Urban Mobility and the Experienced Isolation of Students*

Cody Cook[†] Lindsey Currier[‡] Edward Glaeser[§]

July 2023

Abstract

Cities provide access to stores, public amenities and other people, but that access may provide less benefit for lower-income and younger urbanites who lack money and means of easy mobility. Using detailed GPS location data, we measure the urban mobility and experienced racial and economic isolation of the young and the disadvantaged. We find that students in major metropolitan areas experience more racial and income isolation, spend more time at home, stay closer to home when they do leave, and visit fewer restaurants and retail establishments than adults. Looking across levels of income, students from higher-income families visit more amenities, spend more time outside of the home, and explore more unique locations than low-income students. Combining a number of measures into an index of urban mobility, we find that, conditional on income, urban mobility is positively correlated with home neighborhood characteristics such as distance from the urban core, car ownership, and social capital.

*This project was reviewed and determined exempt by the Stanford Institutional Review Board (protocol IRB-62185). Glaeser acknowledges support from the Star Family Challenge for Promising Scientific Research. We thank Kiran Jain and Alexei Pozdnoukhov at Replica for facilitating access to the data.

[†]Stanford GSB. Email: codycook@stanford.edu.

[‡]Harvard University. Email: lcurrier@harvard.edu.

[§]Harvard University. Email: eglaeser@harvard.edu.

1

Are the benefits of living in a city conditioned by who a person is? Density provides city-dwellers with access to amenities such as commercial centers, public goods, and social infrastructure. Yet proximity may not translate into use, particularly for lower-income youth, who lack funds and cars and whose highly local amenities may differ significantly from those in wealthy neighborhoods. An alternative view is that those with a lower opportunity cost of time are able to travel more and spend more time interacting with their urban environment. In this paper, we use GPS data to test whether students—and lower-income students in particular—seem to take as much advantage of the city as their older and wealthier counterparts.

For over 50 years, social scientists have documented urban residential segregation (Kain, 1968; Taeuber and Taeuber, 1969), and its pernicious effects, particularly for children (Brooks-Gunn et al., 1993; Cutler and Glaeser, 1997; Sampson et al., 2002; Chyn and Katz, 2021; Chyn et al., 2023). More recently, Athey et al. (2021) demonstrate that experiences are not perfectly delineated by place of residence, finding that "experienced isolation" is far lower than residential segregation.[1] Browning et al. (2022) similarly finds that Black youth experience more inter-racial interactions than implied by residential location alone.

Interaction with a diverse set of people is only one potential benefit of urban life. In this paper, we examine a range of outcomes, including racial and income-based experienced isolation, visits to urban amenities, exploration of new places, and distance traveled, using a panel of location data from GPS-enabled devices.[2] We are particularly interested in the lives of younger urbanites, whom recent work has suggested experience less economic upward mobility in denser urban areas (Chetty and Hendren, 2018). At the same time, the past decades have seen a rise in income segregation for

[1] The experienced isolation measure in Athey et al. (2021) builds on the "activity space" literature in sociology (Wong and Shaw, 2011; Shelton et al., 2015; Wang et al., 2018; Beiró et al., 2018). This literature is surveyed in Cagney et al. (2020). Most relevantly, Moro et al. (2021) use GPS data to measure experienced income isolation and explore its connections to device's overall mobility.

[2] Examples of previous applications of GPS location data include mobility during the COVID-19 pandemic (Chang et al., 2020; Allcott et al., 2020; Couture et al., 2021; Chen et al., 2021), waiting times at voting polls (Chen et al., 2019), knowledge spillovers between employees of different firms (Atkin et al., 2020), and demand for amenities (Athey et al., 2021; Miyauchi et al., 2021; Cook, 2023).

households with children, but *not* for childless households (Owens, 2016).

We build a panel of location data from GPS devices and infer three characteristics for each device: household income, race, and student-status. To infer income, we follow Cook (2023) and match each device to its home parcel, then use characteristics of the home (e.g., market value, structure age, location) to predict income. For race, we follow Athey et al. (2021) and use whether or not a device is from a majority non-Hispanic white block group to define "white" and "non-white" devices. Finally, we define "students" as 16-18 year-olds attending school and infer student status from whether the individual's most common weekday location is a high school. For privacy reasons, cellphone record providers remove anyone under 16 years of age. Importantly, this sample contains a mixture of high school students, teachers, and staff, which likely attenuates the differences we measure between students and adults, especially since U.S. teachers are disproportionately white relative to their students.[3]

Nonetheless, we find consistent and significant gaps in the urban mobility and isolation of students versus adults. We also perform several robustness checks, as well as a back-of-the-envelopment bias correction that suggests the impact of teachers is small, in Appendix Sections A and B.

We start by estimating the day-to-day experienced isolation across both race and economic lines. Following the methodology of Athey et al. (2021), we find that students experience more racial and income isolation than adults. Excluding time spent at home, the racial isolation of students is 21% higher than that of adults in the one hundred largest metropolitan areas. The income isolation of students 13% higher than that of adults, driven by the particularly high isolation of high-income students. When we compare the larger and smaller metropolitan areas, we find that the student-adult divide is much starker for the largest metropolitan areas. For example, the racial isolation of students is less than 10% higher than adults in the smallest third of metropolitan areas, but 42% higher in the largest third. One potential explanation for why cities provide benefits for adults (Glaeser and Mare, 2001) but offer less upward mobility for students (Chetty and Hendren, 2018) is that urban children lead more isolated lives than urban adults.

[3] During the 2017-2018 school year, in public schools where the majority of students were Black, 54% of teachers were white (Spiegelman, 2020).

We next turn to broader measures of urban mobility and find a more complex picture. Relative to adults, students spend more time at home and in their neighborhood, stay closer to home when they do leave, and visit fewer restaurants and retail establishments. But students also explore a greater number of unique locations, spend more time in parks and at civil, social, and religious establishments, and spend more time in areas that are richer, more white, less polluted, and have lower crime rates. The differences are often large. For example, students spend nearly 50% more time in their local neighborhood and go to 10-20% fewer restaurants and retail establishments.

The connection between mobility and income is much stronger than the connection between mobility and student-status, with lower-income students appearing far less mobile than wealthier students along every dimension. Student from in the richest quartile of our sample take 57% more visits to entertainment venues, 38% more visits to parks, and experience 54% more total unique locations than students in the lowest income quartile. Higher-income students also spend 3 percentage points less time at home and, when they leave home, travel further afield. These differences attenuate when we control for tract of residence, but even within a tract, higher-income students are more mobile and take advantage of more urban amenities.

Finally, we investigate the correlation between urban mobility and neighborhood characteristics. To simplify our analysis, we aggregate our various measures into a single urban mobility index. Even holding fixed a device's estimated income, urban mobility is rising in median neighborhood income. Urban mobility is lower in places that are more densely populated and closer to the city center, as well as in places with greater transit access. In its current state, public transit in the US does not override other factors that limit urban mobility among the poor. Urban mobility is also higher in areas with greater social capital, as measured by Chetty et al. (2022a,b), suggesting that places with residents who are more connected to urban assets in the physical world also have greater connection across socioeconomic statuses in the virtual world.

While this work cannot speak to long-run costs of urban isolation, we have documented that students appear to live more isolated lives than their adult counterparts, especially in the largest cities. Moreover, lower-income urbanites appear to make far less use of urban amenities, which are themselves core benefits of urban life (Couture et al., 2019; Couture and Handbury, 2023). Wealth appears to be a complement, rather than a substitute, for enjoying the pleasures of urban life.

2 Methods

2.1 GPS Mobility Sample

Location data. Our primary dataset is a panel of GPS locations for a sample of cellphones from 2019. Access to the data is provided by Replica, an urban data platform. For each device, we observe a unique identifier and a sequence of 'stays' at various locations. Each stay includes the geographic coordinates, entry time, and exit time. We have no direct information about the device's user, so must infer whether a device is a student and any demographics, such as race and income, using the location histories of the device. Devices are not uniformly sampled across space, so we use sample weights based on a device's home location to correct for unevenness in sampling. We provided additional details on the data construction in Appendix A. To assuage concerns on representativeness of the GPS sample, we document in Section 3.3 that our sample replicates time use patterns for students in the American Time Use Survey (ATUS) and in Section 3.4 that we replicate travel patterns for youth in the National Household Transportation Survey (NHTS). Both surveys have very small youth samples relative to our data. Our data is also able to add significant detail on mobility patterns.

Identifying students. For each device-quarter in the data, we identify "home" as a device's most common overnight location and 'work' as a device's most common daytime non-home location. We exclude devices for which we observe insufficient data to identify a work location.[4] To label a device as a student, we match "work" locations to the geographic parcels of public schools. We identify the locations of high schools using data from the National Center for Education Statistics (NCES); Appendix Section A.2 describes the matching process. We include only high schools, as our GPS

[4]The vast majority of these devices have insufficient coverage in the data to confidently identify a work location; the median excluded device had only 9 observed stays in all of 2019. Other devices are either unemployed or employed in occupations without a static work location, such as postal workers or taxi drivers. Among devices that had an identified home location, those that were excluded from the sample live in tracts that are slightly more non-Hispanic white (68.8% vs 67.1%) and tracts with higher median household income ($79,653 vs. $75,397).

data are meant to exclude individuals under 16 years old. Appendix Figure A.1 shows that our counts of students at a school are highly correlated with the enrollment reported in the NCES.

Our method of identifying students invariably captures teachers and staff. This adds noise and, to the extent adults working at schools are similar to other adults, biases our measures towards finding no differences between students and adults. However, there does appear to be signal in the classification – for example, when examining the types of establishments they visit, we find that "students" go to far fewer bars and beer/liquor stores than adults. We also do a back-of-the-envelopment bias correction for some results that suggests the impact of teachers is small, in Appendix Section B.

Inferring income and race. We match each device to its home parcel and use parcel level estimates of household income from Cook (2023). The estimates are computed in two stages. First, Cook (2023) matches each parcel to data on housing characteristics from Corelogic, including building age, type of building (e.g., single versus multi-family), and a prediction of its current market value. Based on the relationship between housing characteristics and household income in the American Community Survey (ACS) Public Microdata Sample (PUMS), an initial estimate of household income is formed for each device. One limitation of the ACS PUMS data is that only the Public Use Microdata Area (PUMA) of each home is observed, rather than more granular geographic identifiers such as home block group. To improve estimates using the more granular geographic data available in the GPS data, a second step is performed in which each estimate is updated using an empirical Bayes procedure based on the distribution of household income within the device's home block group.[5]

Following Athey et al. (2021), we classify devices as either a "white device" (WD) or "non-white device" (NWD) based on whether or not their home block group is majority white (non-Hispanic).[6] Using only these two broad classifications of race will mask substantial heterogeneity by race and ethnicity; however, because of the measurement error inherent in using home geography to impute race, we do not attempt to further divide devices based on ethnicity or more granular races. Even

[5]For the full details on the income imputation, refer to Appendix A.2 of Cook (2023).

[6]We use 2019 ACS block groups rather than 2010 Census blocks as the 2010 Census is now significantly out-of-date. The results are similar if we instead use 2010 blocks to identify white and non-white devices.

with this broad classification of race, our measure will frequently mis-classify white individuals as non-white devices and vice versa; the average home block group for WDs is 78.5% white, while the average home block group for NWDs is 21.0% white.

Final sample. To focus on urban environments, we look only at devices living within the one hundred most populous metropolitan Core-Based Statistical Areas (CBSAs). The smallest CBSA that makes this cut is Spokane-Spokane Valley in Washington. The final sample includes 321,955 students and 9.1 million adults.

2.2 Measuring Experienced Isolation and Exposure to Diversity

To estimate experienced income and racial isolation, we follow the methodology introduced in Athey et al. (2021), which we describe in detail in Appendix Section A.4. Experienced income isolation in an MSA measures the difference between the share of lower-income residents' interactions with higher-income residents and the share of higher-income residents interactions that are with other higher-income residents. Using GPS data, we define each device's exposure to members of each race and income group based on the time the spend in each location and the race and income distributions of other devices in that location. An income isolation measure of 0.5 would imply that lower-income devices interact with fifty percentage points fewer higher-income devices than do other higher-income devices.

