ACI Research Paper #07-2021

Segregation Across Neighborhoods in a Small City

Shu En LEE

Jing Zhi LIM

Lucas SHEN

May 2021

Please cite this article as:

Lee, Shu En, Jing Zhi Lim and Lucas Shen, "Segregation Across Neighborhoods in a Small City", Research Paper #07-2021, *Asia Competitiveness Institute Research Paper Series (May 2021)*

Segregation Across Neighborhoods in a Small City

Shu En Lee,[1] Jing Zhi Lim[1] Lucas Shen[1,2*]

[1] Asia Competitiveness Institute, Lee Kuan Yew School of Public Policy, National University of Singapore
[2]School of Social Sciences, Nanyang Technological University
*To whom correspondence should be addressed; E-mail: lucas@lucasshen.com.

Social segregation has profound impacts on socioeconomic outcomes. Using anonymized GPS records for Singapore which we spatially join to census records, we examine daily movement across geographically-refined neighborhoods. We show that the GPS-derived data detect segregation by poverty, even with an imperfect proxy, and in the presence of targeted urban policies aimed at social integration. The findings bode well for the use of GPS data in general to measure social segregation.

A set of established and still growing literature shows that economic and social outcomes are strongly shaped by who people know and the networks they have access to. Social segregation is therefore profoundly important in shaping the paths that individuals take. Past research for instance has found that social connectedness improves startup success (*1*) and labor market opportunities (*2, 3*), and segregation is correlated with greater violence (*4*) and worsens socioeconomic markers such as schooling, employment, and marriage (*5, 6*).

Measuring segregation itself, however, has traditionally been restricted to census records based on where individuals reside (*7–9*). The immediate problem with this measure of residential segregation is that social networks extend far beyond where people reside. Social integration and networks also depend on non-residential exposure (*10*), which includes where people go to work, where they go to school, where they eat, and where they spend their leisure time (*11*).

In this study, we use GPS (global positioning system) ping records obtained from CITYDATA.ai and test whether real-time movement across neighborhoods in the city-state of Singapore, which are geographically refined, is correlated with census-derived poverty levels. Finding that our GPS-derived measure of daily movement patterns is correlated with poverty serves two broad purposes. First, it supports the use of GPS-type data as an emerging tool to understand and study social interactions at a temporal frequency much higher than possible with traditional census-based measures. Second, and as an institutional-specific finding, detecting social segregation from GPS-derived movement patterns suggests that there are limits to urban planning policies that specifically aim to socially integrate a diverse group of population. Examples of social integration policies include those that aim at dispersing members of certain ethnicities and low-income groups to avoid enclaves (*9*), and integrating by ethnic itself nests integration by income.

Even in the presence of policies aimed at social integration, with an imperfect proxy for poverty, and after removing variation in movement patterns that originate from other factors such as spatial frictions and points of interest, we find persistent correlation between poverty and movement patterns. Our analysis is based on the city of Singapore, but the method can be readily adapted to other metropolitan areas, and the efficacy of the GPS-derived data to detect segregation in a city that

Table 1. Descriptive statistics.

	Count/Mean ± s.d.	Min.	Max.
Device hashes	> 17million	.	.
Device hashes in sample (> 30 appearances)	$> 125k$	.	.
Census areas (coarser)	55	.	.
Subzones/neighborhoods (finer)	323	.	.
Census areas, with residential records	52	.	.
Subzones/neighborhoods, with residential records	219	.	.
Census areas, with GPS records	52	.	.
Subzones/neighborhoods, with GPS records	301	.	.
Subzone area (km^2)	2.23 ± 5.67	0.04	69.75
Neighborhood area (km^2) (subzones with residential records)	1.35 ± 1.24	0.05	8.45
Poverty measure (% in 1–2-room public flats)	0.039 ± 0.07	0	0.49
Wealth measure (% in private housing properties)	0.43 ± 0.43	0	1

is meticulous with urban planning and social integration bodes well for the use of GPS data in general.

Geography of Singapore

Singapore is a city-state in Southeast Asia, widely considered as a small but developed economy. The size of the city is approximately 720km^2 or 278 square miles, which is approximately 5 times as large as San Francisco (121km^2), 1.2 times as large as Madrid city (667km^2), and 0.45 times as large as London city (1,570km^2).

For urban planning and census taking, Singapore is delineated, in increasing geographical refinement, into 5 regions, 55 planning areas, and more than 300 subzones. For the purpose of this study, we term the 219 subzones with census residential records as *neighborhood* (Table 1). These neighborhoods are relatively small on average, with the representative neighborhood being about 1.35km^2 (0.52sq mi), which is comparable to the area of two geohash-6 grids (1.2km $\times$ 0.6km) stacked vertically, and with some neighborhoods as small as 0.05km^2 (0.02sq mi), which is comparable to two geohash-8 grids (38m $\times$ 19m). These neighborhoods are the geographical units of analyses (Fig. S1).

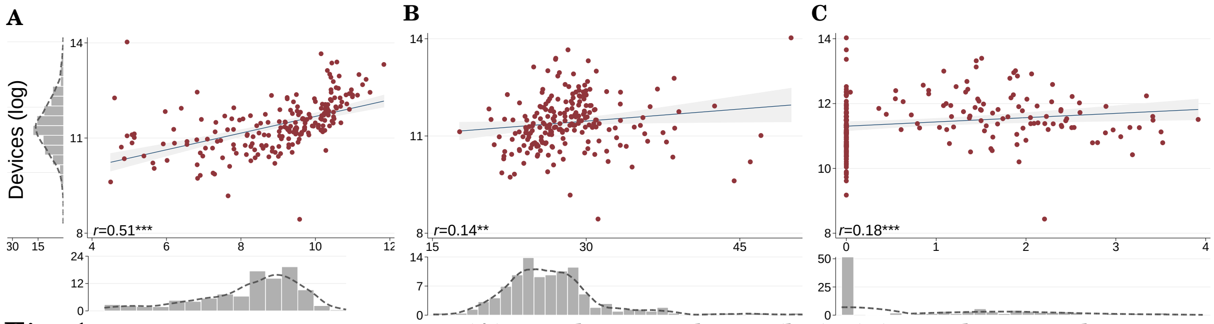

Fig. 1. Device capture in areas. (**A**) Resident population (log). (**B**) Resident population, age 20–39 (%). (**C**) Residents in 1–2-room public flats (log). Plots above include 217 areas where the resident population record is available. The vertical axis is always the log of devices captured in an area. Distribution plots are by percentages. Each panel reports the pairwise correlation coefficient. *** and ** indicates statistical significance at the 1% and 5% level, respectively.

Proxying for neighborhood poverty

The GPS data, being anonymized, is agnostic of individual records, including their location of residence. To impute the origin neighborhood of individuals, we use the most frequented neighborhood inside the sample time frame. The poverty level of a neighborhood we use is the proportion of residents in the neighborhood that live in 1–2-room public high-rise flats. Public flats in Singapore are subsidized, and the 1–2-room variants, being the smallest, are the cheapest type of residence and a higher proportion thus implies higher levels of poverty (*9, 11*). We use this measure instead of actual household income because these are not available at any level finer than the 55 planning areas. This measure should sufficiently detect variation in poverty status because income or wealth levels directly influence house types (*9*), and the vast majority (> 80 percent) of residents reside in the public flats. In the context of Singapore, the public rental scheme, where low-income citizens (whose total household income $<$ \$1,500 per month, or $<$ USD 1,125 per month) can rent 1–2-room public flats at a further and heavily subsidized rate (*12*), further justifies the use of our poverty measure.

