Article

Air pollution exposure disparities across US population and income groups

https://doi.org/10.1038/s41586-021-04190-y

Received: 13 July 2020

Accepted: 26 October 2021

Published online: 12 January 2022

Check for updates

Abdulrahman Jbaily[1]✉, Xiaodan Zhou[2], Jie Liu[2], Ting-Hwan Lee[2], Leila Kamareddine[3], Stéphane Verguet[1] & Francesca Dominici[3,4]✉

Air pollution contributes to the global burden of disease, with ambient exposure to fine particulate matter of diameters smaller than 2.5 μm ($PM_{2.5}$) being identified as the fifth-ranking risk factor for mortality globally[1]. Racial/ethnic minorities and lower-income groups in the USA are at a higher risk of death from exposure to $PM_{2.5}$ than are other population/income groups[2–5]. Moreover, disparities in exposure to air pollution among population and income groups are known to exist[6–17]. Here we develop a data platform that links demographic data (from the US Census Bureau and American Community Survey) and $PM_{2.5}$ data[18] across the USA. We analyse the data at the tabulation area level of US zip codes (*N* is approximately 32,000) between 2000 and 2016. We show that areas with higher-than-average white and Native American populations have been consistently exposed to average $PM_{2.5}$ levels that are lower than areas with higher-than-average Black, Asian and Hispanic or Latino populations. Moreover, areas with low-income populations have been consistently exposed to higher average $PM_{2.5}$ levels than areas with high-income groups for the years 2004–2016. Furthermore, disparities in exposure relative to safety standards set by the US Environmental Protection Agency[19] and the World Health Organization[20] have been increasing over time. Our findings suggest that more-targeted $PM_{2.5}$ reductions are necessary to provide all people with a similar degree of protection from environmental hazards. Our study is observational and cannot provide insight into the drivers of the identified disparities.

Several studies have reported evidence of statistically significant associations between exposure to $PM_{2.5}$ (fine particles with a mass median aerodynamic diameter of less than 2.5 μm) and adverse health outcomes[21–34]. It is also well documented[2–5] that racial/ethnic minorities and people of low socioeconomic status in the US are at a higher risk of death from being exposed to $PM_{2.5}$. Disparities in exposure to air pollution among racial/ethnic and socioeconomic groups in the US are also known to exist[6–15]. Disparities may be represented as either relative or absolute comparisons, where absolute disparities are assessed as absolute differences between groups, while relative disparities are scale invariant[16]. A recent study showed that absolute disparities in $PM_{2.5}$ between more- and less-polluted areas in the US have declined substantially between 1981 and 2001, but that relative disparities persist[17].

Here we advance this line of work by studying relative disparities across income groups (namely income deciles) and racial/ethnic groups for the years 2000 to 2016. We define ethnic groups as those with shared cultural characteristics, and racial groups as those with physical differences that they consider to be socially important[35]. The ethnic group included in this study is the Hispanic or Latino group, and the racial groups are non-Hispanic white, Black, Asian and Native American; for ease of reference, the racial groups are referred to as white, Black, Asian and Native American throughout the text. We stress that our study is descriptive and is not designed to investigate causal aspects related to race. Further novelty of our work includes the investigation of relative disparities related to safety standards (National Ambient Air Quality standards (NAAQS) set by the US Environmental Protection Agency (EPA)[3,19] at 12 μg m^{-3}, and the guideline set by the World Health Organization (WHO)[3,20] at 10 μg m^{-3}) and their trends over the study period.

Our findings regarding relative disparities indicate the importance of strong, targeted air-pollution-reduction strategies, not only to reduce overall air-pollution levels but also to move closer towards the EPA's aim to provide all people with the same degree of protection from environmental hazards. Nonetheless, our evidence should be interpreted in the context of the limitations of the data at hand. Notably, the $PM_{2.5}$ concentrations used here rely on aerosol optical depth (AOD) data (see Methods), but AOD-based particulate estimates tend to underestimate pollution at higher levels and overestimate it at very low levels. In addition we have used US Census data, which are not available for every year of the study period (see Methods), so we have also incorporated interpolation techniques for parts of this period; hence our results are subject to the assumptions made in the interpolation. Finally, average $PM_{2.5}$ concentrations across zip code tabulation areas (ZCTAs) are used,

[1]Department of Global Health and Population, Harvard T. H. Chan School of Public Health, Boston, MA, USA. [2]Environmental Systems Research Institute, Redlands, CA, USA. [3]Department of Biostatistics, Harvard T. H. Chan School of Public Health, Boston, MA, USA. [4]Harvard Data Science Initiative, Cambridge, MA, USA. ✉e-mail: ajbaily@hsph.harvard.edu; fdominic@hsph.harvard.edu

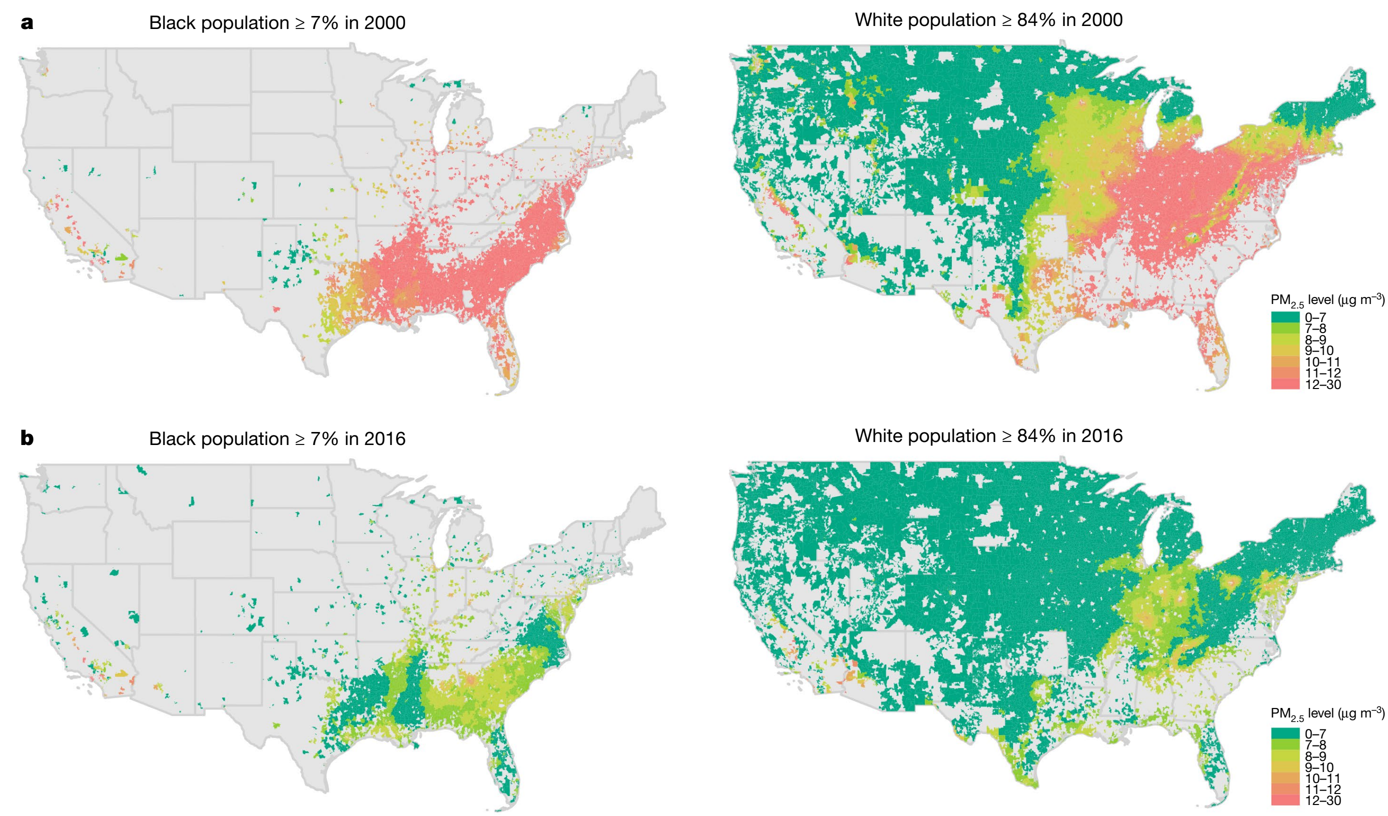


Fig. 1 | Average PM$_{2.5}$ concentration in 2000 and 2016 across ZCTAs in which Black or white populations are overrepresented. We used the white population fraction of the ZCTA population to compute the average white population fraction (aWpf) across all ZCTAs (approximately 84%). Similarly, we computed the average Black population fraction (aBpf) (approximately 7%). **a**, PM$_{2.5}$ levels for the year 2000 in ZCTAs with a Black population fraction of greater than aBpf (left), and in ZCTAs with a white population fraction of greater than aWpf (right). In this year, high PM$_{2.5}$ concentrations exist in almost all ZCTAs with a Black population of more than aBpf, while a wide range of low and high PM$_{2.5}$ concentrations exist in ZCTAs with a white population of more than aWpf. **b**, The same information for the year 2016. Similar maps for other racial/ethnic groups for 2000 and 2016 are shown in Extended Data Fig. 3a, b and Supplementary Videos 2, 3. Note that Hawaii and Alaska are not shown. Imagery provided courtesy of Esri, HERE, Garmin, FAO, NOAA, USGS, ©OpenStreetMap contributors, and the GIS User Community.

which can mask the relationship between income and pollution levels of neighbourhoods within large ZCTAs, and are subject to more error in cases in which substantial within-ZCTA variation in pollution occurs.

Disparities among racial/ethnic groups

The US EPA is required to reexamine the NAAQS every five years, and in 2012, the EPA set the NAAQS for PM$_{2.5}$ to 12 μg m^{-3} (refs. [3,19,36]). On average across the US, we found that PM$_{2.5}$ concentration levels decreased from 2000 to 2016, with the population-weighted average of PM$_{2.5}$ having decreased by 40.4% from the year 2000 (12.6 μg m^{-3}) to 2016 (7.5 μg m^{-3}) (see Supplementary Video 1 and Extended Data Figs. 1a , 2). We also found that the percentage of the population exposed to PM$_{2.5}$ levels higher than 12 μg m^{-3} decreased from 57.3% in 2000 to 4.5% in 2016.

Next, we visualized and examined disparities in exposure to air pollution among racial/ethnic groups. For each racial/ethnic group (white, Black, Asian, Native American and Hispanic or Latino), we constructed a map that shows ZCTAs in which the race/ethnicity is overrepresented. In the case of the white population, for example, we used the white population fraction of the ZCTA population to compute the average white population fraction across all ZCTAs (roughly 84.2%). We then set the margin at 84%, and only show on the map those ZCTAs with a white population fraction that exceeds this margin. The margins for the remaining racial/ethnic groups were computed similarly and are shown in Extended Data Fig. 3. For ease of exposition, we present findings for only two groups in the main text (white and Black groups; see Fig. 1), and include the other racial/ethnic groups in Extended Data Fig. 3 and Supplementary Videos 2, 3. Figure 1a, b shows the PM$_{2.5}$ distributions in ZCTAs in which the Black and white populations are overrepresented for the years 2000 and 2016. We found that ZCTAs in which the Black population is overrepresented (left maps) are dominated by higher PM$_{2.5}$ concentrations relative to ZCTAs with white overrepresentation (right maps), in both 2000 and 2016. Furthermore, we see a steeper decline in PM$_{2.5}$ among the latter.

