0605)	(0.0348)	(0.0640)	(0.0280)	(0.115)	(0.153)	(0.0968)
PERLA 4	-0.768***	-0.521***	0.513***	-0.801***	-0.557***	0.509***	-0.831***	-0.603***	0.461***
	(0.0689)	(0.0916)	(0.0593)	(0.0405)	(0.0946)	(0.0412)	(0.112)	(0.150)	(0.0945)
PERLA 5	-0.745***	-0.319***	0.886***	-0.789***	-0.340***	0.899***	-0.814***	-0.342**	0.883***
	(0.0721)	(0.0958)	(0.0619)	(0.0454)	(0.0988)	(0.0373)	(0.116)	(0.155)	(0.0979)
PERLA 6	-0.732***	0.102	1.469***	-0.770***	0.0939	1.481***	-0.770***	0.118	1.480***
	(0.0819)	(0.109)	(0.0703)	(0.0566)	(0.117)	(0.0632)	(0.131)	(0.175)	(0.110)
PERLA ≥ 7	-0.348**	1.719***	2.245***	-0.612***	0.933***	1.915***	-0.599***	0.970***	2.001***
	(0.137)	(0.182)	(0.118)	(0.0653)	(0.0793)	(0.0340)	(0.151)	(0.202)	(0.126)
Observations	15,976	15,977	15,974	17,280	17,281	17,278	6,642	6,642	6,642
R-squared	0.342	0.261	0.388	0.345	0.260	0.389	0.343	0.243	0.379
Controls	✓	✓	✓	✓	✓	✓	✓	✓	✓

Table S8. Skin tone and $PM_{2.5}$ exposure, additional robustness tests, with controls

Notes: OLS estimation of the relationship between PERLA skin tone and PM2.5 without imputating skin colors (Columns 1 to 3), cluster standard errors to the household level (Columns 4 to 6) and only using the information of the household head (Columns 5 and 6). Controls include sociodemographic (income, an unmet needs indicator, urbanization of the county of residence), migrant status, ethnicity indicators (none, mestizo, indigenous and afrodescendant), voter turnout, and urban and rural upwind probability-weighted fire exposure. Standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Decomposition analysis.

Pooled and threefold decompositions Figure S8 shows the pooled and threefold versions of our main KOBD results displayed in Figure 7. The pooled KOBD entails comparing each group against a pooled reference category, in this case the full sample averages. The threefold KODB splits the overall difference into three components: endowment effects (differences in covariates), coefficient effects (differences in the relationship between covariates and outcomes), and interaction effects. This introduces additional flexibility by separately identifying how differences in characteristics and the returns to those characteristics contribute to disparities, while the the standard twofold decomposition that focuses only on endowments and coefficients(17).

