H + T Index Methods

August, 2017

H+T Index Methods

Table of Contents

Table of Tables

Table of Figures

Table of Equations

Introduction

The Center for Neighborhood Technology's Housing + Transportation (H+T®) Affordability Index (H+T Index) is an innovative tool that measures the true affordability of housing by calculating the transportation costs associated with a home's location. Planners, lenders, and most consumers traditionally consider housing affordable if the cost is 30 percent or less of household income. The H+T Index proposes expanding the definition of housing affordability to include transportation costs at a home's location to better reflect the true cost of households' location choices. Based on research in metro areas ranging from large cities with extensive transit to small metro areas with extremely limited transit options, CNT has found 15 percent of income to be an attainable goal for transportation affordability. By combining this 15 percent level with the 30 percent housing affordability standard, the H+T Index recommends a new view of affordability defined as combined housing and transportation costs consuming no more than 45 percent of household income.

The H+T Index was constructed to estimate three dependent variables (auto ownership, auto use, and transit use) as functions of 14 independent variables (median household income, average household size, average commuters per household, gross household density, Regional Household Intensity, fraction of rental housing units, fraction of single family detached housing, Employment Access Index, Employment Mix Index, block density, Transit Connectivity Index, Total Available Transit Trips per Week, Transit Access Shed, and Jobs within the Transit Access Shed). To hone in on the built environment's influence on transportation costs, the independent household variables (income, household size, and commuters per household) are set at fixed values to control for any variation they might cause. By establishing and running the model for a "typical household" any variation observed in transportation costs is due to place and location, not household characteristics.

Differences in the new Transportation Cost Model

Several improvements have been made to the H+T Index including the addition of rural block groups, an improvement of the transportation cost model and an improved method to derive household transit costs. Previous versions of the H+T only analyzed transportation costs for Census Block Groups in core based statistical areas (metropolitan and micropolitan area). In this version CNT has expanded that coverage to all Census Block Groups in the states and District of Columbia.

In the previous version CNT changed from using a non-linear regression technique that made the model difficult to understand to a model that uses the ordinary least square (OLS) and simple variable transformations to accomplish the regression. In this version CNT still uses the OLS, but has added interactive variables; these interaction variables better capture the relationships between independent variables.

The AllTransit™ database, developed to include all transit agencies allows CNT to have only one model, in previous versions there was model where transit data was available and one where it was not. Using

one model and the interaction variables allows for the transportation behavior modeling in the entire country. Another benefit of the AllTransit database is that it allowed for better assignment of transit cost by county, when it was synced to the National Transit Database from FTA.

Geographic Level and Data Availability

The H+T Index was constructed at the Census block group level. Currently the H+T Index covers all 220,319 Census Block Groups in the US. Due to incompatible and insufficient data Puerto Rico was excluded.

Data Sources

The H+T Index uses data from a combination of Federal sources and transit data compiled by the Center for Neighborhood Technology.

- 2011-2015 American Community Survey 5-year Estimate (2015 ACS) – an ongoing U.S. Census survey that generates data on housing characteristic, transportation use, community demographics, income, and employment.
- U.S. Census TIGER/Line Files – geographical features such as roads, railroads, and rivers, as well as legal and statistical geographic areas.

U.S. Census Longitudinal Employment-Household Dynamics (LEHD) Origin-Destination Employment Statistics (LODES) – detailed spatial distributions of workers' employment and residential locations and the relation between the two at the Census Block level and characteristic detail on age, earnings, industry distributions, and local workforce indicators. LODES data built on 2014 Census data are used here (Wyoming does not have 2014 LEHD data, 2013 data was used, assuming that the differences would be minimal).

- Average annual expenditures and characteristics of all consumer units, from the Consumer Expenditure Survey, 2006-2012 and 2013, used to inflate the cost of auto ownership from the 2010 data above.
- 2015 National Transit Database –fare box revenue and number of transit trips reported by agencies that receive federal assistance.
- AllTransit™ – a database of General Transit Feed Specification (GTFS) data developed by the Center for Neighborhood Technology, including bus, rail, and ferry service for both transit agencies that report their GTFS data publicly and those derived by CNT staff for agencies that do not.
- Odometer readings from The Illinois Department of Natural Resources - odometer data collected by Vehicle Emissions Testing Program.

Housing Costs

To calculate the H in the H+T Index, housing costs are derived from nationally available datasets. Median selected monthly owner costs for owners with a mortgage and median gross rent, both from the 2015

ACS, are averaged and weighted by the ratio of owner- to renter-occupied housing units from the tenure variable for every block group in a CBSA.

Transportation Cost Model

While housing costs are derived from 2015 ACS data, transportation costs, the T in the H+T Index, are modeled based on three components of transportation behavior—auto ownership, auto use, and transit use—which are combined to estimate the cost of transportation.

Basic Structure

The household transportation model is based on a multidimensional regression analysis, in which formulae describe the relationships between three dependent variables (auto ownership, auto use, and transit use) and independent household and local environment variables. Neighborhood level (Census block group) data on median household income, household size, commuters per household, household residential density, walkability and street connectivity, transit connectivity and access, and employment access and diversity were utilized as the independent or predictor variables.

To construct the regression equations, each predictor variable was tested separately; first to determine the distribution of the sample and second to test the strength of the relationship to the criterion variables. The regression analysis was conducted in a comprehensive way, ignoring the distinction between the local environment variables and the household variables in order to obtain the best fit possible from all of the independent variables. The predicted result from each model was multiplied by the appropriate price for each unit—autos, miles, and transit trips—to obtain the cost of that component of transportation. Total transportation costs were calculated as the sum of the three cost components as follows:

$$\text{Household T Costs} = [C_{AO} * F_{AO}(\text{X})] + [C_{AU} * F_{AU}(\text{X})] + [C_{TU} * F_{TU}(\text{X})]$$

Equation 1: Cost of Transportation

Where:

C = cost factor (i.e. dollars per mile)
F = function of the independent variables (F_{AO} is auto ownership, F_{AU} is auto use, and F_{TU} is transit use)

Table 1: Independent Variables Overview

Variable	Description	Data Source	Type
Median HH Income	Median household Income in the block group	2015 ACS	Household
Commuters/HH	Workers per household who do not work at home	2015 ACS	Household
Avg. HH Size	Average number of people per household	2015 ACS	Household
Gross Household Density	Number of households divided by the land area in the Census block group	2015 ACS, TIGER/Line files	Neighborhood Characteristic (housing density)
Regional Household Intensity	Households summed divided by the distance squared in miles between block group by (the households in the block group are not included)	2015 ACS, TIGER/Line files	Neighborhood Characteristic (housing density)
Fraction of Rental Housing units	Fraction of occupied housing units with rental tenure	2015 ACS,	Neighborhood Characteristic (housing density)
Fraction of Single Family Detached Housing	Fraction of single family detached households in the block group	2015 AC	Neighborhood Characteristic (housing density)
Employment Access Index	Jobs summed by blocks divided by the distance squared in miles (if less than one mile not scaled)	Census LEHD-LODES	Neighborhood Characteristic (employment)
Employment Mix Index	Number of block per acre Weighted sum of 13 different employment types each scaled by a coefficient that are optimized using transit use	Census LEHD-LODES	Neighborhood Characteristic (employment)
Block Density	Number of block per acre	TIGER/Line files	Neighborhood Characteristic (walkability)
Transit Connectivity Index	Sum of buses/trains per week scaled by overlap of 1/8 mile rings about every stop that intersects the block group	CNT AllTransit	Neighborhood Characteristic (transit)
Average Available Transit Trips per Week	Number of possible transit rides within the block group and a ¼ mile of its border.	CNT AllTransit	Neighborhood Characteristic (transit)
Transit Access Shed	Total area that transit riders from the block group can access in 30 minutes with 1 or no transfers for all the transit stations within a ¼ mile of the block group	CNT AllTransit	Neighborhood Characteristic (transit)
TAS Jobs	The total number of jobs in the TAS area	CNT AllTransit	Neighborhood Characteristic (transit)

Independent Variables: Household Characteristics

The 2015 ACS, at the block group level, serve as the primary data source for the independent variables pertaining to household characteristics.

Median Household Income:

Median household income is obtained directly from the 2015 ACS.

Average Household Size:

Average household size was calculated using total population in occupied housing units by tenure.

Average Commuters per Household:

Average commuters per household was calculated using the total workers 16 years and over who do not work at home from means of transportation to work and tenure to define occupied housing units. Because means of transportation to work includes workers not living in occupied housing units (i.e. those living in group quarters), the ratio of Total Population in occupied housing units to total population was used to scale the count of commuters to better represent those living in households.

