working paper

20 37

Neighborhood Sorting Obscures Neighborhood Effects in the Opportunity Atlas

Dionissi Aliprantis and Hal Martin

FEDERAL RESERVE BANK OF CLEVELAND

ISSN: 2573-7953

Working Paper 20-37 December 2020

Neighborhood Sorting Obscures Neighborhood Effects in the Opportunity Atlas

Dionissi Aliprantis and Hal Martin

The Opportunity Atlas (OA) is an innovative data set that ranks neighborhoods according to children's adult outcomes in several domains, including income. Conceptually, outcomes offer new evidence about neighborhood effects when measured in isolation from neighborhood sorting. This paper shows that neighborhood sorting contributes to OA estimates. We document cases in which small sample sizes and changes over time can explain disagreements between OA rankings and those based on contemporaneous variables. Our results suggest caution for interpretations of the OA data set at a granular level, particularly for predictions about the outcomes of black children in high-income neighborhoods.

Keywords: neighborhood sorting, neighborhood effect, Opportunity Atlas.

JEL Codes: R23, C81, J15, I38.

Suggested citation: Aliprantis, Dionissi, and Hal Martin. 2020. "Neighborhood Sorting Obscures Neighborhood Effects in the Opportunity Atlas." Federal Reserve Bank of Cleveland, Working Paper No. 20-37. https://doi.org/10.26509/frbc-wp-202037.

Dionissi Aliprantis is at the Federal Reserve Bank of Cleveland (dionissi.aliprantis@clev.frb.org). Hal Martin is at the Federal Reserve Bank of Cleveland (hal.martin@clev.frb.org). The authors thank Kristen Tauber for research assistance. For helpful comments, they thank seminar participants at the Cleveland Fed and the Urban Economics Association 2020 Virtual Meeting.

1 Introduction

Neighborhood effects are the impacts on an individual's outcomes that result from their exposure to living in a given neighborhood.[1] Where do neighborhood effects come from? Approaches in economics and related disciplines suggest that neighborhoods exert an influence on the outcomes of their residents through exposure to the available amenities present in them. For instance, a neighborhood within a particular school's attendance boundary exposes its children to the influence of teachers, peers, and educational resources that contribute to a child's human capital development. Policymakers concerned with poor outcomes have focused some attention for decades on the question: If a family moves to neighborhood x, how will their outcomes change compared to living in neighborhood y? If public funds are spent to support moves to "better" neighborhoods, which neighborhoods are better?

Measuring neighborhood effects is made difficult by the fact that families sort endogenously according to their preferences, information, and resources, and that peer effects are among the amenities over which that sorting occurs. This neighborhood sorting process complicates arriving at a causal answer to policymakers' most pressing questions. Policymakers have generally settled for indicators of neighborhood effects based on the availability of amenities suggested by theory. Amenities like access to jobs, high-performing schools, and a safe environment are inputs to the formation of the human capital that people use to lead productive lives. Researchers and policymakers also look to socio-economic indicators such as low poverty, high adult education levels, etc. to identify promising neighborhoods.

A new approach to identifying neighborhood effects is based on the possibilities offered through big data. Chetty et al. (2020) introduce an innovative new data set called the Opportunity Atlas (OA) that is based on the outcomes of the entire 1978-1983 birth cohort in the US. The OA estimates conditional expectation functions for each neighborhood in terms of parental income, race/ethnicity, and gender using the outcomes of the children who grew up in each neighborhood.

While an outcome-based approach is sure to lead to new insights into neighborhood effects, the phenomenon of neighborhood sorting haunts the OA in multiple ways. First, OA estimates share the problem of endogenous selection common to any reduced-form measure of neighborhood effects: Because parents choose neighborhoods based on their preferences and constraints, a child's exposure to a neighborhood is an endogenous choice outside the control of the researcher. To the extent that parents are the unobserved contributors to a child's outcomes, their choice of neighborhood may simply signal their private and unobserved contributions to a child's development in a more public way.

Second, the OA relies on the presence of children whose parents differ in several dimensions in order to estimate heterogeneous neighborhood effects by race, gender, and income level. Both income and racial segregation are persistent features of neighborhoods across the US. In many

[1] In terms of a model of the intergenerational transmission of human capital, one can define neighborhood effects as neighborhood-specific differences in the productivity of investments in a child's human capital accumulation (Aliprantis and Carroll (2018)).

ways, the sorting of access to amenities along these lines is the primary problem policymakers are interested in solving today by assisting people in lacking communities to access better-resourced ones. Precisely because of *how* residents have sorted across neighborhoods, however, there are many neighborhoods where few or no counterfactuals exist to show how children of different backgrounds performed in the same neighborhood. We show that the degree of sorting by income and race gives reasons to be cautious in accepting the OA's modeled results in some neighborhoods. We discuss the implications of this finding for understanding the contribution of residential segregation to racial income inequality and for interpreting results of the Moving to Opportunity housing mobility experiment.

Third, the OA relies on realized outcomes for children who were born approximately 40 years before today. Neighborhoods and entire cities change over time through gentrification, urban decline, regional migration and community development. While Chetty et al. (2020) provide suggestive evidence that neighborhood change is not significant after 10 years time, this does not necessarily hold in general. We show that the neighborhoods where the OA diverges most from traditional measures of neighborhood effects are those neighborhoods that have experienced the largest changes in their populations, both in terms of raw counts and in terms of observable socio-economic characteristics.

We believe there is important new evidence on neighborhood effects in the Opportunity Atlas, beginning with many of the patterns already documented in Chetty et al. (2020). The results in this note, however, serve as a reminder that OA rankings are not a direct measure of neighborhood effects. We construe our results as suggesting caution for interpretations of the OA data set at a granular level, whether in terms of disagreements with contemporary measures or in terms of counterfactual predictions. Even with access to high-quality administrative data, neighborhood sorting remains the fundamental obstacle to neighborhood effects research.

2 Data

We consider four measures of neighborhood effects, where we define neighborhoods as census tracts.[2] For decennial Census data before 2010, when appropriate we impute count estimates into 2010 tract boundaries using the Longitudinal Tract Data Base (LTDB) described in Logan et al. (2014), Logan et al. (2016), and Logan et al. (2020). For each measure, we rank neighborhoods in terms of the national distribution of individuals.[3]

The first measure we use is the poverty rate in a tract, a common measure since at least Wilson (1987). The index labeled "neighborhood quality" as originally used in Aliprantis and Richter (2020) is the first principal component of six socio-economic factors available in the 1990 decennial census and 2014-2018 American Community Survey (ACS), downloaded from the National Histor-

[2]Assuming that census tracts are the unit over which neighborhood externalities operate is a strong assumption, typically made due to data limitations. See Durlauf (2004) and Galster (2019) for broad discussions and Chetty et al. (2020), Aliprantis (2017b), and McCartney and Shah (2019) for evidence of specific neighborhood effects that are highly-localized.

[3]When interpreting our results, it is important to recall that there are statistical challenges specific to rank measures (Mogstad et al. (2020)).

ical Geographic Information System (NHGIS, Manson et al. (2020)). The Childhood Opportunity Index 2.0 (COI) developed at Brandeis University (Noelke et al. (2020)) includes information from 29 items, many of which come from data sources beyond the Census, like the National Center for Education Statistics (NCES) and the Environmental Protection Agency (EPA).

The Opportunity Atlas (OA) uses the outcomes for individuals born between 1978 and 1983 who spent time growing up in a given neighborhood to predict outcomes for children growing up in those neighborhoods today. This birth cohort corresponds to children aged 6-11 in the 1990 Census. Unless otherwise stated, our analysis focuses on the OA ranking of neighborhoods based on the estimated average family income at age 29 for children with parents at the 25th percentile of income. This ranking is currently being used to guide policy decisions (Bergman et al. (2020)). We sometimes also refer to the OA rankings for high-income kids and low-income kids to denote children with, respectively, 75th and 25th percentile income parents. While we focus on the OA rankings pooled over race/ethnicity and gender, we also consider the OA rankings for black and white males with 25th percentile income parents.

Appendix A describes these measures in greater detail.

We estimate the number of high- and low-income children aged 6-11 in each tract in the 1990 Census. Recall that since the OA is estimated on children born between 1978 and 1983, and the 1990 census asked about the year 1989, the group of children aged 6-11 in the 1990 census is the OA estimation sample.[4] We define quartiles of the household income distribution using the five percent sample of the 1990 Census from the Integrated Public Use Microdata Series (IPUMS-USA, Ruggles et al. (2020)). We estimate the number of high-income kids in a tract as the share of the tract's households that are at or above the 75th percentile of household income times the number of children aged 6-11. We estimate the number of low-income kids in a neighborhood analogously in terms of the households at or below the 25th percentile of household income.[5]

To provide concrete examples of our findings, we consider two case studies from Cuyahoga County, Ohio. Appendix Figure 7a highlights the areas we consider. In blue is the Central Neighborhood, a collection of tracts that is one of the most economically disadvantaged areas in the city of Cleveland. In red is Shaker Heights, an inner ring suburb well-known for its efforts at racial and economic integration (Meckler (2019), Malone (2019), Galster (2019), Ferguson (2001), Ogbu (2003)).