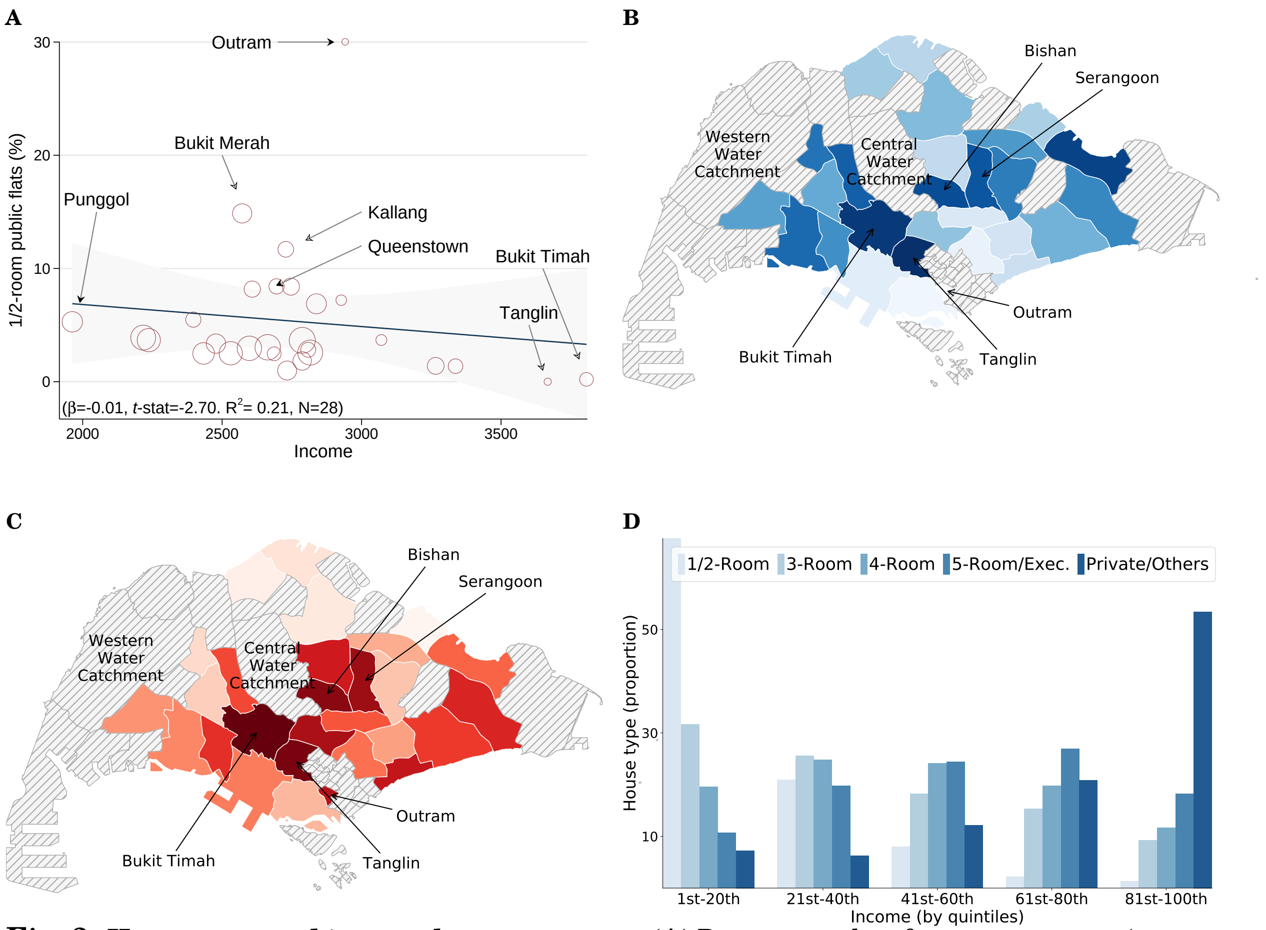

Fig. 2. House type and income by census area. (**A**) Regression plot of poverty measure (proportion of residents living in 1–2-room public flats) against household income, conditional on working population size. Income is income per month per working persons aged 15 and above from the 2015 General Household Survey, and is only available for 28 (of 55) census areas—computed by first taking the midpoint in each income bin and weighting by residents belonging in that bin, then dividing by total number of working population in the area. (**B**) Geographical distribution of 100 - Poverty measure; darker shades encode lower poverty. (**C**) Geographical distribution of Household income; darker shades encode higher income. Available for 28 areas. Shaded areas denote missing data. (**D**) House type, by income quintiles. All house categories are public houses except for the last "Private/Others" category which includes high-rise condominiums and single-family houses.

We start our analyses by corroborating the use of the proposed poverty measure at the coarser geographical level of census planning area, where actual household income census records are available. Fig. 2A plots our poverty measure with household income, available for 28 census areas. As proposed, areas higher in poverty—more proportion of residents in 1–2-room public flats—have lower household income, with household income explaining 21 percent of the variation in poverty. Fig. 2B and Fig. 2C show the geographical coincidence in poverty and household income, also at the census area level, where less poor places are places with higher income. Fig. 2D shows the distribution of housing types by income. The lowest income quintile has by far the largest proportion of residents in 1–2-room public flats while the highest income quintile has by far the largest proportion of residents in private housing. This confirms non-parametrically that our poverty measure closely matches income. Based on how closely the highest income quintile tracks residence in a private housing property, we also use the proportion of residents living in private housing as a *wealth* measure.

Anecdotal Examples from Four Neighborhoods

We first assess movement patterns qualitatively in Fig. 3, which shows the concentration of movement from four neighborhoods of origins, all in the central region, two of which we consider to be on the extreme poverty end and two on the extreme wealth end. First, we observe that individuals are more likely to visit neighborhoods are that nearby rather than farther away. This observation is consistent with spatial frictions affecting segregation (*13*). When they do visit neighborhoods that are farther away, those neighborhoods are non-residential areas, such as industrial parks or business districts. Second, individuals from the poorer areas (darker shades) travel shorter distances than individuals from wealthier areas. Third, and perhaps the most salient observation we make, the farther neighborhoods that individuals travel to are very similar in terms of our poverty measure. In the first row (Fig. 3A and Fig. 3B), the farthest neighborhoods that individuals from poorer neighborhoods travel to are also poor (darker shade), and vice versa. The farthest neighborhoods