Independent Variables: Neighborhood Characteristics

Household Residential Density

In previous versions of the H+T Index household density was found to be one of the most significant variables in explaining the variation in auto use, auto ownership, and transit use. Various definitions of density have been constructed and tested, but net residential density (households per residential acre) was the primary metric used. No national data source of detailed land use data exists so previous versions of the household transportation cost model defined residential density as the average number of households per residential acre for the Census blocks within the block group weighted by count of households. Total households obtained at the block level from the 2010 US Census and TIGER/Line files were used to define blocks. However, since this iteration is using data from the 2015 ACS, the 2010 data is not compatible. Thus, several metrics were developed to estimate how household transportation behavior is driven by household density and concentration.

Gross Household Density

Gross household density is calculated from the 2015 ACS. It is simply the number of households in a census block group divided by the area of land within the block group

Regional Household Intensity

The Regional Household Intensity is constructed using a gravity model which considers both the quantity of, and distance to, all households, relative to any given block group. Using an inverse-square law, intensity is calculated by summing the total number of household divided by the square of the distance to those households, but does not include the households within the block group. This quantity allows us to examine both the intensity of housing development in the region around the block group.

The Regional Household Intensity is calculated as:

$$H \equiv \sum_{i=1}^{n} \frac{hh_i}{r_i^2}$$

Equation 2: Regional Household Intensity Definition

Where:

H is the Regional Household Intensity for a given Census block group
n is the total number of Census blocks (not including the given Census block group)
hh_i is the number of households in the i^{th} Census block
r_i is the distance (in miles) from the center of the given Census block group to the center of the i^{th} Census block

As households get farther away from the Census block group their contribution to the Regional Household Intensity is reduced; for example, one household in a Census block group a mile away adds one, but a household 10 miles away adds 0.01. All households in all US Census blocks groups are included in this measure. However, in order to expedite the calculation, the calculation uses the[1]:

- State totals when the state is not the same as the given block group and is more than 88 miles away,
- County totals when the county is not the same as the given block group and is more than 11.5 miles away, and
- Census tract totals when the tract is not the same at the given block group and is more than 2.5 miles away.

Fraction of Rental House Housing Units

The fraction of occupied housing units that are rental rather than owned by the occupant is calculated using the 2015 ACS data by dividing the number of occupied housing units with tenure defined as rental by the total number of occupied housing units in the Census block group.

Fraction of Single Family Detached Households

The fraction of single family detached households is calculated using the 2015 ACS data by dividing the number of households living in single family detached housing by the total number of households in the Census block group.

Street Connectivity and Walkability

Measures of street connectivity have been found to be good proxies for pedestrian friendliness and walkability. Greater connectivity created from numerous streets and intersections creates smaller blocks and tends to lead to more frequent walking and biking trips, as well as shorter average trips. Three measures of street connectivity — block density, intersection density, and block perimeter — have been found to be important drivers of household travel behavior. However, these three measures are so interrelated only block density was included. The resulting models have essentially equivalent R^2 values compared to when the other measures are included and thus have comparable goodness of fit.

[1] These distance thresholds were developed using the average distance between the geographic entities.

Block Density

Census TIGER/Line files are used to calculate average block density (in acres) using the number of blocks within the block group divided by the total block group land area.

Employment Access and Diversity

Employment numbers are calculated using Longitudinal Employer-Household Dynamics (LEHD) Origin Destination Employment Statistics (LODES) at the Census block group level. The Longitudinal Employer-Household Dynamics (LEHD) program is part of the Center for Economic Studies at the U.S. Census Bureau.

Employment Access Index

The Employment Access Index is constructed using a gravity model that factors in the quantity of, and distance to, all employment destinations, in relation to any given block group. Using an inverse-square law, the Employment Access Index is calculated by summing the total number of jobs divided by the square of the distance to those jobs. This method provides more information than a simple job density measure, in that it includes the accessibility to jobs outside a given Census block group. In addition to measuring access to jobs, it also provides a measure of economic activity created by those jobs.

The Employment Access Index is calculated as:

$$E \equiv \sum_{i=1}^{n} \frac{p_i}{r_i^2}$$

Equation 3: Employment Access Index Definition

Where:

E is the Employment Access for a given Census block group
n is the total number of Census blocks
p_i is the number of jobs in the i^{th} Census block
r_i is the distance (in miles) from the center of the given Census block group to the center of the i^{th} Census block

The proximity of jobs to the Census block group determines their contributive value to the Employment Access Index. For example, one job a mile away adds one, but a job 10 miles away adds 0.01. The measure includes all jobs in all US Census blocks. The index employs the following parameters to accelerate the calculation:[2]

- State totals when the state is not the same as the given block group and is more than 88 miles away,
- County totals when the county is not the same as the given block group and is more than 11.5 miles away, and

[2] These distance thresholds were developed using the average distance between the geographic entities and factor determined such that the calculation remains consistent with using the block groups for a small representative sample.

- Census tract totals when the tract is not the same at the given block group and is more than 2.5 miles away.

Employment Mix Index

The model includes an Employment Mix Index which measures employment diversity in addition to total number of jobs. It is produced by taking the weighted sum of the gravity measure of each type of job (out of 20 total types). The benefit of looking at the mix of employment options can be seen in the R^2 value for the transit use model. The transit use model when transit data is not available produces an R^2 value of 78.8%, but when the employment mix index is included the R^2 increases to 80.7%.

Table 2 lists the 20 employment categories derived from the 2014 LEHD, (all[3] use the Natural Log ($\ln(x)$) transformation function) the weight used, and what the trend is for auto ownership. The variable is the gravity measure of the jobs of the given type (see above). The weight is determined by regressing all of the other independent variables and these 20 against autos per household. The *Index Effect* column indicates what happens to the value of the Employment Mix Index when the fraction of the given employment type increases.

Table 2: Summary of Employment Type and Weighting for Employment Mix Index

Category	NAIC Sector(s)	Weight	Index Effect
Agriculture, Forestry, Fishing and Hunting	11	.0357	Increase
Mining, Quarrying, and Oil and Gas Extraction	21	.0097	Increase
Utilities	22	.0071	Increase
Construction	23	.024	Increase
Manufacturing	31-33	.028	Increase
Wholesale Trade	42	.009	Increase
Retail Trade	44-45	.021	Increase
Transportation and Warehousing	48-49	.013	Increase
Information	51	.035	Increase
Finance and Insurance	52	-.032	Reduce
Real Estate and Rental and Leasing	53	-.022	Reduce
Professional, Scientific, and Technical Services	54	.007	Increase
Management of Companies and Enterprises	55	-.0010	Reduce
Administrative and Support and Waste Management and Remediation Services	56	.020	Increase
Educational Services	61	.029	Increase
Health Care and Social Assistance	62	-.006	Reduce
Arts, Entertainment, and Recreation	71	.001	Increase
Accommodation and Food Services	72	.075	Increase
Other Services [except Public Administration]	81	-.042	Reduce
Public Administration	92	-.003	Reduce

[3] The log transformation (ln(x)) was used in all cases except for NAICS codes 55 and 71 where an inverse transformation (1/x) provided a better fit.

The calculation for the raw employment mix is:

$$R \equiv \sum_{i=1}^{13} W_i \times F_t(e_i)$$

Equation 4: Definition of Raw Employment Mix

Where:

R is the Raw Employment Mix for a given Census block group
i is the employment category
w_i is the weight for the i^{th} employment category
F_i is the linear transformation function for the i^{th} employment category (ln(x) for all except 1/x for NAICS coded 55 and 71).
e_i is the value of the variable in Table 2 for the i^{th} employment category

The full calculation is then evaluated using the following formula.

$$I_{Emix} \equiv 100 \times \frac{R - R_{min}}{R_{max} - R_{min}}$$

Equation 5: Definition of Employment Mix Index

Where:

I_{Emix} is the Employment Mix Index for a given Census block group
R is the Raw Employment Mix for a given Census block group
R_{min} is the minimum value of the Raw Employment Mix for all Census block groups
R_{max} is the maximum value of the Raw Employment Mix for all Census block groups

This index is calculated of all Census block groups in the country as a number from zero to 100; the Figure 1 shows the distribution of values for this index for only the Census block groups in the sample used the H+T index, as does Figure 2 but this shows only the lower values illustrating the long tail. This comes from very remote locations such as in the Aleutian Islands in Alaska, and the outer islands in Hawai'i.