We also computed the population-weighted average PM$_{2.5}$ concentration for every racial/ethnic population (see Methods) (Extended Data Fig. 1b). For all years, we found that the Black, Asian and Hispanic or Latino populations experienced somewhat similar levels of PM$_{2.5}$, which were higher than those experienced by the white population. In 2016, for example, the average PM$_{2.5}$ concentration for the Black population was 13.7% higher than that of the white population and 36.3% higher than that of the Native American population. The Native American population was consistently exposed to the lowest levels of PM$_{2.5}$. Further, we illustrate for the year 2016 how the population-weighted PM$_{2.5}$ average concentration has changed as ZCTAs became more populated by a certain race/ethnicity (Extended Data Fig. 1c). We found that as the Black population increased in a ZCTA, the PM$_{2.5}$ concentration likewise consistently increased, with a steep incline seen for ZCTAs with more than 85% of their population being Black. The trend for the Hispanic or Latino population is similar to that of the Black population. The opposite is seen for the white population: the PM$_{2.5}$ concentration decreased as the density of the white population increased in a ZCTA,

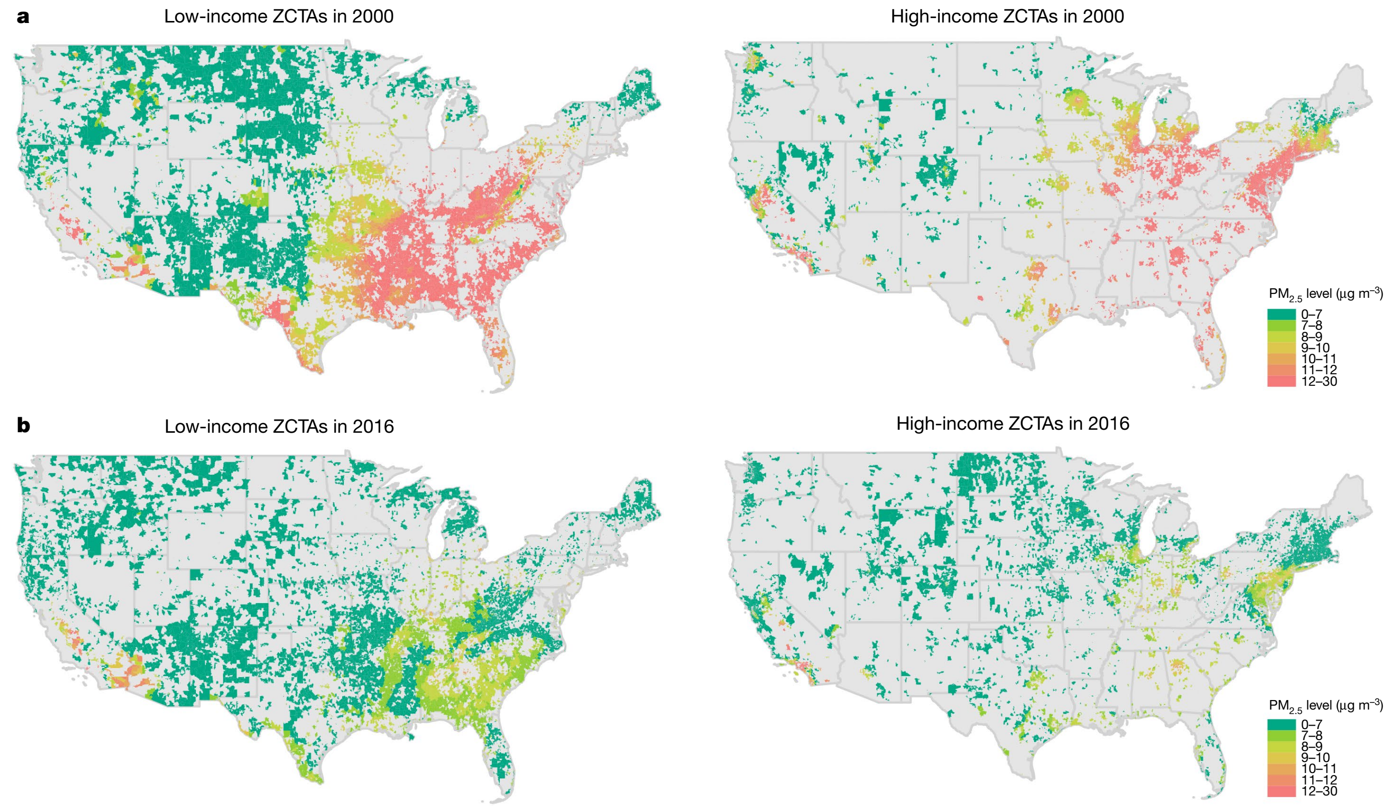


Fig. 2 | Average $PM_{2.5}$ concentration in 2000 and 2016 across low- and high-income ZCTAs. We assigned all ZCTAs percentile ranks from 1 to 100 on the basis of median household income, and categorized them into ten income groups. We designated the lowest and highest three income groups as 'low income' and 'high income', respectively. **a**, $PM_{2.5}$ levels for the year 2000 in low-income (left) and high-income (right) ZCTAs. **b**, The same information for the year 2016. Disparities in exposure to $PM_{2.5}$ between the two groups are apparent, and it can be seen that in both 2000 and 2016, low-income ZCTAs were exposed to higher $PM_{2.5}$ concentrations than were high-income ZCTAs (Supplementary Video 4). Note that Hawaii and Alaska are not shown. Imagery provided courtesy of Esri, HERE, Garmin, FAO, NOAA, USGS, ©OpenStreetMap contributors, and the GIS User Community.

and a steeper decrease is found for those ZCTAs with a white population fraction exceeding 70%. Furthermore, in ZCTAs in which the population of Native Americans is at least 20%, the average $PM_{2.5}$ concentration dropped to less than 4 μg m^{-3}. For the Asian population, a low number of ZCTAs has a population density above 60%, so data beyond this point are not representative and are not shown.

Disparities among income groups

We next visualized and summarized disparities among income groups. We assigned all ZCTAs percentile ranks from 1 to 100 based on median household income, and categorized them into 10 income groups. We designated the lowest and highest three income groups as low income and high income, respectively, and then split the US map into two maps, comprising ZCTAs defined as low or high income (see Methods). We visualized the distribution of $PM_{2.5}$ concentrations on the two maps for 2000 to 2016 (Supplementary Video 4). The map with low-income ZCTAs appears visually to be dominated by an overall higher concentration of $PM_{2.5}$ as compared with the map with high-income ZCTAs, especially in recent years. We include snapshots of 2000 and 2016 in Fig. 2. We summarized the contents of the maps by computing the population-weighted mean of $PM_{2.5}$ concentration in ZCTAs with low- and high-income groups (Extended Data Fig. 1d); ZCTAs of the low-income group were exposed to only slightly higher $PM_{2.5}$ concentrations for most of the study period (2004–2016)—for example, only 4% higher in 2016. Furthermore, we isolated the effects of income on the disparities among racial/ethnic groups in Extended Data Figs. 1e, f. For the low- and high-income groups, the differences in $PM_{2.5}$ concentration across racial/ethnic groups are similar to those in Extended Data Fig. 1b.

Disparities relative to policy standards

We investigated relative disparities in $PM_{2.5}$ exposure in the context of the current NAAQS (12 μg m^{-3}), the guidelines set by the WHO (10 μg m^{-3}), and a lower limit that might be considered in the future (8 μg m^{-3}). To do so, we estimated across the study period the proportion of every racial/ethnic group that was exposed to $PM_{2.5}$ levels higher than one of the listed safety standards. We defined a state of equality (or lack of relative disparities) among various populations as a state of equal proportions above the chosen safety standard across groups.

First, we ranked the US ZCTAs from the least to the most dense with respect to every racial/ethnic group for each year. For the Black population, for example, we used the Black population fraction in every ZCTA to split ZCTAs into 100 quantiles (Extended Data Fig. 4a); the dark-blue region on the map, representing the ZCTA ranking for the Black population, contains the ZCTAs with the highest ratios of Black population to total ZCTA population, and the light-yellow region contains the ZCTAs with the lowest ratios of Black population to total ZCTA population. Similarly, for the remaining populations, the dark-blue and light-yellow regions on their corresponding maps, respectively, signify high and low proportions of that racial/ethnic group. In Fig. 3a, we again focus on two groups for ease of exposition (Black and white groups are chosen for consistency), and we show the ZCTAs with a $PM_{2.5}$ concentration higher than a threshold of 8 μg m^{-3} for the year 2000. Figure 3a shows that almost half of the ZCTAs with $PM_{2.5}$ concentrations above 8 μg m^{-3} are where the Black

