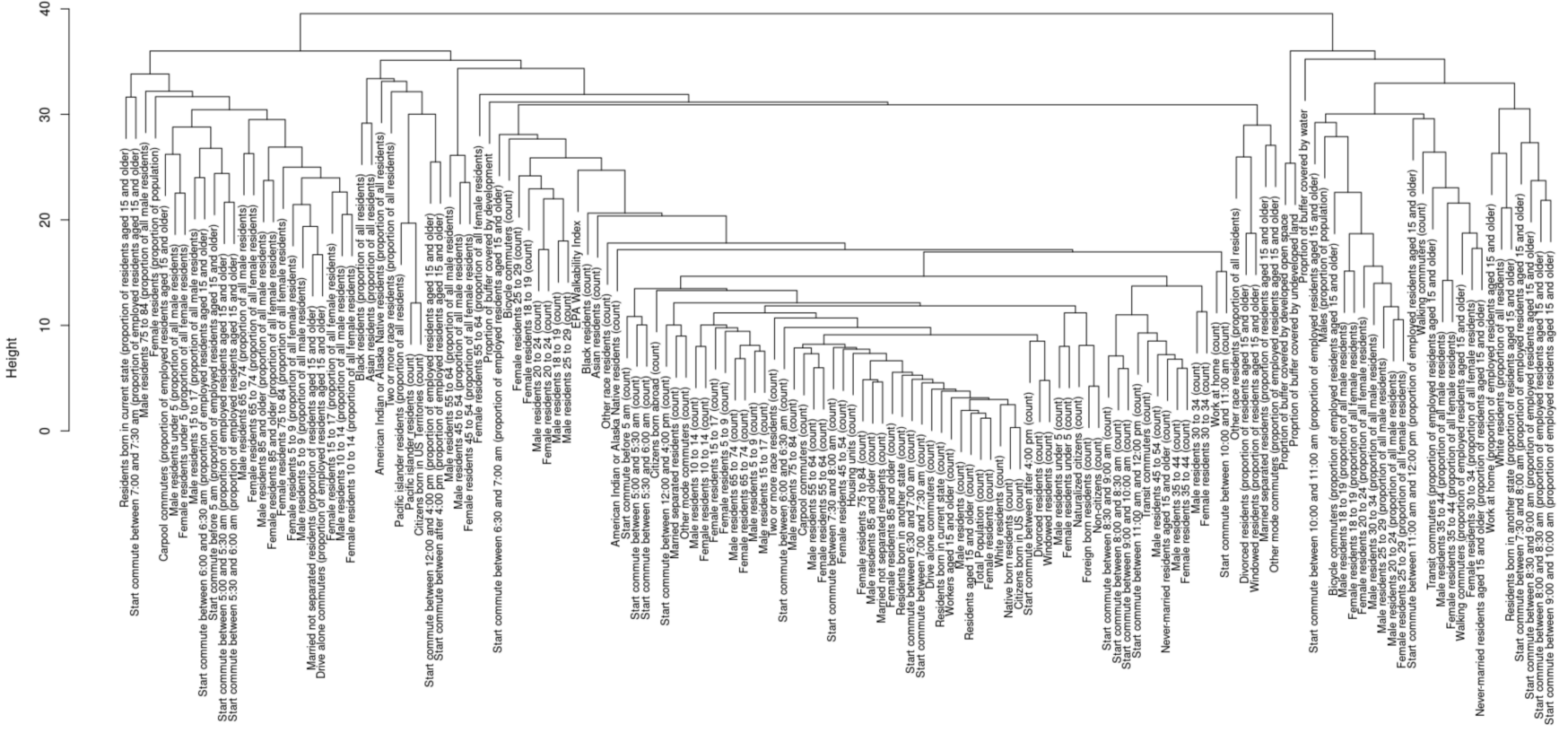


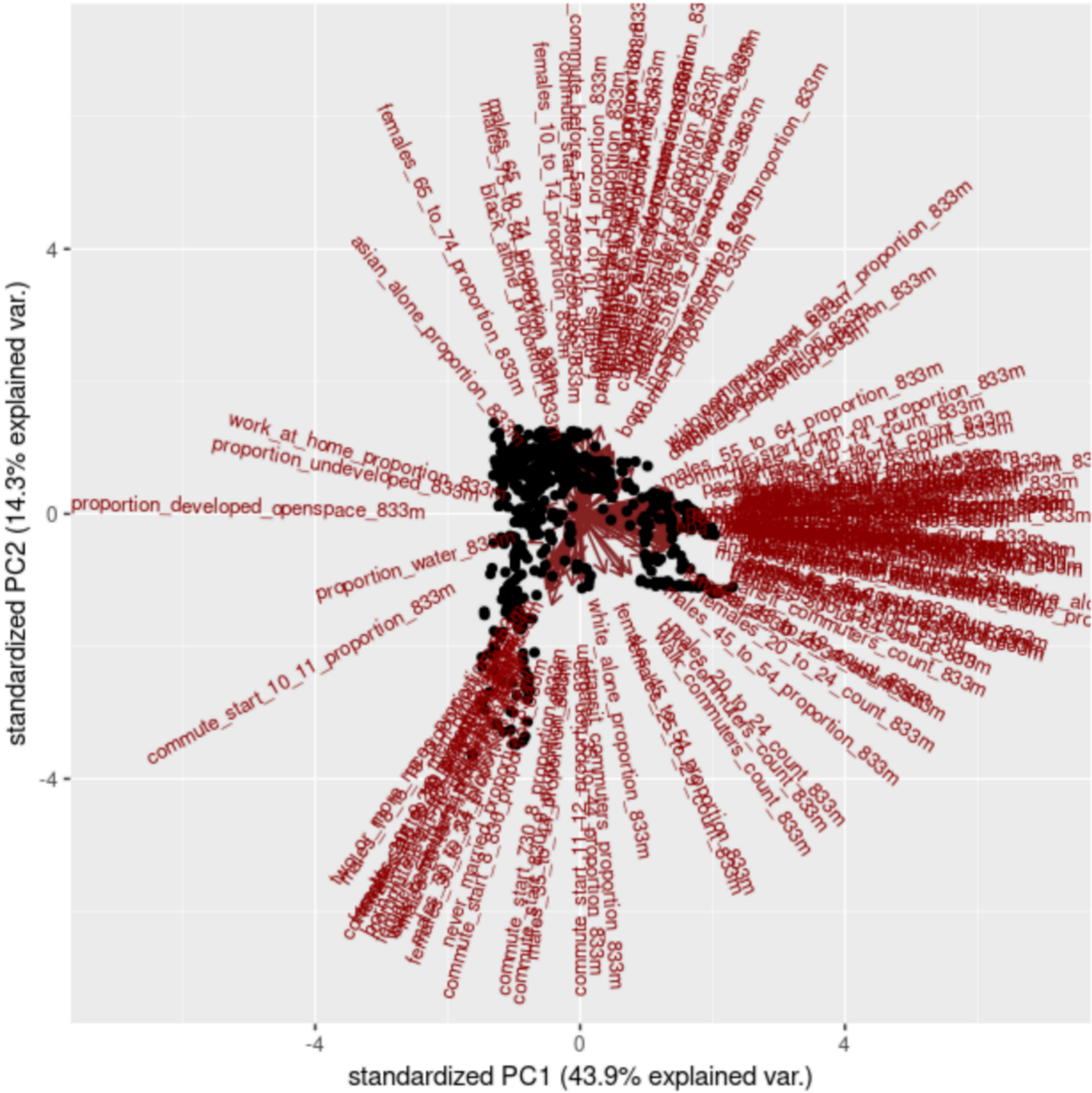
Appendices

Web Figure 1: Hierarchical Clustering of 146 neighborhood measures

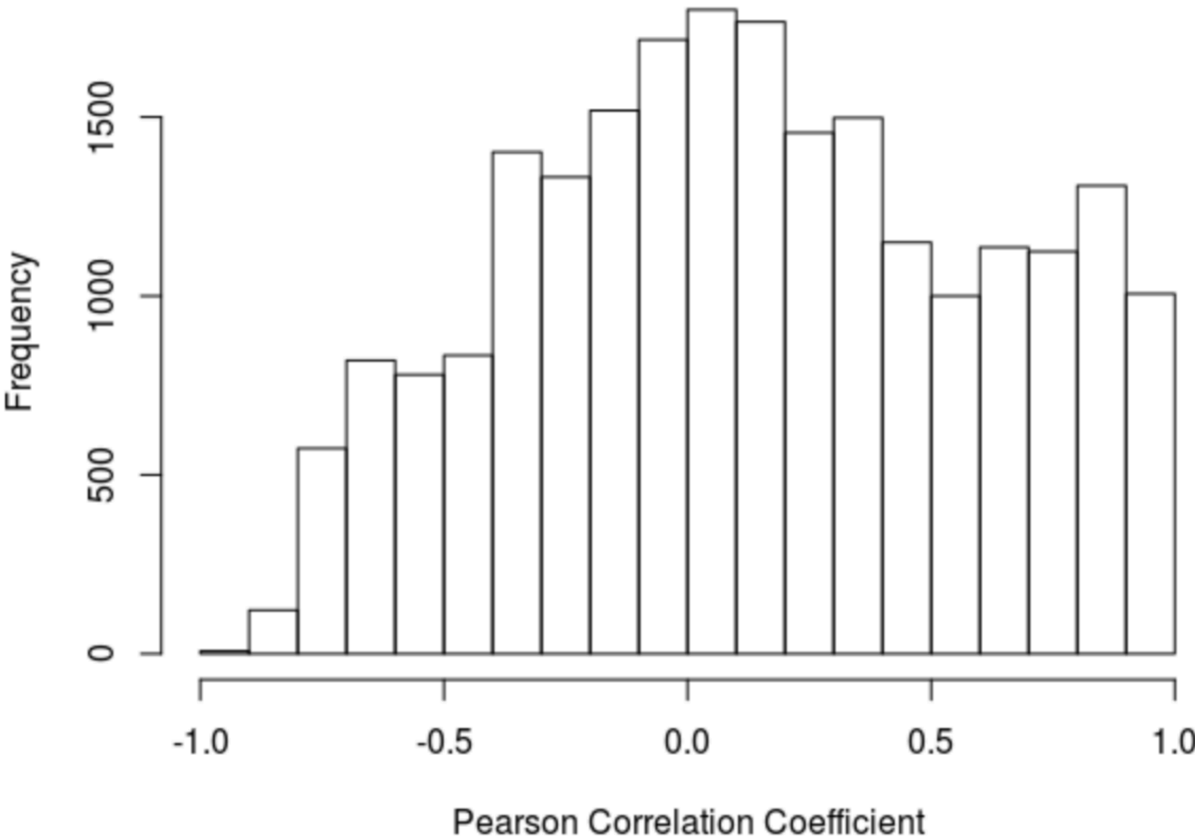


Web Figure 2. Plot of variable loadings on the first two principal components of 146 neighborhood measures

The high degree of inter-correlation among neighborhood measures suggested walking prediction might benefit from dimensionality reduction. Principal component analysis revealed that 43.6% of variation in the neighborhood measures was explained by the first principal component. Measures that loaded positively onto that component (68 measures had loadings above 0.8) generally described features of denser housing conditions. The most strongly positively loading measure on the component was the estimate of people residing in the buffer (loading = 0.995) whereas the most negatively loading was proportion of land dedicated to open space (loading = -0.649). The second principal component, which explained a further 14.3% of the variation, was roughly interpretable as describing the pervasiveness of traditional transportation and household patterns, with proportion of drive alone commuters (loading = 0.87) and proportion of households headed by married couples (loading = 0.81) loading most positively.