These measures capture isolation in space, which may not map one-to-one with social segregation nor meaningful interactions between devices sharing a geographic location (White, 1983). For example, students in diverse schools may occupy similar spaces, but still form social cliques split along racial lines. Similarly, higher-income devices may visit establishments where they share a space with lower-income workers but may not truly interact with these other individuals.

We compute experienced isolation for both race and income. For racial isolation, we follow Athey et al. (2021) and use the WD and NWD designations based on a device's home neighborhood block group demographics. To be precise, these results give us the isolation between people from predominantly white neighborhoods to those from predominantly non-white neighborhoods. For income isolation, we use the individual income estimate outlined above, and split devices based on whether their estimated income is above the median income in the CBSA ("higher-income") or below the median income ("lower-income").

Experienced isolation is a population level statistic based on the interactions of all devices within an MSA, so we introduce a complementary individual level measure which we call "exposure to diversity." Exposure to diversity measures how much a given device's interactions are with devices of the *opposite* group. For example, a WD's exposure to diversity is the share of their interactions—proxied by the places they visit—with NWDs.

2.3 Defining Urban Mobility

In addition to experienced isolation, we look at four other categories of urban mobility: 1) time spent at primary locations (home, work/school, and in the neighborhood); 2) "roaming ranges", or how far devices tend to travel from home, and the number of unique places they visit; 3) use of amenities such as restaurants and shops; and 4) characteristics of tracts visited.

To facilitate an analysis of the connection between neighborhoods and urban mobility, we also collapse our mobility measures down to a single mobility index for each device. Specifically, we first standardize each individual measure using the cross-device mean and standard deviations of the variables, then use Principal Component Analysis (PCA) to collapse the measures into a single measure.[7] The individual measures included are the time at home, work/school and in the neighborhood, number of visit to each category of amenities, number of unique locations visited, and the average miles traveled from home. Finally, for interpretability we transform the first principal component into a z-score, which we use as our final index of urban mobility. Appendix Table B2 shows the correlation between this index and each of the component parts.

2.4 Data Availability

The data that support the findings of this study are available from Replica, but restrictions apply to the availability of the data; access was provided under a Data Use Agreement that prohibits sharing the underlying data.

[7] To put similar weight on amenities, time at locations, and roaming ranges, we first use PCA within each category of mobility and then use PCA across the components estimate for each category.

3 Results

We start by examining students' urban mobility in comparison to adults. We document experienced isolation by race and income and then turn to broader measures of travel, urban amenity consumption, and time use. We then look within the student population to see who benefits from cities and dense urban areas. We focus in particular on household income, both because we hypothesize that income plays a critical role in urban mobility, and because our data allows us uniquely to explore differences in income while holding fixed narrow neighborhoods of residence.[8] In the last section, we look at correlates with the component of urban mobility that is explained by neighborhood.

3.1 Students and Adult Differences in Experienced Isolation

We find that students experience 21% greater racial isolation and 13% greater income isolation than adults. The first row in Table 1 shows that racial isolation outside the home is 0.38 for students, compared to 0.31 for adults, which is similar to the the difference between the 25th and 75th percentile of cities when ranked by overall experienced isolation. Similarly, experienced income isolation outside the home is larger for students (0.26) than for adults (0.23). To facilitate interpretation, imagine everyone always interacts with exactly one other person. Since we have split income groups at the median, an isolation measure of 0.26 means that a higher income student interacts with another higher income student 63% of the time and with a lower-income student 37% of the time (for a difference of 26 percentage points). The levels of experienced isolation rise when we include time within the home, but the student-adult gap remains similar.[9] In Appendix Section B.5, we do a simple back-of-the-envelope calculation to account for bias from including teachers in the "student" sample and find the student estimates change by less than one percentage point.

[8] Another reason to focus primarily on income—rather than race or ethnicity—is that our measure of income based on a device's home parcel is less susceptible to measurement error than our measure of race. Other work in this literature has instead focused more prominently on differences in mobility and neighborhood exposure by race (Levy et al., 2020; Candipan et al., 2021; Brazil, 2022; Xu, 2022).

[9] The higher magnitude of isolation including time at home partially mechanical – recall that for racial segregation, complete isolation at home occurs by definition, as the racial groups are coded at the block group level. While our income measure is more precise than our imputation of race because it uses house level data, because we use the block group income distribution to update our estimates of income the isolation measures including time at home will still be mechanically higher.

Since experienced isolation is defined at the population level, we shift to the individual level using "exposure to diversity," defined in Section 2.2, which captures each individual's exposure to the other group. Outside of the home, adults are in settings where on average 25% of others are from the other racial group and 36% of others are from the other income group (Table 1, panel b). The average exposure to racial diversity is 4.7 percentage points lower for students, which reduces to 3.4 percentage points once controlling for tract of residence. We can split these results by imputed race. We find the typical WD is in a setting where 19% of devices are NWD, while the typical NWD is in a setting where 35% of devices are WD. For students, these numbers are again significantly lower, at 15% and 32%. The result of greater exposure to racial diversity for Black devices parallels the findings in Browning et al. (2022). However, while Browning et al. (2022) emphasizes that difference, our focus is on the fact that all students are more isolated than their adult counterparts. Similarly, students' exposure to income diversity is three percentage points lower than adults' at baseline; perhaps surprisingly, this difference is driven primarily by higher-income students, who are significantly more isolated by income than their adult counterparts.

These estimates suggest that, for students, experienced urban isolation may fall along racial lines more than along lines of income. When we control for home tract, average exposure to income diversity is two percentage points lower for students than adults. The persistence of students' lower levels of exposure to diversity by both race and income when controlling for home tract refutes the notion that the increased isolation of students is mainly driven by the neighborhoods students live in relative to working adults.

3.2 Experienced Isolation by City Size

We now compare the experienced isolation of students and adults in cities of difference sizes. Our main sample pools the one hundred most populous US cities, while Table 2 splits this sample based on whether individuals live in the largest, middle, or smallest third of our sample of cities. The first two columns in Table 2 examine experienced racial isolation for students and adults. Students in our biggest cities experience 41% more racial isolation than students in our smallest cities, while adults in our biggest cities experience just 9% more racial isolation than adults in our

Table 1: Experienced isolation of students and adults

Panel a) Overall experienced isolation (EI)	Aggregate	Students	Adults
Outside of home			
Racial isolation	0.3134	0.3763	0.3110
Income isolation	0.2338	0.2630	0.2320
Including time at home			
Racial isolation	0.7092	0.7597	0.7075
Income isolation	0.6154	0.6563	0.6136

Panel b) Individual exposure to diversity	Average	Coefficient on isStudent (Home CBSA controls)	Coefficient on isStudent (Home tract controls)
Racial diversity			
Exposure to racial diversity	0.2519	-0.0468	-0.0337
		(0.0005)	(0.0003)
Exposure to NWD by WD	0.1943	-0.045	-0.0356
		(0.0004)	(0.0003)
Exposure to WD by NWD	0.3484	-0.0282	-0.0311
		(0.0009)	(0.0006)
Income diversity			
Exposure to income diversity	0.3606	-0.0303	-0.0198
		(0.0005)	(0.0004)
Exposure to L by H	0.3091	-0.0458	-0.0388
		(0.0005)	(0.0004)
Exposure to H by L	0.4331	0.0136	0.0145
		(0.0009)	(0.0005)

Note: This table documents overall experienced isolation measures, computed as a weighted average of MSA level measures, with weights corresponding to the MSA population. 'At home' is defined as within 50 meters of home location. Panel b) runs individual-quarter regressions of exposure to diversity on whether the device is a student with either home MSA or home Census tract fixed effects. Exposure to diversity excludes time spent at home. H and L refer to higher and lower-income, respectively, while WD and NWD refer to white device and non-white device.

smallest cities. Student racial isolation is 42% higher than adult racial segregation in the biggest cities, while student isolation is less than 10% high than adult segregation in the smallest cities. A similar pattern emerges for income isolation. In large cities, income isolation in 32 percent higher for students than for adults, while in small income segregation is only 12% higher for students than for adults. Both measures suggests that urban size increases the isolation of our student sample relative to the population as a whole.

Table 2: Experienced isolation and urban mobility by city size

	Experienced racial isolation		Experienced income isolation	
City size	Students	Adults	Students	Adults
Big – rank 1-33	0.384	0.271	0.314	0.237
Medium – rank 34-66	0.327	0.263	0.294	0.251
Small – rank 67-100	0.272	0.248	0.278	0.249

Note: This table documents average differences in experienced isolation and urban mobility across cities of difference sizes. Cities are ranked based on their 2010 population. Each statistic is computed at the city-level, then averaged across all cities within a given size group.

3.3 Student and Adult Differences in Overall Urban Mobility

Urban neighborhoods provide not just interactions with people, but also interactions with the geographic and economic amenities of cities. We now look at the additional mobility outcomes described in Section 2.3. As in Table 1, Table 3 reports results controlling first for MSA and then for home census tract.

Students and adults differ in the amount of time spent at home, work/school, and in the neighborhood (Table 3, panel a). The average device spends 66% of time at home, 16% of time at work/school and 5% of time in their home neighborhood.[10] Students spend less time at school than adults spend at work, and more time at home and in the neighborhood than adults. In particular, controlling for home tract, students spend nearly 50% more time in their surrounding neighborhood.

[10] We compare this with the time-use results in the American Time Use Survey (ATUS). Averaging across the 2009 to 2019 period to increase sample size (to around 4,000 high school students), we find high school students in the ATUS report spending 66% of their time at home and 18% of their time at school.

Table 3: Urban mobility of students and adults

	Average (not logged)	Coef. on isStudent (Home CBSA FEs)	Coef. on isStudent (Home tract FEs)
Panel a) Time at primary locations			
Frac. of time at home	0.6573	0.0243	0.0203
		(0.0005)	(0.0004)
Frac. of time at work/school	0.161	-0.0165	-0.0127
		(0.0003)	(0.0002)
Frac. time in neighborhood (excl. home)	0.0488	0.0215	0.0231
		(0.0003)	(0.0002)
Panel b) Roaming ranges			
Log avg miles from home	7.5021	-0.3543	-0.3898
		(0.0019)	(0.0014)
Log # unique locations (geohash7)	42.8904	0.0528	0.0138
		(0.002)	(0.0017)
Panel c) Visits to amenities			
Log # restaurant visits	1.5151	-0.0982	-0.1137
		(0.0023)	(0.0019)
Log # retail visits	1.9806	-0.1948	-0.2181
		(0.0025)	(0.002)
Log # park visits	1.1746	0.0778	0.0577
		(0.0022)	(0.0018)
Log # entertainment visits	1.399	0.0672	0.0318
		(0.0022)	(0.0019)
Log # civil, social, religion visits	0.2523	0.0326	0.025
		(0.0011)	(0.001)
Panel d) Characteristics of tracts visited			
Log median HH income	76386	0.0829	0.0381
		(0.0008)	(0.0006)
Frac. college graduate	0.3924	0.0105	0.0017
		(0.0003)	(0.0002)
Frac. White alone	0.5733	0.0285	0.0071
		(0.0004)	(0.0002)
Air quality (PM25)	8.8084	-0.026	0.0079
		(0.0016)	(0.0012)
Log crimes per sq. mi.	9.711	-0.1973	-0.0709
(Chicago & Los Angeles, 2010-2018)		(0.0159)	(0.0085)

Note: The table documents coefficients from regressions of mobility metrics on whether the device is a student with fixed effects for either the device's home MSA or Census tract. Miles from home is the average distance of stays outside of the home on days the device stayed within 50 miles of home, weighted by the stay duration. "At home/work/school" is defined as within 50 meters of the location's coordinates, while 'in the neighborhood' is defined as within 1 mile of home. We use data on the average estimated tract level air pollution in 2019 from the Environmental Protection Agency (EPA). For the crime outcomes, we subset to just devices that live within those city boundaries and measure crime as the sum of all crimes reported between 2010-2018 in a tract. The characteristics of tracts visited results exclude time spent at home or work/school. To handle zeros, we use the inverse hyperbolic sine instead of the logarithm. Both the averages and regressions use the device weights; for tract outcomes, the regressions are also weighted by time spent in the tract. Standard errors are clustered at the device-quarter level.