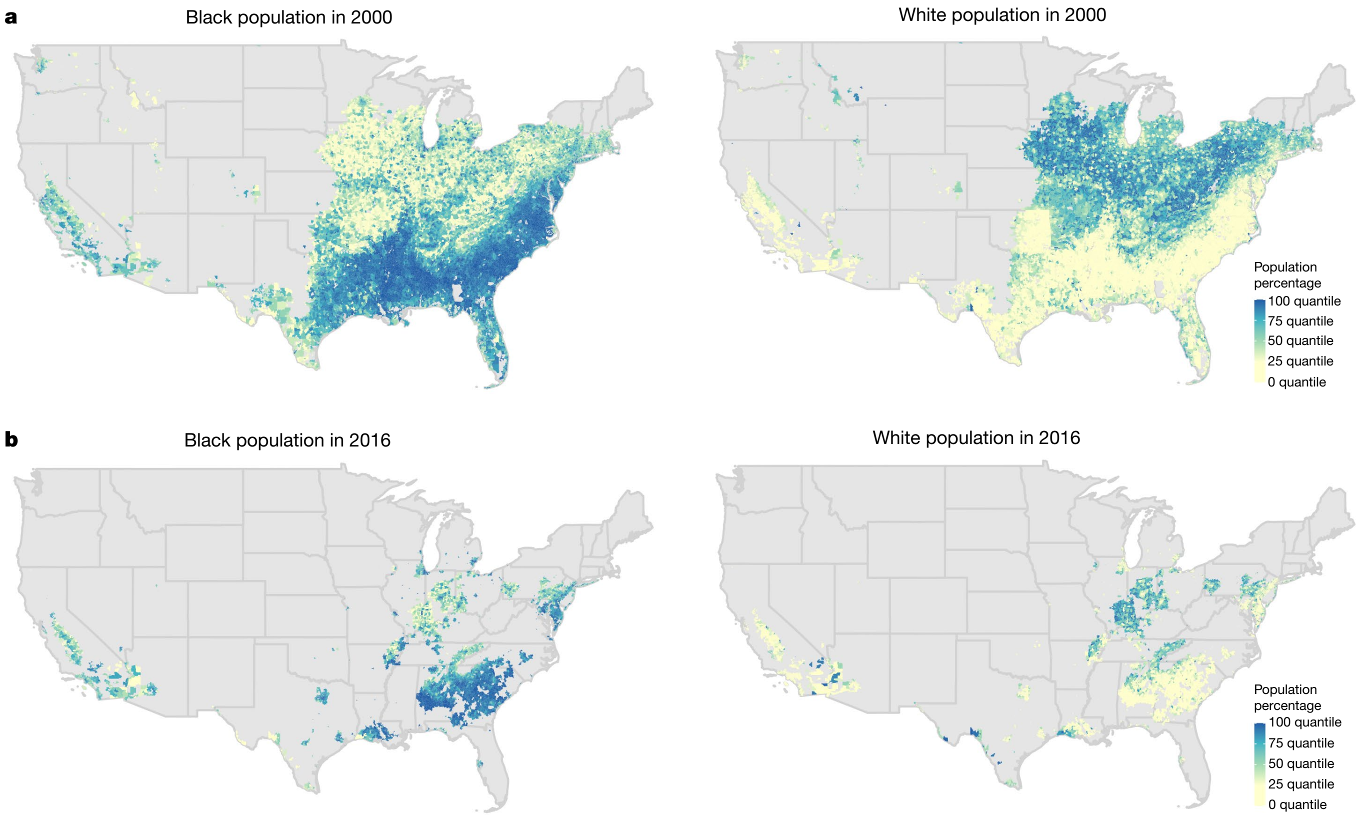


Fig. 3 | US ZCTAs with average $PM_{2.5}$ concentrations of more than 8 μg m^{-3} for Black and white populations in 2000 and 2016. a, **b**, The maps show only those US ZCTAs with $PM_{2.5}$ levels of more than 8 μg m^{-3}, for 2000 (**a**) and 2016 (**b**). The maps on the left are colour-coded on the basis of the fraction of the Black population in ZCTAs; those on the right are colour-coded on the basis of the white population fraction. For example, in the left map in **a**, the dark-blue and light-yellow colours correspond to ZCTAs with, respectively, the largest and smallest Black population percentages of the total ZCTA population in 2000 (or, equivalently, where the Black population is respectively overrepresented and underrepresented in 2000). This map shows that, in 2000, almost half of the ZCTAs with $PM_{2.5}$ levels above 8 μg m^{-3} corresponded to those in which the Black population most lived (almost half of the map is dark blue). By 2016, ZCTAs that remained higher than 8 μg m^{-3} were mainly those that were dominated by the Black population (left map in **b**), with the white population being underrepresented (right map in **b**). Supplementary Videos 5–8 show the distribution of the different racial/ethnic groups across multiple ranges of $PM_{2.5}$ concentrations in 2000 and 2016. Note that Hawaii and Alaska are not shown. Imagery provided courtesy of Esri, HERE, Garmin, FAO, NOAA, USGS, ©OpenStreetMap contributors, and the GIS User Community.

population is concentrated (the southern part of the map, as indicated by the dark-blue region on the Black population map), and the other half is where the white population is concentrated (the northern part of the map, as indicated by the dark-blue region on the white population map). We reproduced the scenario for 2016 (Fig. 3b), finding that the majority of ZCTAs still above 8 μg m^{-3} were those with a concentrated Black population (the majority of the 2016 map representing the Black population (left) is dark blue and the majority of that representing the white population map (right) is light yellow). This visualization shows that $PM_{2.5}$ reductions between 2000 and 2016 have not benefited all areas of the US equally, and consequently resulted in an increase in relative disparities in exposure to air pollution (as will be shown numerically later). We also extended Fig. 3 to include the Asian, Native American and Hispanic or Latino populations, and present the results in Extended Data Fig. 4. Here we used a threshold of 8 μg m^{-3} to allow clearer visualization of the disparities. A lower number of ZCTAs is exposed to $PM_{2.5}$ concentrations above 10 μg m^{-3} and 12 μg m^{-3}, so visualizations at these thresholds are not as clear. Nevertheless, the same visualization is repeated for multiple thresholds, including 10 μg m^{-3} and 12 μg m^{-3}, in Supplementary Videos 5–8.

Second, we provide numerical summaries of disparities in the proportions of racial/ethnics groups exposed to $PM_{2.5}$ levels above the chosen standard by using the coefficient of variation (CoV) (Fig. 4 and Methods). The solid blue line in Fig. 4 shows that 89% of the population was exposed to $PM_{2.5}$ levels higher than 8 μg m^{-3} in 2000, but only 41% in 2016. However, Fig. 4 also reveals that relative disparities among racial/ethnic groups in exposure to $PM_{2.5}$ levels higher than 8 μg m^{-3} (solid blue bars) have increased from 2000 to 2016 (see Methods for more details on interpreting change in disparities across years using CoV). This result aligns with the relative pollution reductions shown in Fig. 3; while the trend of the increasing relative disparities in Fig. 4 may be partially driven by the overall decrease in pollution levels, targeted air pollution reduction strategies (affecting different areas of the maps in Fig. 3 by varying amounts) may be needed to cause a decrease in relative disparities. Figure 4 also shows the analysis for thresholds $T = 10$ μg m^{-3} and $T = 12$ μg m^{-3}. A consistent trend in disparities over time is seen across the different thresholds. Moreover, as expected from our definition of relative disparities, as the set threshold increases, relative differences across racial/ethnic groups become more pertinent for a given year. In addition to using the easily interpretable CoV, we repeated the disparities analysis of Fig. 4 with Atkinson and Gini indices—alternative metrics used in the literature[7,37,38]. These findings can be seen in Extended Data Figs. 5 and 6 and are similar to those of Fig. 4.

Discussion

We have built a data set that includes around 32,000 US ZCTAs with detailed information on demographic and pollution data for the period

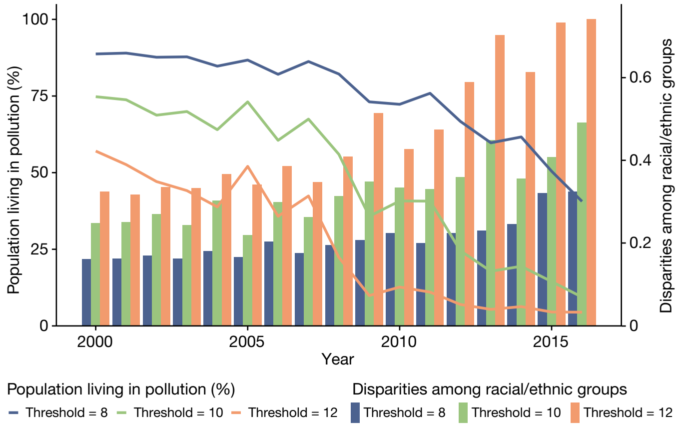


Fig. 4 | Relative disparities in exposure to $PM_{2.5}$ among racial/ethnic groups for 2000–2016. The bars show disparities in exposure (as measured by CoV) to $PM_{2.5}$ concentrations above thresholds of 8 μg m^{-3}, 10 μg m^{-3} and 12 μg m^{-3} among different racial/ethnic groups (Black, white, Asian, Native American and Hispanic or Latino). The line graphs show the percentage of the US population living above the thresholds of 8 μg m^{-3}, 10 μg m^{-3} and 12 μg m^{-3}. The trend reveals that the decrease in air pollution over the years has been accompanied by an increase in relative disparities in exposure to air pollution among different communities.

2000 to 2016. Our study provides a transparent and reproducible data-science perspective and unique visualizations of the exposure to $PM_{2.5}$ in the US and the associated disparities among racial/ethnic and income groups. Our work is descriptive in nature and is not meant to investigate causal aspects of $PM_{2.5}$ reductions and disparities in the US. When possible, we have applied sensitivity analyses to confirm our findings. For example, our results were consistent across both urban and rural areas of the US (Extended Data Fig. 7). In addition, we have applied our analyses to two independent data sets of predicted $PM_{2.5}$ levels for the US, and our findings were consistent (Extended Data Fig. 8). Nonetheless, our study could be strengthened by addressing some caveats.

First, we used average $PM_{2.5}$ concentrations across ZCTAs. This is an important limitation, because there could be substantial within-ZCTA variation in pollution. A smaller unit of analysis such as a Census block group might have further strengthened our findings, but at the cost of higher uncertainty in the estimated levels of $PM_{2.5}$ for this smaller spatial scale. Also, $PM_{2.5}$ concentrations rely on AOD estimates and are therefore subject to error. The performance of this approach has been evaluated[18], finding that $PM_{2.5}$ values estimated in this way are generally consistent with direct ground-based $PM_{2.5}$ values. Still, it is important to interpret these values with caution.

Second, we have used US Census data that span the years 2000 to 2016, during which period demographic changes might have occurred that could (together with changes in pollution levels) have contributed to our findings. To mitigate this challenge, we recalculated the distributions of different populations across the US for every year, before computing pollution-exposure values for those populations, but we have not carried out tests related to demographic changes such as residential sorting.

Third, because US Census data are not available for every year (see Methods), we have used interpolation techniques for parts of the study period, and hence inequalities between years (especially in the earlier years) are subject to the assumptions made in the interpolation. Fourth, the CoV has been used frequently for economic applications, but we are not aware of its prior application to pollution studies. Although the CoV captures our definition of disparities, caution should be applied before applying it to other disparity measures. Researchers use the Atkinson index widely, but it is a measure that suffers from low interpretability and user subjectivity owing to its dependence on an inequality aversion parameter set by the user[7,37–39]. We have computed the Atkinson index for a full range of values of the inequality aversion parameter (Extended Data Fig. 6), as well as the Gini index, and compared the results with those obtained using the CoV. The implications of using the Atkinson and Gini indices on a small data set such as these exposure data ($n = 5$ for racial/ethnic groups) are not well documented in the literature. Nonetheless, we found similar trends in disparities across the three metrics.

Finally, determining whether disparities in air pollution have been increasing or decreasing is a cumbersome task, owing to the various units of analysis one can investigate. The population-weighted $PM_{2.5}$ mean is one possible unit[40], but here, our interest in the implications of our findings for pollution-related regulations in the US led us to set the unit of analysis as the exposure of populations to $PM_{2.5}$ levels above pollution thresholds defined by the EPA and WHO. Additionally, disparities may be defined as an absolute or relative concept[16], and each scenario may lead to different interpretations. For example, other studies have reported that the pollution decrease tends to be targeted around the dirtiest monitor in counties in nonattainment with NAAQS[41], and a related study found that these areas are regions within a nonattainment county that are poorer and have a higher share of non-white residents[42].

Further research could explore the underlying drivers of the observed disparities, and investigate how future national air-quality standards could encourage more equitable attainment. This could help to inform the EPA's strategies for reducing air pollution, in order to decrease nationwide $PM_{2.5}$ concentration levels as well as relative disparities, to better address environmental injustice.

Online content

Any methods, additional references, Nature Research reporting summaries, source data, extended data, supplementary information, acknowledgements, peer review information; details of author contributions and competing interests; and statements of data and code availability are available at https://doi.org/10.1038/s41586-021-04190-y.