Figure 1: Histogram of Employment Mix Index

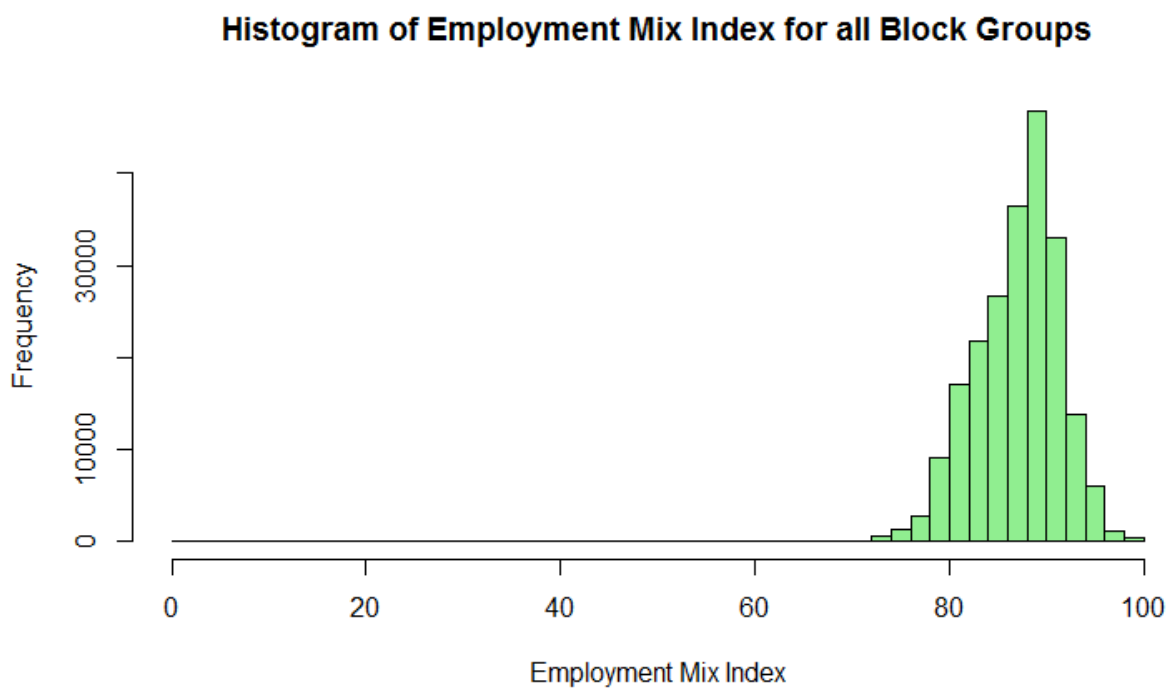
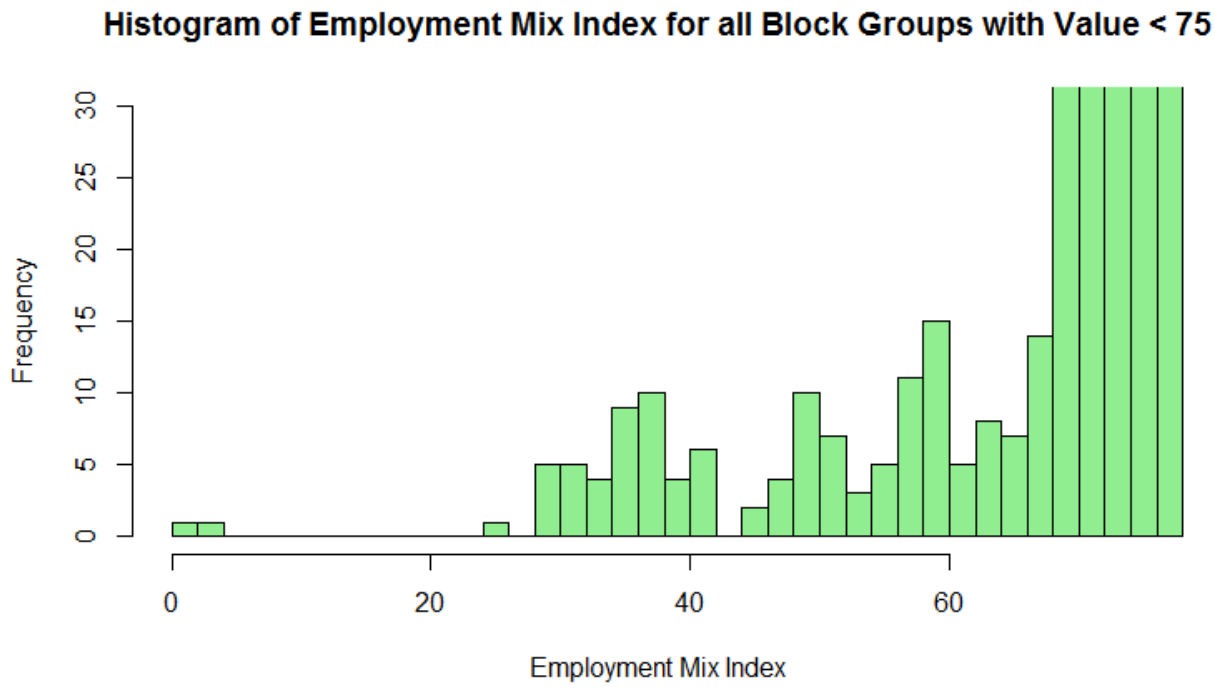


Figure 2: Histogram of Employment Mix Index for Block Groups with Value < 75



Transit Access and Connectivity

Transit access is measured through *General Transit Feed Specification* (GTFS) data collected and synthesized by CNT. In addition to the publicly available GTFS data (provided by many, but not all, transit agencies) CNT has created GTFS structured datasets utilizing online transit maps and schedules. In many cases, CNT has directly contacted transit agencies to obtain more specific information on stop locations and schedules. All GTFS data is merged into a proprietary dataset through a CNT tool known as AllTransit™ Data Builder. AllTransit is an online tool that facilitates the collection, normalization, aggregation, and analysis of GTFS data to determine fixed-route transit service.

To date, CNT has compiled stop, station, and frequency data for bus, rail, and ferry service for all major transit agencies in regions with populations greater than 100,000. Attachment A lists the transit agencies for which data has been compiled. In regions where data is not available, CNT has assumed that the transit service is not large enough to affect the fits, thus zero is assumed for all transit inputs.

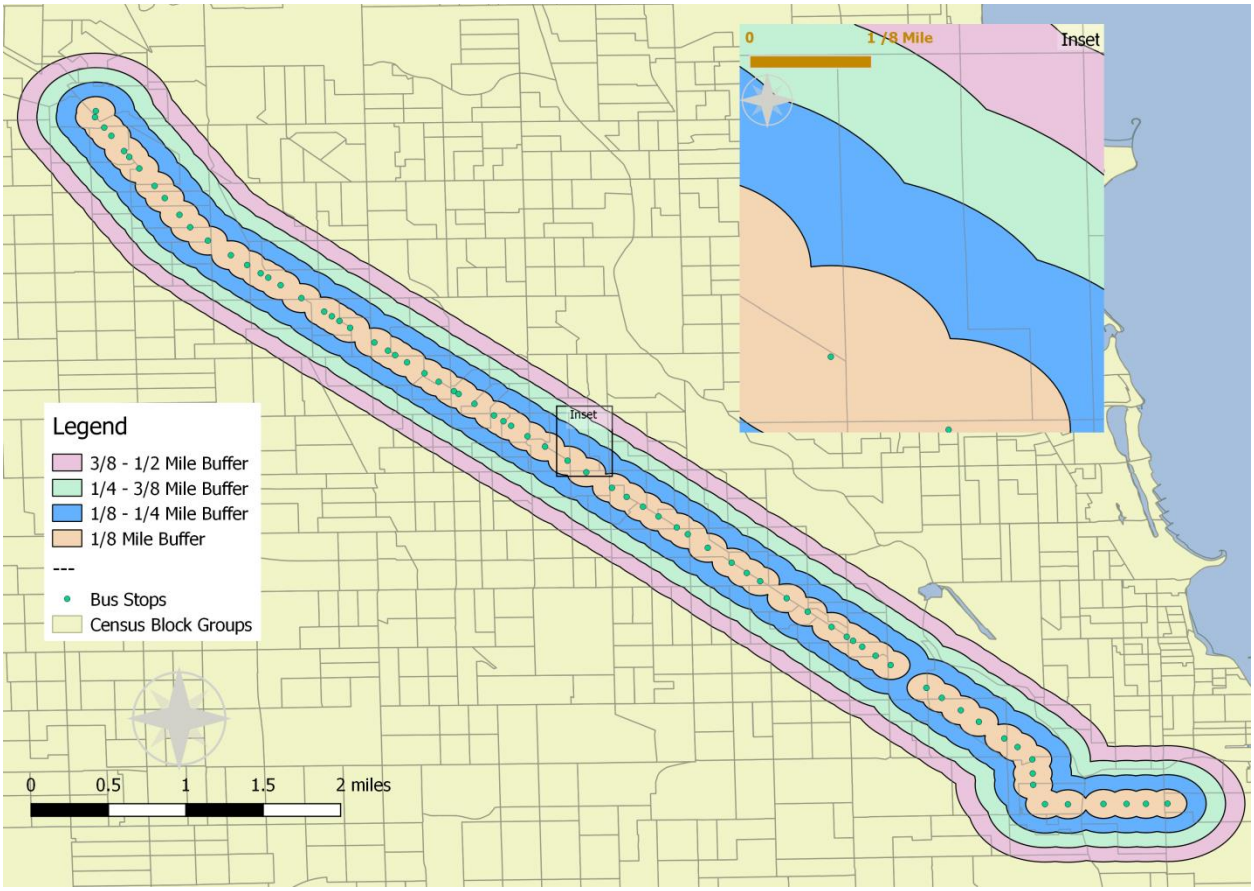
Four measures of transit access are used in the model: the Transit Connectivity Index (TCI), Transit Access Shed (TAS), Transit Access Shed Jobs (TAS Jobs), and Average Available Transit Trips per Week. The TCI estimates how many transit opportunities are within walking distance of a census block group. The TAS is a proxy measure for how far one can travel in 30 minutes on transit, while the TAS Jobs is the sum of the total number of jobs within the TAS.