3 Neighborhood Effects in the Opportunity Atlas

We first establish evidence that OA rankings do indeed reflect evidence on neighborhood effects relative to measures based on demographic characteristics. One approach to doing so is to relate

[4] Moreover, 1990 is likely to be the most important decennial census for use in comparison with the OA, since the age range of 6-11 is likely when neighborhoods most influenced children's outcomes relative to the alternative ranges of 0-1 or 16-21.

[5] Appendix Figure 6 replicates this approach to measuring the number of poor kids in a tract and finds that this measure is generally accurate.

each tract to neighborhood effects likely to operate outside of the tract. To do this, for each tract we calculate the mean neighborhood quality of other tracts in the same school district. If the OA ranking of neighborhoods has information on neighborhood effects beyond the neighborhood quality measure, we expect that this information would be correlated with the influence of other tracts sharing a given tract's school district.

We find that this is indeed the case: Tracts linked with higher quality tracts through their school district are systematically ranked higher by the OA than by neighborhood quality. Appendix Table 3 and Appendix Figure 9 show the details.

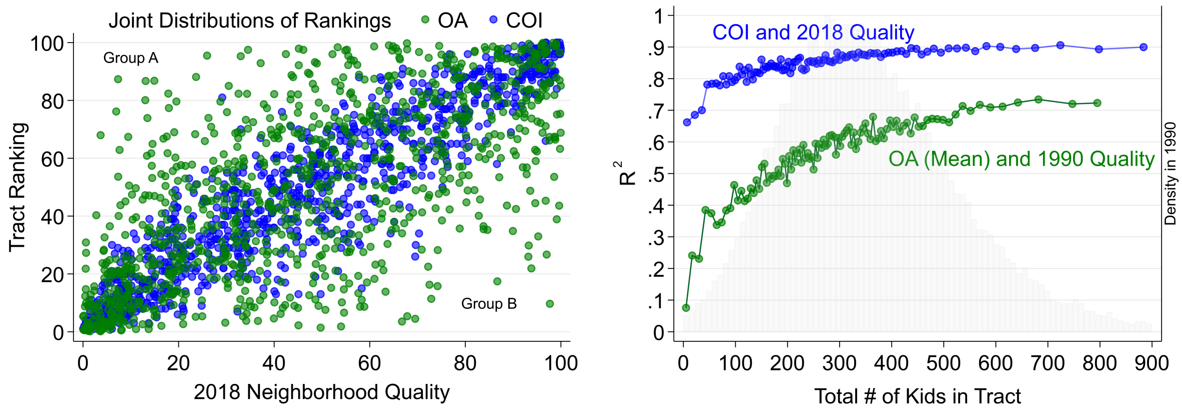
4 Neighborhood Sorting in the Opportunity Atlas

We now investigate the role of neighborhood sorting in determining OA rankings. We begin by documenting the fact that OA rankings have substantial variation that is not explained by observed characteristics. For example, when a tract's ranking in terms of 2018 neighborhood quality or the COI is regressed on the tract's ranking in terms of 2018 poverty, one obtains values of R^2 of 0.74 and 0.70, respectively. The additional variation in the COI is expected, due to its inclusion of many more variables than neighborhood quality (Lens (2017)). Subsequent additional variation in the OA may also be expected, but the magnitude of the additional variation is surprisingly low: When a tract's ranking in terms of the OA is regressed on the tract's ranking in terms of 2018 poverty, one obtains an R^2 of 0.35. See Appendix Table 4 and Figure 10 for more details.

Another way of showing the additional variation in the OA ranking is in Figure 1, which displays scatter plots of 1,000 randomly-selected tracts. The left panel shows that the COI and 2018 quality rankings of neighborhoods are highly correlated: A low-quality tract receives a low COI ranking, and a high-quality tract receives a high COI ranking. In contrast, the OA ranking of neighborhoods is closer to uniform conditional on quality. A low-quality tract may receive a very high OA ranking, and a high-quality tract may receive a very low OA ranking.

The right panel in Figure 1 provides suggestive evidence that the sample sizes generated by neighborhood sorting contribute to disagreements between the OA and neighborhood quality. The green dots in Figure 1b plot the R^2 of a regression of the OA ranking of a tract on its 1990 neighborhood quality ranking conditional on having a small range of children in the 1990 Census.[6] We see that the R^2 can rise above 0.7 for tracts with many children, but that the R^2 starts below 0.1 in tracts with the fewest children. This pattern is suggestive that an important share of the additional variation in the OA is due to small sample sizes resulting from neighborhood sorting.

[6]Note that in this case we are using the OA ranking for children with mean-income parents.



(a) COI, OA, and 2018 Quality

(b) COI, OA (Mean), and Quality

Figure 1: Joint Distributions of Rankings
Note: The left panel displays a scatterplot of 1,000 randomly-selected Census tracts. Green dots show the joint distribution of the OA and 2018 neighborhood quality rankings of neighborhoods, and blue dots show the joint distribution of the COI and 2018 neighborhood quality rankings of neighborhoods. The right panel reports the R^2 from population-weighted regressions within single percentiles of the total number of children aged 6-11 in the tract in the 1990 Census. The green dots show the R^2 for regressions of OA (mean) ranking on 1990 neighborhood quality and the blue dots show the R^2 for regressions of COI ranking on 2018 neighborhood quality.

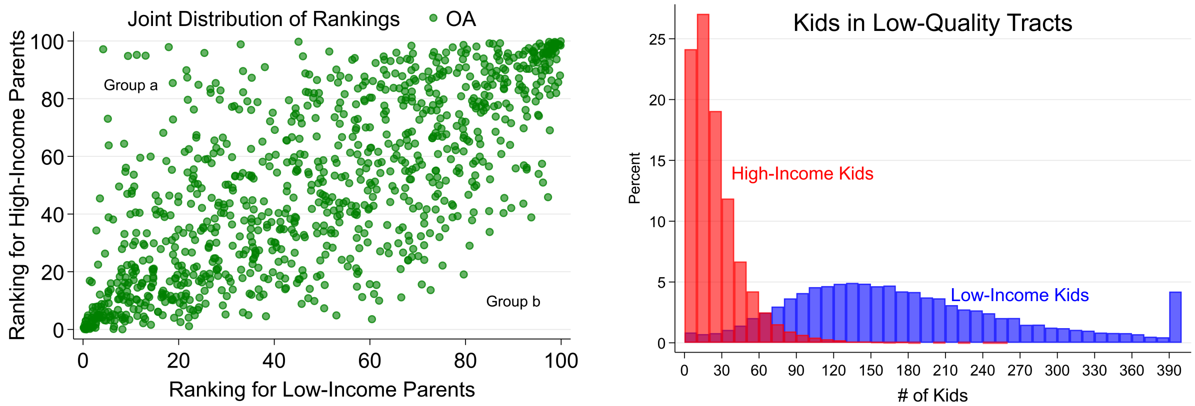
4.1 Neighborhood Sorting by Income Contributes to Disagreements

The left panel in Figure 2 shows that the high level of variation in the OA ranking of neighborhoods is also present when ranking neighborhoods for children from rich versus poor parents. The neighborhoods in "Group a" in the figure are ranked very low for children of poor parents, but very high for children of rich parents. Conversely, the neighborhoods in "Group b" are ranked very high for children of poor parents, but very low for children of rich parents.

How much does the variation in Figure 2a reflect neighborhood effects rather than neighborhood sorting? In other words, could the variation in Figure 2a be driven by small sample sizes due to neighborhood sorting by income?

Figure 2b shows that there was strong neighborhood sorting by income in the US in the 1990 Census: There are very few observations of high-income kids aged 6-11 in low-quality tracts. Half of low-quality tracts have less than 20 high-income children. Likewise, Appendix Figure 13b shows that there are few observations of low-income kids aged 6-11 in high-quality tracts. Almost half of high-quality tracts have less than 30 low-income kids.

Consider the implications of these figures for the neighborhoods in Group a in Figure 2a. OA estimates for a low-quality neighborhood where high-income kids do well but low-income kids do poorly are likely to reflect statistical noise, rather than neighborhood effects, due to the absence of high-income kids. Similarly, OA estimates for a high-quality neighborhood where high-income kids do well but low-income kids do poorly are likely to reflect statistical noise, rather than neighborhood effects, due to the absence of low-income kids.