1. Cohen, A. J. et al. Estimates and 25-year trends of the global burden of disease attributable to ambient air pollution: an analysis of data from the global burden of diseases study 2015. *Lancet* **389**, 1907–1918 (2017).
2. Di, Q. et al. Air pollution and mortality in the Medicare population. *N. Engl. J. Med.* **376**, 2513–2522 (2017).
3. Bell, M. L., Zanobetti, A. & Dominici, F. Evidence on vulnerability and susceptibility to health risks associated with short-term exposure to particulate matter: a systematic review and meta-analysis. *Am. J. Epidemiol.* **178**, 865–876 (2013).
4. Wang, Y. et al. Long-term exposure to PM2.5 and mortality among older adults in the southeastern US. *Epidemiology* **28**, 207 (2017).
5. Kioumourtzoglou, M.-A., Schwartz, J., James, P., Dominici, F. & Zanobetti, A. PM2. 5 and mortality in 207 US cities: modification by temperature and city characteristics. *Epidemiology* **27**, 221 (2016).
6. Bell, M. L. & Ebisu, K. Environmental inequality in exposures to airborne particulate matter components in the United States. *Environ. Health Perspect.* **120**, 1699–1704 (2012).
7. Rosofsky, A., Levy, J. I., Zanobetti, A., Janulewicz, P. & Fabian, M. P. Temporal trends in air pollution exposure inequality in Massachusetts. *Environ. Res.* **161**, 76–86 (2018).
8. Mikati, I., Benson, A. F., Luben, T. J., Sacks, J. D. & Richmond-Bryant, J. Disparities in distribution of particulate matter emission sources by race and poverty status. *Am. J. Public Health* **108**, 480–485 (2018).
9. Miranda, M. L., Edwards, S. E., Keating, M. H. & Paul, C. J. Making the environmental justice grade: the relative burden of air pollution exposure in the United States. *Int. J. Environ. Res. Public Health* **8**, 1755–1771 (2011).
10. Mohai, P., Pellow, D. & Roberts, J. T. Environmental justice. *Annu. Rev. Environ. Resour.* **34**, 405–430 (2009).
11. Agyeman, J., Schlosberg, D., Craven, L. & Matthews, C. Trends and directions in environmental justice: from inequity to everyday life, community, and just sustainabilities. *Annu. Rev. Environ. Resour.* **41**, 321–340 (2016).
12. Banzhaf, S., Ma, L. & Timmins, C. Environmental justice: the economics of race, place, and pollution. *J. Econ. Perspect.* **33**, 185–208 (2019).
13. Kelly, J. T. et al. Examining $PM_{2.5}$ concentrations and exposure using multiple models. *Environ. Res.* **196**, 110432 (2020).
14. Fann, N., Coffman, E., Timin, B. & Kelly, J. T. The estimated change in the level and distribution of $PM_{2.5}$-attributable health impacts in the United States: 2005–2014. *Environ. Res.* **167**, 506–514 (2018).
15. Tessum, C. W. et al. $PM_{2.5}$ polluters disproportionately and systemically affect people of color in the United States. *Sci. Adv.* **7**, (2021).
16. Harper, S. et al. Using inequality measures to incorporate environmental justice into regulatory analyses. *Int. J. Environ. Res. Public Health* **10**, 4039–4059 (2013).
17. Colmer, J., Hardman, I., Shimshack, J. & Voorheis, J. Disparities in $PM_{2.5}$ air pollution in the United States. *Science* **369**, 575–578 (2020).
18. Meng, J. et al. Estimated long-term (1981–2016) concentrations of ambient fine particulate matter across North America from chemical transport modeling, satellite remote sensing, and ground-based measurements. *Environ. Sci. Technol.* **53**, 5071–5079 (2019).

19. US Environmental Protection Agency. Process of reviewing the National Ambient Air Quality Standards. https://www.epa.gov/criteria-air-pollutants/process-reviewing-national-ambient-air-quality-standards
20. World Health Organization. *Air Quality Guidelines. Global Update 2005. Particulate Matter, Ozone, Nitrogen Dioxide, and Sulfur Dioxide* (World Health Organization, 2005).
21. Dominici, F. et al. Fine particulate air pollution and hospital admission for cardiovascular and respiratory diseases. *J. Am. Med. Assoc.* **295**, 1127–1134 (2006).
22. Bell, M. L. et al. Seasonal and regional short-term effects of fine particles on hospital admissions in 202 US counties, 1999–2005. *Am. J. Epidemiol.* **168**, 1301–1310 (2008).
23. Kloog, I. et al. Short term effects of particle exposure on hospital admissions in the mid-atlantic states: a population estimate. *PLoS One* **9**, (2014).
24. Bravo, M. A. et al. Airborne fine particles and risk of hospital admissions for understudied populations: effects by urbanicity and short-term cumulative exposures in 708 US counties. *Environ. Health Perspect.* **125**, 594–601 (2017).
25. Dominici, F., McDermott, A., Zeger, S. L. & Samet, J. M. National maps of the effects of particulate matter on mortality: exploring geographical variation. *Environ. Health Perspect.* **111**, 39–44 (2003).
26. Beelen, R. et al. Effects of long-term exposure to air pollution on natural-cause mortality: an analysis of 22 European cohorts within the multicentre escape project. *Lancet* **383**, 785–795 (2014).
27. Crouse, D. L. et al. Ambient PM2.5, O_3, and NO_2 exposures and associations with mortality over 16 years of follow-up in the Canadian census health and environment cohort (CanCHEC). *Environ. Health Perspect.* **123**, 1180–1186 (2015).
28. Makar, M. et al. Estimating the causal effect of fine particulate matter levels on death and hospitalization: are levels below the safety standards harmful? *Epidemiology* **28**, 627 (2017).
29. Di, Q. et al. Association of short-term exposure to air pollution with mortality in older adults. *J. Am. Med. Assoc.* **318**, 2446–2456 (2017).
30. Wu, X., Braun, D., Schwartz, J., Kioumourtzoglou, M. & Dominici, F. Evaluating the impact of long-term exposure to fine particulate matter on mortality among the elderly. *Sci. Adv.* **6**, eaba5692 (2020).
31. Liu, C. et al. Ambient particulate air pollution and daily mortality in 652 cities. *N. Engl. J. Med.* **381**, 705–715 (2019).
32. Samoli, E. et al. Estimating the exposure–response relationships between particulate matter and mortality within the aphea multicity project. *Environ. Health Perspect.* **113**, 88–95 (2005).
33. Samet, J. M., Dominici, F., Curriero, F. C., Coursac, I. & Zeger, S. L. Fine particulate air pollution and mortality in 20 us cities, 1987–1994. *N. Engl. J. Med.* **343**, 1742–1749 (2000).
34. Shah, A. S. et al. Global association of air pollution and heart failure: a systematic review and meta-analysis. *Lancet* **382**, 1039–1048 (2013).
35. American Psychological Association. Racial and ethnic identity. https://apastyle.apa.org/style-grammar-guidelines/bias-free-language/racial-ethnic-minorities
36. EPA proposes to retain NAAQS for particulate matter. https://www.epa.gov/newsreleases/epa-proposes-retain-naaqs-particulate-matter
37. Clark, L. P., Millet, D. B. & Marshall, J. D. National patterns in environmental injustice and inequality: outdoor NO_2 air pollution in the United States. *PLoS One* **9**, e94431 (2014).
38. Levy, J. I., Chemerynski, S. M. & Tuchmann, J. L. Incorporating concepts of inequality and inequity into health benefits analysis. *Int. J. Equity Health* **5**, 2 (2006).
39. Lambert, P. J., Millimet, D. L. & Slottje, D. Inequality aversion and the natural rate of subjective inequality. *J. Public Econ.* **87**, 1061–1090 (2003).
40. Currie, J., Voorheis, J. & Walker, R. *What Caused Racial Disparities in Particulate Exposure to Fall? New Evidence From the Clean Air Act and Satellite-Based Measures of Air Quality*. Technical report (National Bureau of Economic Research, 2020).
41. Auffhammer, M., Bento, A. M. & Lowe, S. E. Measuring the effects of the clean air act amendments on ambient PM_{10} concentrations: the critical importance of a spatially disaggregated analysis. *J. Environ. Econ. Manage.* **58**, 15–26 (2009).
42. Grainger, C. A. The distributional effects of pollution regulations: do renters fully pay for cleaner air? *J. Public Econ.* **96**, 840–852 (2012).

Methods

Our data set includes US ZCTAs for 2000–2016 (N is approximately 32,000). For each ZCTA, we obtained demographic and socioeconomic variables from the US Census Bureau where available, and used interpolation techniques (moving average) to determine those of the missing years. More specifically, for the years 2000 and 2010, we used ZCTA estimates from the decennial census. For the period 2001–2009, we interpolated the data using moving averages for each census variable and for each ZCTA using the 'ImputeTS' R package. For the period 2011–2016, we used the five-year data from the American Community Survey (ACS5). Documentation of all calculations and source data used is available in the following GitHub repository: https://github.com/NSAPH/National-Causal-Analysis/tree/master/Confounders/census. Variables of interest comprised median household income, proportions of Native Americans, Asian, white, Black, and Hispanic or Latino residents, and population density. For each year, we assigned all ZCTAs percentile ranks from 1 to 100 on the basis of median household income, and categorized them into ten income groups. Throughout the paper, we use 'low income' and 'high income' to label the lowest three and highest three income groups, respectively.

We also used a publicly available data set containing predicted $PM_{2.5}$ concentration levels in the US[18]. This study estimated ground-level total $PM_{2.5}$ concentrations over North America by combining aerosol optical depth (AOD) retrievals from the NASA MODIS, MISR and SeaWIFS instruments with the GEOS-Chem chemical transport model, subsequently calibrated to regional ground-based observations of total mass using geographically weighted regression (GWR). The authors[18] evaluated the performance of their approach and reported that the estimated $PM_{2.5}$ concentrations were generally consistent with direct ground-based $PM_{2.5}$ measurements (R^2 varying between 0.6 and 0.8). Their collocated comparison of the trends of population-weighted annual average $PM_{2.5}$ from their estimates and ground-based measurements was highly consistent. They also reported that the accuracy of the $PM_{2.5}$ prediction models was similar for low and high levels of exposure, implying no large differences in performance between urban and rural areas. Using this data set, for each ZCTA we computed annual averages of $PM_{2.5}$. We built one data set by combining the demographic and $PM_{2.5}$ variables across all US ZCTAs for 2000–2016. Our data set analysis reveals time patterns in air pollution across the US and disparities in exposure to air pollution among racial/ethnic and income groups. We use dynamic videos to communicate our findings, along with plots that summarize and clarify the information embedded in our visualizations.

We first defined a group population-weighted $PM_{2.5}$ concentration, in which a group can be an income group such as the first decile, or an ethnic group such as the Hispanic or Latino population. In the case of racial/ethnic groups, the population-weighted $PM_{2.5}$ concentration in racial/ethnic group k is given by:

$$\overline{\mathrm{PM}}_{2.5k} = \frac{\sum_j \mathrm{PM}_{2.5j} p_{k,j}}{\sum_j p_{k,j}},$$

where summation occurs over all ZCTAs. $p_{k,j}$ is the number of people in racial group k living in the ZCTA j, and $PM_{2.5j}$ is the $PM_{2.5}$ level in the ZCTA j. In the case of income groups, the population-weighted $PM_{2.5}$ concentration of income group i is:

$$\overline{\mathrm{PM}}_{2.5i} = \frac{\sum_{j \in i} \mathrm{PM}_{2.5j} p_j}{\sum_{j \in i} p_j},$$

where summation occurs only over ZCTAs j belonging to income group i. p_j is the total population of ZCTA j, and $PM_{2.5j}$ is the $PM_{2.5}$ level in ZCTA j.