When outside the home, students and adults differ in how far they "roam" (Table 3, panel b). The average distance from home is 7.5 miles across all devices; this distance is about 35% smaller for students, a gap that persists even when comparing students and adults living in the same tract. Although students travel shorter distances, their travel patterns are less routine. The average student visits 5% more unique locations—defined by 500ft × 500ft squares—than the average adult, although the much of this difference in exploration can be explained by differences in where students live.

The somewhat modest connection between student-status and unique locations masks a large shift in the nature of the locations visited (Table 3, panel c). Controlling for home tract, students visit 11% fewer retail shops and 22% fewer restaurants, but 3% more entertainment venues and 6% more parks. Students also visit more civic, religious, and social venues, but the overall number of such visits among both students and adults are small.[11]

Finally, when outside of home and work/school, students visit Census tracts that are on average richer, better educated, have a higher white population share, less pollution, and less crime (Table 3, panel d). To evaluate this, we compute the time a device spends in each Census tract and regress characteristics of the tract on student-status, weighting by the time spent in the tract. While the gaps between average education, pollution, and white share of the population of tracts visited by students compared to adults are quite small, the gap in the crime rate is large. Our crime data are limited to Chicago and Los Angeles due to data availability, but in those cities, students visit tracts with 20% fewer crimes per square mile. The gap in crime rate falls to 7% when controlling for home tract, suggesting that students live in lower-crime neighborhoods on average, perhaps because childless adults may take more locational risks when deciding where to live than adults with children.

Overall, these results suggest that while students experience more isolation than adults—especially along racial lines—in other ways the overall experience of urban youth differs from that of urban adults predictably. Students generally go to somewhat nicer neighborhoods and are exposed to slightly less crime. They go to fewer restaurants and shops, but more parks and entertainment

[11] These visits exclude those to any location identified as a device's workplace, so the differences are not driven by adults being more likely to work at restaurants and shops.

venues. Yet this overall picture of teenage life in cities masks considerable heterogeneity by levels of income among the student population.

3.4 Household Income and the Urban Mobility of Students

We now look at the relationship between urban mobility and household income and home neighborhood characteristics. We find that both play a role – higher-income students have greater urban mobility across a range of measures holding neighborhood fixed, but also, even for students with similar income, neighborhood characteristics matter. We continue to focus on students in the main text, but Appendix Figure C.4 reproduce these results for the adult population.

To document differences by income, we divide devices into quartiles of predicted household income and compare a range of mobility measures across each quartile.[12] For each measure, we present a versions controlling only for a device's home MSA, and a version controlling for home tract.

First, relative to the baselines from Table 3, students in the highest income quartile spend about 5% less time at home and 18% more time in the neighborhood than students in the lowest income quartile (Figure 1, panels a and b). The effect declines substantially when we control for home tract, suggesting a large part of the reason lower-income students spend more time at home and less time in their neighborhood is due to the neighborhood itself.

Second, richer students both visit more unique locations and tend to travel further when they leave the house (Figure 1, panels c and d). The relationship between income and number of locations visited is strong and monotonic. Students in the highest income quartile visit 54% more unique locations than students in the lowest income quartile. The link between distance from home and income is non-monotonic, although the bottom quartile of income stays the closest to home. When controlling for home tract, the coefficients drop by about half, but the relationship between income and number of unique places visited remains strong.

[12] We believe that measurement error in income would likely to attenuate the estimated relationship between income and mobility; however, families who spend more on their housing, relative to permanent income, may be more interested in consuming urban amenities.

Third, there are stark differences by income in students' consumption of various local amenities such as restaurants, shops, and parks (Figure 1, panels e and f). The strongest relationship is between income and visits to entertainment venues – controlling for home MSA, students from the top income quartile visit 57% more entertainment locations than those in the bottom income quartile. The impact of income on park and restaurant visits is smaller, but still large. Students in the highest-income quartile take 35% more visits to restaurants and 38% more visits to parks than students in the lowest-income quartile. On average, the gap between the top and bottom quartiles attenuates by 45% when controlling for home tract, but again remain large.

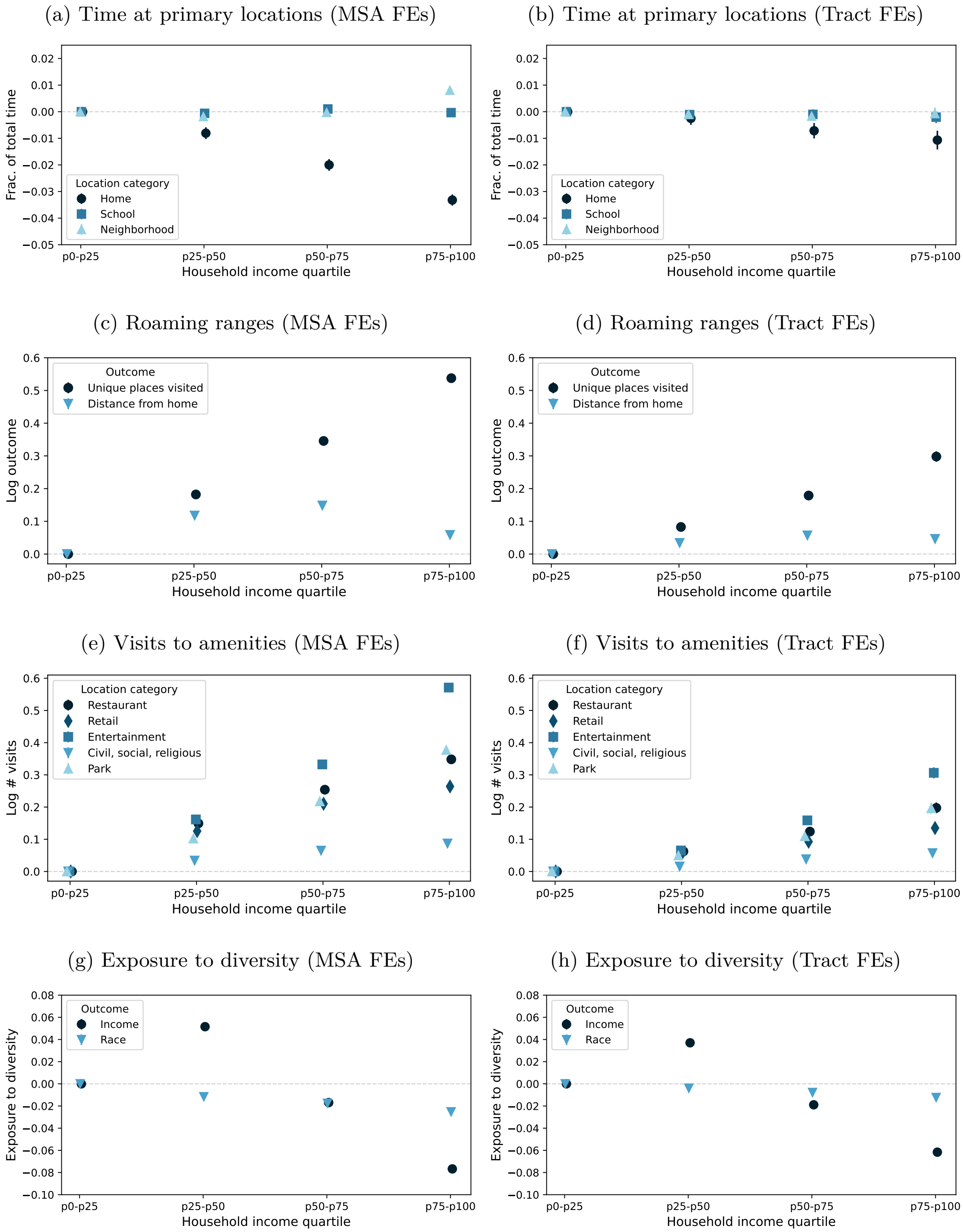
Finally, there is a weak relationship between household income and exposure to both income and racial diversity (Figure 1, panels g and h). Exposure to income diversity is unsurprisingly lowest for the middle of the distribution, as people with incomes slightly above the median income are likely to interact with people whose incomes are below the median. As we move out to the top and bottom quartiles, however, we see some asymmetries: students from the highest-income households are more isolated from below median income households than the lowest-income households are from above median income households, echoing our results from Table 1. The connection between income exposure and income remains large when we control for home tract, suggesting this trend is not a function of where families live in a metro area, but a more fundamental characteristic of household travel patterns. Racial exposure to diversity declines with income, though the effects are small and, once we control for tract, the relationship becomes economically insignificant.

As a validation check, we conduct a similar exercise using the 2017 National Household Transportation Survey (NHTS) in Appendix Table B4. Although, the NHTS does not allow us to look at detailed destination types or to control for neighborhood differences, we can look at some trends by income. Similar to our results, we find that both travel for amenities and time away from the home is rising in household income.

3.5 Correlation of Urban Mobility with Neighborhood Characteristics

How does the urban mobility of students relate to the characteristics of their home neighborhood? To facilitate an analysis of the connection between neighborhoods and urban mobility, we

Figure 1: Urban mobility of students by income



Note: This figure plots the relationship between student income and various measures of urban mobility. Each point is a coefficient from a regression of a device-quarter level mobility measure (e.g., fraction time at home) on indicators for whether the quartile of a devices predicted income and MSA fixed effects. 95% confidence intervals are represented by bars around each point (although are frequently covered by the point itself).

collapse our mobility measures down to a single index for each device, defined in Section 2.3. The first panel of Figure 2 shows that, as shown above for the component parts, urban mobility increases steeply with the predicted household income of the device. The linear coefficient when the mobility index is regressed on log predicted income is 0.41, suggesting that a twofold increase in predicted income increases urban mobility by nearly half of a standard deviation. In all subsequent panels, we show results with controls for just MSA as well as with controls for the log of predicted income, to isolate the correlation that persists beyond the impact of household income. Appendix Table B3 reports the coefficients from analogous linear regressions, which we report in the main text to summarize the graphs.

Even controlling for household income, there is a persistent positive relationship between urban mobility and neighborhood median household income (panel b). Without controls, the coefficient on neighborhood income is 0.40, which is almost as large in magnitude as the coefficient on individual income. When we include both variables in a regression, the coefficient on neighborhood income attenuates to 0.17; even for devices of a similar income level, those in higher income neighborhoods exhibit greater urban mobility.

Urban mobility is decreasing in measures of "urbanity," such as population density and proximity to city hall (Figure 2, panels c and d). The linear coefficient for urban mobility regressed on log population density is -0.09, which falls in magnitude to -0.07 when we control for individual income. While denser areas have more nearby amenities to visit, devices living in these areas have lower urban mobility on average. Similarly, mobility rises with the log distance from city hall. The coefficient without income controls is 0.15, which falls to 0.10 when we control for income. One obvious reason students in dense neighborhoods might have lower mobility is that they might be less likely to have access to a car. Panel e) looks at the share of households in the tract that own a car. There is a strong relationship between urban mobility and car ownership: for each ten percentage point increase in car ownership (approximately one standard deviation), urban mobility increases by 0.16 standard deviations. In contrast, urban mobility is slightly lower in neighborhoods that the Department of Housing and Urban Development (HUD) identify as having greater transit access. This relationship between urban mobility and transit access, however, is primarily driven by population density; once controlling for population density, there is no meaningful relationship between urban mobility and transit access.