(a) OA Rankings for 25th and 75th Percentile Households

(b) Kids in Low-Quality Neighborhoods

Figure 2: Joint Distribution of Rankings and Sample Sizes
Note: The left panel displays a scatterplot of the same 1,000 randomly-chosen Census tracts from Figure 1a. The green dots show the joint distribution of tracts' rankings in terms of the adult incomes of children with 25th percentile (low-) income parents and with 75th percentile (high-) income parents. The right panel displays the estimated number of children aged 6-11 in the 1990 Census with parents in the top and bottom quartiles of household income residing in tracts in the bottom quartile of 1990 neighborhood quality.

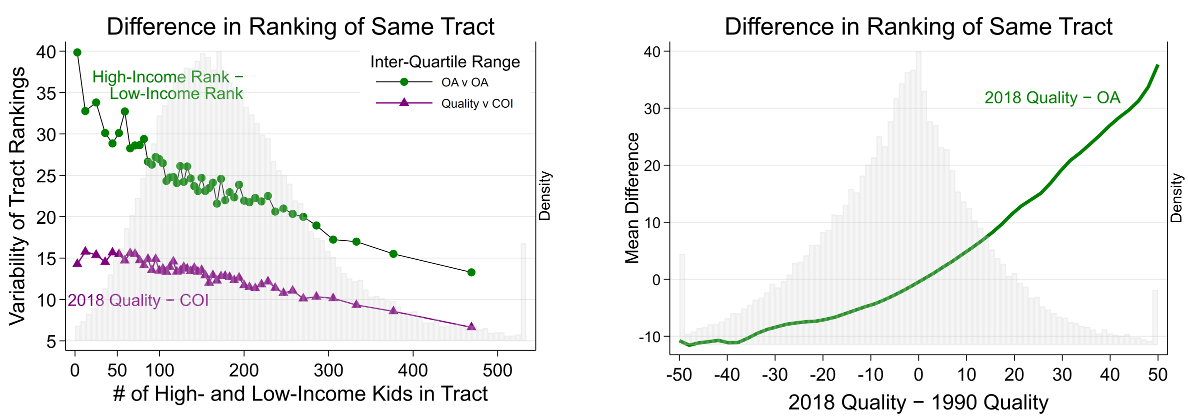
Looking at the case study of Shaker Heights, Ohio in Appendix Figure 7b helps to illustrate what neighborhood sorting by income means for OA estimates. This figure reports the expected individual income for children growing up in tracts in Shaker Heights together with the number of children aged 6-11 in the 1990 Census with parents at or below the 25th percentile of household income. The tracts highlighted in white show that large differences in estimates could be driven by a lack of low-income children. The white tract in the west predicts income that is 40 percent higher than the white tract in the east. There is reason to believe that this difference captures neighborhood effects: 1990 neighborhood quality in the western tract is 98, compared to 1990 quality of 83 for the eastern tract. However, there is also reason to believe that this difference in predicted income reflects small sample sizes. The tract in the west had 22 low-income children in the 1990 Census, while the tract in the east had 15.

The question of when differences in quality and OA rankings are driven by sample size is difficult because there is a mechanical relationship between neighborhood quality and the number of high- or low-income children in a census tract. However, we can look at the variation within the OA rankings by parental income to get some sense of the role of sample size in generating variation in the OA. Figure 3a shows that the variability in the difference in OA rankings conditional on parental income by showing the interquartile range (75th percentile minus the 25th percentile) of the distribution conditional on the sum of high- and low-income kids. We see that variability in the difference of OA rankings is highest when there are few children in a tract and lowest in tracts with many children. This suggests that differences in OA estimates for high- and low-income children are driven by statistical noise in addition to neighborhood effects. Neighborhood sorting by income is likely to generate the same kinds of statistical noise in the difference in OA and contemporary rankings of the same tract.

When interpreting the above figures, it is important to note that the OA rankings have been purposefully perturbed to protect the confidentiality of subjects. Adding statistical noise to the OA guided by this ethical constraint, as described in Chetty and Friedman (2019), means that larger perturbations are added precisely where the sample sizes are the smallest. This makes large deviations most likely to occur in tracts with few observations, just as they are more likely to occur for groups with smaller samples.[7]

4.2 Neighborhood Sorting over Time Contributes to Disagreements

A known issue with the OA rankings is that they measure outcomes for children who grew up in each tract decades ago. There are reasons to doubt that this would affect OA rankings of tracts today, since the ranking of tracts within metros tends to be stable over time (Malone and Redfearn (2018)). Figure 12b shows, perhaps surprisingly then, that changes in quality over time are highly predictive of disagreements between neighborhood quality and OA rankings. The figure shows a local linear regression of the difference in ranking between 2018 neighborhood quality and the OA ranking as a function of the difference between 2018 and 1990 neighborhood quality rankings. Changes in neighborhoods over time predict large differences between the neighborhood quality ranking based on current inhabitants and the OA ranking based on previous inhabitants. Appendix Figures 11 and 12 show this finding in terms of population growth as well.



(a) Variability of Rankings of the Same Tract

(b) Disagreements by Change in Quality

Figure 3: Explaining Differences in the Rankings of the Same Tract
Note: The left panel plots the interquartile range (75th percentile minus the 25th percentile) of the distribution of differences between the OA high-income and OA low-income ranking of a tract conditional on the sum of high- and low-income kids in the tract. The right panel plots a local linear regression of the mean difference in a tract's 2018 neighborhood quality minus its OA ranking.

Returning to Figure 1a, we again consider the issue of how to interpret differences between OA rankings of tracts and rankings based on contemporary measures of observable characteristics. Our

[7]See, for example, that the relative share of added noise to income for Blacks is higher than it is for Whites as reported in Table II of Chetty et al. (2020).

results suggest that the presence of tracts in Group B could be explained in terms of neighborhood sorting over time rather than neighborhood effects. When combined with the results on neighborhood sorting by income, we conclude that both a small number of low-income children and changes over time could explain tracts in Group B.

4.3 Neighborhood Sorting by Race Leads to Limited Overlap

One of the first questions we might want to ask with the new OA data is the following: How much would intergenerational income mobility converge if black and white boys grew up in the same neighborhoods? This question is motivated by the facts that (i) a child's expected income is different across race, even conditional on parental income (Mazumder (2012)), and (ii) this racial inequality in intergenerational mobility is driven by differences in boys' outcomes (Chetty et al. (2020)).

Using the OA data set, Chetty et al. (2020) estimate that black males under-perform the individual income of white males by 10 percentile points across the income distribution of parents. In regressions that includes tract and block fixed effects this gap is only reduced to 8 and 7 percentile points. A natural interpretation of these results is that the intergenerational income mobility gap would fall by at most 30 percent if black and white boys grew up in the same neighborhoods.

There are three important caveats to interpreting this exercise, with the third being newly presented in this note.

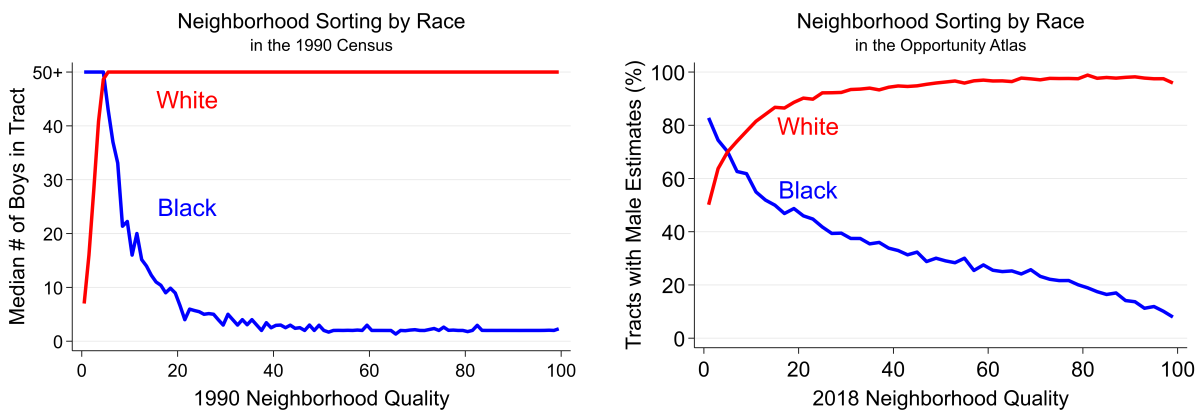
First, neighborhood effects may not be invariant to the societal changes required for residential integration. There are reasons to believe that residential integration in the US would create "a national reckoning that would lead to spiritual renewal" (Coates (2014)) and a renewed "self-conception as a democratic society" (Rothstein (2017)), generating the type of social change required for black and white people to be treated more equally when living in the same neighborhoods than is currently the case (Aliprantis et al. (2020)).