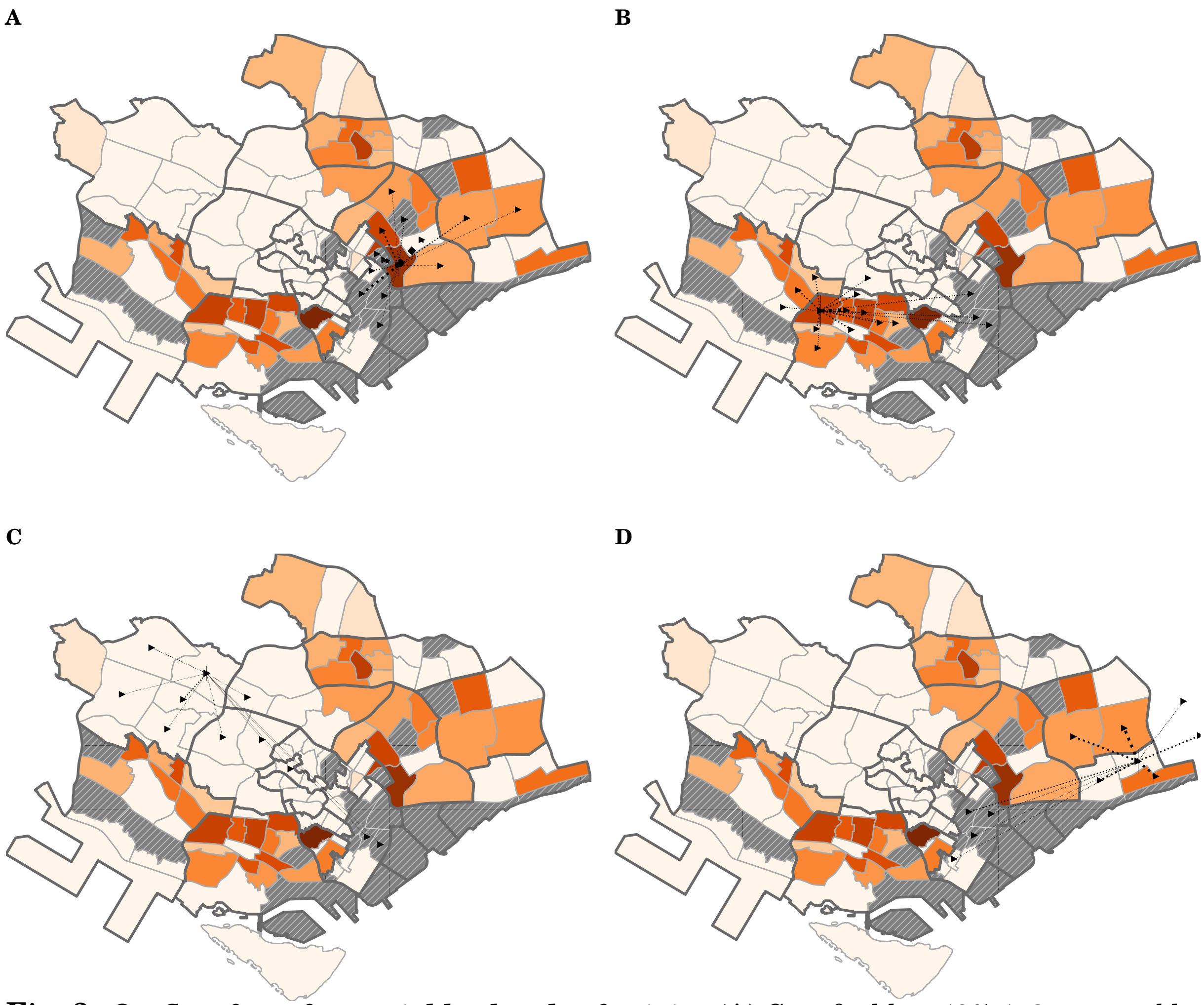

Fig. 3. Outflow from four neighborhoods of origin. (**A**) Crawford has 49% 1–2-room public flats and 7% private properties. (**B**) Alexandra Hill has 29% 1–2-room public flats and 1% private properties. (**C**) Hillcrest has 100% private properties. (**D**) Katong has 100% private properties. Table S1 shows the distribution of residence type for these four neighborhoods. Black lines show flow of movement from four selected areas of origin; two that are poor (first row) and two that are rich (second row). Thicker lines indicate higher frequency. Inflows less than 0.8 (5th percentile) are not plotted. Map is a cutout of the central region. Darker shades in map indicate poorer areas. Thin gray lines indicate subzones while thicker gray lines indicate the coarser census area borders Shaded areas indicate non-residential areas.

that individuals from the wealthier neighborhoods (Fig. 3C and Fig. 3D) travel to are wealthy areas (lighter shade). This behavior is consistent with social frictions, where segregation occurs because individuals socially expose themselves less often to individuals from other groups because of the type of group per se.

Effect of Poverty on Movement

To formalize how movement patterns across neighborhoods is correlated their poverty levels, we estimate the following equation:

$$\log(\text{inflow})_{odt} = \alpha + \beta_d P_d + \beta_o P_o + \Gamma_t X_{odt} + \varepsilon_{odt}, \tag{1}$$

where inflow_{odt} is the percentage of people from origin neighborhood o who visit destination neighborhood d during day t. P_d and P_o are poverty levels in the destinations and origins, measured using the proportion of residents who live in the cheapest 1–2-room public flats. The βs are our coefficients of interest, which tell us the extent to which our poverty measure predicts movement patterns.

Since our interest is only in correlating movement patterns to poverty, and not to push for a causal interpretation, confounding is not a first-order problem. We nonetheless include controls that may explain away part of the variation in movement patterns, so that the β coefficients are capturing social frictions. Specifically, Eq. (1) includes the (centroid-based) origin-to-destination neighborhood distances, a contiguity dummy, and neighborhood size, to account for spatial frictions in segregation. Demographics controls include resident age groups (which are correlated with both residence type and distance to workplace (*14*)), population size, and a measure of urban population density using the prevalence of devices capture from the CITYDATA.ai records itself. This is our baseline model.

To test sensitivity, more demanding specifications of Eq. (1) in turn include controls for the area fixed effects (one geographical level higher than neighborhoods), density of places of interests (e.g. schools, transit stations, tourist attractions, etc.), neighborhood-specific rainfall (Fig. S3), and the prevalence of construction, manufacturing, and services-related businesses located in the neighborhoods. All of

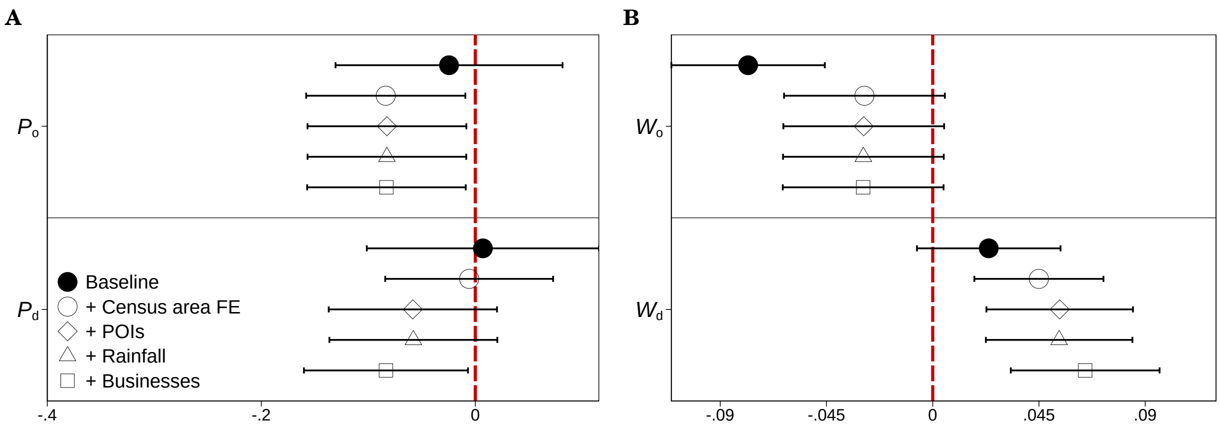

Fig. 4. Effect of poverty and wealth on movement. (**A**) Plots of the estimated β coefficients of the poverty measure from estimating Eq. (1). (**B**) Similar, except that the poverty measure is now swapped out for the wealth measure. Dependent variable is log of inflow from origin o to destination d. Capped horizontal lines are 90% confidence intervals from standard errors clustered at the origin-by-destination area level.

these are included semi-parametrically with the coefficients allowed to change over time. All estimations also include date fixed effects, so that day-specific movement patterns (e.g. more leisure travels on weekends) are removed. Standard errors are clustered at the origin-by-destination census area level.