Finally, urban mobility is positively correlated with measures of neighborhood social capital as measured by connections on social media (Figure 2, panels g and h). Panel g) looks at the link between geographic mobility and economic connectedness, which is a measure of Facebook connections between lower and higher-income individuals (Chetty et al., 2022a). Specifically, it is the average share of above-median socioeconomic status friends among below-median socioeconomic status residents of the zip code. The linear coefficient is 0.76, which falls to 0.32 when we control for individual income. Panel h) shows a weaker positive link between urban mobility and how clustered social connections are for neighborhood residents ("network clustering"), another measure produced by Chetty et al. (2022a) that captures the rate at which two friends of a resident are also friends with each other. The relationship between mobility and network clustering is attenuated when controlling for population density, but the relationship with economic connectedness persists. This finding is consistent with the hypothesis that virtual connections rely on physical ones; links between different groups happen when people traverse their city, such that areas with higher urban mobility generate more bridging social connections. Of course, it is equally possible that causality goes the other way, and that greater social capital leads to greater urban mobility, as our findings are only correlations.

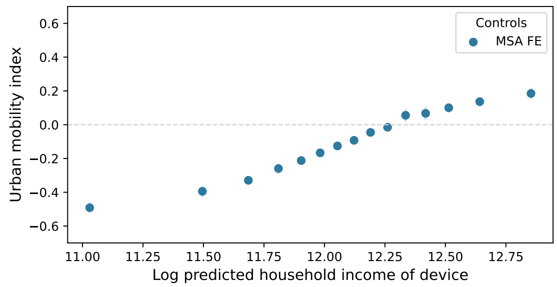
4 Conclusion

Students experience greater income and racial isolation than adults, and this gap is much larger in the biggest metropolitan areas. Students also spend more time at home and in their neighborhood, stay closer when they leave the home, and lead less routinized lives, visiting more unique locations.

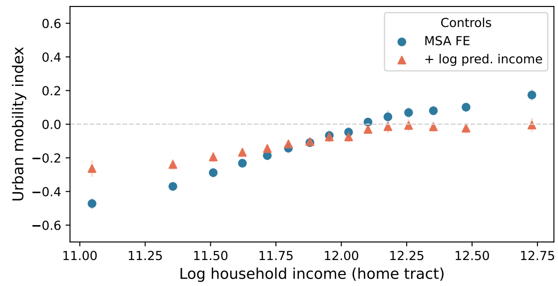
Differences in urban mobility are even larger *within* the population of students across different levels of income. Higher-income students are much more likely to visit every form of local amenity, explore more unique locations, spend less time at home, and roam further from home. In each case, the differences attenuate when comparing students who live within the same neighborhood,

Figure 2: Correlates of urban mobility

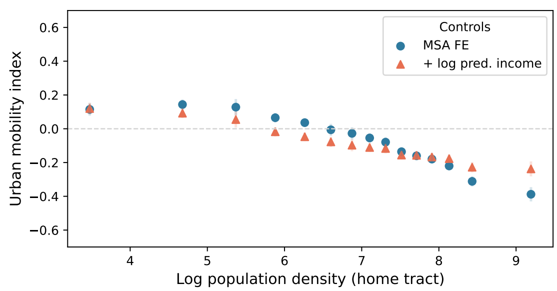
(a) Predicted household income (device level))



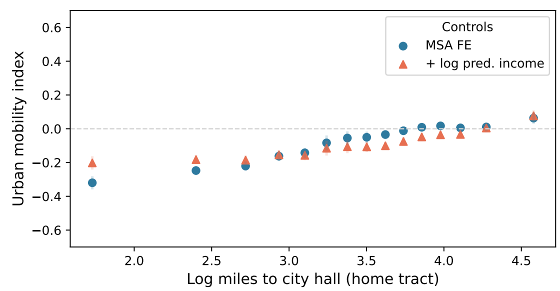
(b) Median household income (home tract)



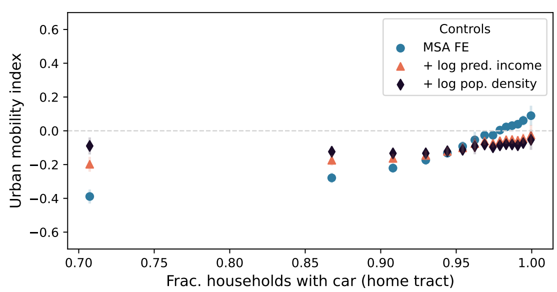
(c) Population density



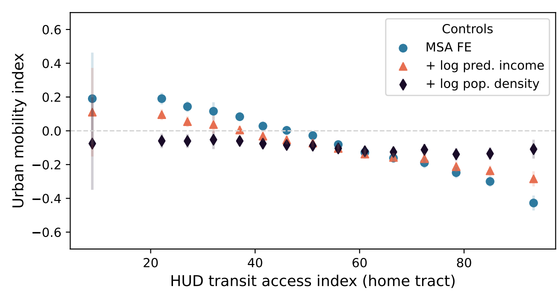
(d) Distance to city hall



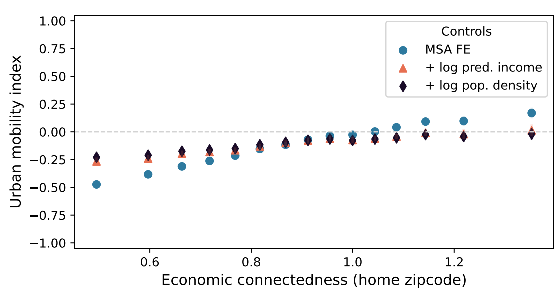
(e) Frac. households with car



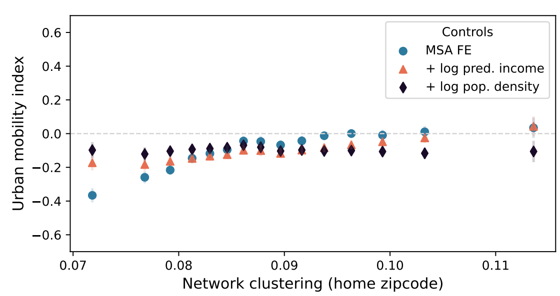
(f) Index of transit access



(g) Economic connectedness



(h) Network clustering



Note: This figure plots the relationship between a device-quarter level index of urban mobility and various device and home neighborhood characteristics using a series of binscatters, controlling for MSA fixed effects and, for some, the log of a device's predicted income. Transit access is based on an index published by the Department of Housing and Urban Development (HUD). Economic connectedness and network clustering are measures of social capital based on the Facebook social graph (Chetty et al., 2022a). Economic connectedness measures friendships across income groups while network clustering measures how likely it is that any two friends of a given person are themselves friends on Facebook. 95% confidence intervals are represented by bars around each point (although are frequently covered by the point itself).

but often remain large. On average, home neighborhood can explain about half of the gap between the mobility differences of higher versus lower-income students.

Urban mobility is correlated with a range of neighborhood characteristics, even controlling for a device's income. Areas that have higher car ownership, are less dense, and higher neighborhood income all have higher levels of urban mobility. Urban mobility is also higher for students living in tracts that have greater social capital, perhaps because physical connection increases social capital.

This work highlights a central paradox of urban America. Lower-income youth living in urban areas, where amenities and public goods are dense, appear to be getting the least out of urban life. Income seems to condition the benefits of urban living. We hope that future work will help us to understand why lower-income residents seem to get less out of cities and to identify the long-run consequences of reduced urban mobility.

References

Allcott, Hunt, Levi Boxell, Jacob Conway, Billy Ferguson, Matthew Gentzkow, and Benny Goldman, "What Explains Temporal and Geographic Variation in the Early US Coronavirus Pandemic?," Technical Report w27965, National Bureau of Economic Research, Cambridge, MA October 2020.

Athey, Susan, Billy Ferguson, Matthew Gentzkow, and Tobias Schmidt, "Experienced Segregation," Technical Report w27572, National Bureau of Economic Research, Cambridge, MA July 2020.

_ , _ , _ , and _ , "Estimating experienced racial segregation in US cities using large-scale GPS data," *Proceedings of the National Academy of Sciences*, 2021, *118* (46).

Atkin, David, Keith Chen, and Anton Popov, "The Returns to Serendipity: Knowledge Spillovers in Silicon Valley," 2020, p. 49.

Beiró, Mariano G, Loreto Bravo, Diego Caro, Ciro Cattuto, Leo Ferres, and Eduardo Graells-Garrido, "Shopping mall attraction and social mixing at a city scale," *EPJ Data Science*, 2018, *7*, 1–21.

Birant, Derya and Alp Kut, "ST-DBSCAN: An algorithm for clustering spatial–temporal data," *Data & Knowledge Engineering*, January 2007, *60* (1), 208–221.

Brazil, Noli, "Environmental inequality in the neighborhood networks of urban mobility in US cities," *Proceedings of the National Academy of Sciences*, 2022, *119* (17), e2117776119.

Brooks-Gunn, Jeanne, Greg J Duncan, Pamela Kato Klebanov, and Naomi Sealand, "Do neighborhoods influence child and adolescent development?," *American journal of sociology*, 1993, *99* (2), 353–395.

Browning, Christopher R., Nicolo P. Pinchak, Bethany Boettner, Catherine A. Calder, and Jake Tarrence, "Geographic Isolation, Compelled Mobility, and Everyday Exposure to Neighborhood Racial Composition among Urban Youth1," *https://doi.org/10.1086/721666*, 11 2022, *128*, 914–961.

Cagney, Kathleen A, Erin York Cornwell, Alyssa W Goldman, and Liang Cai, "Urban mobility and activity space," *Annual Review of Sociology*, 2020, *46*, 623–648.

Candipan, Jennifer, Nolan Edward Phillips, Robert J Sampson, and Mario Small, "From residence to movement: The nature of racial segregation in everyday urban mobility," *Urban Studies*, 2021, *58* (15), 3095–3117.

Chang, Serina, Emma Pierson, Pang Wei Koh, Jaline Gerardin, Beth Redbird, David Grusky, and Jure Leskovec, "Mobility network models of COVID-19 explain inequities and inform reopening," *Nature*, November 2020.

Chen, M Keith, Judith A Chevalier, and Elisa F Long, "Nursing home staff networks and COVID-19," *Proceedings of the National Academy of Sciences*, 2021, *118* (1).

_ **, Kareem Haggag, Devin G Pope, and Ryne Rohla**, "Racial disparities in voting wait times: evidence from smartphone data," *The Review of Economics and Statistics*, 2019, pp. 1–27.

Chetty, Raj and Nathaniel Hendren, "The Impacts of Neighborhoods on Intergenerational Mobility I: Childhood Exposure Effects*," *The Quarterly Journal of Economics*, August 2018, *133* (3), 1107–1162.

_ **, Matthew O. Jackson, Theresa Kuchler, Johannes Stroebel, Nathaniel Hendren, Robert B. Fluegge, Sara Gong, Federico Gonzalez, Armelle Grondin, Matthew Jacob, Drew Johnston, Martin Koenen, Eduardo Laguna-Muggenburg, Florian Mudekereza, Tom Rutter, Nicolaj Thor, Wilbur Townsend, Ruby Zhang, Mike Bailey, Pablo Barberá, Monica Bhole, and Nils Wernerfelt**, "Social capital I: measurement and associations with economic mobility," *Nature*, August 2022.

_ **, _ , and _** , "Social capital II: determinants of economic connectedness," *Nature*, August 2022.

Chyn, Eric and Lawrence F. Katz, "Neighborhoods Matter: Assessing the Evidence for Place Effects," *Journal of Economic Perspectives*, 9 2021, *35*, 197–222.

_ **, Robert Collinson, and Danielle Sandler**, "The Long-Run Effects of Residential Racial Desegregation Programs: Evidence from Gautreaux," 2023, *Working paper.*

Cook, Cody, "Heterogeneous Preferences for Neighborhood Amenities: Evidence from GPS data," *Available at SSRN 4212524*, 2023.

Couture, Victor and Jessie Handbury, "Neighborhood Change, Gentrification, and the Urbanization of College Graduates," *Journal of Economic Perspectives*, 2023.

_ **, Cecile Gaubert, Jessie Handbury, and Erik Hurst**, "Income Growth and the Distributional Effects of Urban Spatial Sorting," Technical Report w26142, National Bureau of Economic Research, Cambridge, MA August 2019.