Second, the OA data are for adults in their 30s. While the OA takes lifecycle bias into account by characterizing neighborhoods in terms of rankings rather than raw outcomes (Chetty et al. (2020)), this issue cannot be entirely accounted for by research design. Using data with a longer time horizon, like the Panel Study of Income Dynamics (PSID), would likely lead us to conclude that the gap in mobility is even larger than estimated in the OA data (Mazumder (2018)).

Third, the regressions estimated in Chetty et al. (2020) only apply to a select set of neighborhoods. The strength of residential segregation in the US means that we simply have not observed many black children growing up in tracts where we would expect them to experience large positive neighborhood effects on their economic outcomes.

In the 1990 Census, the median tract in the top half of neighborhood quality had 2 black boys in the OA sample age range (6-11). Figure 4a shows how quickly the median number of black boys in 1990 Census tracts falls as 1990 neighborhood quality increases. At the lowest levels of quality, most tracts have 50 black boys or more with which to estimate outcomes. But once quality gets out of the bottom decile, the number of black boys is already too low to reliably estimate outcomes in many neighborhoods. Outside the bottom third of tracts, most neighborhoods simply do not have enough observations to reliably predict how black boys would do if residing there. Appendix Figure 14a shows that this is not a matter of black boys being concentrated in urban areas; the same pattern holds in metros with populations of at least 1 million inhabitants.

Figure 4b shows how this neighborhood sorting by race in the 1990 Census passes through to the Opportunity Atlas. The share of tracts with publicly-reported outcome estimates for black males drops rapidly as 2018 neighborhood quality rises. In the top half of tracts, 21 percent of tracts have estimates for black males.[8] The strong neighborhood sorting by race in the 1990 Census could also have implications for the patterns documented in Section VII.B.2 of Chetty et al. (2020).



(a) Number of Boys by Quality

(b) Tracts with OA Estimates by Quality

Figure 4: Neighborhood Sorting by Race
Note: The left panel plots the median number of black and white boys in the OA sample age range, 6-11, in census tracts at each level of 1990 neighborhood quality. The right panel plots the percent of tracts at each level of 2018 neighborhood quality that have OA rankings published for black and white males.

Returning to the case study of Shaker Heights, Ohio, we can see a concrete example of what neighborhood sorting by race means for OA estimates.

Appendix Figure 8a shows the expected individual income in tracts along with the number of black boys aged 6-11 in the 1990 Census. Again comparing the tracts highlighted in white, we see that the western tract has predicted income that is nearly 50 percent higher than the eastern tract. Given the sample sizes of 22 and 10 in these tracts, though, it is difficult to judge how much of this difference reflects neighborhood effects and how much reflects statistical noise due to small samples. We also note that these sample sizes in Shaker Heights reflect tracts *below* the blue line in Figure 4b. Only one tract in Shaker Heights does not have OA estimates for black males, and that tract had 5 black boys aged 6-11 in the 1990 Census. Tracts with 10 and 11 boys in the 1990 Census do have publicly-reported estimates.

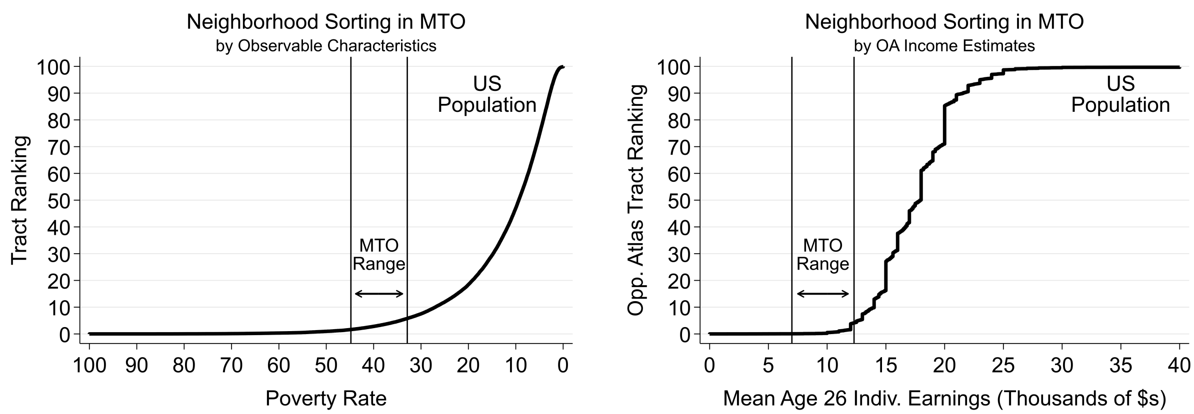
[8] Chetty et al. (2020) report a sample size cutoff of 20 observations for publicly releasing a tract's estimate, and the distributions of within-tract gaps shown in Chetty et al. (2020) Online Appendix Figure XIVa excludes tracts with fewer than 50 black or white male children.

5 Policy Implications

5.1 Housing Mobility Programs

Aliprantis et al. (2020) show that in most cases the choice between measures of neighborhood effects will not have large implications for the success of housing mobility programs. Nevertheless, we are still interested in using the OA to gain new insights into housing mobility programs.

Neighborhood sorting affects how we can use the OA to think about housing mobility programs like Moving to Opportunity (MTO). Figure 5a illustrates the dichotomy in interpretations of MTO based on observable socio-economic characteristics. When viewed in terms of changes in the raw poverty rate, MTO can be interpreted as having induced large changes in participants' neighborhood poverty (Kling et al. (2007), Fryer Jr and Katz (2013), Ludwig et al. (2008)). When viewed in terms of changes in the distribution of poverty and other observable characteristics, MTO can be interpreted as having induced small changes in participants' neighborhoods. The latter view sees MTO as having moved participants around segregated neighborhoods still likely to be disconnected from the mainstream economy (Sampson (2008), Clampet-Lundquist and Massey (2008), Aliprantis (2017a)).



(a) Sorting in MTO by Poverty

(b) Sorting in MTO by OA Income Estimates

Figure 5: Neighborhood Sorting in Moving to Opportunity (MTO)

Note: The left panel presents data from the 2000 US Census along with MTO treatment and control group means as reported in Kling et al. (2007). The right panel uses those age 26 in the 5% IPUMS-USA sample of the 2004-2008 American Community Survey, when the OA sample was aged 26, to compute the percentiles of the individual-level earnings distribution. OA tract outcomes for children with 10th percentile parents estimated and reported in terms of these percentiles are linked with the US population in the 2000 Census to then provide an OA ranking of tracts in terms of mean age 26 individual earnings. The MTO treatment and control group means are taken to be \$7,000 and \$12,289 for mean individual earnings for children with parents at p=10 in the Opportunity Atlas based on Chetty et al. (2020) Figure X.

Figure 5b shows that the dichotomy in interpretations of MTO based on observable characteristics extends to the interpretation of MTO based on the OA. If viewed in terms of the raw change in mean individual income, the change induced by MTO was large. If viewed in terms of the OA ranking of neighborhoods in terms of mean individual income, the change induced by MTO was small. The range of neighborhoods shown in Figure X of Chetty et al. (2020) is restricted to the left tail of tracts in the US according to their OA rankings. The case study of the Central Neighborhood in Cleveland, Ohio helps us to understand this range. Appendix Figure 8b shows that the treatment mean in MTO does not reach beyond the Central Neighborhood: The experimental group in MTO tended to remain in the most economically-disadvantaged tracts in the country.

Appendix Figure 15 shows the implications of the OA dichotomy for extrapolating results from MTO using the OA data. The left panel shows that when extrapolating based on tracts' raw outcomes across the range in Chetty et al. (2020) Figure XIV, the support of the data is large relative to the support of extrapolation. The right panel shows that when extrapolating based on tracts' rankings, the support of the data is small relative to the support of extrapolation. Our conclusion is that identifying neighborhood effects using MTO data will require using a model, whether a linear regression as in Chetty et al. (2020) or a model of neighborhood selection as in Aliprantis and Richter (2020).

6 Conclusion

The Opportunity Atlas makes substantial contributions to our understanding of intergenerational mobility and provides a foundation for the future of neighborhood effects research. We find that caution is warranted, however, in using the Opportunity Atlas (OA) as a literal atlas of opportunity. Its authors point to the OA's disagreement with traditional measures of neighborhood effects as evidence that the OA is more informative due to its direct measurement of outcomes. We find some suggestive evidence to support that assertion. However, we also find that the promising estimates policymakers might care about most are often thin on underlying data (ie, high opportunity areas for low-income Black people). The resulting estimates are likely the product of parametric estimation more than a robust sample of parents and their children.