Transit Connectivity Index

The Transit Connectivity Index is a measure of access to bus stops and rail stations that CNT developed specifically for use in the household transportation cost model. To calculate this measure, four concentric rings one-eighth of a mile in width (access zones) were plotted around each bus stop, and eight such rings around each rail[4] station, and then merged together around a common route (see **Error! eference source not found.** for a typical bus route).

[4] By rail we include, light rail, heavy rail, commuter rail, ferry terminals, and other stations such as vernaculars and cable cars.

Figure 3: Illustration of four 1/8th mile Rings around Transit Route (For a Single Bus Route in a Chicago)



Using these access zones, the following are defined for each block group:

Table 3: Bus Transit Connectivity Variables

Variable	Description
L	Land area of the block group covered by access zone
F	Service frequency value (Trips per week)
B	Total block group land area

At each block group, eight transit access values were calculated for each transit route where at least one of the access zones intersects the block group. The following formula is used for each transit route to obtain the scaled frequency (SF) by zone for each block group.

$$SF_d = \sum_{i=1}^{n} \frac{L_{i,d} F_{i,d}}{B}$$

Equation 6: Bus SF_d Calculation

Where:

d is the index across the eight concentric circular access zones,
n is the total intersecting transit routes access zone ring d.

These values are calculated for every block group that a given zone intersects; meaning that in well-served block groups there will be values for zones corresponding to multiple bus stops.

The farther an access zone is from its transit node, the less of a contribution it should make to the level of access in any block group it intersects, however the relative area covered by these distant access zones is larger because of their shape. In order to account for the decreasing access benefits at greater distances and the increased area coverage a weight is given to each value of SF_d calculated using regression. Measured values of percent journey to work by transit were regressed against the 8 SF_d values (as defined above) using an ordinary least square to define the weight of each of the eight rings.

The sum of the weights times the SD_d are calculated for each block group. This quantity is not easily translated since it is the combination of many factors, so the final value for this index is a number from 0-100 representing the first stage of a fit for the use of transit for commuter's journey to work by the following formula:

$$Bus\ TCI \equiv 100 \times \frac{STD - STD_{min}}{STD_{max} - STD_{min}}$$

Equation 7: Bus Connectivity Index Calculation

.

Where:

STD is the sum of all of the SF_d i.e. $STD = \sum_{d=1}^{8} Wt_d SF_d$,
STD_{min} is the minimum value for all block groups and
STD_{max} is the maximum value for all block groups.

The same method was used to run the Transit Connectivity Index for rail; however this mode uses eight instead of four one-eighth mile rings (1 mile). These components of the Transit Connectivity Index are added to the similar bus component to make the final Transit Connectivity Index. The following table shows the final weighting for the ring (both bus and rail) that create the final Transit Connectivity Index. Figure 4 is a histogram of the value of TCI for all block groups in the US, and Figure 5 shows that there is a long tail in this distribution going, by definition, to the max value of 100.

Table 4: Coefficients use in Transit Connectivity Index

Description	Transformation Function	Weight	Index Effect
Bus Ring 1	$\sqrt{x}$	0.160	Increase
Bus Ring 2	$\sqrt{x}$	0.032	Increase
Bus Ring 3	$\sqrt{x}$	0.023	Increase
Bus Ring 4	$\sqrt{x}$	0.031	Increase
Rail Ring 1	$\sqrt{x}$	0.104	Increase
Rail Ring 2	$\sqrt{x}$	0.109	Increase
Rail Ring 3	$\sqrt{x}$	0.091	Increase
Rail Ring 4	$\sqrt{x}$	0.068	Increase
Rail Ring 5	$\sqrt{x}$	0.042	Increase
Rail Ring 6	$\sqrt{x}$	0.059	Increase
Rail Ring 7	$\sqrt{x}$	0.030	Increase
Rail Ring 8	$\sqrt{x}$	0.083	Increase

Figure 4: Histogram of TCI for all Block Groups

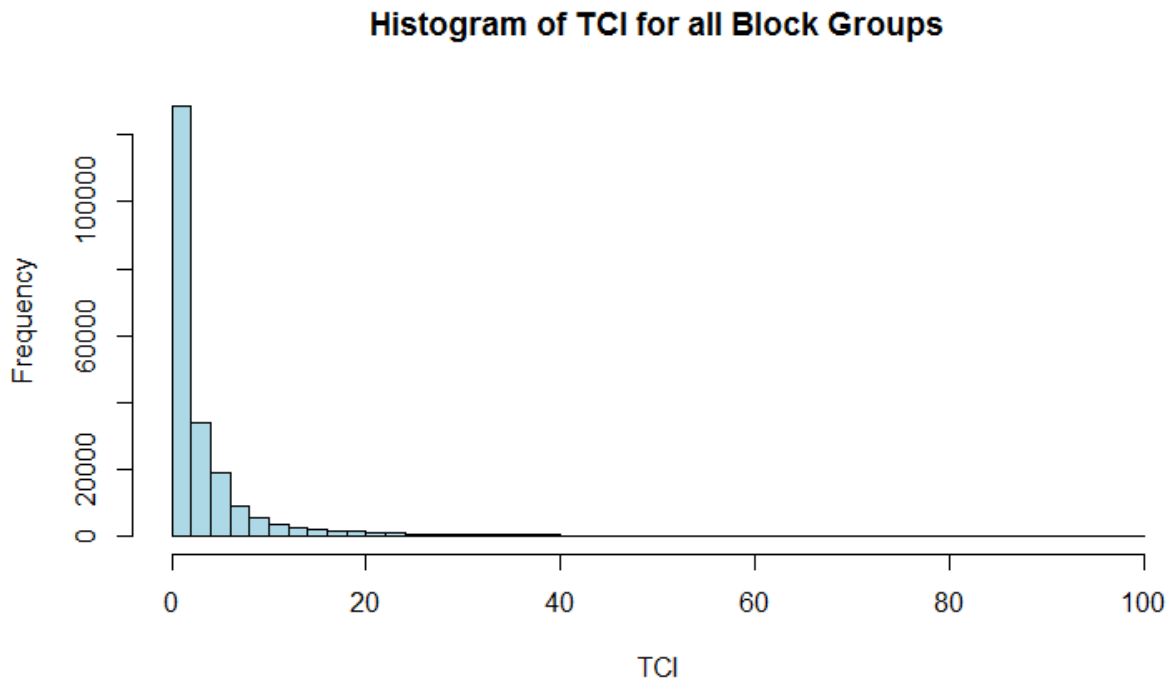
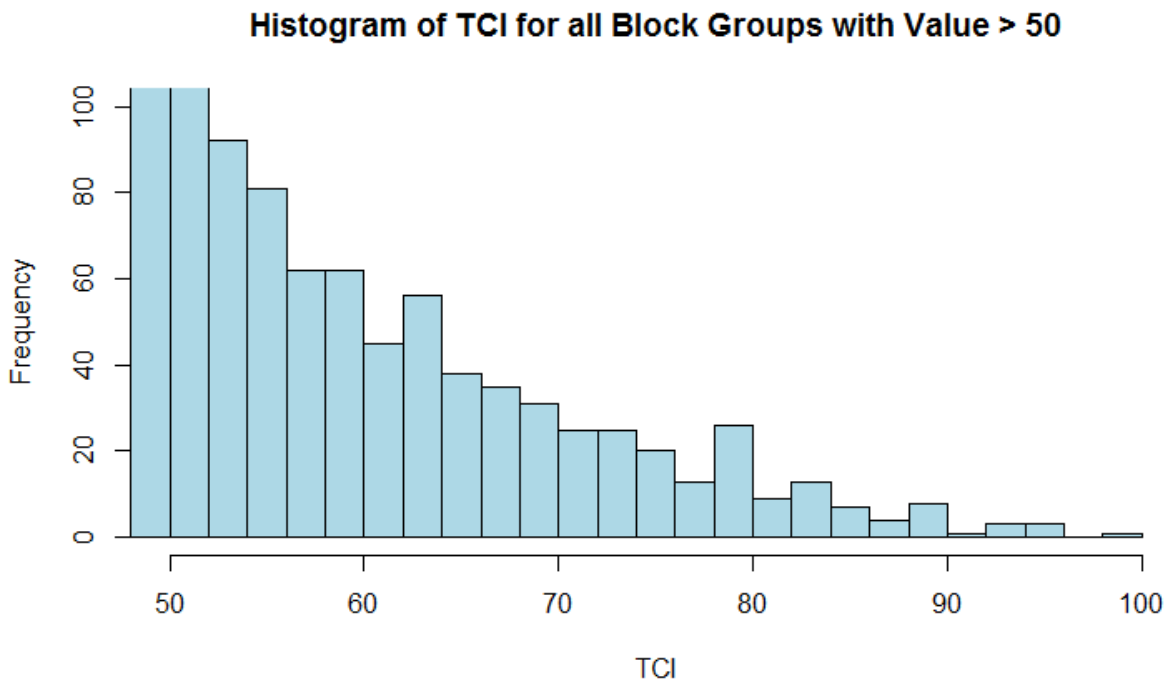