Web Figure 3. Histogram of Pearson correlation coefficients between each pair of neighborhood measures collected for TRAC participants at baseline. The quantity of measure pairs with correlation coefficients between 0.5 and 1.0 indicates there is high risk of over-fitting artifacts due to predictor multi-collinearity in multivariable models including more than one of these neighborhood measures.



Web Table 1: Measures incorporated in the ACMT, sorted by association with count of walk bouts

Measure	Direction of Association with walk bouts	Negative Log of p-value	Principal component ¹
White residents (count)	+	21.9	1
Male residents (count)	+	21.4	1
Residents aged 15 and older (count)	+	21.3	1
Citizens born in US (count)	+	21.2	1
Native born residents (count)	+	21.1	1
Residents born in another state (count)	+	21.1	1
Total Population (count)	+	21.1	1
Residents born in current state (count)	+	20.9	1
Workers aged 15 and older (count)	+	20.9	1
Housing units (count)	+	20.7	1
Female residents (count)	+	20.7	1
Start commute between 7:00 and 7:30 am (count)	+	20.4	1
Start commute between 6:30 and 7:00 am (count)	+	20.4	1
Drive alone commuters (count)	+	20.4	1
Married not separated residents (count)	+	20.3	1
Start commute between 7:30 and 8:00 am (count)	+	20.0	1
Male residents 85 and older (count)	+	19.9	1
Female residents 85 and older (count)	+	19.7	1
Male residents 5 to 9 (count)	+	19.7	1
Male residents 65 to 74 (count)	+	19.4	1
Start commute between 9:00 and 10:00 am (count)	+	19.3	1
Female residents 65 to 74 (count)	+	19.2	1
Female residents 45 to 54 (count)	+	19.2	1
Start commute between 8:00 and 8:30 am (count)	+	19.2	1
Never-married residents aged 15 and older (count)	+	19.1	1
Start commute between after 4:00 pm (count)	+	19.0	1
Divorced residents (count)	+	18.8	1
Male residents 45 to 54 (count)	+	18.7	1
Start commute between 8:30 and 9:00 am (count)	+	18.6	1
Female residents 75 to 84 (count)	+	18.6	1
Carpool commuters (count)	+	18.5	1
Citizens born abroad (count)	+	18.4	1
Windowed residents (count)	+	18.2	1
Male residents 15 to 17 (count)	+	18.1	1
Male residents 75 to 84 (count)	+	17.9	1
Male residents 55 to 64 (count)	+	17.9	1
Work at home (count)	+	17.7	1

Foreign born residents (count)	+	17.6	1
Female residents 55 to 64 (count)	+	17.6	1
EPA Walkability Index	+	17.5	1
Start commute between 12:00 and 4:00 pm (count)	+	17.5	1
Female residents 15 to 17 (count)	+	17.3	1
Non-citizens (count)	+	17.2	1
Start commute between 6:00 and 6:30 am (count)	+	17.2	1
Female residents 5 to 9 (count)	+	17.1	1
Two or more race residents (count)	+	17.0	1
Naturalized citizens (count)	+	16.8	1
Male residents 35 to 44 (count)	+	16.6	1
Male residents under 5 (count)	+	16.4	1
Start commute between 5:00 and 5:30 am (count)	+	16.4	1
Female residents 30 to 34 (count)	+	16.1	1
Female residents under 5 (count)	+	16.1	1
Start commute between 5:30 and 6:00 am (count)	+	15.9	1
Male residents 10 to 14 (count)	+	15.8	1
Married separated residents (count)	+	15.7	1
Start commute between 11:00 am and 12:00 pm (count)	+	15.5	1
Female residents 10 to 14 (count)	+	15.4	1
Start commute between 10:00 and 11:00 am (count)	+	15.3	1
Male residents 18 to 19 (count)	+	15.2	1
Other mode commuters (count)	+	15.1	1
Female residents 35 to 44 (count)	+	15.0	1
Male residents 30 to 34 (count)	+	14.6	1
Transit commuters (count)	+	14.6	1
American Indian or Alaska Native residents (count)	+	13.9	1
Male residents 20 to 24 (count)	+	13.9	2
Start commute before 5 am (count)	+	13.7	1
Other race residents (count)	+	12.4	1
Asian residents (count)	+	11.9	1
Proportion of buffer covered by developed open space	-	11.9	1
Male residents 25 to 29 (count)	+	10.2	1
Black residents (count)	+	8.5	1
Bicycle commuters (count)	+	7.2	1
Female residents 20 to 24 (count)	+	6.6	1
Female residents 25 to 29 (count)	+	5.9	2
Walking commuters (count)	+	5.5	2
Female residents 18 to 19 (count)	+	5.5	3
Start commute between 6:30 and 7:00 am (proportion of employed residents aged 15 and older)	+	5.3	1
Start commute between 10:00 and 11:00 am (proportion of employed residents aged 15 and older)	-	5.2	1
Proportion of buffer covered by undeveloped land	-	4.0	1