_ **, Jonathan I Dingel, Allison Green, Jessie Handbury, and Kevin R Williams**, "JUE Insight: Measuring movement and social contact with smartphone data: a real-time application to COVID-19," *Journal of Urban Economics*, 2021, p. 103328.

Cutler, David M. and Edward L. Glaeser, "Are Ghettos Good or Bad?*," *The Quarterly Journal of Economics*, August 1997, *112* (3), 827–872.

Glaeser, Edward L and David C Mare, "Cities and skills," *Journal of labor economics*, 2001, *19* (2), 316–342.

Kain, John F., "Housing Segregation, Negro Employment, and Metropolitan Decentralization*," *The Quarterly Journal of Economics*, May 1968, *82* (2), 175–197.

Levy, Brian L, Nolan E Phillips, and Robert J Sampson, "Triple disadvantage: neighborhood networks of everyday urban mobility and violence in US cities," *American Sociological Review*, 2020, *85* (6), 925–956.

Logan, John R, Elisabeta Minca, and Sinem Adar, "The geography of inequality: Why separate means unequal in American public schools," *Sociology of education*, 2012, *85* (3), 287–301.

_ **, Weiwei Zhang, and Deirdre Oakley**, "Court orders, white flight, and school district segregation, 1970–2010," *Social Forces*, 2017, *95* (3), 1049–1075.

Miyauchi, Yuhei, Kentaro Nakajima, and Stephen J Redding, "Consumption access and agglomeration: evidence from smartphone data," Technical Report, National Bureau of Economic Research 2021.

Moro, Esteban, Dan Calacci, Xiaowen Dong, and Alex Pentland, "Mobility patterns are associated with experienced income segregation in large US cities," *Nature communications*, 2021, *12* (1), 1–10.

Owens, Ann, "Inequality in Children's Contexts," *https://doi.org/10.1177/0003122416642430*, 4 2016, *81*, 549–574.

Sampson, Robert J., Jeffrey D. Morenoff, and Thomas Gannon-Rowley, "Assessing "neighborhood effects": Social processes and new directions in research," *Annual review of sociology*, 2002, *28* (1), 443–478. Publisher: Annual Reviews 4139 El Camino Way, PO Box 10139, Palo Alto, CA 94303-0139, USA.

Shelton, Taylor, Ate Poorthuis, and Matthew Zook, "Social media and the city: Rethinking urban socio-spatial inequality using user-generated geographic information," *Landscape and urban planning*, 2015, *142*, 198–211.

Spiegelman, Maura, "Race and Ethnicity of Public School Teachers and Their Students," 2020.

Taeuber, Karl E. and Alma F. Taeuber, *Negroes in Cities: Residential Segregation and Neighborhood Change*, Atheneum, 1969. Google-Books-ID: x3MqAAAAYAAJ.

Wang, Qi, Nolan Edward Phillips, Mario L Small, and Robert J Sampson, "Urban mobility and neighborhood isolation in America's 50 largest cities," *Proceedings of the National Academy of Sciences*, 2018, *115* (30), 7735–7740.

White, Michael J, "The measurement of spatial segregation," *American journal of sociology*, 1983, *88* (5), 1008–1018.

Wong, David WS and Shih-Lung Shaw, "Measuring segregation: An activity space approach," *Journal of geographical systems*, 2011, *13* (2), 127–145.

Xu, Wenfei, "The contingency of neighbourhood diversity: Variation of social context using mobile phone application data," *Urban Studies*, 2022, *59* (4), 851–869.

A Supplementary Information

A.1 GPS Data

The GPS data come from an unbalanced panel of GPS-enabled devices in 2019. The underlying data are a series of 'pings' that consist of a device identifier, coordinates, and timestamp. Pings can be triggered either by a user opening specific apps or via apps that share location data in the background. Our data from Replica are aggregates of pings into "stays" at locations, which each stay includes the device identifier, an entry/exit time, and GPS coordinates. Replica combines pings into stays using a spatio-temporal clustering algorithm similar to ST-DBSCAN (Birant and Kut, 2007). For example, a visit to the supermarket may result in a hundred pings, all of which would become a single stay. One limitation of using stays rather than raw pings is that we do not observe any time in transit; pings that occur between stays (e.g., on the highway) are discarded as part of the stay construction process.

Home and work locations are identified by Replica based on heuristics for when individuals tend to be at home versus work. Home location is generally the most common overnight location in a quarter, while work locations is generally the most common non-home daytime location.[13] A recurring issue with GPS data is what to do with devices that appear only sparingly. There are 132 million unique devices in the raw data, but the median device had only 9 observed stays and was seen in only 2 months of 2019. Only 3% of devices average over 60 stays a month.

To identify visits to different types of establishments, we use data from SafeGraph on the locations of various Points of Interest (POIs). The establishments data include the polygon describing the establishment's footprint. We use this polygon to identify when a device visits a given establishments. For establishments located within a larger, parent location (e.g., a restaurant within a mall), we assign the parent location rather than trying to disambiguate the individual establishment.

We categorize establishments according to their North American Industry Classification System (NAICS) code. Restaurants are those with NAICs codes beginning with "722". Retail locations are those establishments with NAICS codes beginning with "44" or "45". Parks, while not estab-

[13]We require devices to have at least 8 overnights and 5 days at work in the quarter to make the sample.

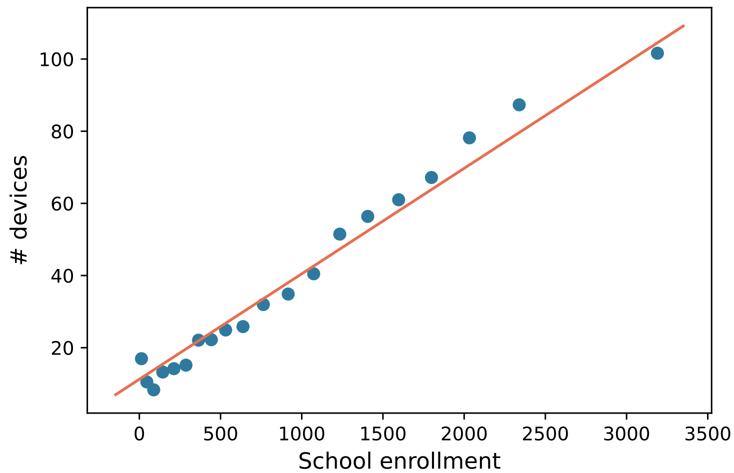
lishments, are identified in SafeGraph with a NAICS codes of "712190". Entertainment locations are all non-park POIs with a NAICS code beginning with "71". Finally, religious organizations are those with NAICS codes beginning with "8131".

A.2 Building a sample of schools

To build a sample of schools, we start with a directory of school locations from the National Center for Education Statistics (NCES) and include only public high schools. The NCES data only includes the school address, name, grades served, and enrollment. Moreover, the address is often a PO box or simply the town center rather than the actual school location. To match NCES schools to parcels, we first match each school to Safegraph data on schools using the school's name and location. The Safegraph data includes precise coordinates for each school as well as polygons. Unfortunately, the Safegraph polygons—which are often automatically generated from satellite imagery—are often inaccurate for schools; for large schools with multiple building, the polygon will generally only include a single building. We instead match the Safegraph coordinates to parcel-level data from LandGrid. These parcels come from local municipalities and provide more accurate boundaries for school locations. In total, 79% of schools representing 89% of enrollment are successfully matched to a parcel. Table A1 documents differences in characteristics of in-sample and out-of-sample schools; most notably, in-sample schools have substantially higher average enrollment.

Figure A.1 plots the relationship between a school's enrollment and the number of devices we label as a student at that school. The two counts of students are highly correlated. Figure A.2 plots the percent of residents who are students in the GPS sample as well as the percent who are in grades 9-12 in the 2019 5-year ACS. Students predominantly live in the less dense areas of the city. The overall trend is true in both the ACS and GPS, although we consistently find fewer high school (HS) students in the GPS data than in the ACS. This could be due to a number of factors, including that many HS students are under 16 years old and that we cannot match all schools to parcel polygons.

Figure A.1: NCES school enrollment & number of devices in sample



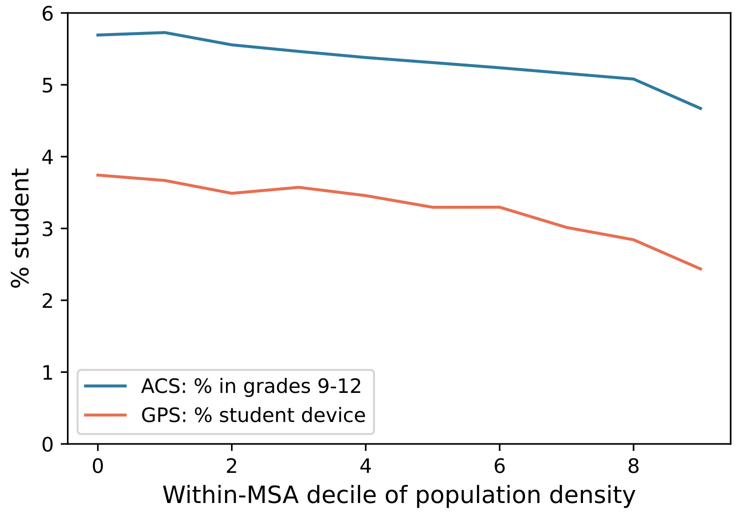
Note: This figure plots the relationship between the enrollment of a school as reported by the NCES and the number of devices labeled as a student of that school.

Table A1: Characteristics of in-sample and out-of-sample schools

	In-sample	Out-of-sample
Enrollment	558.66	258.56
	(4.84)	(5.97)
Number of teachers	35.85	17.35
	(0.27)	(0.38)
Tract median income	77534	74124
	(259)	(492)
Tract fraction white (non-Hispanic)	0.698	0.688
	(0.002)	(0.003)
N	22661	6098

Note: The table documents the average of a range of characteristics for schools that are in-sample and out-of-sample. Tract characteristics are based on the 2019 5-year ACS. Standard errors are reported in parentheses.

Figure A.2: Percent of residents who are students by density



Note: This figure plots the percent of residents who are students by within-CBSA home block group density. We separately plot the percent student in the GPS sample as well as the percent in grades 9-12 in the 2019 5-year ACS.

A.3 Sample Quality and Sample Weights

Figure A.4 plots the distribution of block group fraction white-alone (non-Hispanic). The average block group of white devices is 78.5% white, while the average block group of non-white devices is 21.0% white.

Figure A.5 plots the fraction of devices who live in each decile of tract characteristics. If sampling were orthogonal to tract characteristics, 10% of devices would be sampled from each decile. Instead, we can see that devices are over-sampled from poorer, less white, less educated, and more dense areas.

To address the uneven sampling of devices, we re-weight home locations to match the distribution of the 2019 5-year ACS by using the following sample weights

$$\lambda_{iq} = \frac{N_{g(iq)}}{\tilde{N}_{g(iq)}}$$

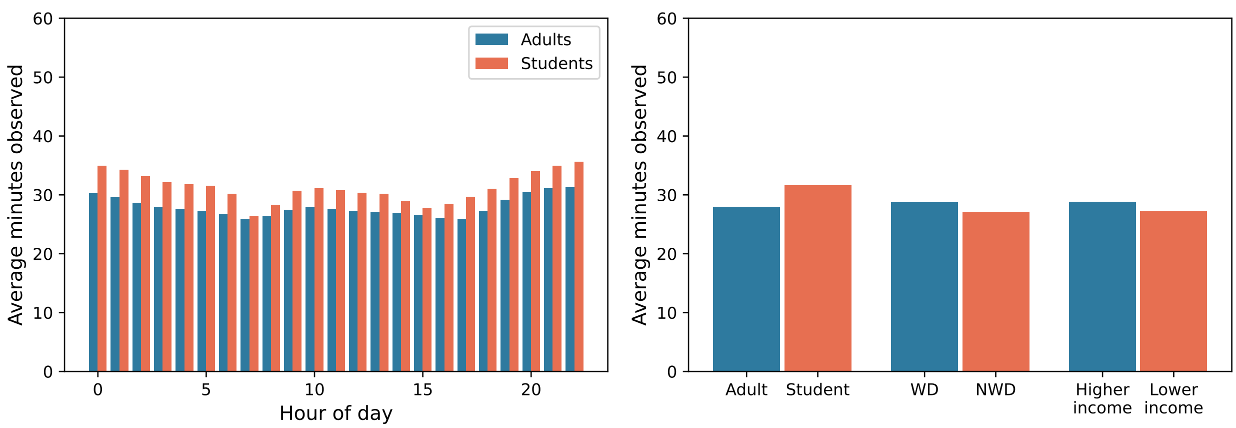
where $N_{g(iq)}$ is the ACS population of a device's home tract g and $\tilde{N}_{g(iq)}$ is the total number of devices observed in tract g in quarter q. The average device weight is 20.2 (standard deviation of 12.9). We use these weights for all results.