Figure 5: Histogram of TCI for all Block Groups with TCI Greater than 50



Transit Access Shed

The Transit Access Shed (TAS) is defined as a geographic area accessible within 30 minutes by public transportation. This measure was derived from the AllTransit GTFS data. For each transit stop, all stops that can be reached within 30 minutes were identified. One transfer within a quarter-mile of a stop was allowed, and all transfers were padded with 10 minutes of walking and/or waiting. The stops reachable within 30 minutes were based on the minimum travel time between the two stops, allowing the inclusion of more distant stops that are reachable within 30 minutes via express service. For each origination stop, a quarter-mile buffer was created around the destination stops. Based on the location of the originating stop, the access shed was then aggregated for each stop to the block group by including stops that were within the block group or within a quarter of a mile of its boundary. Finally, the accessible area or Transit Access Shed is calculated by summing the areas of the quarter-mile buffers around every stop that is within 30 minutes as defined above. In order to assign a value to a Census block group, the Transit Access Shed for all stops within walking distance of the block group are merged into one grand shed. This area is then assigned as the block group's Transit Access Shed.

Transit Access Shed Jobs

Transit Access Shed Jobs is the total number of jobs within the TAS. The count of jobs was obtained from the Census LEHD-LODES data.

Average Available Transit Trips per Week

Average Available Transit Trips per Week is the average frequency of service from the AllTransit GTFS data, for all stops within the Census block group or within a half mile of it borders.

Dependent Variables

Auto Ownership

For the dependent variable auto ownership, the regression analysis was fit using measured data on auto ownership obtained from the 2015 ACS. Aggregate number of vehicles available by tenure defined the total number of vehicles, and tenure defined the universe of occupied housing units. Average vehicles per occupied housing unit were calculated at the block group level.

Auto Use

For the dependent variable auto use, the regression analysis was fit using measured data on the amount households drive, vehicle miles traveled (VMT) per automobile. Odometer readings from 2010 through 2012 odometer readings were acquired in Illinois for the Chicago and St. Louis metro areas. Data were matched for over 660,000 records (two records for each individual vehicle identification number (VIN)) and the change provided VMT estimates. The dataset represents a diverse set of place types from rural areas to large cities, and provides a very good data set to calibrate the model. Data obtained were geographically identified with ZIP+4™ and then assigned to Census block groups.

The final value of VMT includes an additional factor of eight percent to compensate for the fact that the vehicles in this sample were all five years old or older. This factor is obtained from the research commissioned and published by US HUD and US DOT to develop the Location Affordability Index.[5]

Transit Use

Because no direct measure of transit use was available at the block group level, a proxy was utilized for the measured data representing the dependent variable of transit use. From the 2015 ACS, Means of transportation to work was used to calculate a percent of commuters utilizing public transit.

[5] See http://www.locationaffordability.info/LAPMethodsV2.pdf page 24.

Household Transportation Regression Analysis

For this version of the H+T Index the model has been simplified. The most important relationships between independent and dependent variables are non-linear, so in previous iterations of the household transportation model a non-linear regression was used. In this iteration this non-linearity is compensated for by using simple transformation functions. These functions (Linear (x), Square Root ($\sqrt{x}$), Natural Log ($\ln(x)$), and Inverse ($1/x$)) are used to give the best fit using an ordinary least square fit. The final fit uses eleven independent variables and is broken up into six independent models. A model is constructed for each dependent variable (auto ownership, auto use and transit use) using the four transit variables (Bus Access Index, Rail Access Index, Transit Shed and Jobs in Transit Shed) where these measures are available. Another model is also constructed leaving these transit measures out in order to make a good model for regions where transit data is missing.

An ordinary least square regression including all interaction terms (where statistically significant) was used to estimate the fit coefficients; the equation that will estimate the dependent variable from independent variables is:

$$D = I + \sum_{i=1}^{n} C_i \times f_i(x_i) + \sum_{i=1}^{n} \sum_{j=i}^{n} C_{ij} \times f_i(x_i) \times f_j(x_j)$$

Equation 8: Equation for Estimating Dependent Variable from Coefficients and Independent Variables

Where:

D is the dependent variable for a given Census Block Group i.e. Autos per Household
I is the Intercept – obtained in the regression
i is the index of the independent variable i.e. i goes from 1 to 10 for a regression that had 10 independent variables
j is the index of the independent variable i.e. j goes from i to 10 for a regression that had 10 independent variables and so as to include all interaction terms only once
C_i is the fit coefficient for the i^{th} independent variable
C_{ij} is the fit coefficient for the i^{th} j^{th} interaction variable
f_i is the linearization transformation function
x_i is the value of the i^{th} independent variable

Choosing Linear Transformation Functions for the Independent Variables

In order to address the nonlinear nature of the relationships between the independent and the dependent variables, a linear transformation function was chosen. These were limited to Linear (x), Square Root ($\sqrt{x}$), Natural Log ($\ln(x)$), and Inverse ($1/x$). For variables that could have a legitimate value of zero, the functions Safe Natural Log $\ln(x+1)$ and Safe Inverse $1/(x+1)$ were also used. In order to

choose the best transformation each variable was tested to determine which transformation resulted in the best fit.

The example below considers how median household income drives auto ownership. Figure 6 and **Error! Reference source not found.**[6] show that the relationship between auto ownership and median household income is nonlinear. However, the relationship between auto ownership and the natural log of the median household income shows a more linear relationship; note the increase in R^2 by over eight percentage points. This technique was then repeated for every variable for every model to select the optimal transformation.