We also computed relative disparities in exposure to $PM_{2.5}$ among different populations. We define a state of equality (or lack of relative disparities) among various populations as a state in which equal proportions of the various populations are exposed to pollution levels higher than a threshold of interest, chosen in relation to the EPA standard and WHO guidelines for $PM_{2.5}$. To estimate such disparities, we first defined an additional $PM_{2.5}$-related variable (q) and used it to quantify the level of disparities in exposure to $PM_{2.5}$ concentrations among the different racial/ethnic groups.The variable q is defined as the percentage of a population exposed to $PM_{2.5}$ levels above a certain threshold, T. We can calculate q for specific population subgroups. For example, we can compute the percentage of a racial/ethnic group (such as Native Americans) that is exposed to $PM_{2.5}$ levels above $T = 8\ \mu g\ m^{-3}$.

To measure the degree of disparities across racial/ethnic groups in exposure to $PM_{2.5}$ concentrations above T for a specific year, we first computed q for every racial/ethnic group. We then computed the coefficient of variation (CoV), also referred to as the 'between group' variance:

$$\mathrm{CoV} = \frac{\sqrt{\mathrm{Var}(q)}}{\mu(q)}$$

where Var is the variance of q, and μ is the mean of q. CoV measures the variability of a series of numbers independently of the data magnitude, so it captures the variation in q among racial/ethnic groups in a given year relative to the mean exposure levels to pollution during that year. The choice of CoV is supported by its multiple attributes, such as its independence of ordered social groups and of an inequality aversion parameter[16]. It is also easily interpretable and sensitive to large differences from the average.

For example, consider three years, Y_1, Y_2 and Y_3, where the percentages of five racial/ethnic groups being exposed to $PM_{2.5}$ levels above a threshold T are, respectively: $q_1 = (11\%, 13\%, 14\%, 15\%, 17\%)$; $q_2 = (10\%, 12\%, 14\%, 16\%, 18\%)$; $q_3 = (1\%, 1.2\%, 1.4\%, 1.6\%, 1.8\%)$. From Y_1 to Y_2, the coefficient of variation increases from $\mathrm{CoV}_1 = 0.160$ to $\mathrm{CoV}_2 = 0.226$, which indicates that the variation in exposure to air pollution relative to the mean (and, equivalently, relative disparities among the racial/ethnic groups) increased by a factor of 1.41. On the other hand, although the pollution levels decreased drastically between Y_2 and Y_3, as can be seen by the different orders of magnitude of q_2 and q_3, the coefficient of variation is unchanged ($\mathrm{CoV}_3 = 0.226$), indicating that the relative disparities in exposure to air pollution among the racial/ethnic groups is the same between Y_2 and Y_3. These examples highlight the power of using CoV to capture relative variation in the data, independently of its magnitude. This is very important for our application, because the level of pollution changes considerably over the years. Further, if the exposure across all groups in Y_3 decreases by the same absolute amount, for example to $q_4 = (0.8\%, 1\%, 1.2\%, 1.4\%, 1.6\%)$ in year Y_4, the coefficient of variation increases from $\mathrm{CoV}_3 = 0.226$ to $\mathrm{CoV}_4 = 0.264$, indicating an increase in relative disparities relative to the new lower mean. Finally, a state of total equality or absence of disparities would exist when the exposure across all groups is identical; for example $q_5 = (0.8\%, 0.8\%, 0.8\%, 0.8\%, 0.8\%)$ in year Y_5. Because of the preexisting disparities in year Y_4, targeted pollution-reduction strategies that affect the various groups differently may be required to achieve a state of total equality with no disparities such as q_5 in year Y_5.

The procedure outlined for quantifying disparities through CoV can be applied for any $PM_{2.5}$ threshold (T), and can be repeated for all years to track the evolution of disparities in exposure to air pollution among different racial/ethnic groups. We supplemented the computation of relative disparities with the CoV by using the Atkinson and Gini indices.

Data availability

Data are available in the following GitHub repositories: https://github.com/NSAPH/National-Causal-Analysis/tree/master/Confounders/census and https://github.com/xiaodan-zhou/pm25_and_disparity.

Code availability

Code is available in the following GitHub repository: https://github.com/xiaodan-zhou/pm25_and_disparity.

43. Di, Q. et al. Assessing $PM_{2.5}$ exposures with high spatiotemporal resolution across the continental United States. *Environ. Sci. Technol.* **50**, 4712–4721 (2016).
44. Di, Q. et al. An ensemble-based model of $PM_{2.5}$ concentration across the contiguous United States with high spatiotemporal resolution. *Environ. Int.* **130**, 104909 (2019).

Acknowledgements We thank R. Martin and J. D. Schwartz for providing the air-pollution data; B. Sabbath for cleaning and preparing the data sets; and L. Bennett for comments and discussions. We also thank J. Kodros for his comments on an earlier draft. This work was supported financially by grants from the Health Effects Institute (4953- RFA14-3/16-4), the National Institutes of Health (DP2MD012722, P50MD010428), the National Institutes of Health and Yale University (R01MD012769), the National Institutes of Health and National Institute of Environmental Health Sciences (R01 ES028033, R01ES026217, R01AG066793-01, R01ES029950, R01ES028033-S1), the National Institutes of Health and Columbia University (1R01ES030616), the Environmental Protection Agency (83587201-0), The Climate Change Solutions Fund, and a Harvard Star Friedman Award.

Author contributions A.J., S.V. and F.D. contributed to the study design. A.J. led the research, with support from X.Z. and supervision from F.D. Maps and videos were prepared by X.Z., J.L. and T.-H.L. A.J. drafted the manuscript, with support from L.K., S.V. and F.D. All authors read and approved the final manuscript for submission.

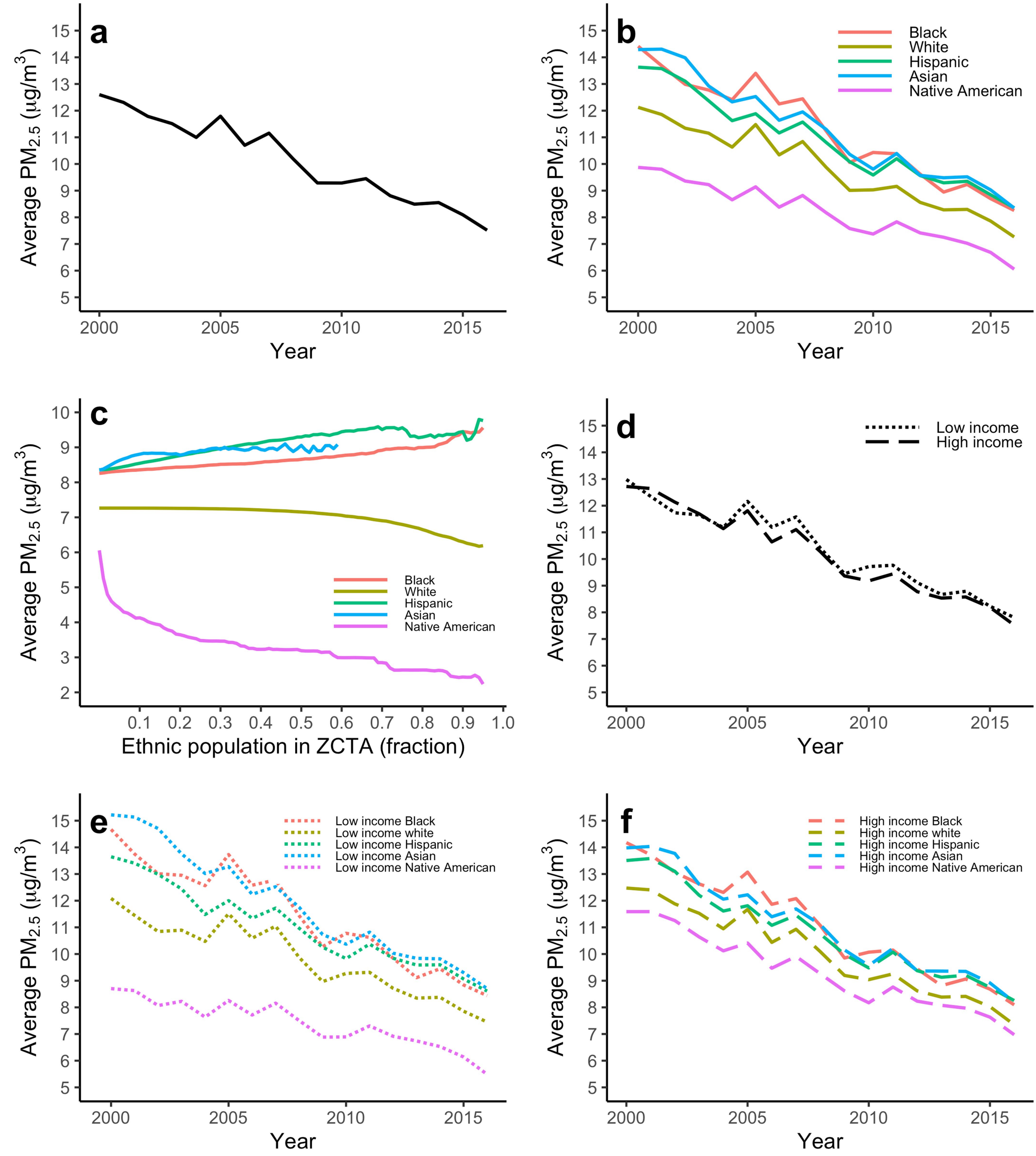
Competing interests The authors declare no competing interests.

Additional information

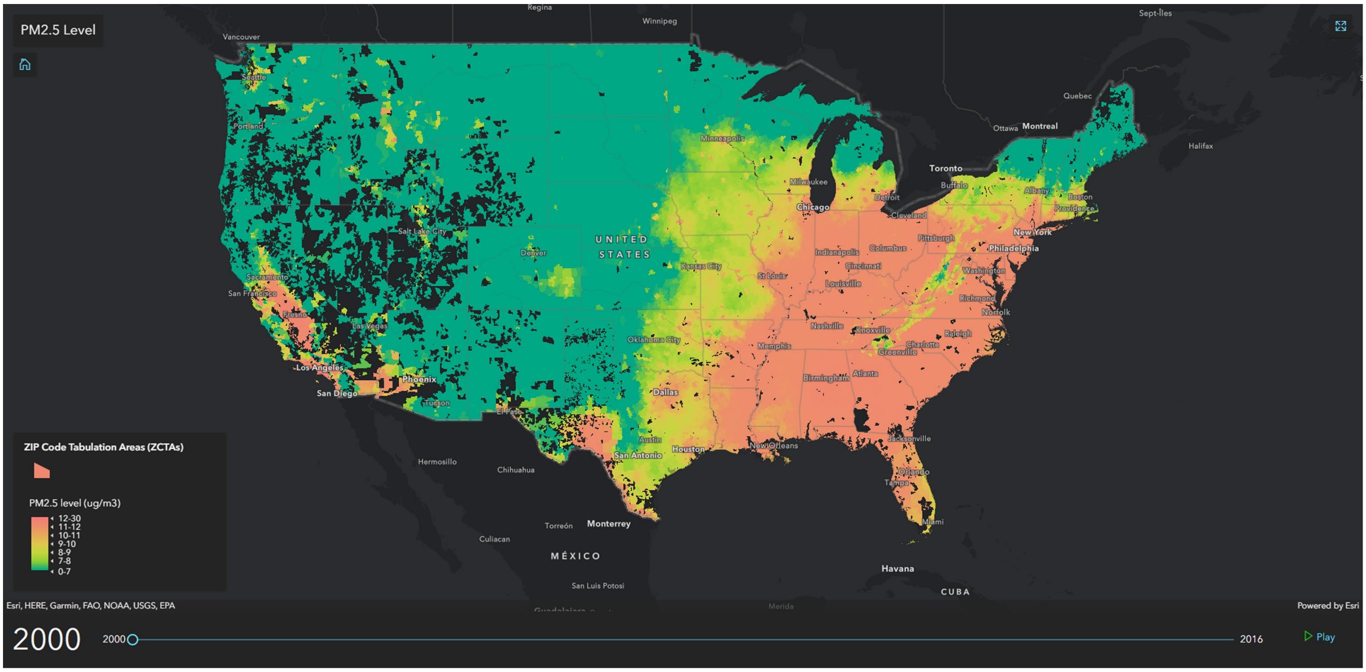
Supplementary information The online version contains supplementary material available at https://doi.org/10.1038/s41586-021-04190-y.