The GPS data is both a sample of devices and a sample of each device's overall movement; we do not observe a device's full location history, but only times at which an app recorded their location and shared it with data aggregators, which can happen both in the background or when a user actively opens an app. One concern is that students and adults may vary in the both the apps they use and when they interact with each app. Figure A.3 Panel (a) plots average number of minutes observed in each hour of the day for each day a device is in the data, where "observed" is defined as the time between the entry and exit times of a stay. On average we observe 29 minutes of each hour. The number of minutes observed varies slightly by the hour of the day, with greater coverage during the morning and evening hours. The time not covered by stays can be due to either travel or poor coverage; unfortunately, we are not able to test whether devices were pinging in between the observed stays (e.g., on a highway), although the missing time during the night is almost certainly poor coverage rather than travel. Figure A.3 Panel (b) plots the average minutes observed each day for different types of devices; students, white-devices (WD), and higher income devices are each observed slightly more than adults, non-white devices, and lower income devices (respectively).

Figure A.3: Time observed for devices

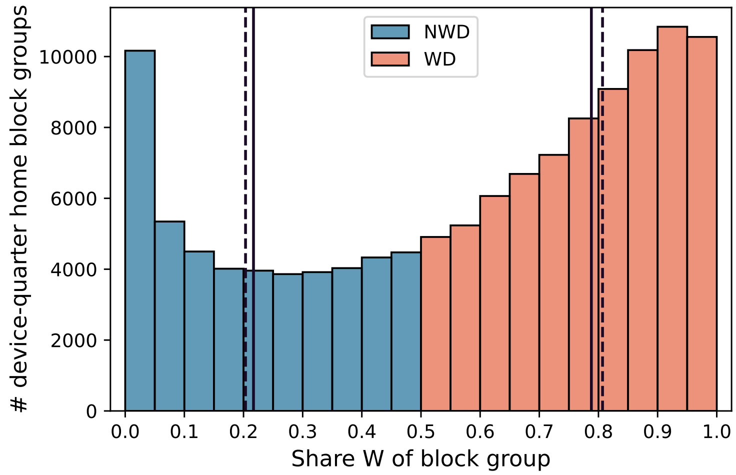
(a) Stay coverage by hour of day

(b) Stay coverage by device type (hourly)



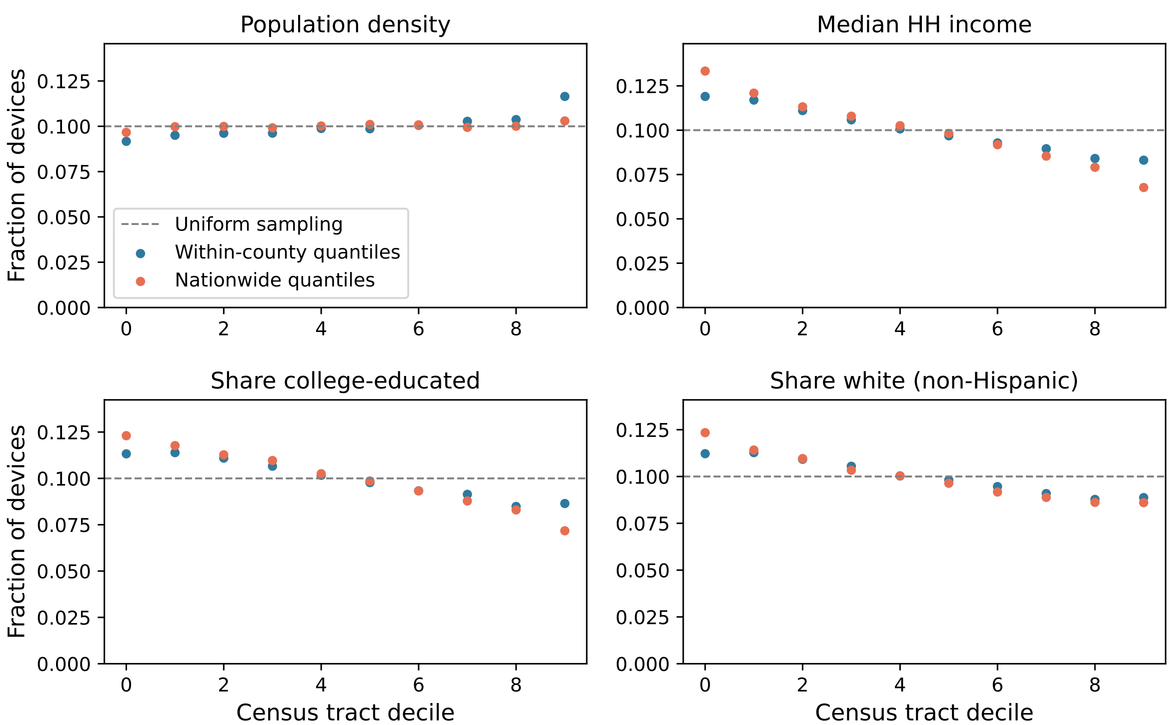
Note: This figure plots the average number of minutes observed in an hour by time of day and by device type for days when a device is observed at least once. We define "observed" as the time between the entry and exit time of a stay.

Figure A.4: Distribution of block group level fraction white



Note: This figure plots a histogram of the number of the number of devices sampled from different bins of home block group fraction white alone (non-Hispanic). The solid line represents the mean fraction white for each WD and NWD; the dashed line is the median.

Figure A.5: Sampling of devices by block group characteristics



Note: This figure plots the fraction of devices who live in each decile of tract characteristics. If sampling were orthogonal to tract characteristics, 10% of devices would be sampled from each decile.

A.4 Experienced Income and Racial Isolation Measures

Aggregate experienced isolation between any two groups A and B is defined as the difference between the average exposure of group A to other members of group A and the average exposure of group B to members of group A. Specifically, we define

$$\mathrm{EI}_g = \frac{1}{|\mathrm{A}_g|} \sum_{i \in \mathrm{A}_g} \int_{t=0}^{1} s(l(i,t),t)dt - \frac{1}{|\mathrm{B}_g|} \sum_{i \in \mathrm{B}_g} \int_{t=0}^{1} s(l(i,t),t)dt, \tag{A.1}$$

where g indicates a geographic unit such as CBSA, A_g is the set of devices in group A, B_g is the set of devices in group B, and $s(l(i,t),t)$ is the share of devices in individual i's location l at time t who are from group A.

We make several assumptions in order to estimate Equation A.1. First, we assume that $s(l(i,t),t)$ does not vary by time for a given location.[14] Second, we assume that the full population of visits can be approximated using the device sample, re-weighted accordingly. Finally, we discretize locations by geohash7s (approximately 500×500 feet).

To estimate Equation A.1, we first construct leave-one-out estimates of $s(\cdot)$ for each individual and location pair as:

$$\hat{s}_l^{-i} = \frac{\sum_{j \in P_l^{\mathrm{A},-i}} d_j}{\sum_{k \in P_l^{-i}} d_k}, \tag{A.2}$$

where P_l^{-i} is the set of stays in location l by devices excluding i, $P_l^{\mathrm{A},-i}$ is the set of stays in location l by members of group A, and d_j is the duration of stay j. Weighting by duration in the location approximates the ping-level measure used in Athey et al. (2021). Next, for each individual-quarter in our sample we measure aggregate exposure as the duration weighted average $\hat{s}_l^{-i}$:

$$\hat{S}_{iq} = \frac{\sum_{j \in P_{iq}} \hat{s}_{l(j)}^{-i} d_j}{\sum_{k \in P_{iq}} d_k} \tag{A.3}$$

where P_{iq} is the set of all stays by device i in quarter q and $l(j)$ is the location corresponding to stay j.

[14] This is most clearly violated in residential areas, where daytime and nighttime populations will differ substantially.

Finally, we estimate experienced isolation for CBSA g as:

$$\widehat{\mathrm{EI}}_g = \frac{1}{|\mathrm{A}_g|}\sum_{i\in \mathrm{A}_g}\sum_{q\in Q_i}\hat{S}_{iq} - \frac{1}{|\mathrm{B}_g|}\sum_{i\in \mathrm{B}_g}\sum_{q\in Q_i}\hat{S}_{iq} \tag{A.4}$$

where Q_i is the set of quarters in which device i is observed and we abuse notation slightly to let A_g be the set of all device-quarters (rather than just devices) of group A in geography g and B_g be the corresponding set of device-quarters of group B.[15] To estimate experienced isolation for students and adults, we estimate Equation A.4 separately for each type. Note that the leave-one-out estimates of exposure remain the same for both groups – consequently, our experienced isolation measure for students measures exposure to both students and adults.

We also define a new companion measure to facilitate individual-level regressions, which we call experienced diversity. Experienced isolation is a population-level statistic; for example, an experienced isolation score of 0.1 implies a gap of 10 percentage points in the exposure to members of group A by other members of group A relative to members of group B. Experienced diversity, in contrast, is the average exposure to devices of the *opposite* group and is estimated as

$$\widehat{\mathrm{ED}}_{iq} = \frac{1}{\sum_{j\in P_{iq}} d_j}\sum_{j\in P_{iq}} \mathbf{1}\{i\in \mathrm{B}\}\times \hat{s}^{-i}_{l(j)}d_j + \mathbf{1}\{i\in \mathrm{A}\}\times (1-\hat{s}^{-i}_{l(j)})d_j \tag{A.5}$$

The two measures are closely linked – experienced isolation is a transformation of the average experienced diversity by each group in a city.

B Additional results

B.1 Experienced Isolation with Continuous Race

Our baseline measure of experienced isolation uses a binary measure of race—as in Athey et al. (2020)—based on whether a device's home block group is majority white non-Hispanic or not. We also explore assigning devices a continuous measure of race using the percent of their home block

[15] We weight all statistics using device-quarter weights that correct for unevenness in the home locations of the GPS sample compared to the ACS. We provide more details on these weights, which are used for all results in the paper, in Appendix Section A.3.

group that is white. Using this continuous measure of race, we can estimate experienced racial isolation as

$$\hat{EI}_g^C = \frac{1}{|\mathrm{WD}_g|}\sum_i \sum_{q\in Q_i} \rho_{iq}\lambda_{iq}\hat{S}_{iq} - \frac{1}{|\mathrm{NWD}_g|}\sum_i \sum_{q\in Q_i} (1-\rho_{iq})\lambda_{iq}\hat{S}_{iq} \tag{B.1}$$

where ρ_{iq} is the continuous measure of race, $|\mathrm{WD}_g| = \sum_i \sum_{q\in Q_i} \lambda_{iq}\rho_{iq}$, and $|\mathrm{WD}_g| = \sum_i \sum_{q\in Q_i} \lambda_{iq}(1-\rho_{iq})$

The results are in Table B1. Using this measure, students are 4.12% more isolated in aggregate and 17.6% more isolated when outside of the home. However, the levels are dramatically different.