The results from estimating Eq. (1) are reported in Fig. 4, which confirm that social frictions drive segregation across neighborhoods on a daily basis. Fig. 4A shows the results focusing on poverty. The solid black circle denotes the baseline result, and the hollow markers show how the estimated coefficients change as the model includes additional controls. The estimate of -0.083 for P_o (Table S2) implies that a 50 p.p. increase in the origin poverty measure (the max in the sample, Table 1) decreases inflow by approximately 4 percent ($= 100 \times -0.08 \times 0.5$). The estimate of -0.084 for P_d, on the other hand, implies that a 50 p.p. increase in the destination poverty measure (the max in the sample, Table 1) decreases inflow by 4 percent as well. The estimates are marginally and statistically significant at $p < 0.1$, but one-sided tests that the coefficients are less than zero cannot be rejected (Table S2).

These estimates confirm that the poverty status of neighborhoods, both as origins and as destinations, strongly correlates with movement patterns, even after removing variations in geographical frictions and demographics. Importantly, our estimates that individuals from poorer places are more likely to visit poorer areas are conditional on poorer places being poorer—that is, with fewer businesses and points of interest, which is a fact of urban planning.

In Fig. 4B, we use the wealth measure (proportion of residents living in private housing properties) instead and re-estimate Eq. (1). The coefficient for wealth in the origin neighborhood, W_o, is -0.03 (Table S3), which implies a decrease in inflow of 1.3 percent ($= 100 \times -0.03 \times .43$), while the coefficient for wealth in the destination neighborhood, W_d, is 0.065, which implies an increase in inflow of about 2.8 percent ($= 100 \times 0.065 \times .43$). These estimates suggest that neighborhood wealth status determines movement inwards rather outwards, or, that where people are going matters more than where they are from along the wealth dimension. Including both poverty and wealth measures together leads to similar conclusions, but with the effects concentrated by poverty of origin neighborhood and the wealth of destination neighborhood (Table S4).

Discussion

Our findings bear implications for both the literature and the efficacy of socio-engineering policies in general. In the context of the city-state of Singapore with a history in policies that aim at social integration via public housing allocation along ethnic lines (*9*), which itself nests integration by income, detecting social segregation by real-time exposure of individuals suggests that segregation cannot simply be socio-engineered away. This is of relevance since we know that segregation can have profound consequences along both social and economic dimensions (*4, 15, 16*) and create schisms in everyday social life (*11*).

In general, the findings also fit in the literature that uses non-conventional approaches to quantify social segregation (*13, 17*), as well as an emerging set of studies where the availability of GPS data provides further understanding of human behavior. Examples include how partisanship affects family ties (*18*), how policing spatially affects criminal behavior (*19*), how GPS can be used to predict poverty and wealth when census records are sparse (*20*), and in contemporaneous studies using GPS data to examine safe-distancing in the Covid-19 pandemic (*21, 22*).

References

1. D. Banerji, T. Reimer, *Computers in Human Behavior* **90**, 46 (2019).

2. M. Banerjee, P. Ingram, *HEC Paris Research Paper No. SPE-2018-1305; Columbia Business Research Paper No. 18–74* (2018).

3. L. Hensvik, O. N. Skans, *Journal of Labor Economics* **34**, 825 (2016).

4. L. D. Cook, T. D. Logan, J. M. Parman, *Social Science History* **42**, 635 (2018).

5. K. Y. Chay, J. Guryan, B. Mazumder, Early Life Environment and Racial Inequality in Education and Earnings in the United States, *Working Paper 20539*, National Bureau of Economic Research (2014).

6. R. Chetty, N. Hendren, L. F. Katz, *American Economic Review* **106**, 855 (2016).

7. R. Hutchens, *Mathematical social sciences* **42**, 13 (2001).

8. A. Rodriguez-Moral, M. Vorsatz, *An Overview of the Measurement of Segregation: Classical Approaches and Social Network Analysis* (Springer International Publishing, Cham, 2016), pp. 93–119.
9. C. H. Sin, *Housing Studies* **17**, 267 (2002).
10. D. S. Massey, N. A. Denton, *Social forces* **67**, 281 (1988).
11. M. Low, *Channel News Asia* (2018).
12. Housing & Development Board, Public Rental Scheme - Eligibility.
13. D. R. Davis, J. I. Dingel, J. Monras, E. Morales, *Journal of Political Economy* **127**, 1684 (2019).
14. Housing & Development Board, HDB Sample Household Survey 2018 (2021).
15. E. O. Ananat, *American Economic Journal: Applied Economics* **3**, 34 (2011).
16. D. M. Cutler, E. L. Glaeser, *The Quarterly Journal of Economics* **112**, 827 (1997).
17. S. Athey, B. A. Ferguson, M. Gentzkow, T. Schmidt, Experienced segregation, *Tech. rep.*, National Bureau of Economic Research (2020).
18. M. K. Chen, R. Rohla, *Science* **360**, 1020 (2018).
19. C. Blattman, D. P. Green, D. Ortega, S. Tobón, *Journal of the European Economic Association* (2021). Jvab002.
20. J. Blumenstock, G. Cadamuro, R. On, *Science* **350**, 1073 (2015).
21. H. Allcott, *et al.*, *Journal of Public Economics* **191**, 104254 (2020).
22. A. Gollwitzer, *et al.*, *Nature human behaviour* **4**, 1186 (2020).

Acknowledgments

This study would not have been possible without an early data grant from CITYDATA.ai, a San Franciso-based GovTech 100 company. We thank Apurva ”Apu” Kumar and Lee Yew Leong from CITYDATA.ai for advice, as well as Huang Yuting for helpful comments. All interpretations in this manuscript are those of the authors and not necessarily those of affiliated institutions.

Supplementary Materials

To build the gravity panel by `origin-destination-date`, we start by using the device list records from CITYDATA.ai which are stored as flat files. For each URA census subzone and each day, we have a recorded list of captured devices in that subzone using the device hash.

1. To get cross-area movement flows, we treat each device hash as an individual and aggregate the CITYDATA.ai records up to the `origin-destination-date` level. This gives for any `origin-destination-date` the count of inflow from the `origin` area to the `destination` area on a given `date`. We then divide this inflow count by the number of devices for that `origin-date` to get the inflow measure.

2. To infer the ”origin” area of an individual, we aggregate the CITYDATA.ai records up to the `device-area`, and make the simple assumption that the area with the highest appearance count for a given device hash is the ”origin” area. Devices that appear <30 times in the 91-days sample period are dropped. Certain devices have ties in the rank and we treat them as different individuals. This yields records from approximately 125k devices.

3. To measure poverty/wealth at the smaller subzones using the census records, we use the the proportion of residents living in 1–2-room public housing flats—this is the count (rounded off to tens in the official records) of local residents living in 1–2-room HDB (Housing & Development Board) flats divided by the total number of local residents in the subzone. Records for "HUDC Flats (excluding those privatised)" and "Others" are excluded from both numerator and denominator. All house type records are for the year 2019.