Correspondence and requests for materials should be addressed to Abdulrahman Jbaily or Francesca Dominici.

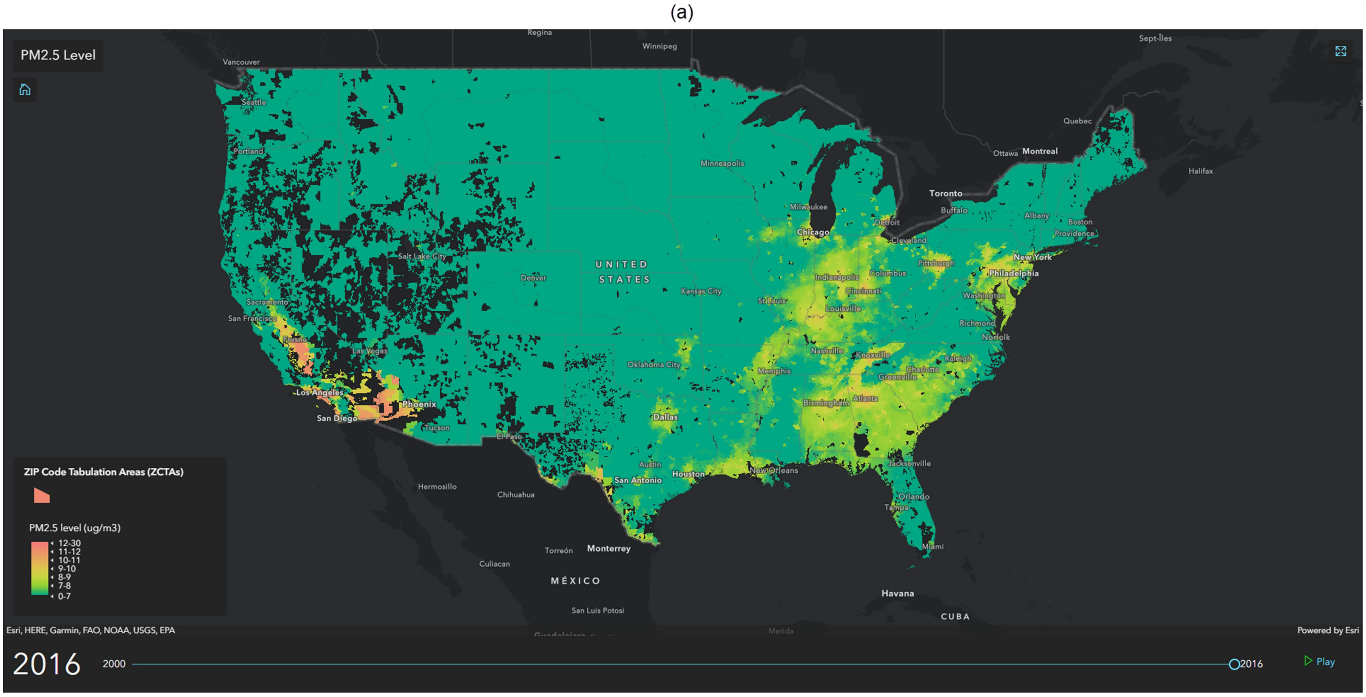
Peer review information *Nature* thanks Corbett Grainger, Jonathan Levy, Arden Pope III and the other, anonymous, reviewer(s) for their contribution to the peer review of this work. Peer reviewer reports are available.



Extended Data Fig. 1 | Summary $PM_{2.5}$ metrics across racial/ethnic and income groups. a, The population-weighted average of $PM_{2.5}$ decreased by 40.4% from the year 2000 to 2016. **b**, Population-weighted average $PM_{2.5}$ concentration across the different racial/ethnic communities for 2000 to 2016, showing that Black and Native American populations live in the most- and least-polluted areas, respectively. **c**, Population-weighted average $PM_{2.5}$ concentration across racial/ethnic communities as a function of ZCTA racial/ethnic population (%) for 2016. For example, when the racial/ethnic population percentage is equal to 0.2, the red curve includes every ZCTA where the Black population is 20% or more, and the blue curve includes every ZCTA where the white population is 20% or more. As a ZCTA's Black and Hispanic or Latino populations increase, the $PM_{2.5}$ concentration levels increase. The opposite effect is seen for the white and Native American communities. **d**, The population-weighted average $PM_{2.5}$ concentration across the income groups reveals that the low-income group has been exposed to only slightly higher $PM_{2.5}$ levels than the high-income groups since 2004. **e**, Population-weighted average $PM_{2.5}$ concentrations across the different racial/ethnic communities that are in the low-income group, for 2000–2016. **f**, Population-weighted average $PM_{2.5}$ concentrations across the different racial/ethnic communities that are in the high-income group, for 2000–2016. Panels **e**, **f** show similar differences in average $PM_{2.5}$ concentrations across the racial/ethnic groups as seen in **b**.

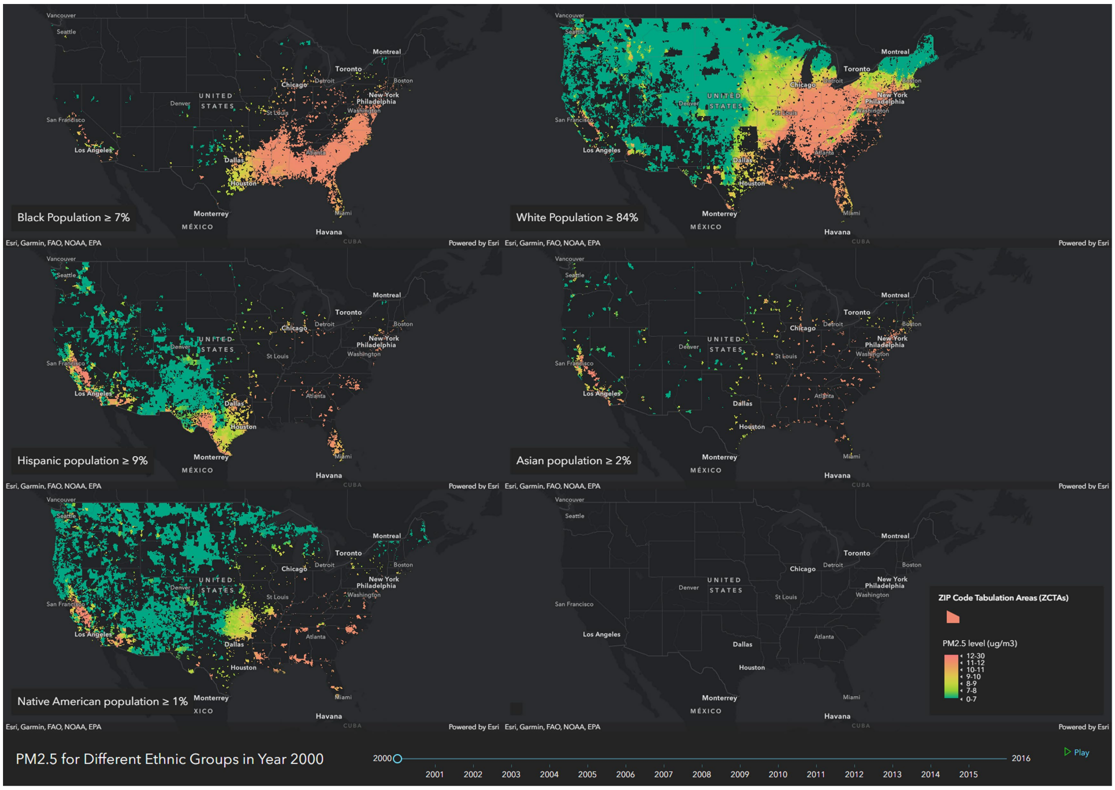


(a)

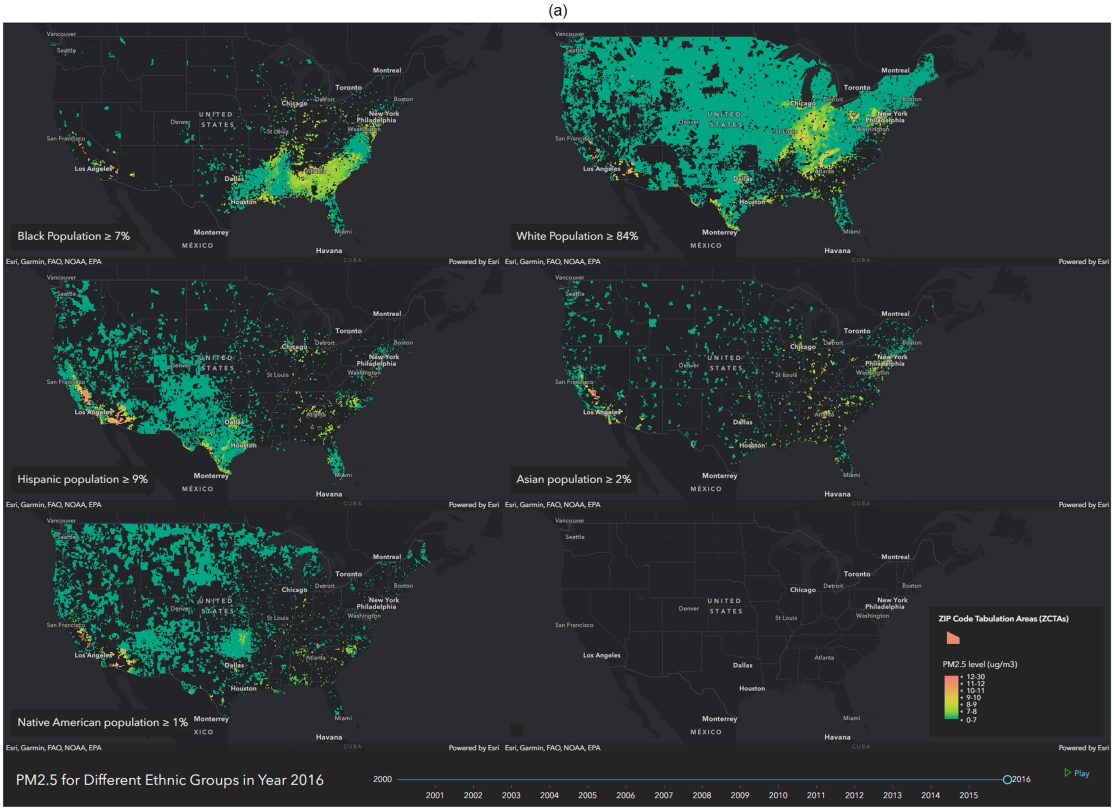


(b)

Extended Data Fig. 2 | Average $PM_{2.5}$ concentrations across the US. **a**, Distribution of $PM_{2.5}$ in 2000. **b**, Distribution of $PM_{2.5}$ in 2016. Supplementary Video 1 shows the change in the distribution of $PM_{2.5}$ concentration levels in the US from 2000 to 2016. Note that Hawaii and Alaska are not shown. Imagery provided courtesy of Esri, HERE, Garmin, FAO, NOAA, USGS, ©OpenStreetMap contributors, and the GIS User Community.