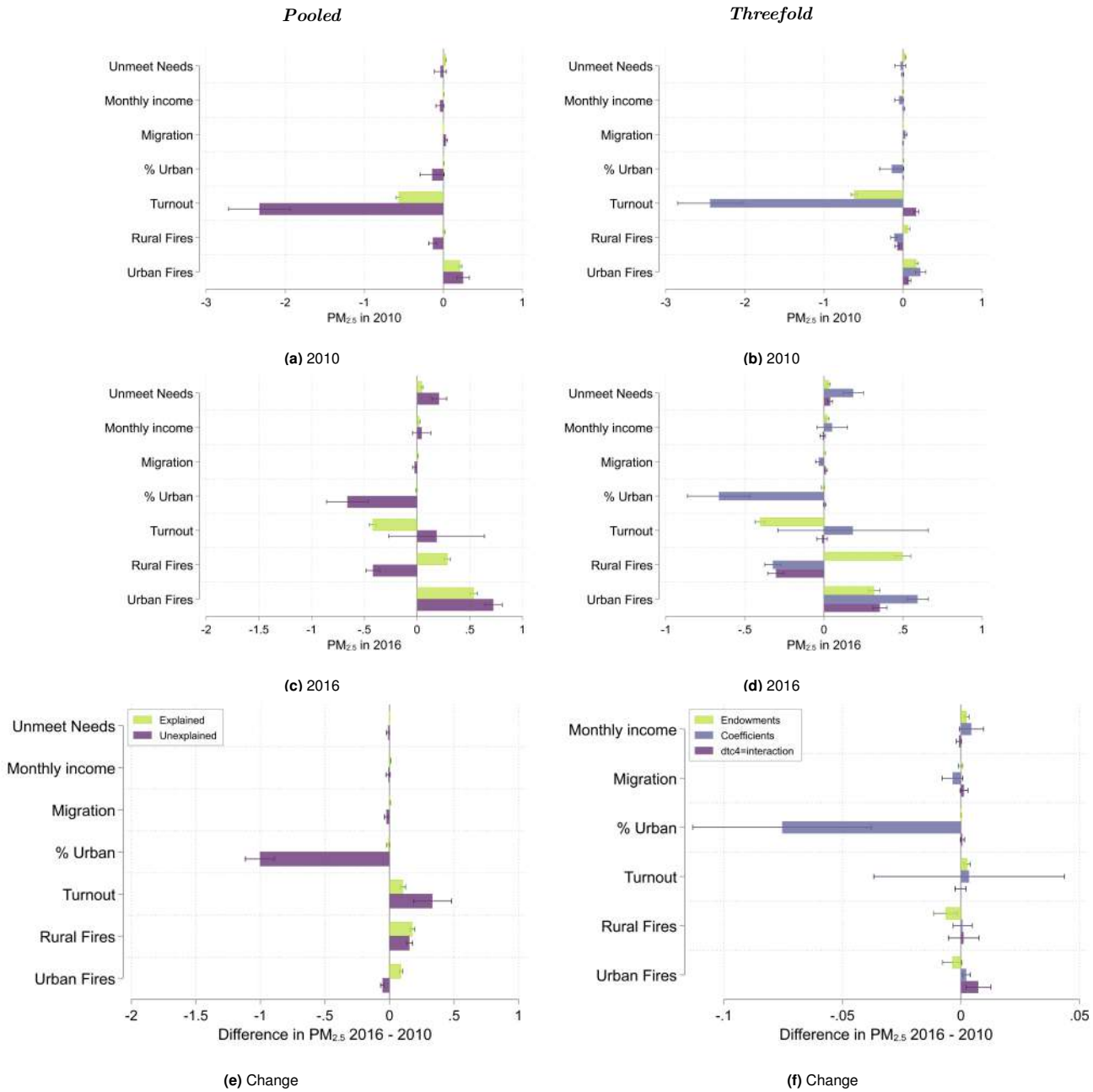


Fig. S8. Explained and unexplained part of the $PM_{2.5}$ skin color gap, Pooled and Threefold KOB Decompositions

Multiple KOBD cutoffs Figure S9 shows the results of conducting KOBDs using different cutoffs to define groups L and D. In particular, one for each cutoff at tone $j \in [3, 7]$.