   Census variables for basic demographics at the subzone level are derived in the same manner. These include age demographics (% residents aged 65 and above) and resident population size.

4. To measure poverty/wealth at the larger census planning area unit, we simply use the census records of both residential house types and household income. Household income is a per-population weighted-average: for each income bin reported we take the midpoint and multiply it by the number of residents in that income bin, then we aggregate up to the area level and divide by the total residents in that area.

5. To get distance measures, we use the 2014 masterplan data from `https://data.gov.sg/dataset/master-plan-2014-subzone-boundary-web`. The geographic data is stored in a projected coordinate system - SVY21, which allows for a more accurate representation of the Singapore area. For each subzone, we calculate the coordinates of its centroid. We then take the distance between all pairs of centroids to obtain the inter-subzone distance, `subzone-centroid-distance`. Additionally we obtain the edge-to-edge distance between all pairs of subzones. If the edge distance equals 0, the 2 subzones

are considered to be contiguous.

6. To crosswalk from addresses to subzones, we do a simple point-in-polygon query to see if the coordinate falls in a subzone. If a coordinate falls in two different subzones, which might happen for coordinates on the edge of subzones, we choose the modal subzone.

7. For the area-specific POIs (places of interests), we mostly default to the official records found in `https://data.gov.sg`. For POIs records stored in a shapefiles or its equivalent, we simply do a point-in-polygon query to match POIs to areas. Number of POIs in a subzone includes: libraries, supermarkets, parks, preschools, schools (primary and secondary), silverzones, sport facilities, train stations, and tourist attractions.

   For businesses, from the entity status description of the official records we retain those that are ”live”, and then focus on three main industry divisions, based on the SSIC (Singapore Standard Industrial Classification), that account for a large portion of the employment force: construction (SSIC 41), manufacturing (SSIC 10), and services (SSIC 46, 47, 49). For records with street address, we map them to postcodes and then crosswalk to subzones as described in note 6.

8. Daily rainfall (mm) records come from the MSS (Meteorological Service Singapore). To derive rainfall for each `subzone-date` we map each subzone to the nearest recorded weather station using the centroid of the subzone. Our data includes 46 weather stations on record (Fig. S3). We apply linear interpolation for days where the rainfall record is missing. Temperature and wind are available only for certain weather stations are thus not included.

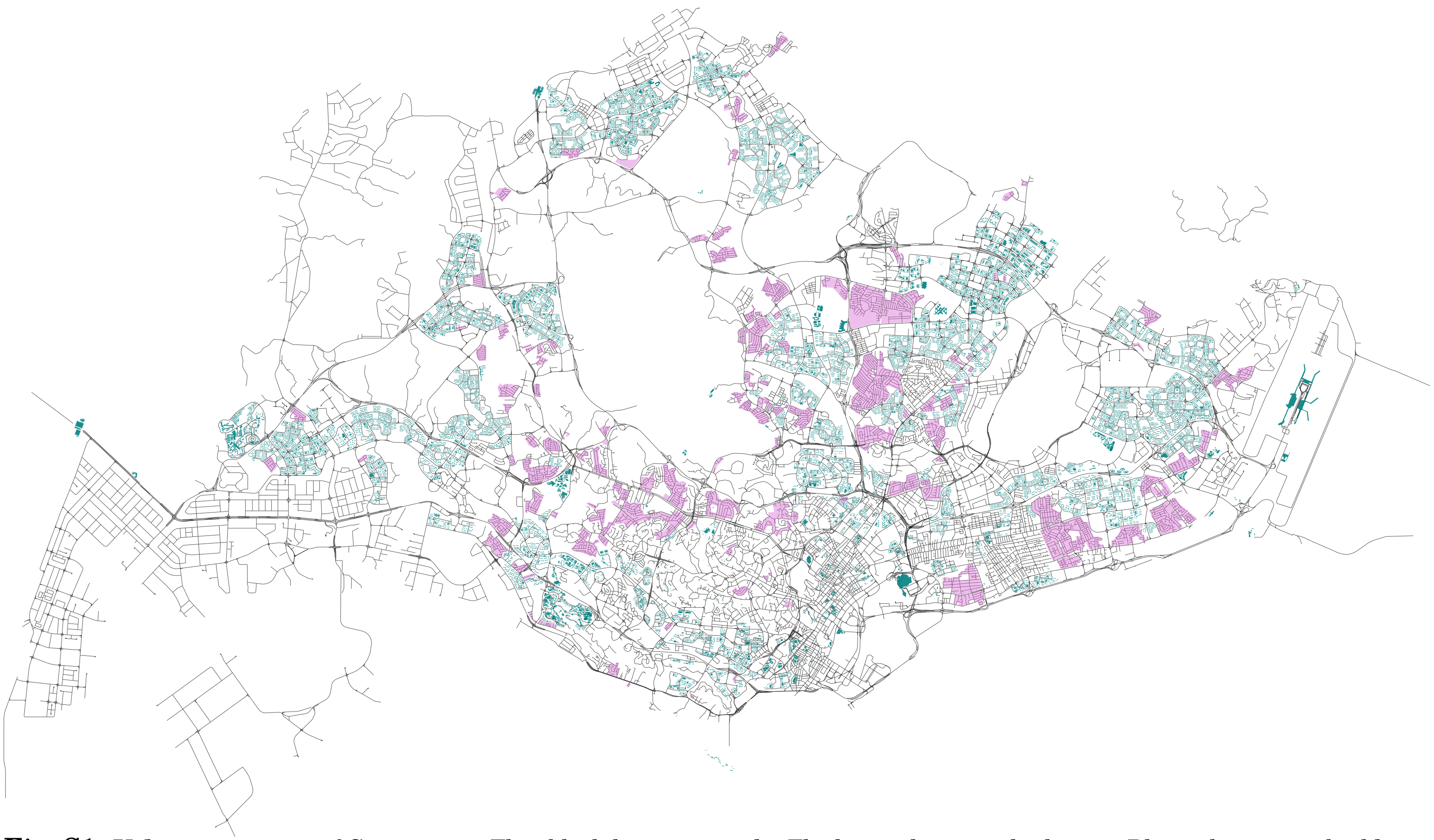

Fig. S1. Urban overview of Singapore. Thin black lines are roads. Thick gray lines are highways. Blue polygons are buildings, which include the public housing flats. Pink polygons are the single-family houses with dedicated landspace.