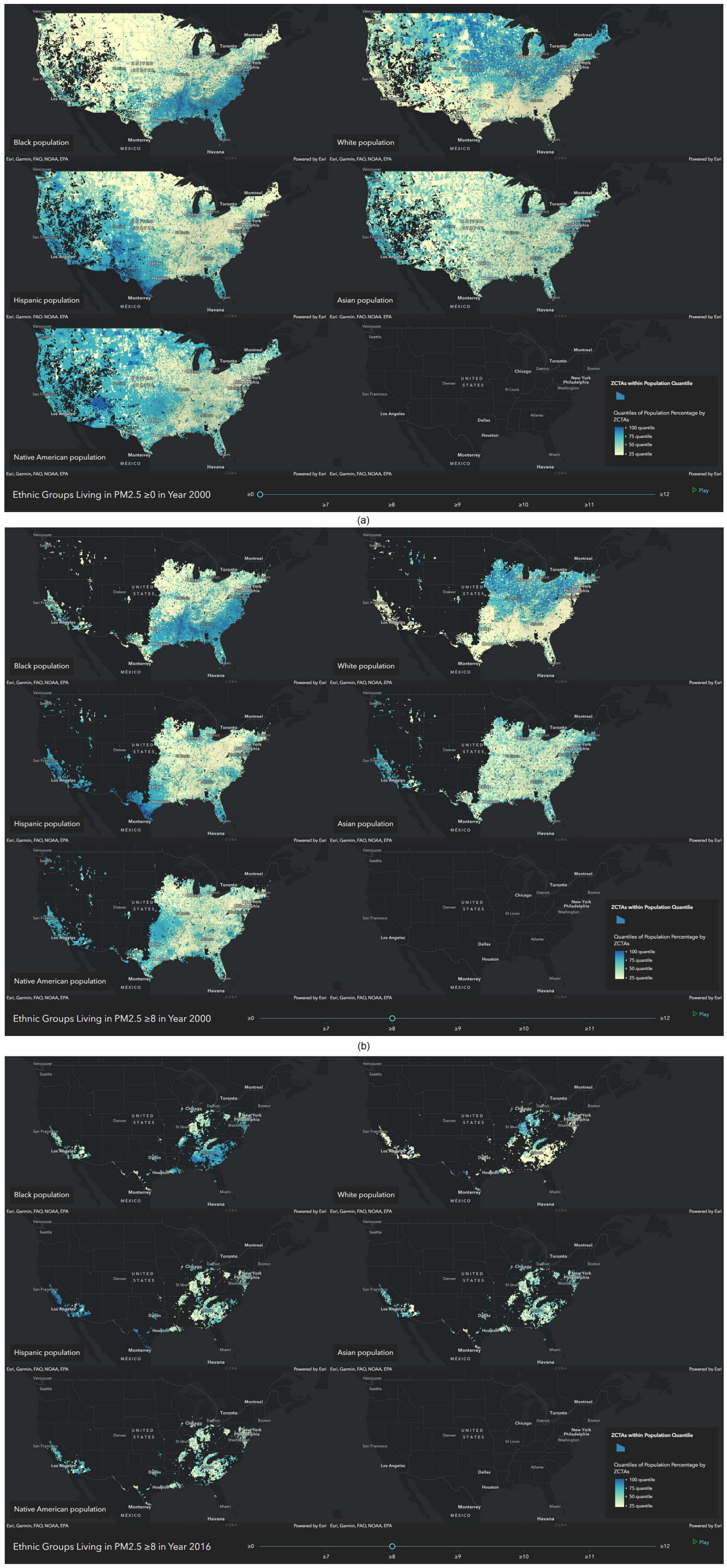


(a)

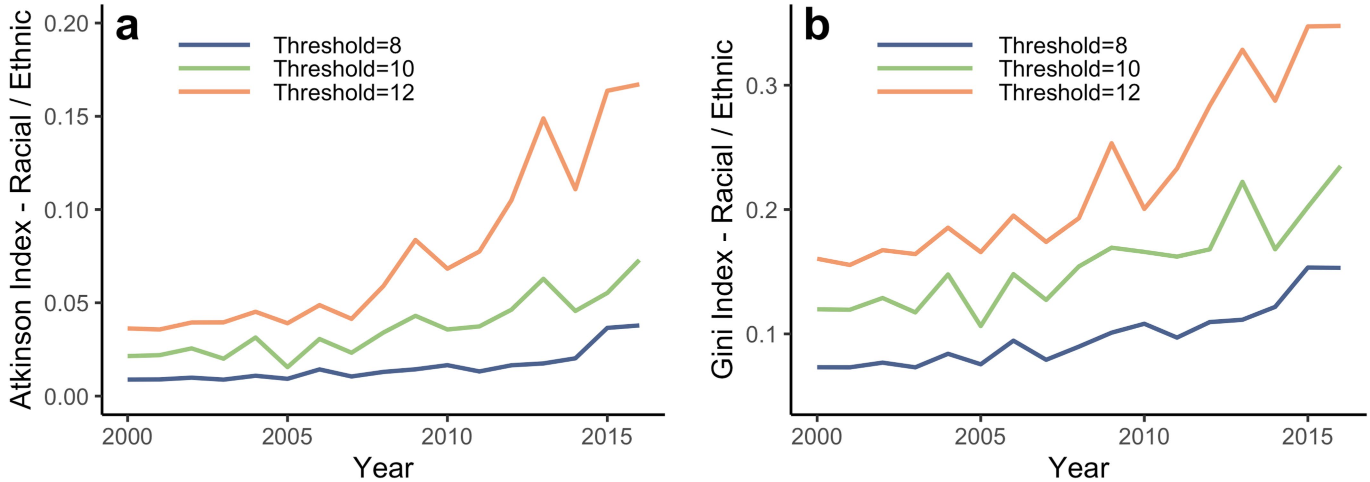


(b)

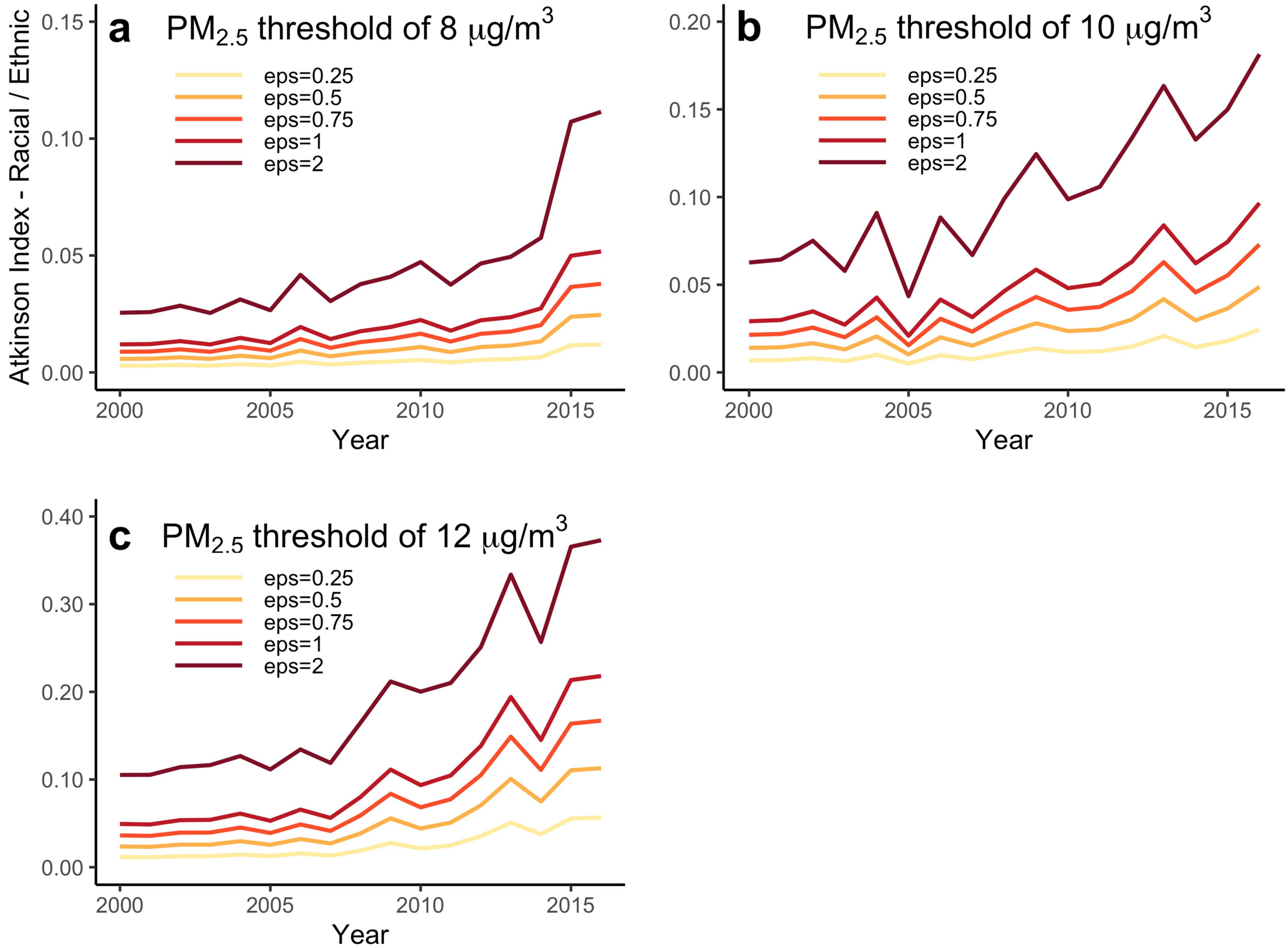
Extended Data Fig. 3 | Average $PM_{2.5}$ concentrations across ZCTAs in which different racial/ethnic groups are overrepresented. a, Distribution of $PM_{2.5}$ across five different maps for 2000, each showing the ZCTAs in which one race/ethnicity group is overrepresented. **b**, Distribution of $PM_{2.5}$ across five different maps for 2016, each showing the ZCTAs in which one race/ethnicity group is overrepresented. Supplementary Videos 2, 3 show the change in the distribution of $PM_{2.5}$ concentrations across the five maps from 2000 to 2016. Note that Hawaii and Alaska are not shown. Imagery provided courtesy of Esri, HERE, Garmin, FAO, NOAA, USGS, ©OpenStreetMap contributors, and the GIS User Community.