The challenges we document in this note are not unique to the OA; it is difficult both to identify neighborhood effects (Graham (2018)) and to communicate uncertainty (Manski (2015)). Nevertheless, our results suggest practitioners should neither start nor end their search for opportunity neighborhoods with OA data. We recommend users of the OA data pay significant attention to the credibility of estimates when focusing on neighborhoods that have historically had few families of a particular group on which to base the estimates and neighborhoods that have experienced significant change over time.

References

Aliprantis, D. (2017a). Assessing the evidence on neighborhood effects from Moving to Opportunity. *Empirical Economics 52*(3), 925–954. DOI: 10.1007/s00181-016-1186-1.

Aliprantis, D. (2017b). Human capital in the inner city. *Empirical Economics 53*(3), 1125–1169. DOI: 10.1007/s00181-016-1160-y.

Aliprantis, D. and D. Carroll (2018). Neighborhood dynamics and the distribution of opportunity. *Quantitative Economics 9*(1), 247–303. DOI: 10.3982/QE785.

Aliprantis, D., D. Carroll, and E. Young (2020). What explains neighborhood sorting by income and race? *FRB of Cleveland WP 18-08R*. DOI: 10.26509/frbc-wp-201808r.

Aliprantis, D., H. Martin, and K. Tauber (2020). What determines the success of housing mobility programs? *FRB of Cleveland WP 20-36*. DOI: 10.26509/frbc-wp-202036.

Aliprantis, D. and F. G.-C. Richter (2020). Evidence of neighborhood effects from Moving to Opportunity: LATEs of neighborhood quality. *The Review of Economics and Statistics 102*(4), 633–647. DOI: 10.1162/rest_a_00933.

Bergman, P., R. Chetty, S. DeLuca, N. Hendren, L. F. Katz, and C. Palmer (2020). Creating Moves to Opportunity: Experimental evidence on barriers to neighborhood choice. *NBER Working Paper 26164*. DOI: 10.3386/w26164.

Chetty, R. and J. N. Friedman (2019). A practical method to reduce privacy loss when disclosing statistics based on small samples. *American Economic Review Papers & Proceedings 109*, 414–420. DOI: 10.1257/pandp.20191109.

Chetty, R., J. N. Friedman, N. Hendren, M. R. Jones, and S. R. Porter (2020). The Opportunity Atlas: Mapping the childhood roots of social mobility. *Mimeo., Opportunity Insights*.

Chetty, R., N. Hendren, M. R. Jones, and S. R. Porter (2020). Race and economic opportunity in the United States: An intergenerational perspective. *Quarterly Journal of Economics 135*(2), 711–783. DOI: 10.1093/qje/qjz042.

Clampet-Lundquist, S. and D. S. Massey (2008). Neighborhood effects on economic self-sufficiency: A reconsideration of the Moving to Opportunity experiment. *American Journal of Sociology 114*(1), 107–143. DOI: 10.1086/588740.

Coates, T.-N. (2014, June). The case for reparations. *The Atlantic*.

Durlauf, S. N. (2004). Neighborhood Effects. In J. V. Henderson and J.-F. Thisse (Eds.), *Cities and Geography*, Volume 4 of *Handbook of Regional and Urban Economics*, Chapter 50, pp. 2173–2242. Elsevier. DOI: 10.1016/S1574-0080(04)80007-5.

Ferguson, R. F. (2001, April 8). A diagnostic analysis of Black-White GPA disparities in Shaker Heights, Ohio. *Brookings Papers on Education Policy*, 347–414. DOI: 10.1353/pep.2001.0003.

Fryer Jr, R. G. and L. F. Katz (2013). Achieving escape velocity: Neighborhood and school interventions to reduce persistent inequality. *American Economic Review 103*(3), 232–237. DOI: 10.1257/aer.103.3.232.

Galster, G. C. (2019). *Making Our Neighborhoods, Making Ourselves.* Chicago: University of Chicago Press. DOI: 10.7208/chicago/9780226599991.001.0001.

Graham, B. S. (2018). Identifying and estimating neighborhood effects. *Journal of Economic Literature 56*(2), 450–500. DOI: 10.1257/jel.20160854.

Kling, J. R., J. B. Liebman, and L. F. Katz (2007). Experimental analysis of neighborhood effects. *Econometrica 75*(1), 83–119. DOI: 10.1111/j.1468-0262.2007.00733.x.

Lens, M. C. (2017). Measuring the geography of opportunity. *Progress in Human Geography 41*(1), 3–25. DOI: 10.1177/0309132515618104.

Logan, J. R., B. Stults, and Z. Xu (2016). Validating population estimates for harmonized census tract data, 2000-2010. *Annals of the American Association of Geographers 106*(5), 1013–1029. DOI: 10.1080/24694452.2016.1187060.

Logan, J. R., Z. Xu, and B. Stults (2014). Interpolating US decennial census tract data from as early as 1970 to 2010: A longitudinal tract database. *Professional Geographer 66*(3), 412–420. DOI: 10.1080/00330124.2014.905156.

Logan, J. R., C. Zhang, B. Stults, and T. Gardner (2020). Improving estimates of neighborhood change with constant tract boundaries. *Mimeo., Brown University*.

Ludwig, J., J. B. Liebman, J. R. Kling, G. J. Duncan, L. F. Katz, R. C. Kessler, and L. Sanbonmatsu (2008). What can we learn about neighborhood effects from the Moving to Opportunity experiment? *American Journal of Sociology 114*(1), 144–188. DOI: 10.1086/588741.

Malone, C. (2019, April 8). A tale of two suburbs. *FiveThirtyEight*.

Malone, T. and C. L. Redfearn (2018). Shocks & Ossification: The durable hierarchy of neighborhoods in US metropolitan areas from 1970 to 2010. *Regional Science and Urban Economics 69*, 94–121. DOI: 10.1016/j.regsciurbeco.2018.01.002.

Manski, C. F. (2015). Communicating uncertainty in official economic statistics: An appraisal fifty years after Morgenstern. *Journal of Economic Literature 53*(3), 631–53. DOI: 10.1257/jel.53.3.631.

Manson, S., J. Schroeder, D. V. Riper, and S. Ruggles (2020). *IPUMS National Historical Geographic Information System* (15.0 ed.). Minneapolis: University of Minnesota. [Database] DOI: 10.18128/D050.V15.0.

Mazumder, B. (2012). Black-white differences in intergenerational mobility in the US. *Chicago Fed WP 2011-10R*.

Mazumder, B. (2018). Intergenerational mobility in the United States: What we have learned from the PSID. *The Annals of the American Academy of Political and Social Science 680*(1), 213–234. DOI: 10.1177/0002716218794129.

McCartney, W. B. and A. M. Shah (2019). The economic importance of neighbors: Evidence from hyperlocal social influence effects in mortgage markets. *Mimeo., Purdue University*. DOI: 10.2139/ssrn.2882317.

Meckler, L. (2019, October 11). This trail-blazing suburb has tried for 60 years to tackle race. What if trying isn't enough? *The Washington Post*.

Mogstad, M., J. P. Romano, A. Shaikh, and D. Wilhelm (2020). Inference for ranks with applications to mobility across neighborhoods and academic achievement across countries. *NBER Working Paper 26883*. DOI: 10.3386/w26883.

Noelke, C., N. McArdle, M. Baek, N. Huntington, R. Huber, E. Hardy, and D. Acevedo-Garcia (2020). *Child Opportunity Index 2.0 Technical Documentation.* Brandeis University. Retrieved from diversitydatakids.org/researchlibrary/research-brief/how-we-built-it.

Ogbu, J. U. (2003). *Black American students in an affluent suburb: A study of academic disengagement.* Routledge. DOI: 10.4324/9781410607188.

Rothstein, R. (2017). *The Color of Law.* New York and London: W. W. Norton & Company.

Ruggles, S., S. Flood, R. Goeken, J. Grover, E. Meyer, J. Pacas, and M. Sobek (2020). *IPUMS USA: Version 10.0 [dataset].* Minneapolis, MN: IPUMS. DOI: 10.18128/D010.V10.0.

Sampson, R. J. (2008). Moving to inequality: Neighborhood effects and experiments meet social structure. *American Journal of Sociology 114*(1), 189–231. DOI: 10.1086/589843.

Wilson, W. J. (1987). *The Truly Disadvantaged: The Inner City, the Underclass, and Public Policy.* University of Chicago.

A Strengths and Weaknesses of Each Measure

The strength of each measure of neighborhood effects we consider tends to be a weakness of the other measures. Table 2 summarizes the strengths and weakness of each measure.