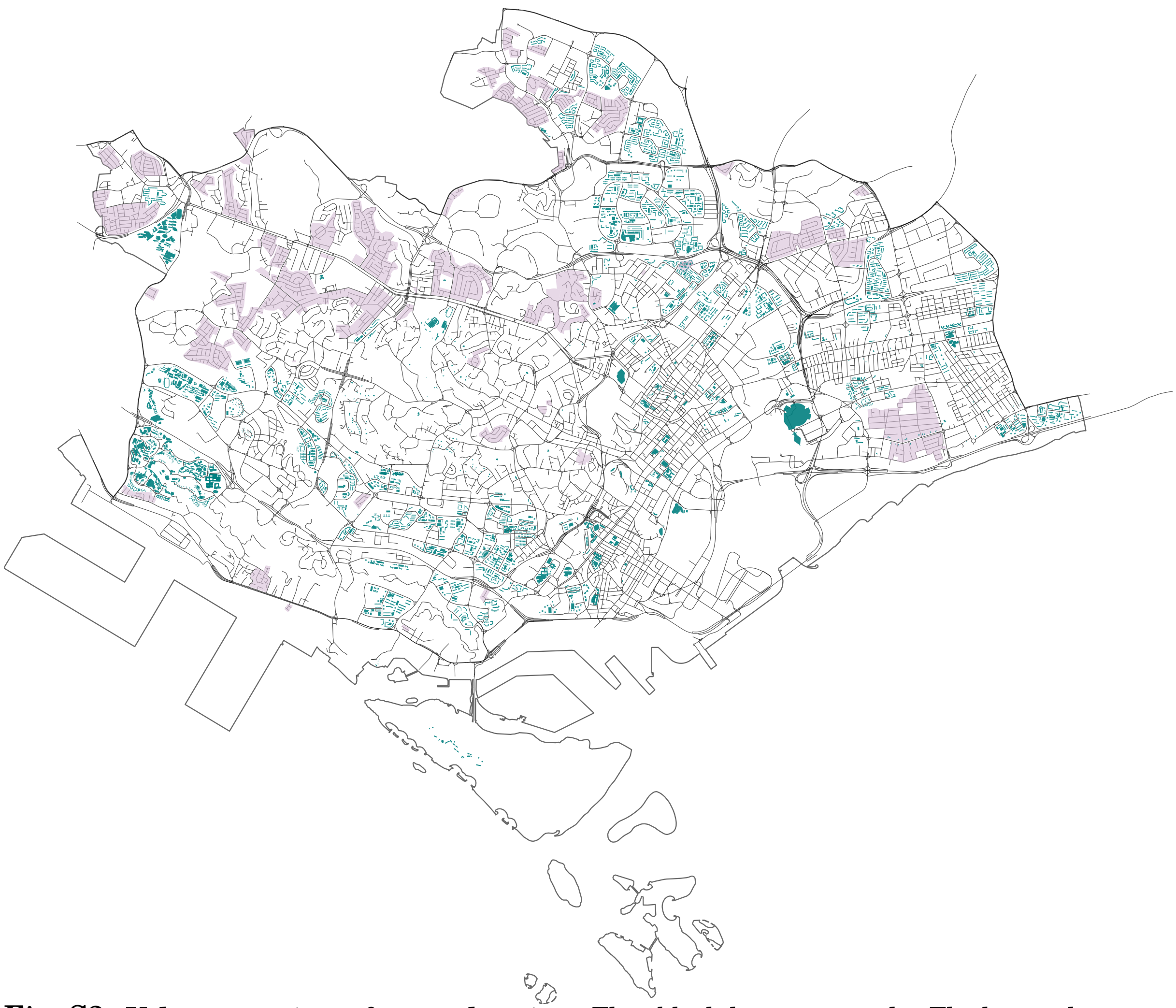

Fig. S2. Urban overview of central region. Thin black lines are roads. Thick gray lines are highways. Blue polygons are buildings, which include the public housing flats. Pink polygons are the single-family houses with dedicated landspace. Corresponds to Fig. 3 and Fig. S1

Table S1. Distribution of house types for the selected neighborhoods. Resident count and percentage of residents living in the different types of houses. Corresponds to Fig. 3.

	Housing type (%)					
	Public					
Neighborhoods	**1–2-room flat**	**3-room flat**	**4-room flat**	**5-room flat**	**Private**	**Residents**
Crawford	49.21	26.9	8.94	8.26	6.67	8,840
Alexandra Hill	29.29	18.88	28.46	23.37	0	2,210
Hillcrest	0	0	0	0	100	9,010
Katong	0	0	0	0	100	9,160

Table S2. Effect of poverty on movement patterns. Coefficients estimated from Eq. (1). Standard errors are clustered at the origin-by-destination census planning area level. One-sided tests with the stated alternative hypothesis H_a are also reported. *** Significant at the 1 per cent level. ** Significant at the 5 per cent level. * Significant at the 10 per cent level. Table corresponds to Fig. 4.

	Dependent variable: log(inflow) into neighborhood				
Independent variable	**1**	**2**	**3**	**4**	**5**
P_o (Poverty, origin neighborhood)	-0.0246	-0.0837^{*}	-0.0826^{*}	-0.0826^{*}	-0.0830^{*}
	(0.0644)	(0.0452)	(0.0451)	(0.0451)	(0.0450)
P_d (Poverty, destination neighborhood)	0.0071	-0.0057	-0.0584	-0.0579	-0.0835^{*}
	(0.0659)	(0.0477)	(0.0478)	(0.0477)	(0.0466)
$H_a : P_o < 0$, p-val	.352	$.032^{**}$	$.034^{**}$	$.033^{**}$	$.033^{**}$
$H_a : P_d < 0$, p-val	.543	.452	.111	.112	$.037^{**}$
Day fixed effects	✓	✓	✓	✓	✓
Census area-by-WoY fixed effects		✓	✓	✓	✓
Places of interests			✓	✓	✓
Neighborhood-specific rainfall				✓	✓
Businesses in neighborhood					✓
R^2	0.7222	0.7329	0.7348	0.7348	0.7350
Days	91	91	91	91	91
Clusters	1,434	1,434	1,434	1,434	1,434
Observations	1,261,171	1,261,171	1,261,171	1,261,171	1,261,171

Table S3. Effect of wealth on movement patterns. Coefficients estimated from Eq. (1). Standard errors are clustered at the origin-by-destination census planning area level. One-sided tests with the stated alternative hypothesis H_a are also reported. *** Significant at the 1 per cent level. ** Significant at the 5 per cent level. * Significant at the 10 per cent level. Table corresponds to Fig. 4.

	Dependent variable: log(inflow) into neighborhood				
Independent variable	**1**	**2**	**3**	**4**	**5**
W_o (Wealth, origin neighborhood)	-0.0782^{***}	-0.0289^{*}	-0.0292^{*}	-0.0294^{*}	-0.0295^{*}
	(0.0166)	(0.0174)	(0.0174)	(0.0174)	(0.0174)
W_d (Wealth, destination neighborhood)	0.0237	0.0450^{***}	0.0538^{***}	0.0535^{***}	0.0646^{***}
	(0.0155)	(0.0139)	(0.0158)	(0.0158)	(0.0161)
$H_a : W_o > 0$, p-val	1	.952	.9537	.9546	.9551
$H_a : W_d > 0$, p-val	$.0631^{*}$	$.0006^{***}$	$.0003^{***}$	$.0004^{***}$	0^{***}
Day fixed effects	✓	✓	✓	✓	✓
Census area-by-WoY fixed effects		✓	✓	✓	✓
Places of interests			✓	✓	✓
Neighborhood-specific rainfall				✓	✓
Businesses in neighborhood					✓
R^2	0.7225	0.7330	0.7349	0.7349	0.7352
Days	91	91	91	91	91
Clusters	1,434	1,434	1,434	1,434	1,434
Observations	1,261,171	1,261,171	1,261,171	1,261,171	1,261,171

Table S4. Effect of poverty *and* wealth on movement patterns. Coefficients estimated from Eq. (1). Standard errors are clustered at the origin-by-destination census planning area level. One-sided tests with the stated alternative hypothesis H_a are also reported. *** Significant at the 1 per cent level. ** Significant at the 5 per cent level. * Significant at the 10 per cent level.