Figure 6: Autos/HH vs. Median Household Income

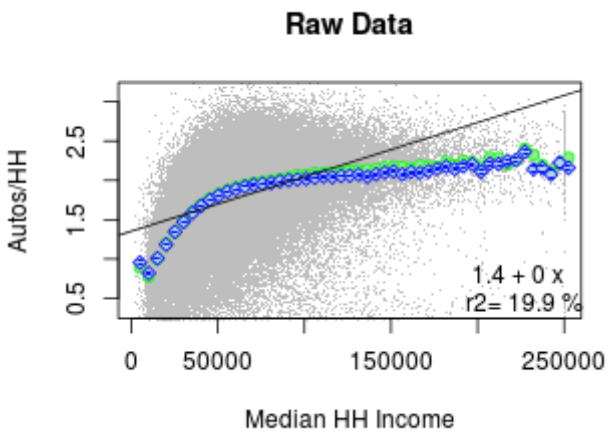
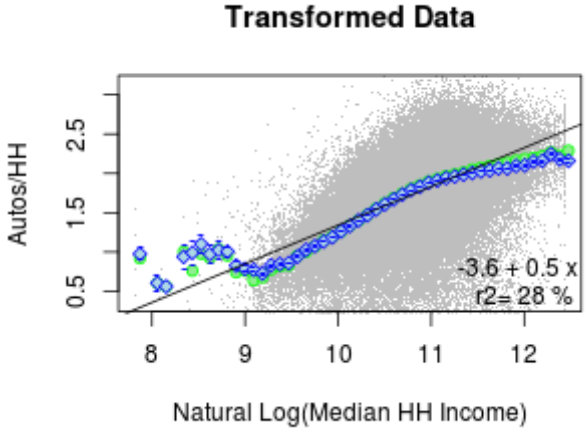


Figure 7: Autos/HH vs. the Natural Log of the Median Household Income



Choosing Independent Variables

The method to test the statistical significance of variables and to determine those that would reduce the multicollinearity of the set of independent variables, and all the possible interaction terms is described below:

- Initially the fit is done including all possible variable combinations. Then an iterative process is conducted where the least statistically significant variable or combination of variables is dropped, examining the probability that the value is consistent with zero (Pr(>|t|)).
- Then the fit is redone, and the least statistically significant variable or combination of variables is dropped, this continues until the least statistically significant variable or combination of variables' probability that it is consistent with zero is less than 5%.
- This is done in an iterative manner because simply deleting all the variable or combination of variables that have Pr(>|t|))>5% would eliminate more variables than necessary, since all this set of variables and combination of variables display a very high level of multicollinearity.

[6] These plots contain the data from every Census Block Group (in grey dots), the mean of those data in fifty bins (blue diamonds) and the median those data in fifty bins (green circles). The line represents the linear fit to the data.

- Once this process is completed, another is applied that adds each eliminated variable back into the regression equation to see if by eliminating all other variables it may become statistically significant again.

See Table 6 in the Model Findings section below for the final set of variables used.

Table 5: List of all variables and combination of variables for the autos/HH regression before eliminating the insignificant ones, listed in order of least significant to most significant

Variable 1	Variable 2	Value	Error	Value/Error
TCI		.00	.01	91.79%
Level of Service	**TAS**	.00000000	.00000001	91.35%
Median HH Income	**TAS Jobs**	-.0001	.0007	87.34%
Gross HH Density	**Level of Service**	.0000001	.0000005	85.72%
Employment Gravity	**TCI**	.0002	.0007	77.66%
Employment Mix	**TCI**	-.0001	.0002	69.89%
Block Density	**Commuters per HH**	.002	.004	64.78%
Employment Mix		-.01	.01	59.83%
Fraction Rental Housing Units	**TCI**	-.001	.001	58.47%
TAS Jobs		-.01	.01	55.07%
Household Gravity	**TAS**	.000000001	.000000001	51.52%
Employment Mix	**Fraction Rental Housing Units**	.002	.002	48.16%
Fraction Single Family Detached	**Level of Service**	-.000002	.000003	45.16%
Level of Service	**Median HH Income**	-.0000008	.0000010	42.66%
HH Size	**Level of Service**	-.0000008	.0000009	38.04%
Fraction Single Family Detached	**Household Gravity**	.0000002	.0000002	36.42%
Level of Service	**TAS Jobs**	-.0000007	.0000007	33.13%
Employment Gravity	**Level of Service**	-.000001	.000001	32.97%
Block Density		.04	.04	31.40%
Household Gravity	**TAS Jobs**	.00000004	.00000003	23.84%
TAS	**TAS**	.0000007	.0000006	20.00%
Block Density	**Employment Gravity**	.003	.002	20.04%
HH Size	**Household Gravity**	.00000009	.00000006	17.67%
Employment Mix	**Level of Service**	-.0000004	.0000003	16.47%
Block Density	**TAS Jobs**	.0006	.0004	13.73%
Median HH Income	**TCI**	-.0008	.0005	12.62%
Employment Gravity	**TAS**	-.00009	.00006	10.65%
Fraction Rental Housing Units	**Level of Service**	.000004	.000003	9.44%
HH Size	**TCI**	-.0008	.0005	8.71%
Fraction Single Family	**Employment Gravity**	.013	.008	7.91%

Variable 1	Variable 2	Value	Error	Value/Error
Detached				
Block Density	**TAS**	-.00006	.00003	7.38%
Intercept		-1.5	.8	5.29%
Gross HH Density	**TCI**	-.0006	.0003	4.84%
Median HH Income	**Median HH Income**	.005	.003	4.71%
Gross HH Density	**TAS**	-.00006	.00003	4.43%
Fraction Rental Housing Units	**Gross HH Density**	-.009	.004	4.16%
Gross HH Density	**Employment Gravity**	-.004	.002	3.96%
Employment Mix	**Gross HH Density**	.0010	.0005	3.75%
Household Gravity	**Level of Service**	.00000000002	.00000000001	2.89%
Employment Mix	**TAS Jobs**	.0004	.0002	2.84%
Commuters per HH	**TCI**	-.0021	.0009	2.52%
Employment Mix	**Fraction Single Family Detached**	.005	.002	1.93%
Block Density	**Fraction Single Family Detached**	-.011	.005	1.60%
TAS Jobs	**TCI**	-.0006	.0003	1.55%
Gross HH Density	**Gross HH Density**	-.0015	.0006	1.47%
HH Size	**TAS Jobs**	.0013	.0005	1.18%
Fraction Rental Housing Units	**Employment Gravity**	.027	.010	0.57%
Commuters per HH	**Fraction Rental Housing Units**	-.05	.01	0.16%
Commuters per HH	**Gross HH Density**	.009	.003	0.12%
Level of Service	**Level of Service**	.00000000018	.00000000005	0.10%
Gross HH Density	**Household Gravity**	-.00000015	.00000004	0.07%
Commuters per HH	**Level of Service**	-.000006	.000002	0.05%
Fraction Rental Housing Units	**Fraction Rental Housing Units**	.06	.02	0.05%
Household Gravity	**Median HH Income**	.00000029	.00000008	0.03%
Employment Mix	**Median HH Income**	-.004	.001	0.03%
Level of Service		.00008	.00002	0.03%
TAS	**TCI**	.000031	.000008	0.02%
Commuters per HH	**Fraction Single Family Detached**	-.04	.01	0.01%
Fraction Rental Housing Units	**Fraction Single Family Detached**	-.07	.02	0.00%
Employment Gravity		.24	.06	0.00%
Commuters per HH	**Commuters per HH**	.025	.006	0.00%
Household Gravity	**TCI**	-.000000027	.000000007	0.00%
Fraction Single Family Detached	**TCI**	.005	.001	0.00%

Variable 1	Variable 2	Value	Error	Value/Error
Fraction Rental Housing Units	**TAS**	-.0006	.0001	0.00%
Employment Gravity	**Median HH Income**	-.018	.004	0.00%
HH Size	**Block Density**	-.009	.002	0.00%
Level of Service	**TCI**	-.00000036	.00000008	0.00%
Commuters per HH	**Employment Mix**	-.007	.001	0.00%
Block Density	**Level of Service**	.0000036	.0000007	0.00%
HH Size	**TAS**	.00020	.00004	0.00%
Fraction Rental Housing Units	**TAS Jobs**	.008	.002	0.00%
Fraction Single Family Detached	**Gross HH Density**	-.020	.004	0.00%
HH Size	**Fraction Rental Housing Units**	-.038	.007	0.00%
Block Density	**Block Density**	-.0048	.0009	0.00%
Gross HH Density		.19	.03	0.00%
Fraction Rental Housing Units	**Household Gravity**	-.0000011	.0000002	0.00%
Median HH Income		.46	.08	0.00%
Block Density	**Household Gravity**	.00000036	.00000006	0.00%
TAS	**TAS Jobs**	-.00022	.00004	0.00%
Commuters per HH	**Employment Gravity**	.037	.006	0.00%
Block Density	**Fraction Rental Housing Units**	.036	.006	0.00%
Fraction Single Family Detached	**TAS Jobs**	.008	.001	0.00%
Block Density	**Gross HH Density**	.009	.001	0.00%
Gross HH Density	**TAS Jobs**	.0024	.0003	0.00%
Employment Gravity	**TAS Jobs**	-.0048	.0007	0.00%
TAS Jobs	**TAS Jobs**	.0020	.0003	0.00%
Commuters per HH	**TAS**	.00062	.00008	0.00%
Employment Mix	**Employment Mix**	.00046	.00006	0.00%
Block Density	**Employment Mix**	-.0044	.0006	0.00%
Fraction Single Family Detached		-1.2	.1	0.00%
HH Size	**Commuters per HH**	-.040	.005	0.00%
HH Size	**Gross HH Density**	.012	.001	0.00%
Commuters per HH	**TAS Jobs**	-.009	.001	0.00%
Fraction Single Family Detached	**TAS**	-.0009	.0001	0.00%
Employment Mix	**TAS**	.00013	.00001	0.00%
Block Density	**TCI**	-.0036	.0004	0.00%
Employment Mix	**Employment Gravity**	-.0077	.0008	0.00%