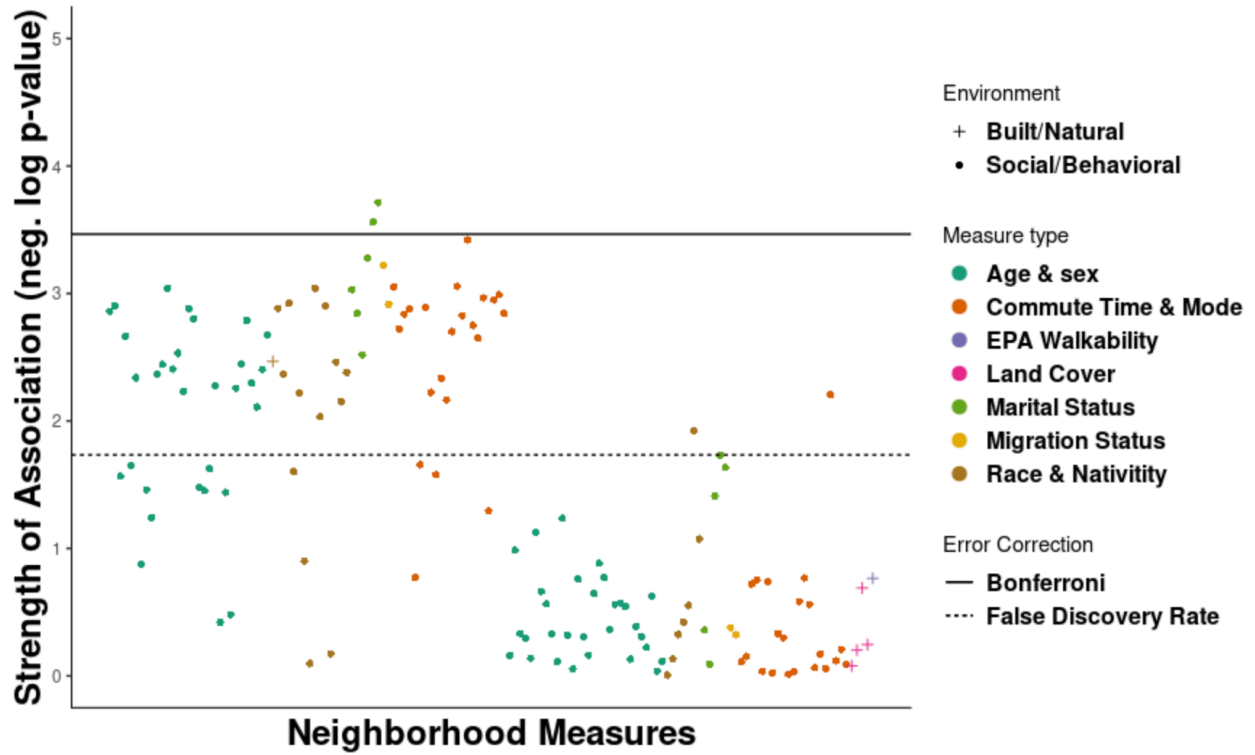
Asian residents (proportion of all residents)	-	3.5	4
White residents (proportion of all residents)	+	3.3	1
Male residents 75 to 84 (proportion of all male residents)	-	2.9	NA
Black residents (proportion of all residents)	-	2.5	NA
Two or more race residents (proportion of all residents)	-	2.4	NA
Citizens born in US Territories (count)	+	2.3	NA
Start commute before 5 am (proportion of employed residents aged 15 and older)	-	2.1	NA
Female residents 65 to 74 (proportion of all female residents)	-	1.8	NA
Residents born in current state (proportion of residents aged 15 and older)	+	1.7	NA
Male residents 45 to 54 (proportion of all male residents)	+	1.6	NA
Female residents 10 to 14 (proportion of all female residents)	-	1.6	NA
Other mode commuters (proportion of employed residents aged 15 and older)	+	1.6	NA
Female residents under 5 (proportion of all female residents)	-	1.6	NA
Windowed residents (proportion of residents aged 15 and older)	+	1.5	NA
Other race residents (proportion of all residents)	+	1.5	NA
Pacific islander residents (proportion of all residents)	-	1.5	NA
Female residents 20 to 24 (proportion of all female residents)	-	1.4	NA
Male residents 5 to 9 (proportion of all male residents)	+	1.4	NA
Male residents under 5 (proportion of all male residents)	-	1.3	NA
Start commute between 5:00 and 5:30 am (proportion of employed residents aged 15 and older)	+	1.2	NA
Start commute between 12:00 and 4:00 pm (proportion of employed residents aged 15 and older)	-	1.2	NA
Married separated residents (proportion of residents aged 15 and older)	+	1.1	NA
Divorced residents (proportion of residents aged 15 and older)	+	1.1	NA
Female residents 85 and older (proportion of all female residents)	+	1.1	NA
Start commute between 6:00 and 6:30 am (proportion of employed residents aged 15 and older)	-	1.0	NA
Pacific islander residents (count)	+	1.0	NA
Proportion of buffer covered by development	+	1.0	NA