Experienced isolation with continuous race will be biased downwards relative to "true" exposure to diversity. Imagine a device whose true exposure to white individuals is 100%. When using a continuous measure of race, each exposure to a white individual will not count as a fully segregated exposure but instead will assume the device is exposed to the average percent white in the individual's home block group, making it look like the device has more diverse exposures than it does. Similarly, a device whose true exposure is 0% white will have a positive estimated exposure.

Table B1: Experienced isolation using continuous measure of race

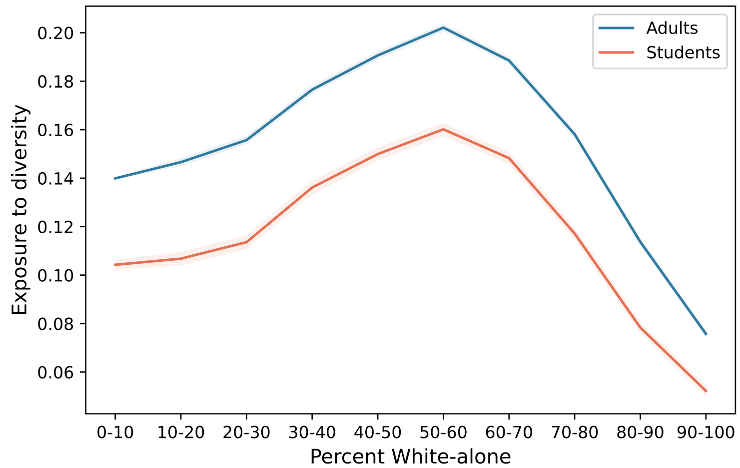
Panel a: experienced isolation	Aggregate	Student	Adult
All	0.2134	0.2248	0.2123
Excluding time at home	0.1141	0.1370	0.1129
Excluding time in home tract	0.1060	0.1274	0.1050

Note: This table documents a few basic measures of experienced isolation using a continuous measure of race based on the percent non-white in a device's home block group

B.2 Exposure to Diversity and Home Block Group Race

Figure B.1 plots the relationship between percent white in a device's home block group and the device's exposure to diversity. This figure is constructed by regressing exposure to diversity for a given device-quarter on whether the device is a student, interacted with the percent of their home block group's residents who are white (truncated to nearest 10%). Recall that exposure to diversity uses WD and NWD, rather than true race; as such, devices in racially mixed block groups still have exposure to diversity well below 50%.

Figure B.1: Exposure to diversity by home block group percent white

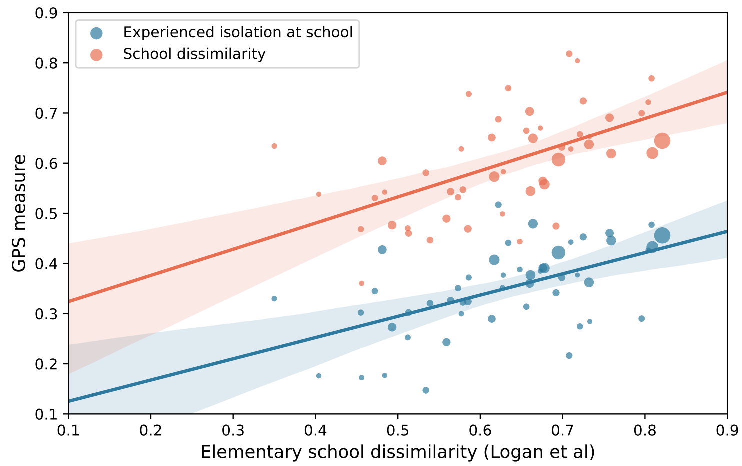


Note: This figure documents how exposure to diversity varies by the racial composition of a device's home block group. The figure is constructed by regressing exposure to diversity for a given device-quarter on whether the device is a student, interacted with the percent of their home block group's residents who are white (truncated to nearest 10%). The shaded region represents a 95% confidence interval

B.3 Experienced Isolation and School Dissimilarity

Figure B.2 compares measures of segregation from GPS data at the MSA-level to the dissimilarity indices from Logan et al. (2012, 2017) for the top 50 MSAs. The first measure is the experienced isolation of students while at school, while the second is the dissimilarity of WD/NWD school populations within a CBSA, intended to be analogous to the measure in Logan et al. (2012, 2017). Dissimilarity is defined as the percentage of students in one group who would have to move to a different school to achieve a racial balance representative of the MSA. The GPS measures and the Logan measure are positively correlated, which is encouraging, but there are also clear differences. These differences are perhaps expected for a number of reasons: 1) the dissimilarity indices are based on 2010 elementary school enrollment while the experienced isolation is based on 2019 high school students; 2) the dissimilarity indices compare white students and either Black or Hispanic, while we are comparing students from majority white and non-white neighborhoods; 3) our data will include some teachers; 4) the dissimilarity indices are defined either for MSAs or PMSAs instead of CBSAs, so the geographic match is imperfect (e.g., the results for Chicago are comparing the Chicago PMSA and the Chicago-Joiliet-Naperville CBSA).

Figure B.2: GPS measures vs. elementary school dissimilarity



Note: Figure B.2 compares MSA-level experienced isolation of students while at school and dissimilarity of school populations of WD/NWD to the dissimilarity indices from Logan et al. (2012, 2017) . The dissimilarity indices are based on 2010 elementary school enrollment from NCES. Correlation for EI is 0.55 and for dissimilarity is 0.56.

B.4 Experienced Isolation and Residential Isolation

Figure B.3 plots the relationship between an CBSA's experienced and residential isolation, splitting experienced isolation by whether or not a device is within its home tract. The relationship is plotted for both binary and continuous race and can help highlight the differences of each type of race assignment. For binary race, the time at home is estimated to be extremely isolated—generally far above residential isolation—because devices are spending the majority of their time in a home parcel where all other devices have the same, binary race. Meanwhile, for continuous race, experienced isolation is approximately equal to residential isolation while in the home tract. This is because experienced isolation with continuous race and residential are making a similar underlying assumption – interactions in the home tract are (approximately) with the average race of that location.[16]

For the majority of CBSAs, we estimate that experienced isolation is *higher* than residential isolation. This is counter to the findings in Athey et al. (2021). We believe that the discrepancies stem from differences in data construction. Athey et al. (2021) use raw GPS pings, which are

[16]The measures differ for two reasons: 1) with experienced isolation, race is measured at the block group rather than tract level and 2) interactions in the home block group include outside visitors, although for residential tracts the majority of interactions are with other residents.

recorded each time an app on the device connects to GPS, while we use staypoints, which are aggregations of pings into "stays" in a given location. In Athey et al. (2021), a unit of "exposure" to a neighborhood is therefore at the ping level, while in our measure we weight by staypoint minutes in a location.

The use of stays instead of pings leads to larger estimates of experienced isolation for two reasons, both related to how time spent at home affects the aggregate measure. First, using pings puts less weight on time spent at home. Devices are less likely to ping overnight while the holder is sleeping; however, a staypoint is still formed from the evening to the morning. In their earlier working paper version, Athey et al. (2020) report that 42.5% of pings are at home for the average device (their Table A3), while devices in our sample are at home for 63.9% of minutes observed. Excluding night time stays, for example, substantially reduces our measure of experienced isolation, although it remains higher than residential isolation (Table 1).

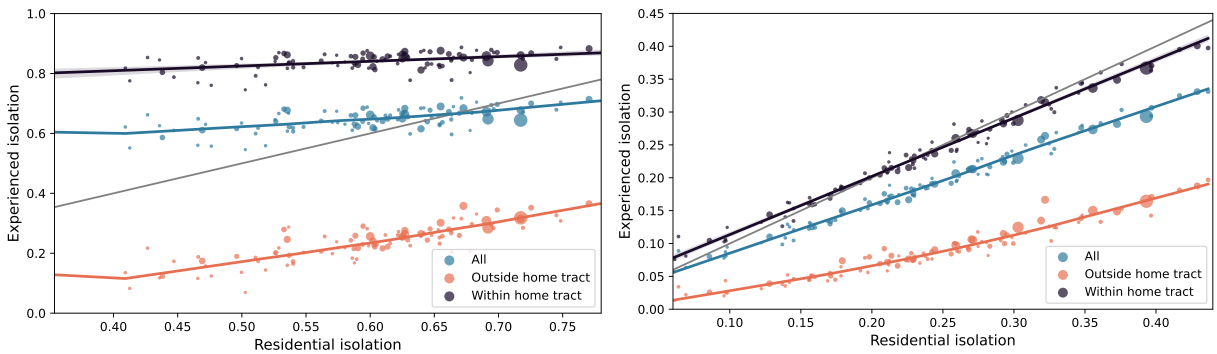
The second reason is that exposure when using pings will include devices walking/driving through the location, while staypoints include only those devices that stop for at least 5 minutes in a location. This difference is particularly noticeable for residential locations – the experienced isolation in Athey et al. (2020) of "at home" pings is 0.672 (Table A9), which is lower than expected given that, by definition, all devices who live in that geohash7 are assigned the same race. This low number is likely due to devices who walk/drive through the location during the day, who are more diverse than the residents. The assumption that the time people visit a location is independent of their race is violated in the case of residences where night-time and day-time populations differ substantially.

While the home-based assignment of race makes it difficult to compare experienced and residential isolation directly, much like Athey et al. (2021) we find that individuals experience far less isolation outside of the home tract than they do within their home tract and that experienced isolation outside of the home tract is also substantially lower than residential isolation. In many ways, this is the core insight of their paper: residential isolation overstates the isolation that individual experience in their day-to-day lives, as time outside of the home is empirically less isolating. Finally, while comparisons of residential and experienced isolation are complicated by the home-based assignment of race, we do not belief the issue affects the relative differences in either measure of isolation for students and adults.

Figure B.3: Experienced vs. residential isolation

(a) Binary race

(b) Continuous race



Note: Each figure plots the relationship between experienced isolation and residential isolation at the MSA level. Residential isolation is estimated to be consistent with the method of estimating experienced isolation; it uses either binary block-group level race or continuous race.

B.5 Bias From Student Imputation

As detailed in Section 2.1, we define "students" by matching "work" locations to the geographic parcels of public schools, which results in our student sample inevitably including some teachers and staff. We expect this to attenuate measured differences between students and adults. It may also bias comparisons within the student population by race, since teachers are disproportionately white relative to their student body. If, for instance, teachers have high urban mobility, we may over-estimate the gap in urban mobility for the white versus non-white student populations.

In this section, we do a back of the envelope calculation to estimate this bias. In particular, we consider the exposure to diversity (ED) results presented in Table 1, panel b. Our estimates for white adults, non-white adults, white students, and non-white students are 0.194, 0.348, 0.149, and 0.320, respectively. Taking the average student and teacher populations in sample from A1, we assume 6% of our "students" are really teachers, 79% of which we assume are white based on the NCES estimate (Spiegelman, 2020).

Let the subscript $i \in \{W, NW\}$ indicate white or non-white, the superscripts $\{s, t\}$ indicate student and teacher status, N^s indicate our original sample size of students, and N^t indicate our new estimated sample size of teachers. We can update ED for white and non-white students as:

$$ED_i^{s*} = \frac{N_i^s \cdot ED_i^s - N_i^t \cdot ED_i^t}{N_i^s - N_i^t}$$

Plugging in the four estimates of ED above updates the estimate for white students to $ED_W^{s*} = .145$ from $ED_W^s = .149$ and for non-white students to $ED_{NW}^{s*} = .319$ from $ED_{NW}^s = .320$. This suggests the bias from differences in race between students and adults is fairly limited.

C Supplementary tables and figures

Table B2: Correlation between urban mobility index and its components

Measure (z-score)	Corr. with urban mobility index
Urban mobility index	1.0
# restaurant visits	0.649
# retail visits	0.638
# park visits	0.49
# entertainment visits	0.561
# civil, social, religous visits	0.255
Frac. time at home	-0.229
Frac. time at work	-0.093
Frac. time in neighborhood	-0.036
# unique locations	0.773
Avg. miles from home (excl. home)	0.394

Note: This table plots the correlation between the single urban mobility index and each of its component parts. Each component is z-scored. To construct the index, we first do PCA on each of the three groups: amenity visits, fraction of time in primary locations, and roaming ranges. We then do a PCA on the resulting principal components and take the z-score to get our final measure.