	Dependent variable: log(inflow) into neighborhood				
Independent variable	**1**	**2**	**3**	**4**	**5**
P_o (Poverty, origin neighborhood)	-0.122^{*}	-0.105^{**}	-0.104^{**}	-0.104^{**}	-0.145^{***}
	(0.064)	(0.045)	(0.045)	(0.045)	(0.045)
P_d (Poverty, destination neighborhood)	0.032	0.017	-0.024	-0.024	-0.046
	(0.065)	(0.047)	(0.047)	(0.047)	(0.045)
W_o (Wealth, origin neighborhood)	-0.086^{***}	-0.034^{*}	-0.034^{*}	-0.034^{*}	-0.033^{*}
	(0.016)	(0.017)	(0.017)	(0.017)	(0.019)
W_d (Wealth, destination neighborhood)	0.026	0.045^{***}	0.052^{***}	0.052^{***}	0.077^{***}
	(0.016)	(0.014)	(0.016)	(0.016)	(0.017)
$H_a : P_o < 0$, p-val	$.028^{**}$	$.011^{**}$	$.011^{**}$	$.011^{**}$	$.001^{***}$
$H_a : P_d < 0$, p-val	.69	.644	.304	.306	.153
$H_a : W_o > 0$, p-val	1	.973	.974	.975	.96
$H_a : W_d > 0$, p-val	$.051^{*}$	$.001^{***}$	$.001^{***}$	$.001^{***}$	0^{***}
Day fixed effects	✓	✓	✓	✓	✓
Census area-by-WoY fixed effects		✓	✓	✓	✓
Places of interests			✓	✓	✓
Neighborhood-specific rainfall				✓	✓
Businesses in neighborhood					✓
R^2	0.723	0.733	0.735	0.735	0.770
Days	91	91	91	91	91
Clusters	1,434	1,434	1,434	1,434	1,414
Observations	1,261,171	1,261,171	1,261,171	1,261,171	1,260,594

A

B

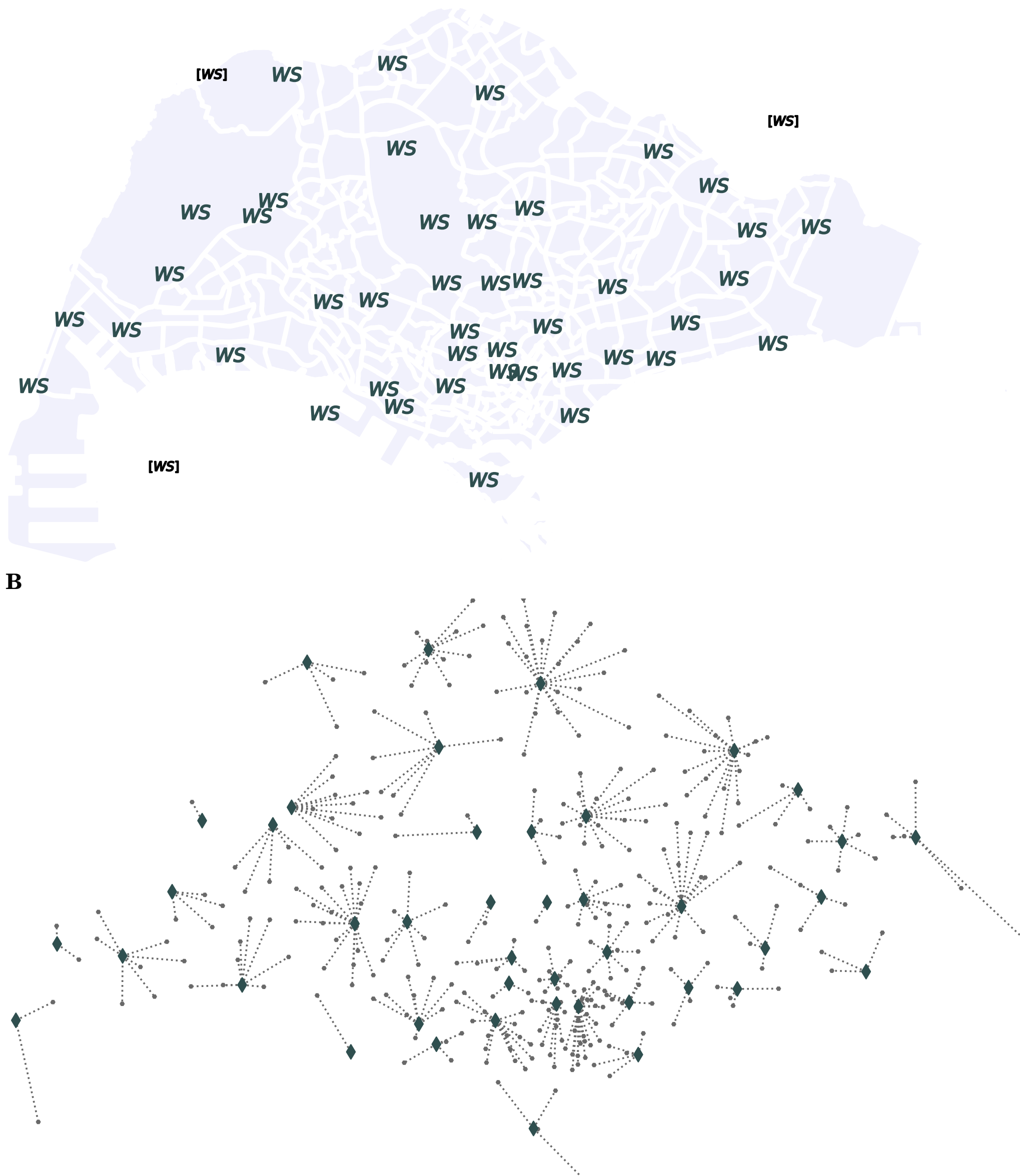

Fig. S3. Assignment of weather stations to neighborhoods (**A**) Location of weather stations in Singapore. The three weather stations that end up with no area mapping is shown in brackets. (**B**) Matches encoded by dotted lines, connecting weather stations to neighborhoods.

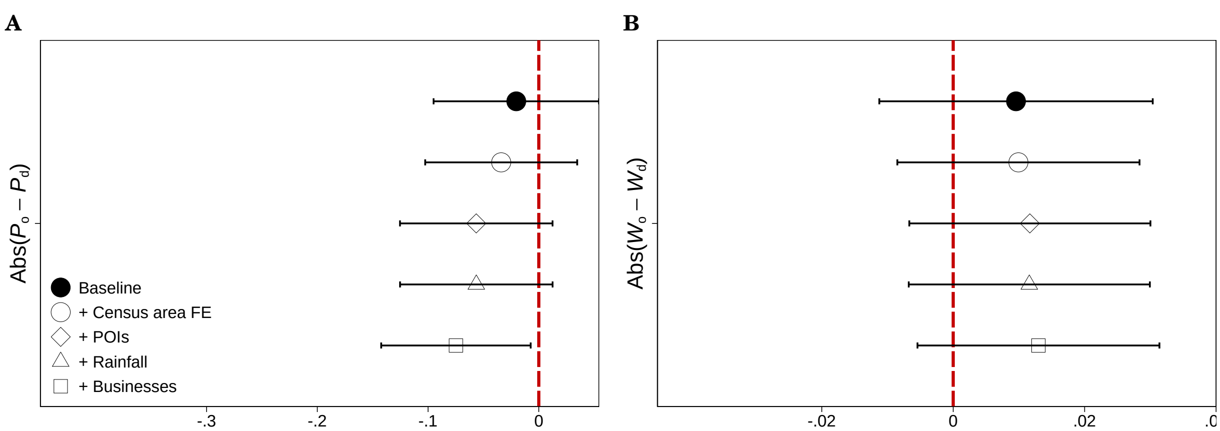

Fig. S4. Effect of absolute differences of poverty and wealth on movement. (**A**) Plots of the estimated coefficient of the absolute difference in poverty measure between origin and destination neighborhoods from estimating Eq. (1). (**B**) Similar, except that the poverty measure is now swapped out for the wealth measure. Dependent variable is log of inflow from origin o to destination d. Capped horizontal lines are 90% confidence intervals from standard errors clustered at the origin-by-destination area level.