American Indian or Alaska Native residents (proportion of all residents)	+	1.0	NA
Walking commuters (proportion of employed residents aged 15 and older)	-	0.9	NA
Male residents 10 to 14 (proportion of all male residents)	-	0.9	NA
Start commute between 7:30 and 8:00 am (proportion of employed residents aged 15 and older)	+	0.9	NA
Female residents 18 to 19 (proportion of all female residents)	-	0.8	NA
Carpool commuters (proportion of employed residents aged 15 and older)	-	0.8	NA
Male residents 15 to 17 (proportion of all male residents)	+	0.7	NA
Never-married residents aged 15 and older (proportion of residents aged 15 and older)	-	0.6	NA
Start commute between 8:30 and 9:00 am (proportion of employed residents aged 15 and older)	+	0.6	NA
Start commute between 9:00 and 10:00 am (proportion of employed residents aged 15 and older)	-	0.6	NA
Work at home (proportion of employed residents aged 15 and older)	+	0.6	NA
Male residents 55 to 64 (proportion of all male residents)	-	0.6	NA
Female residents 30 to 34 (proportion of all female residents)	-	0.6	NA
Transit commuters (proportion of employed residents aged 15 and older)	+	0.6	NA
Female residents 55 to 64 (proportion of all female residents)	+	0.5	NA
Male residents 18 to 19 (proportion of all male residents)	-	0.5	NA
Male residents 25 to 29 (proportion of all male residents)	-	0.4	NA
Male residents 30 to 34 (proportion of all male residents)	-	0.4	NA
Start commute between 7:00 and 7:30 am (proportion of employed residents aged 15 and older)	+	0.4	NA
Residents born in another state (proportion of residents aged 15 and older)	-	0.3	NA
Male residents 65 to 74 (proportion of all male residents)	-	0.3	NA
Start commute between 5:30 and 6:00 am (proportion of employed residents aged 15 and older)	-	0.2	NA