The strength of neighborhood poverty is that (P-i) the neighborhood effects mechanisms it is thought to capture are clear and intuitive; and that (P-ii) the data are both timely and made publicly-available by the NHGIS. We believe that concentrated poverty has negative effects on residents in a neighborhood. The weakness of poverty as a measure of neighborhood effects is that (P-a) we also believe there should be considerable variation in neighborhood effects depending on other characteristics of a neighborhood.

The strength of neighborhood quality as a measure of neighborhood effects is that (Q-i) it can be easily calculated from timely census data that is made publicly-available by the NHGIS. (Q-ii) While only negligibly more difficult to calculate than neighborhood poverty, neighborhood quality captures many of the additional variables thought to determine how a neighborhood affects its residents above and beyond poverty alone. Weaknesses of quality are that (Q-a) it may not capture all relevant neighborhood characteristics; and (Q-b) the characteristics included may not affect outcomes in a straight-forward (ie, linear/additive) way.

Two strengths of the COI as a measure of neighborhood effects are that (COI-i) it can be calculated from timely data; and (COI-ii) it incorporates even more neighborhood characteristics thought to affect residents' outcomes, like school district outcomes and pollution, from disparate datasets that contain information not available in the Census. Strength COI-ii helps to address weakness Q-a of quality by doing more to capture all relevant neighborhood characteristics, but this comes with the tradeoff of (COI-a), that the COI is more difficult to calculate, requiring the assembly of multiple datasets. Another weakness is a holdover from quality: (COI-b) the characteristics included may not affect outcomes in a straight-forward (ie, linear/additive) way.

Three strengths of the OA as a measure of neighborhood effects are that it allows us to (OA-i) measure actual outcomes; (OA-ii) measure said outcomes conditional on individual characteristics like race/ethnicity and gender; and (OA-iii) measure a wide variety of outcomes like incarceration, teenage pregnancy, and marriage. Three weaknesses of the OA as a measure of neighborhood effects are (OA-a) neighborhood sorting by individual-level demographic characteristics X_i resulting in small sample sizes; (OA-b) neighborhood sorting over time resulting in bias; and (OA-c) interpretation due to the fact that outcomes are a result of neighborhood effects and individual characteristics. Weakness OA-a is an issue because OA rankings are estimated. Thus, to capture strength OA-ii, we may be concerned for cases where sample sizes are small enough to make estimates noisy. Strength OA-i is particularly exciting because it could allow us to address weaknesses Q-b and COI-b. However, weakness OA-c makes strength OA-i difficult to gauge. Measuring outcomes does not overcome the fundamental issue in neighborhood effects research, neighborhood sorting. We simply do not know if realized outcomes reflect neighborhood effects or neighborhood sorting.

Table 1: Measures of Neighborhood Effects

Neighborhood Poverty: Tract-level poverty rate

Years: 2014-2018, 2013-2017, ..., 2005-2009, 2000, 1990, 1980, 1970

Area: All census tracts

Sources: American Community Survey (ACS) from 2005–; Decennial censuses from 1970-2000; Longitudinal Tract Database (LTDB)

Construction: Created by dividing the number of people in poverty in a tract by the total number of people in a tract (for whom poverty status is determined). LTDB is used to interpolate earlier years into 2010 census tract boundaries.

Neighborhood Quality: Aliprantis and Richter (2020)'s tract-level neighborhood quality index

Years: 2014-2018, 2013-2017, ..., 2005-2009, 2000, 1990, 1980, 1970

Area: All census tracts

Sources: American Community Survey (ACS) from 2005–; Decennial censuses from 1970-2000; Longitudinal Tract Database (LTDB)

Construction: Created by using principal components analysis to combine tract-level ranks of six neighborhood characteristics into a single tract-level ranking of neighborhoods. Those six characteristics are the poverty rate, the share of adults 25+ with a high school diploma, the share of adults 25+ with a BA, the Employment to Population Ration for adults 16+, the labor force participation rate for adults 16+, and the share of families with children under 18 with only a mother or father present. LTDB is used to interpolate earlier years into 2010 census tract boundaries.

COI: Brandeis University's tract-level Child Opportunity Index 2.0

Years: 2013-2017 and 2008-2012

Area: All census tracts

Sources: 29 indicators from numerous sources including the American Community Survey (ACS), National Center for Education Statistics (NCES), Stanford Education Data Archive (SEDA), GreatSchools (GS) proprietary data, US Department of Education EDFacts, US Department of Education Office for Civil Rights Data Collection (CRDC), Environmental Protection Agency Risk-Screening Environmental Indicators (EPA RSEI), Centers for Disease Control and Prevention (CDC), Opportunity Atlas, RW Johnson Foundation 500 Cities Project

Construction: Created by combining tract-level measures of many neighborhood characteristics into a single tract-level ranking of neighborhoods.

OA: Opportunity Atlas tract-level income estimates

Years: 1978-83 birth cohorts

Area: Census tracts with sufficient observations

Sources: Census 2000 and 2010; Federal income tax returns in 1989, 1994, 1995, and 1998-2015

Construction: Estimate child's expected income conditional on their parents' household income

Table 2: Strengths and Weaknesses of Neighborhood Effects Measures

Neighborhood Poverty Strengths

P-i: the neighborhood effects mechanisms it is thought to capture are clear

P-ii: the data are both timely and made publicly-available by the NHGIS

Neighborhood Poverty Weaknesses

P-a: We believe there should be considerable variation in neighborhood effects depending on other characteristics of a neighborhood

Neighborhood Quality Strengths

Q-i: Measure can be easily calculated from timely census data that is made publicly-available by the NHGIS

Q-ii: Measure captures 6 key variables thought to determine neighborhood effects

Neighborhood Quality Weaknesses

Q-a: Measure may not capture all relevant neighborhood characteristics

Q-b: The characteristics included may not affect outcomes in a straight-forward (ie, linear/additive) way

COI Strengths

COI-i: Measure can be calculated from timely data

COI-ii: Measure captures many measurable variables thought to determine neighborhood effects

COI Weaknesses

COI-a: Measure is more difficult to calculate than poverty or quality

COI-b: The characteristics included may not affect outcomes in a straight-forward (ie, linear/additive) way

OA Strengths

OA-i: Measure is based on realized outcomes rather than neighborhood characteristics

OA-ii: Measure can be made conditional on individual characteristics like race/ethnicity and gender

OA-iii: Measure can be made for a wide variety of outcomes like incarceration, teenage pregnancy, and marriage

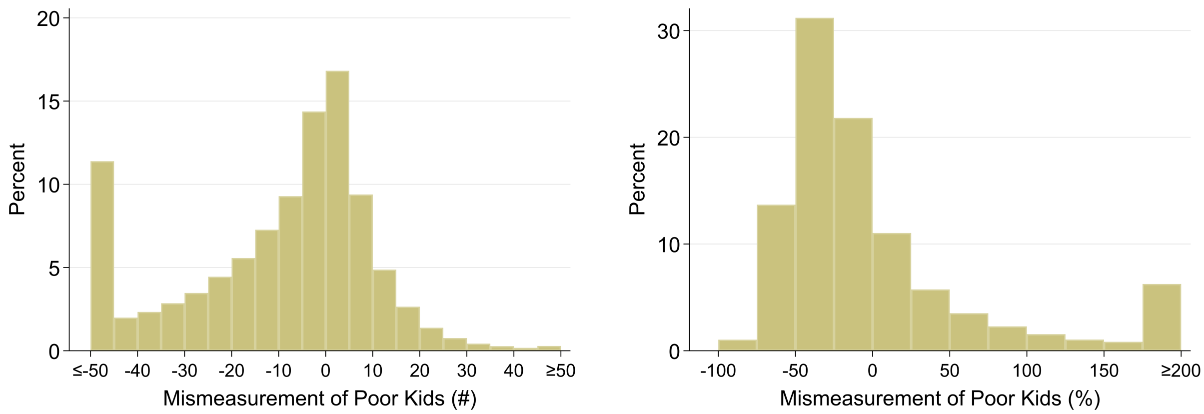
OA Weaknesses

OA-a: Neighborhood sorting by X_i's resulting in small sample sizes

OA-b: Neighborhood sorting over time resulting in bias

OA-c: Neighborhood sorting means realized outcomes do not necessarily reflect neighborhood effects (that is, $Y_i = f_i(X_i, X_j) \neq f_i(X_j)$)