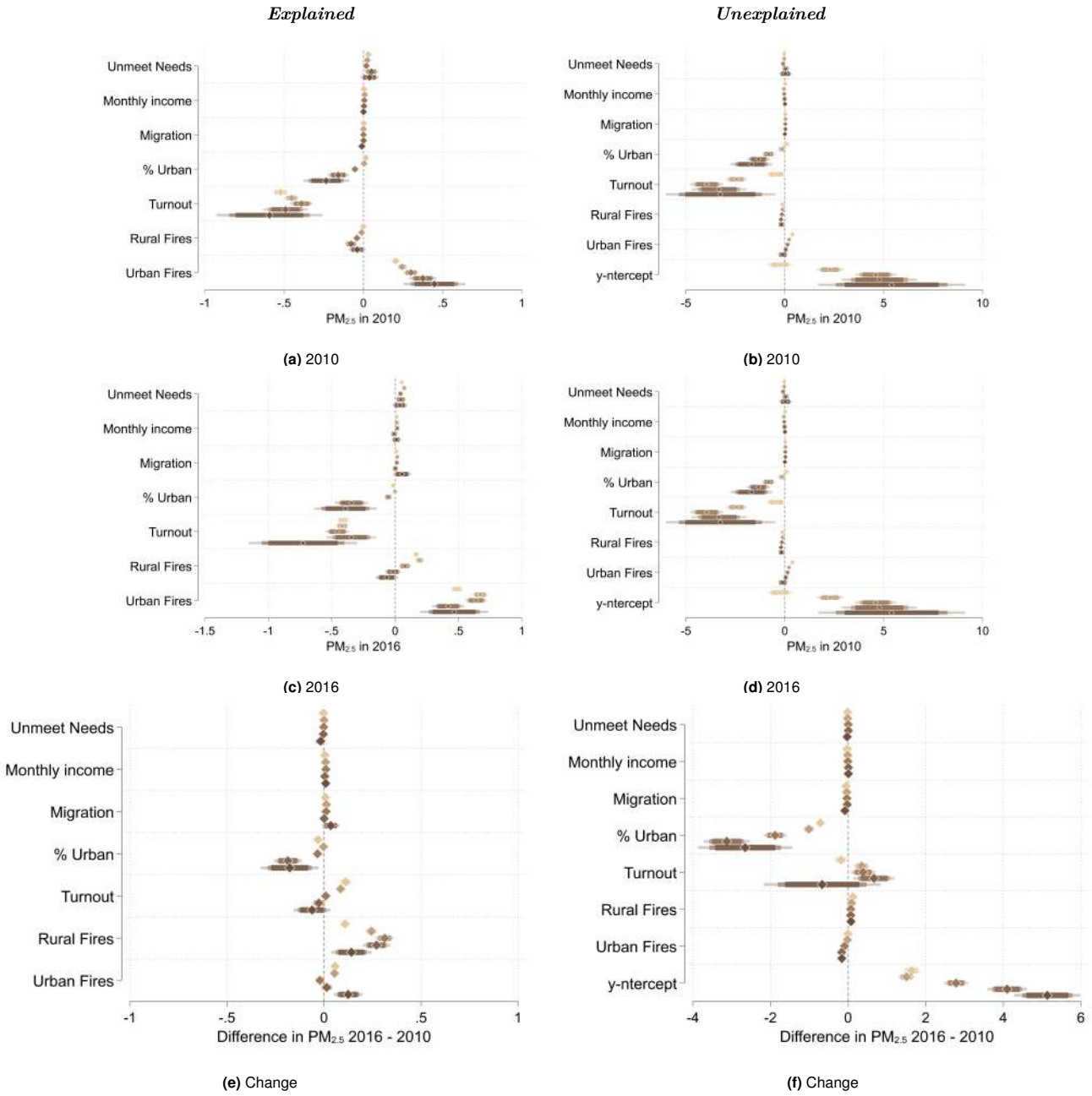


Fig. S9. Explained and unexplained part of the $PM_{2.5}$ skin color gap, KOB Decompositions, by various skin color cutoffs

 Note: Urban fires are defined broadly, including peri-urban and rural areas, as long as the fire occurred within the boundaries of a populated center, defined as any conglomerate of at least 20 contiguous residences.

	(1)	(2)
	$PM_{2.5}$	$PM_{2.5}$
	2010	2016
Urban fires	-0.0177**	0.125***
	(0.00817)	(0.00596)
Rural fires	1.045***	2.099***
	(0.0676)	(0.0532)
Observations	23,410	23,408
R-squared	0.016	0.189

Table S9. Urban and Rural Fires and PM$_{2.5}$

Notes: OLS estimation of the relationship between urban and rural upwind probability- and intensity- weighted fire exposure (K) and PM2.5. Standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

References

1. LM Castaño Mesa, *Colombia in motion. 2010-2013-2016: The changes in the life of households based on the Colombian Longitudinal Survey (ELCA) by Universidad de los Andes.* (Ediciones Uniandes-Universidad de los Andes), (2018).
2. L Bonilla-Mejía, LF Morales, D Hermida, LA Flórez, The Labor Market Effect of South-to-South Migration: Evidence From the Venezuelan Crisis. *Int. Migr. Rev.* **0**, 01979183231162626 (2023).
3. A Van Donkelaar, et al., Monthly Global Estimates of Fine Particulate Matter and Their Uncertainty. *Environ. Sci. & Technol.* **55**, 15287–15300 (2021).
4. M Liu, et al., Quantifying pm2.5 mass concentration and particle radius using satellite data and an optical-mass conversion algorithm. *ISPRS J. Photogramm. Remote. Sens.* **158**, 90–98 (2019).
5. A Donkelaar, et al., Optimal estimation for global ground-level fine particulate matter concentrations. *J. Geophys. Res. Atmospheres* **118** (2013).
6. Y Wei, et al., The impact of exposure measurement error on the estimated concentration–response relationship between long-term exposure to pm2.5 and mortality. *Environ. Heal. Perspectives* **130**, 077006 (2022).
7. J Richmond-Bryant, TC Long, Influence of exposure measurement errors on results from epidemiologic studies of different designs. *J. Expo. Sci. & Environ. Epidemiol.* **30**, 420–429 (2020).
8. NASA, Viirs-modis data (https://www.earthdata.nasa.gov/learn/find-data/near-real-time/viirs) (year?) Accessed: 2024-07-18.
9. NASA, Firms data (https://www.earthdata.nasa.gov/learn/find-data/near-real-time/firms/about-firms) (year?) Accessed: 2024-07-18.
10. European Centre for Medium-Range Weather Forecasts, Era5-land: Ecmwf's global reanalysis for land applications (2024) Accessed: 2024-08-21.
11. European Centre for Medium-Range Weather Forecasts, How to calculate wind speed and wind direction from u and v components of the wind (2024) Accessed: 2024-07-22.
12. MA Rangel, TS Vogl, Agricultural fires and health at birth. *Rev. Econ. Stat.* **101**, 616–630 (2019).
13. NASA, Firms faq (https://www.earthdata.nasa.gov/faq/firms-faq) (year?) Accessed: 2024-07-18.
14. HK Pullabhotla, M Souza, Air pollution from agricultural fires increases hypertension risk. *J. Environ. Econ. Manag.* **115**, 102723 (2022).
15. Departamento Administrativo Nacional de Estadística (DANE), Geoportal dane (https://www.dane.gov.co/files/geoportal-provisional/index.html) (2018) Accessed: 2024-08-21.
16. WH Organization, *WHO global air quality guidelines: particulate matter (PM2.5 and PM10), ozone, nitrogen dioxide, sulfur dioxide and carbon monoxide.* (World Health Organization), (2021).
17. N Fortin, T Lemieux, S Firpo, Decomposition methods in economics in *Handbook of labor economics.* (Elsevier) Vol. 4, pp. 1–102 (2011).
18. L Cain, D Hernandez-Cortes, C Timmins, P Weber, Recent findings and methodologies in economics research in environmental justice. *Rev. Environ. Econ. Policy* **18**, 116–142 (2024).