Variable 1	Variable 2	Value	Error	Value/Error
Median HH Income	**TAS**	.00049	.00005	0.00%
Fraction Rental Housing Units		-1.8	.2	0.00%
Fraction Single Family Detached	**Median HH Income**	.078	.008	0.00%
TCI	**TCI**	.00030	.00003	0.00%
Block Density	**Median HH Income**	.026	.002	0.00%
Commuters per HH	**Household Gravity**	-.0000014	.0000001	0.00%
Fraction Rental Housing Units	**Median HH Income**	.12	.01	0.00%
Fraction Single Family Detached	**Fraction Single Family Detached**	-.12	.01	0.00%
Gross HH Density	**Median HH Income**	-.026	.002	0.00%
Household Gravity	**Household Gravity**	.0000000000114	.0000000000008	0.00%
HH Size	**HH Size**	-.018	.001	0.00%
HH Size	**Fraction Single Family Detached**	.080	.006	0.00%
TAS		-.015	.001	0.00%
Employment Gravity	**Employment Gravity**	.036	.002	0.00%
Commuters per HH	**Median HH Income**	-.095	.006	0.00%
Household Gravity		-.000028	.000002	0.00%
HH Size	**Employment Gravity**	-.045	.003	0.00%
Commuters per HH		1.7	.1	0.00%
Household Gravity	**Employment Gravity**	-.0000024	.0000001	0.00%
Employment Mix	**Household Gravity**	.00000057	.00000003	0.00%
HH Size	**Median HH Income**	.075	.003	0.00%
HH Size	**Employment Mix**	.0157	.0006	0.00%
HH Size		-1.59	.05	0.00%

Transportation Cost Calculation

The transportation model in the H+T Index estimates three components of travel behavior: auto ownership, auto use, and transit use. To calculate total transportation costs, each of these modeled outputs is multiplied by a cost per unit (e.g., cost per mile) and then summed to provide average values for each block group.

Auto Ownership and Auto Use Costs

Auto ownership and use costs are derived from research conducted by HUD and DOT using the Consumer Expenditure Survey (CES) from the US Bureau of Labor Statistics. The research is based on the

2005-2010 waves of the CES, and costs are estimated for autos up to ten years old. Because expenditures are represented in inflation-adjusted 2010 dollars using the Consumer Price Index for all Urban Consumers (CPI-U), an inflation factor is applied to estimate the cost of auto ownership into 2015 dollars. The factor used is derived from the CES; the average expenditure in 2010 is $2,588 and in 2015 it is $3,997, thus the factor applied is 1.544.

Expenses are then segmented by five ranges of household income ($0-$20,000; $20,000-$40,000; $40,000-$60,000; $60,000-$100,000; and, $100,000 and above) and applied to the modeled autos per household and annual VMT for the appropriate income range.

Transit Use Costs

The 2015 National Transit Database (NTD) served as the source for transit cost data. Specifically, directly operated and purchased transportation revenue were used. The transit revenue, as reported by each of the transit agencies in the 2015 NTD, was assigned to agencies and related geographies where GTFS data were collected. This transit revenue was allocated to the counties served based on the percentage of each transit agency's bus and rail stations weighted by the number of trips provided within each county served. For example, if a transit agency had a total of 500 bus stops and 425 of those stops were located in county A, and 75 stops extend into a neighboring county B, and all stops are served at the same level of frequency, county A received 85 percent of the transit revenue and county B received 15 percent.

To estimate average household transit costs, the modeled percentage of transit commuters and total households in each block group was used. Each county's estimated transit revenue was assigned to block groups on this basis. The block group number of transit commuters is calculated and summed to estimate the total number of transit commuters in the county. The county-wide transit revenue is then allocated to block groups based on the proportion of the county's commuters living there. The average household transit cost for each block group is calculated by dividing the block group's allocation of transit revenue by number of households.

This same method was used to estimate the average number of household transit trips for each block group. Using the total unlinked trips from the 2015 NTD, this measure was estimated using allocation the total number of annual trips in each metropolitan area proportionally to block groups based on number of households and the percent of journey to work trips.

There are a number of counties for which GTFS data are not available and/or there was no revenue listed in the 2015 NTD. In these cases, the national averages from previous paragraphs were used for these counties. The average transit costs and trips were then allocated to the block group level based on the percentage of transit commutes and household commuter counts. The end result was an average household transit cost and transit trips for all block groups.

Constructing the H+T Index

Because the H+T Index was constructed to estimate the three dependent variables (auto ownership, auto use, and transit use) as functions of independent variables, any set of independent variables can be altered to see how the outputs are affected. In order to focus on the effects of the built environment, the independent household variables (income, household size, and commuters per household) were set at fixed values. This controls for any variation in the dependent variables that is a function of household characteristics, leaving the remaining variation a sole function of the built environment. In other words, by establishing and running the model for a "typical household," (one defined as earning the regional area median income, having the regional average household size, and having the regional average number of commuters per household) any variation observed in transportation costs is due to place and location, not household characteristics.

The Regional Typical Household takes into account all types of households in the region, and does not represent a specific household, but an average of all households. Every region has a unique mix of households: two-commuter households, single-earner households, adults with no children, single people, etc. - so the Regional Typical Household represents a composite of the broad range of households within a region.

Model Findings

The following three tables show the results of the regressions. The *Function* columns indicate what linearization function was used, the *Value* column give the value of the fit coefficient, the *Error* column gives the value of the standard error on the coefficient.

Table 6: Results of Auto Ownership Regression (R^2 = 80.65%)

Variable 1	Fun 1	Variable 2	Fun 2	Val	Error
Gross HH Density	ln(x)	**Household Gravity**	x	-.00000009	.00000003
Block Density	ln(x)	**Household Gravity**	x	.00000021	.00000004
Block Density	ln(x)	**TAS Jobs**	ln(1+x)	.0007	.0001
Gross HH Density	ln(x)	**TAS**	$\sqrt{x}$	.00006	.00001
Fraction Rental Housing Units	x	**TCI**	x	-.0042	.0007
Fraction Single Family Detached	x	**TAS Jobs**	ln(1+x)	.0037	.0006
Fraction Rental Housing Units	x	**Household Gravity**	x	.0000009	.0000001
Fraction Single Family Detached	x	**TAS**	$\sqrt{x}$	-.00048	.00006
Fraction Rental Housing Units	x	**Level of Service**	x	.000010	.000001
HH Size	x	**Gross HH Density**	ln(x)	.0046	.0004
Block Density	ln(x)	**TCI**	x	-.0020	.0002
Fraction Single Family Detached	x	**Fraction Single Family Detached**	x	-.084	.008
Commuters per HH	x	**Commuters per HH**	x	-.033	.003
Commuters per HH	x	**Fraction Single Family Detached**	x	.094	.007
Commuters per HH	x	**Level of Service**	x	-.0000086	.0000006
Block Density	ln(x)	**Commuters per HH**	x	.021	.001
Gross HH Density	ln(x)	**Gross HH Density**	ln(x)	.0024	.0001