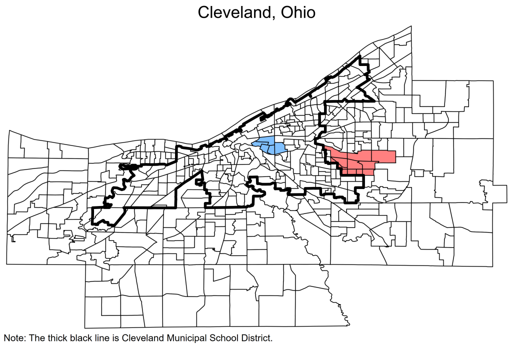
B Measuring Kids by Poverty Status



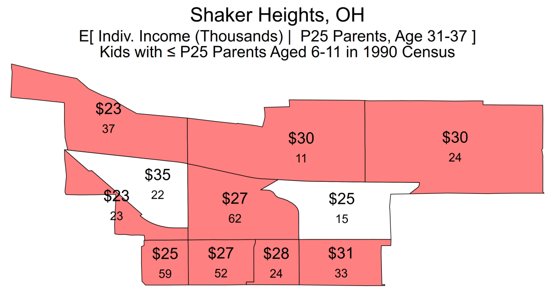
(a) Raw (b) Percentage

Figure 6: Examining Mismeasurement using Kids in Poverty
Note: We estimate the number of poor children in a tract, $\widehat{p}$, as the share of a tract's families that are poor times the number of kids age 6-11 in the tract. In the NHGIS publicly-released 1990 Census data we can observe the true number of children aged 6-11 who are poor, p. In the left panel we compute mismeasurement as $\widehat{p}-p$, and in the right panel we compute mismeasurement as $100 \times \frac{\widehat{p}-p}{p}$. Note the asymmetric tails of mismeasurement: Poor children in a tract imply poor adults in a tract, but poor adults in a tract are not necessarily accompanied by children.

C Case Studies

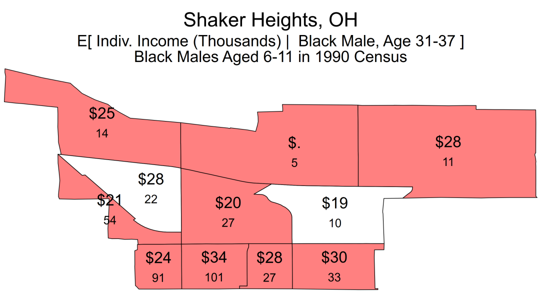


(a) Cuyahoga County, OH

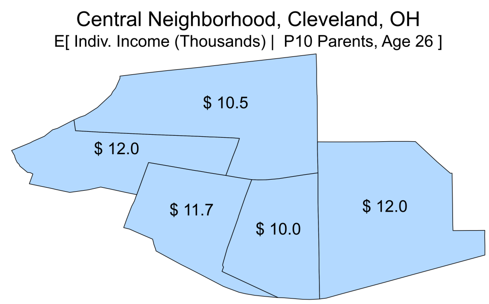


(b) Shaker Heights, OH

Figure 7: Case Studies from Cleveland, Ohio
Note: The left panel shows Cuyahoga County, Ohio. Highlighted in blue are tracts in the Central Neighborhood, and highlighted in red are tracts in the city of Shaker Heights. The right panel shows OA estimates of individual income for low-income children together with the number of low-income children in the OA sample age range in the 1990 Census.



(a) Shaker Heights, OH



(b) Central Neighborhood, Cleveland, OH

Figure 8: Case Studies from Cleveland, Ohio
Note: The left panel shows OA estimates of individual income for black males together with the number of black boys in the OA sample age range in the 1990 Census. The right panel shows OA estimates of individual income at age 26 for children with parents at the 10th percentile of household income.

D Neighborhood Effects

Table 3: Test for Neighborhood Effects

Dep. Variable	Indep. Variable	β	R^2
1990 Quality$_i$ – OA$_i$	1990 Quality$_i$ – 1990 Quality$_{-i}$	0.42	0.10
		(0.00)	

Note: The dependent variable is the 1990 quality rank of each tract i minus the OA rank of each tract i. The independent variable is the 1990 quality rank of each tract i minus the mean 1990 quality rank of tracts, excluding tract i, located in the same school district as tract i. 1990 school district boundaries are obtained from the National Center for Education Statistics (NCES) website. We use the boundary file from 1995 which represents the 1989-1990 school district boundaries, and focus on unified school district boundaries.

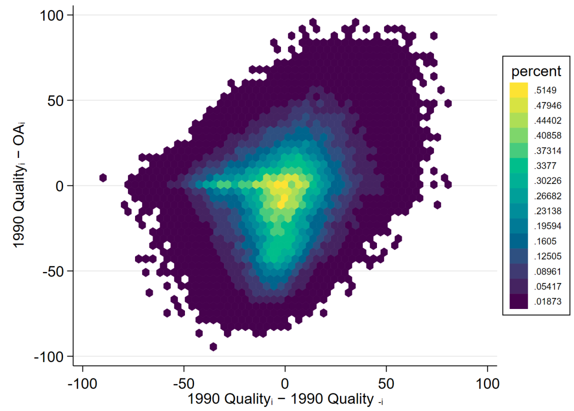


Figure 9: OA and School District Disagreements
Note: This figure shows the data reported in the Table 3. OA disagrees with the 1990 quality ranking of a tract in the same direction as the rest of the tract's school district disagrees with the 1990 quality ranking of the tract.

E Variation in Neighborhood Effects Measures

Table 4: Variation Explained

Neighborhood Effects Measure	Independent Variable	R^2
2018 Quality	2018 Poverty	0.74
COI	2018 Poverty	0.70
OA	2018 Poverty	0.35
OA	1990 Poverty	0.33
OA	1990 Quality	0.39
2018 Quality	1990 Quality	0.67
2018 Quality	COI	0.86

Note: All measures are in terms of percentile ranks. The top three rows are the relationships shown in the figure on the right. All regressions are weighted by the population at the time of measurement for the independent variable. The OA rank is in terms of the income estimates pooled over race/ethnicity for children from parents with 25th percentile incomes.

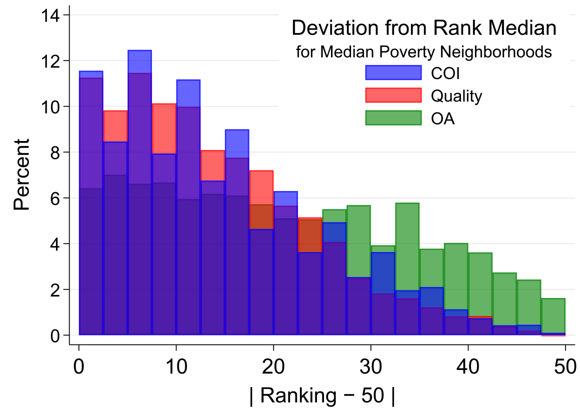
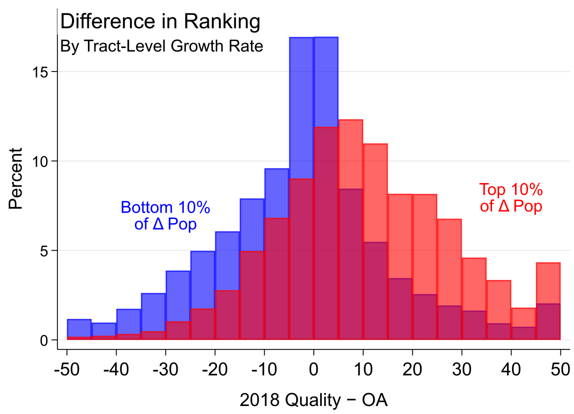
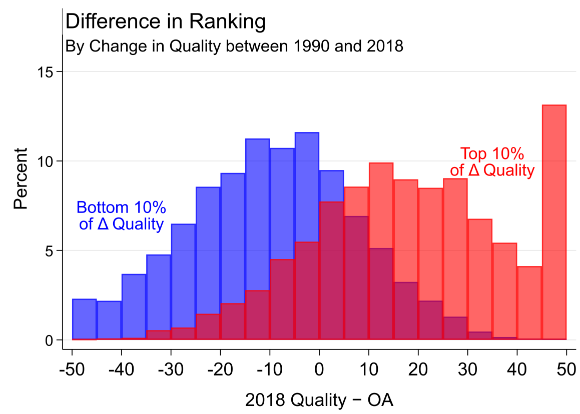


Figure 10: Variation in Other Measures for Median Poverty Tracts
Note: The figure shows tracts that are between the 47.5th and 52.5th percentiles of the individual-level distribution of tract-level poverty rates in 2014-2018.