Table S5. Effect of absolute difference in poverty and wealth on movement patterns. Coefficients estimated from Eq. (1). Standard errors are clustered at the origin-by-destination census planning area level. One-sided tests with the stated alternative hypothesis H_a are also reported. *** Significant at the 1 per cent level. ** Significant at the 5 per cent level. * Significant at the 10 per cent level. Table corresponds to Fig. S4.

	Dependent variable: log(inflow) into neighborhood				
Independent variable	**1**	**2**	**3**	**4**	**5**
Abs $(P_o - P_d)$	-0.0203	-0.0340	-0.0565	-0.0564	-0.0748^{*}
	(0.0454)	(0.0417)	(0.0418)	(0.0418)	(0.0410)
H_a : Abs($P_o - P_d$) < 0, p-val	.327	.208	$.089^{*}$	$.089^{*}$	$.034^{**}$
Day fixed effects	✓	✓	✓	✓	✓
Census area-by-WoY fixed effects		✓	✓	✓	✓
Places of interests			✓	✓	✓
Neighborhood-specific rainfall				✓	✓
Businesses in neighborhood					✓
R^2	0.7222	0.7329	0.7348	0.7348	0.7350
Days	91	91	91	91	91
Clusters	1,434	1,434	1,434	1,434	1,434
Observations	1,261,171	1,261,171	1,261,171	1,261,171	1,261,171
Abs $(W_o - W_d)$	0.0096	0.0099	0.0117	0.0116	0.0130
	(0.0126)	(0.0112)	(0.0111)	(0.0111)	(0.0112)
H_a : Abs($W_o - W_d$) < 0, p-val	.775	.812	.852	.85	.877
Day fixed effects	✓	✓	✓	✓	✓
Census area-by-WoY fixed effects		✓	✓	✓	✓
Places of interests			✓	✓	✓
Neighborhood-specific rainfall				✓	✓
Businesses in neighborhood					✓
R^2	0.7222	0.7329	0.7348	0.7348	0.7350
Days	91	91	91	91	91
Clusters	1,434	1,434	1,434	1,434	1,434
Observations	1,261,171	1,261,171	1,261,171	1,261,171	1,261,171

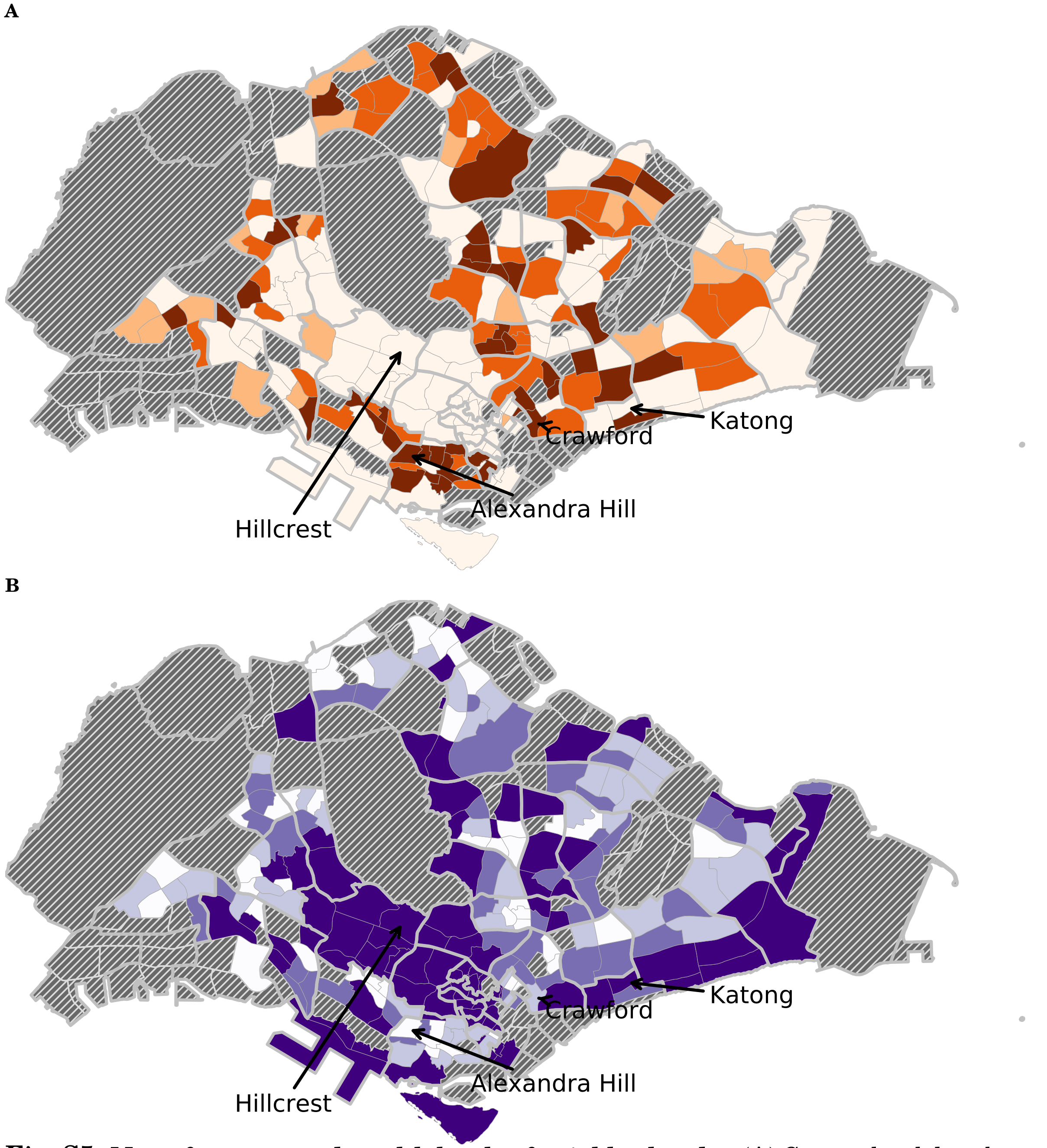

Fig. S5. Map of poverty and wealth levels of neighborhoods. (**A**) Geographical distribution of poverty levels, measured by proportion of residents living in 1–2-room public flats. (**B**) Geographical distribution of wealth levels, measured by proportion of residents living in private residential properties. Includes 219 of the nbh/census subzones. Non-residential subzones, or subzones without residential records, are shaded out. The four neighborhoods annotations corresponds to Fig. 3.