Female residents 25 to 29 (proportion of all female residents)	-	0.2	NA
Female residents 75 to 84 (proportion of all female residents)	+	0.2	NA
Bicycle commuters (proportion of employed residents aged 15 and older)	-	0.2	NA
Female residents 35 to 44 (proportion of all female residents)	+	0.2	NA
Male residents 20 to 24 (proportion of all male residents)	-	0.2	NA
Proportion of buffer covered by water	+	0.2	NA
Drive alone commuters (proportion of employed residents aged 15 and older)	+	0.2	NA
Start commute between after 4:00 pm (proportion of employed residents aged 15 and older)	+	0.2	NA
Married not separated residents (proportion of residents aged 15 and older)	+	0.1	NA
Female residents 15 to 17 (proportion of all female residents)	-	0.1	NA
Start commute between 11:00 am and 12:00 pm (proportion of employed residents aged 15 and older)	-	0.1	NA
Start commute between 8:00 and 8:30 am (proportion of employed residents aged 15 and older)	-	0.1	NA
Female residents (proportion of population)	+	0.1	NA
Males (proportion of population)	-	0.1	NA
Female residents 5 to 9 (proportion of all female residents)	-	0.0	NA
Female residents 45 to 54 (proportion of all female residents)	-	0.0	NA
Male residents 35 to 44 (proportion of all male residents)	-	0.0	NA

¹ Component, if any, to which this measure was assigned in the NWAS-inspired Principal Component Analysis described in Web Appendix #1

Web Figure 4: Manhattan Plot showing NE-WAS results for analysis relating neighborhood exposures to total time spent in walking bouts



Web Appendix 1. Alternate (NWS) variable selection approach

As an alternate analytic strategy, we followed Lynch, et al., by performing a principal component analysis of all neighborhood variables to identifying variable clusters (analogous to haplotype blocks in a genetic analysis) that together explain neighborhood measure variance.

Next, we assigned each variable to the component on which its coefficient was highest. We then identified the ‘top hits’ – the individual variable whose bivariate coefficient significance test most conclusively ruled out the null association (i.e. had the highest negative log p-value) – for the first four components.

These four components we chosen to explain nearly 90% of the variance. Our inspiration, the Lynch, et al. analysis, selected variables that explained 90% of the variance. We stopped at the four components that explained 88% of the variance because no variable contributed more to component 5 or 6 than to one of the first four components.

The first component, which included 75 of the 80 measures, was defined by strong positive loadings related to population counts. The second component exhibited a distinct loading pattern among the age group counts wherein counts of residents in groups comprising 17 and under and over 55 loaded positively whereas those from 18-54 loaded negatively. Many of those measures loaded more strongly onto the population count measure and so were assigned there for this analysis; nonetheless, we named the second component age distribution to reflect this pattern. The third and fourth components each were defined by a single group.

The components ranked by variance explained were

Neighborhood Component	Number of measures	Variance Explained	Cumulative Variance Explained
Population Counts	75	76%	76%
Age Distribution	3	5%	82%
Young Female Population	1	4%	85%
Asian Population	1	2%	88%

And the variables selected to represent each component, with associations, were

Neighborhood Component	Representative Variable	(Confidence Interval)
Population Counts	Housing Unit Count	1.38 (1.29, 1.47)
Age Distribution	Count Females 25-29	1.18 (1.10, 1.26)
Young Female Population	Count Females 18-19	1.17 (1.09, 1.25)
Asian Population	Proportion Asian-American	0.87 (0.81, 0.94)

We conclude from this approach that housing unit counts remain the best overall predictor of walking, and that, generally, other factors strongly associated with residential density, including density of population subgroups and commute mode choice, are good candidate proxy measures for walkability as well.