F Changes over Time

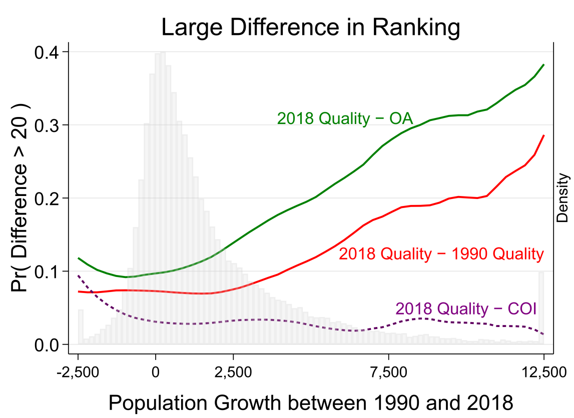


(a) Change in Population

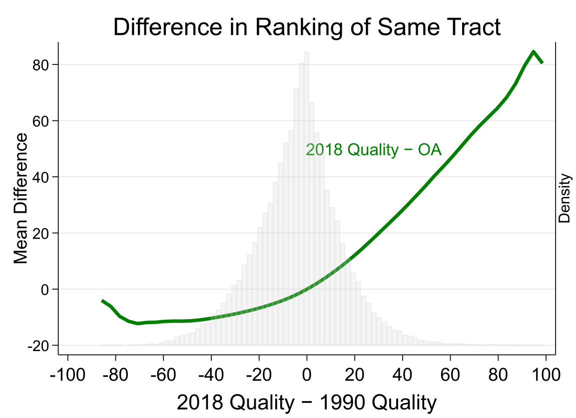


(b) 2017 Quality – 1990 Quality

Figure 11: Predicting Disagreement in 2018 Quality and OA Rankings
Note: The left panel shows the distributions of disagreement between the 2018 quality and OA rankings of tracts for those tracts in the top and bottom 10 percent of population growth between 1990 and 2018. The right panel shows the distributions of disagreement between the 2018 quality and OA rankings of tracts for those tracts in the top and bottom 10 percent of the change in quality between 1990 and 2018.



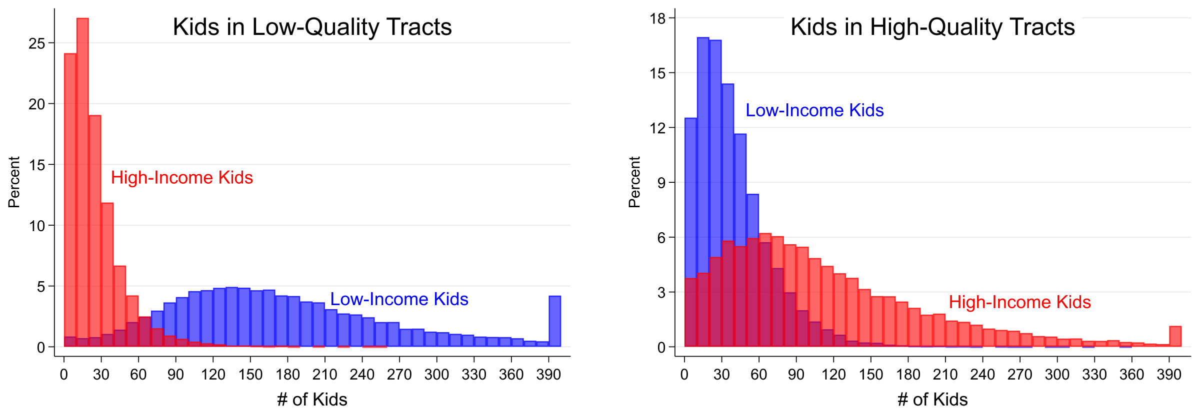
(a) Large Disagreements by Population Growth



(b) Expected Disagreement by Change in Quality

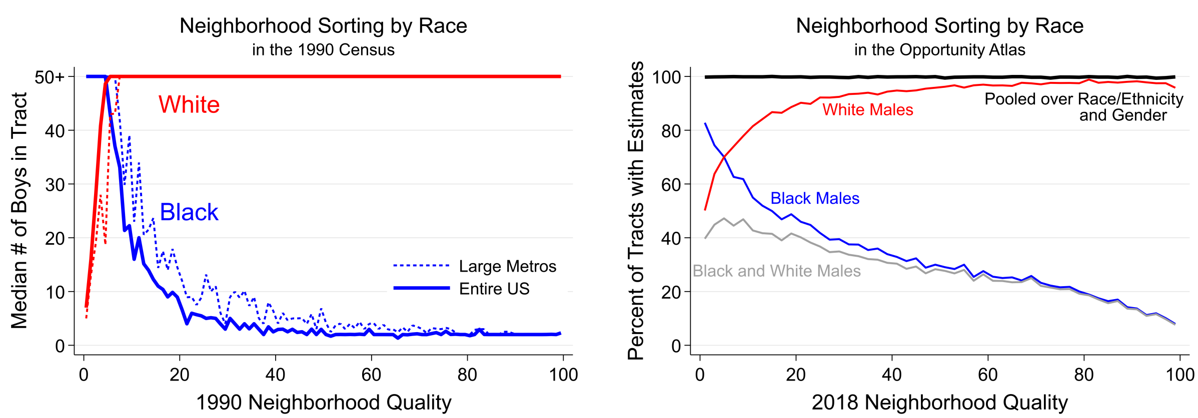
Figure 12: Predicting Large Disagreements in 2018 Quality and OA Rankings
Note: The left panel shows local linear regressions of the probability that 2018 quality ranks a tract at least 20 percentile points higher than another measure as a function of population growth in the tract between 1990 and 2018. The other rankings shown are OA in green, 1990 quality in red, and COI in purple. The right panel shows the mean difference in 2018 quality and OA rankings of a tract as a function of the change in quality between 1990 and 2018.

G Common Support



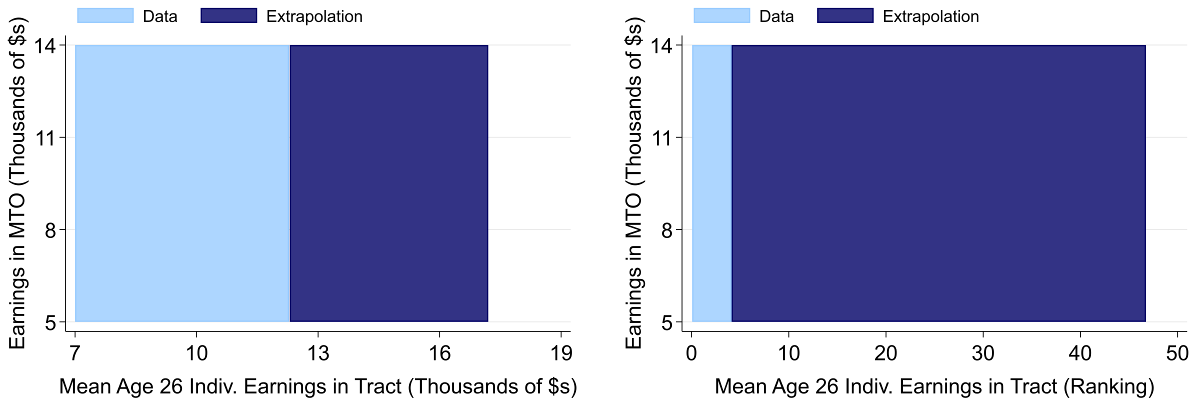
(a) Kids in Low-Quality Neighborhoods (b) Kids in High-Quality Neighborhoods

Figure 13: Sample Sizes by Income
Note: The left panel displays the estimated number of children aged 6-11 in the 1990 Census with parents in the top and bottom quartiles of household income residing in tracts in the bottom quartile of 1990 neighborhood quality. The right panel displays the estimated number of children aged 6-11 in the 19990 Census with parents in the top and bottom quartiles of household income residing in tracts in the top quartile of 1990 neighborhood quality.



(a) Number of Boys by Quality (b) Overlap of Estimates

Figure 14: Sample Sizes and Common Support by Race
Note: The left panel shows the median number of black and whites boys in a tract conditional on being in a given percentile of 1990 neighborhood quality. The dashed lines show the medians when calculated only for tracts in the 54 largest metros in the 2017 American Community Survey, with each metro have at least 1 million inhabitants. The right panel shows the percent of tracts with OA estimates conditional on black males and white males at each percentile of neighborhood quality.



(a) By Raw Neighborhood Characteristic

(b) By Neighborhood Ranking

Figure 15: Extrapolating with the MTO Data

Note: Both panels assume MTO treatment and control group means of $7,000 and $12,289 for mean individual earnings for children with parents at p=10 in the Opportunity Atlas based on Chetty et al. (2020) Figure X. Both panels assume extrapolation to $17,207 for neighborhood mean individual earnings for children with parents at p=10 in the Opportunity Atlas based on Chetty et al. (2020) Figure XIV. The left panel shows the support of the MTO data in light blue and the range of extrapolation in dark blue in terms of the raw neighborhood outcome. The right panel shows the support of the MTO data in light blue and the range of extrapolation in dark blue in terms of the OA ranking of neighborhoods.