Table B3: Correlates of urban mobility

Outcome: urban mobility index	Univariate	Controlling for income	
	Covariate	Covariate	Log pred. income
Device covariates			
Log pred. device household income	0.414		
	(0.006)		
Home tract covariates			
Log median household income	0.402	0.174	0.271
	(0.006)	(0.012)	(0.012)
Log pop. density	-0.0942	-0.0685	0.369
	(0.002)	(0.002)	(0.006)
Log miles to city hall	0.145	0.095	0.381
	(0.004)	(0.004)	(0.006)
Frac. households with car	1.165	0.373	0.377
	(0.032)	(0.035)	(0.007)
HUD transit index	-0.0073	-0.0048	0.37
	(0.0002)	(0.0002)	(0.006)
Home zipcode covariates			
Economic connectedness	0.761	0.328	0.314
	(0.017)	(0.023)	(0.009)
Network clustering	7.831	4.964	0.401
	(0.477)	(0.386)	(0.007)

Note: This table plots linear coefficients corresponding to the relationships shown in Figure 2. The data device-quarter level (students only) and include 371,243 observations. The first column regresses the urban mobility index on the listed covariate and documents. The second and third column report coefficients on the covariate and on log predicted device household income from a regression of urban mobility on both. All regressions additionally control for city fixed effects and are weighted by the device sample weights. Standard errors are clustered at the home tract level, except for the zipcode covariates where they are instead clustered at the home zipcode level.

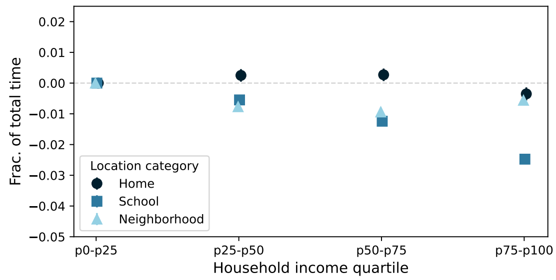
Table B4: NHTS Travel Patterns for 16-18 Year Olds

Income quartile	Shopping Trips	Social Trips	Meal Trips	Dwelling Time
1	0.277	0.292	0.095	292.542
	(0.021)	(0.021)	(0.011)	(8.586)
2	0.35	0.32	0.179	320.575
	(0.025)	(0.019)	(0.013)	(6.884)
3	0.22	0.352	0.198	311.193
	(0.013)	(0.014)	(0.011)	(5.273)
4	0.254	0.428	0.226	334.674
	(0.012)	(0.014)	(0.01)	(4.589)

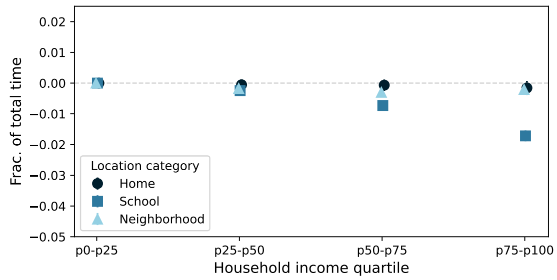
Note: This table reports mean number of trips and mean total dwelling time for individuals aged 16-18 in the 2017 National Household Transportation Survey (NHTS). The NHTS is a federal survey on travel patterns conducted roughly once every eight years, and it includes a small sample of individuals age 16-18 (7,239 in the 2017 survey), as well as some coarse information on travel destinations. Income quantiles are defined by the household income of the entire NHTS 2017 sample. Shopping, social, and meal-related trips are defined by the reported destination of the trip. Dwelling time is the total minute the individual reported dwelling at any destination that day. Similar to results with our data, the NHTS shows that both travel for amenities and time away from the home are rising in household income.

Figure C.4: Urban mobility of adults by income

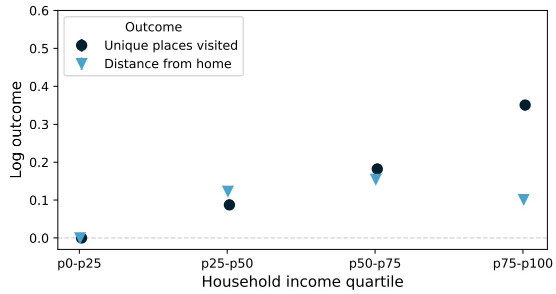
(a) Time at primary locations (MSA FEs)



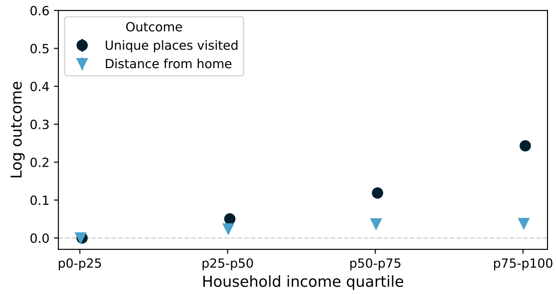
(b) Time at primary locations (Tract FEs)



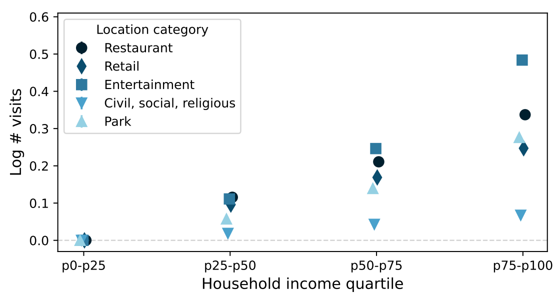
(c) Roaming ranges (MSA FEs)



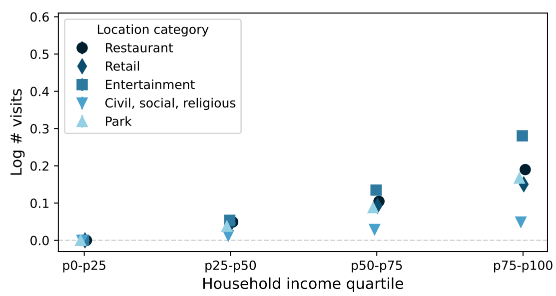
(d) Roaming ranges (Tract FEs)



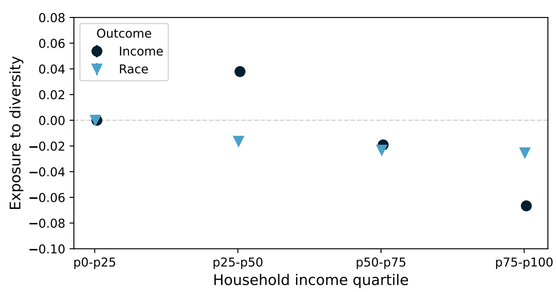
(e) Visits to amenities (MSA FEs)



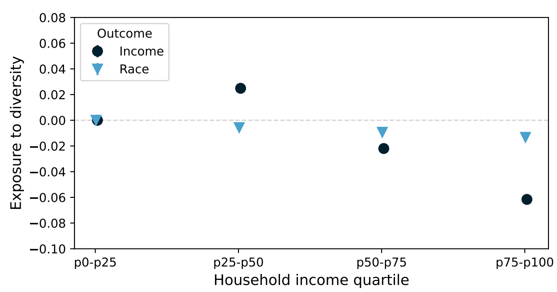
(f) Visits to amenities (Tract FEs)



(g) Exposure to diversity (MSA FEs)

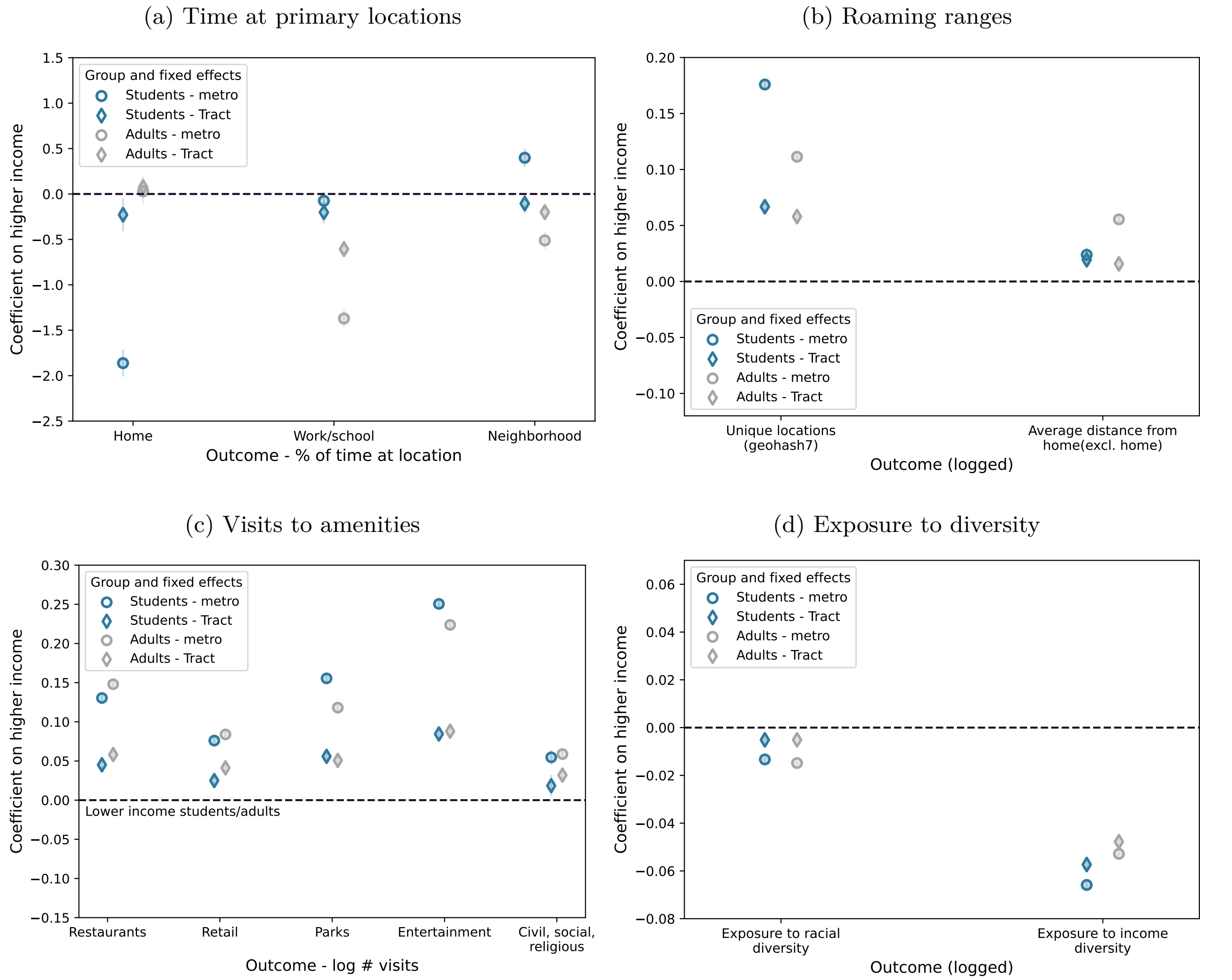


(h) Exposure to diversity (Tract FEs)



Note: This figure replicates Figure 1 using adult rather than student devices. Each point is a coefficient from a regression of a device-quarter level mobility measure (e.g., fraction time at home) on indicators for whether the quartile of a devices predicted income and MSA fixed effects. 95% confidence intervals are represented by bars around each point (although are frequently covered by the point itself).

Figure C.5: Differences by income for students versus adults



Note: This figure plots average differences between above and below-median income devices, split by students and adults. Each point is a coefficient from a regression using device-quarter data of some outcome (e.g., percent of time at home) on whether the device is above median income, with controls for either home MSA or tract. 95% confidence intervals are represented by bars around each point (although are frequently covered by the point itself).