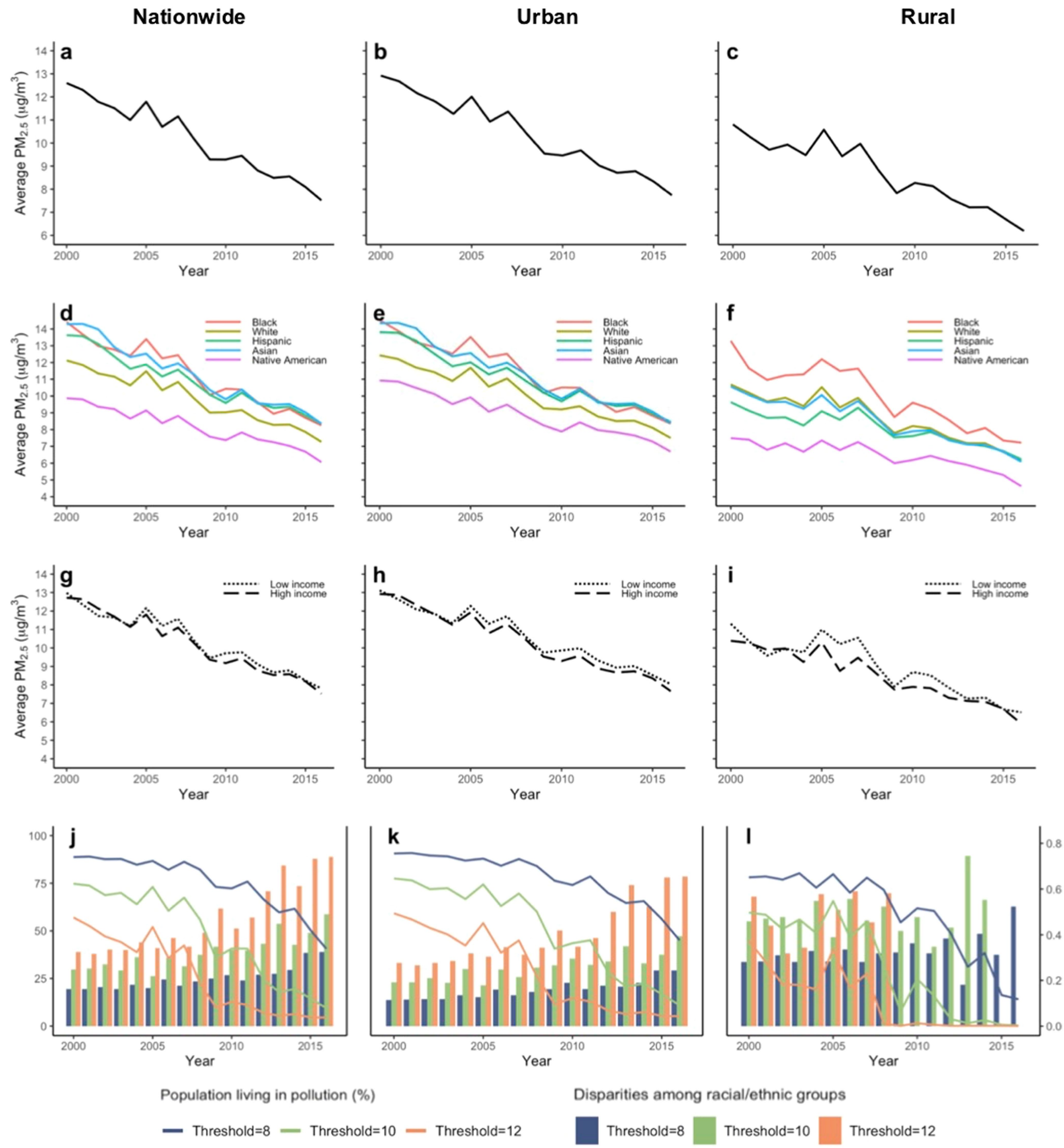
Extended Data Fig. 4 | Distribution of racial/ethnic populations above a $PM_{2.5}$ threshold of 8 μg m^{-3} for 2000 and 2016. a, US ZCTAs for each race/ethnicity are ranked on the basis of the ratio of the race/ethnicity population to the total ZCTA population. Dark blue indicates fractions close to 1 (ZCTAs in which the corresponding race/ethnicity most lives), and light yellow indicates fractions close to 0 (ZCTAs in which the corresponding race/ethnicity least lives). **b**, US ZCTAs with $PM_{2.5}$ concentrations higher than 8 μg m^{-3} in 2000. **c**, US ZCTAs with $PM_{2.5}$ concentrations higher than 8 μg m^{-3} in 2016. Supplementary Videos 5–8 show the distribution of the different racial/ethnic groups across multiple ranges of $PM_{2.5}$ concentrations for 2000 and 2016. Note that Hawaii and Alaska are not shown. Imagery provided courtesy of Esri, HERE, Garmin, FAO, NOAA, USGS, ©OpenStreetMap contributors, and the GIS User Community.



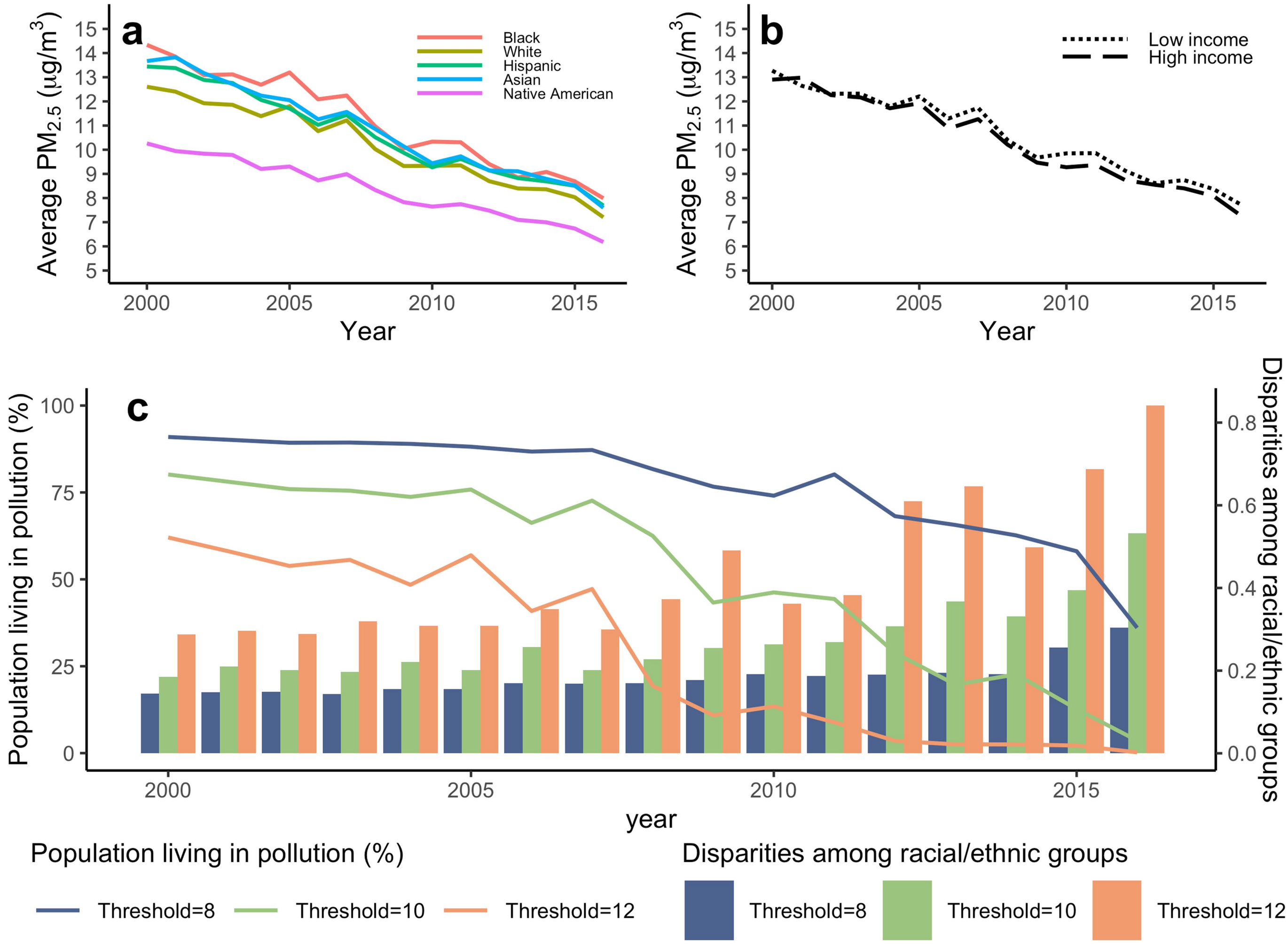
Extended Data Fig. 5 | Supplementary measures of relative disparities in exposure to $PM_{2.5}$ among racial/ethnic groups for 2000–2016. a, The Atkinson index is computed to measure relative disparities among the racial/ethnic groups (Black, white, Asian, Native American and Hispanic or Latino). **b**, The Gini index is computed to measure relative disparities among the racial/ethnic groups (Black, white, Asian, Native American and Hispanic or Latino). The trends in both indices are similar to that measured by CoV (Fig. 4): racial/ethnic disparities in exposure to air pollution relative to pollution levels at or below the EPA standard are increasing. The Atkinson and Gini indices were computed using the inequality package 'ineq' in R software. The input is the proportion of the racial/ethnic (or income) groups living above the set $PM_{2.5}$ threshold. We set the Atkinson aversion parameter, ε, to 0.75 (ref. [7]); the sensitivity of the index to different values of ε is shown in Extended Data Fig. 6.



Extended Data Fig. 6 | Sensitivity of the Atkinson index to the inequality aversion parameter *ε*. **a**, Sensitivity of the Atkinson index relative to a $PM_{2.5}$ threshold of 8 μg m^{-3}. **b**, Sensitivity of the Atkinson index relative to a $PM_{2.5}$ threshold of 10 μg m^{-3}. **c**, Sensitivity of the Atkinson index relative to a $PM_{2.5}$ threshold of 12 μg m^{-3}. A consistent trend is shown across the parameter values.



Extended Data Fig. 7 | Replication of the main findings across urban and rural areas. A ZCTA's population density is used as a metric to control for urbanicity in our study. We classify urban and rural areas on the basis of the percentage of the urban population in each ZCTA; such percentages are available from the US Census Bureau for 2010. ZCTAs with an urban population of more than 50% are classified as urban, whereas those with an urban population of less than 50% are classified as rural. For nationwide, urban and rural US, we reproduce our main results: namely, the average $PM_{2.5}$ concentrations for the total population (**a**–**c**), for racial/ethnic groups (**d**–**f**) and for income groups (**g**–**i**), as well as disparities among racial/ethnic groups (**j**–**l**). Similarities in the results across the national, urban and rural US are apparent and findings are consistent regardless of the urbanicity of ZCTAs. Note that in the case of the rural US, we only compute disparities (**l**) for the years in which the proportion of the population exposed to $PM_{2.5}$ concentrations above the thresholds of interest is non-zero. For example, the proportion of the population in the rural US that is exposed to $PM_{2.5}$ concentrations above $T = 12\ \mu g\ m^{-3}$ reaches near-zero levels in 2009, and hence disparities after this year are not computed.



Extended Data Fig. 8 | Sensitivity of our main findings to estimates of $PM_{2.5}$. We replicated our analysis using an independent pollution data set[43,44], and we show here the sensitivity of our findings to the new $PM_{2.5}$ estimates. **a**, Replication of Extended Data Fig. 1b using the alternative data set. **b**, Replication of Extended Data Fig. 1d using the alternative data set. **c**, Replication of Fig. 4 using the alternative data set. Our main findings are robust and consistent across the two data sets. (Minor differences resulting from the different pollution estimates can be spotted, as expected.).