Variable 1	Fun 1	Variable 2	Fun 2	Val	Error
Gross HH Density	ln(x)	**TCI**	x	-.0028	.0002
Employment Mix	x	**Fraction Single Family Detached**	x	.0024	.0001
HH Size	x	**HH Size**	x	-.0168	.0008
TCI	x	**TCI**	x	.000137	.000006
Commuters per HH	x	**Fraction Rental Housing Units**	x	.197	.009
HH Size	x	**TCI**	x	-.0029	.0001
Block Density	ln(x)	**Fraction Rental Housing Units**	x	.048	.002
Fraction Single Family Detached	x	**Household Gravity**	x	.0000029	.0000001
TAS	$\sqrt{x}$			-.00133	.00005
Commuters per HH	x	**Household Gravity**	x	-.00000155	.00000005
Fraction Single Family Detached	x	**Gross HH Density**	ln(x)	-.047	.002
Commuters per HH	x			.34	.01
Intercept				-3.10	.03
Block Density	ln(x)	**Employment Mix**	x	-.00086	.00002
HH Size	x	**TAS**	$\sqrt{x}$	.00059	.00001
HH Size	x	**Median HH Income**	ln(x)	.0196	.0005
HH Size	x	**Fraction Rental Housing Units**	x	-.169	.004
Employment Gravity	ln(x)			-.149	.001
Median HH Income	ln(x)			.196	.002
Employment Mix	x			.0365	.0004

Table 7: Results of Auto Use (VMT) Regression (R^2 = 83.23%)

Variable 1	Fun 1	Variable 2	Fun 2	Val	Error
HH Size	1/x	**Block Density**	$\sqrt{x}$	5092	2139
Fraction Rental Housing Units	x	**Fraction Rental Housing Units**	x	-1603	624
Fraction Single Family Detached	ln(1+x)	**TAS Jobs**	$\sqrt{x}$	-3	1
Gross HH Density	ln(1+x)	**TAS**	x	-.009	.003
Level of Service	$\sqrt{x}$	**TCI**	$\sqrt{x}$	5	1
Fraction Single Family Detached	ln(1+x)	**Level of Service**	$\sqrt{x}$	-39	8
Median HH Income	ln(x)	**TAS Jobs**	$\sqrt{x}$	-.39	.08
TCI	$\sqrt{x}$	**TCI**	$\sqrt{x}$	-94	18
Commuters per HH	1/(1+x)	**TAS Jobs**	$\sqrt{x}$	9	2
Fraction Rental Housing Units	x	**Level of Service**	$\sqrt{x}$	-40	7
HH Size	1/x	**TAS**	x	.10	.02
HH Size	1/x	**HH Size**	1/x	17184	2511
HH Size	1/x	**Fraction Rental Housing Units**	x	14159	1787
Fraction Rental Housing Units	x			-6692	785
Block Density	$\sqrt{x}$			-10202	990
Block Density	$\sqrt{x}$	**Household Gravity**	$\sqrt{x}$	31	3
HH Size	1/x	**Fraction Single Family Detached**	ln(1+x)	12222	978
HH Size	1/x	**Median HH Income**	ln(x)	-3624	206
Intercept				21791	1228

Variable 1	Fun 1	Variable 2	Fun 2	Val	Error
Employment Mix	x	**Median HH Income**	ln(x)	25	1
Commuters per HH	1/(1+x)	**Employment Mix**	x	-230	10
Household Gravity	$\sqrt{x}$	**Employment Gravity**	ln(1+x)	-3.4	.1

Table 8: Results of Transit Use (R^2 = 74.7%)

Variable 1	Fun 1	Variable 2	Fun 2	Val	Error
Gross HH Density	$\sqrt{x}$	**Gross HH Density**	$\sqrt{x}$	-.016	.007
Block Density	x	**Level of Service**	$\sqrt{x}$	-.018	.007
Block Density	x	**Block Density**	x	-1.5	.6
Gross HH Density	$\sqrt{x}$	**Median HH Income**	x	-.0000035	.0000006
Block Density	x	**Gross HH Density**	$\sqrt{x}$	.9	.1
Median HH Income	x	**TAS Jobs**	$\sqrt{x}$	-.000000022	.000000004
Block Density	x	**Employment Mix**	x	-.071	.008
Commuters per HH	1/(1+x)	**Commuters per HH**	1/(1+x)	-4.3	.5
Fraction Single Family Detached	$\sqrt{x}$	**Employment Gravity**	x	.000020	.000002
Fraction Single Family Detached	$\sqrt{x}$	**Gross HH Density**	$\sqrt{x}$	-1.00	.09
Block Density	x	**Employment Gravity**	x	-.000022	.000002
HH Size	x	**HH Size**	x	-.27	.02
HH Size	x	**Household Gravity**	x	-.0000101	.0000009
Median HH Income	x	**TCI**	x	.0000017	.0000001
Employment Gravity	x	**Level of Service**	$\sqrt{x}$	.000000092	.000000008
Block Density	x	**Median HH Income**	x	-.000069	.000005
Commuters per HH	1/(1+x)	**Fraction Single Family Detached**	$\sqrt{x}$	5.3	.4
Level of Service	$\sqrt{x}$	**Median HH Income**	x	-.00000052	.00000004
Gross HH Density	$\sqrt{x}$	**TCI**	x	.027	.002
Fraction Single Family Detached	$\sqrt{x}$	**Household Gravity**	x	.000065	.000005
HH Size	x			2.5	.1
Fraction Single Family Detached	$\sqrt{x}$	**Level of Service**	$\sqrt{x}$	-.083	.005
Block Density	x	**Fraction Single Family Detached**	$\sqrt{x}$	11.7	.7
HH Size	x	**Gross HH Density**	$\sqrt{x}$	.41	.02
Gross HH Density	$\sqrt{x}$	**TAS**	$\sqrt{x}$	-.0057	.0003
Block Density	x	**Household Gravity**	x	.000120	.000006
HH Size	x	**TAS**	$\sqrt{x}$	-.0092	.0004
Median HH Income	x	**Median HH Income**	x	.000000000177	.000000000008
HH Size	x	**Median HH Income**	x	.0000146	.0000007
Employment Gravity	x	**Employment Gravity**	x	.0000000000201	.0000000000009
HH Size	x	**Fraction Single Family Detached**	$\sqrt{x}$	-1.91	.08

Variable 1	Fun 1	Variable 2	Fun 2	Val	Error
Commuters per HH	1/(1+x)	**Median HH Income**	x	-.000093	.000004
TAS	$\sqrt{x}$	**TAS**	$\sqrt{x}$	-.000093	.000004
Fraction Single Family Detached	$\sqrt{x}$	**TAS**	$\sqrt{x}$	.042	.002
Employment Mix	x	**Fraction Rental Housing Units**	x	-.119	.004
HH Size	x	**Level of Service**	$\sqrt{x}$	.037	.001
Employment Gravity	x	**Median HH Income**	x	-.000000000268	.000000000009
Commuters per HH	1/(1+x)	**Fraction Rental Housing Units**	x	20.3	.7
Fraction Rental Housing Units	x	**Household Gravity**	x	.000062	.000002
Household Gravity	x	**Household Gravity**	x	-.000000000304	.000000000009
Household Gravity	x	**Median HH Income**	x	.00000000097	.00000000003
Fraction Single Family Detached	$\sqrt{x}$	**TAS Jobs**	$\sqrt{x}$	-.0225	.0006
TCI	x			.96	.01
TCI	x	**TCI**	x	-.0120	.0003
Employment Mix	x			-.329	.007
Commuters per HH	1/(1+x)	**Employment Gravity**	x	-.000098	.000002
Commuters per HH	1/(1+x)	**Household Gravity**	x	.000236	.000005
Intercept				25.1	.6
TAS Jobs	$\sqrt{x}$			.0220	.0004

Neighborhood Characteristic Scores

The H+T is based on the idea that some places are more efficient than others, a concept known as location efficiency. One way to measure this efficiency is to examine the extent to which a place is auto dependent. By looking at the place driven components of the regression equation to predict auto ownership (and in one case the transit use equation), comparisons between places can be made. Location efficiency can be scored by controlling for household characteristics and examining at how block groups compare with one another with regard to compact development, access to employment and variety of jobs, and level of transit service. Three scores were developed to make such comparisons: the Compact Neighborhood Score, Job Access Score and AllTransit™ Access Score. All are available on the H+T mapping tool, data download, and the H+T Fact Sheet.

They are all scores in the sense that they do not have a direct value of location efficiency to them, but are the rank of the block group relative to all other block groups in the H+T Index. This is accomplished by first evaluating the components of the equation of the subset of independent variables (for example, the Job Access Score uses Employment Access, and Job Mix Index), then this number (V_r) is scaled from 0 to 100 (I_r), and then all the block groups are ranked and given a number from 0 to 10 (S_{10}) reflecting their rank. The final score is one tenth of the percentile they fall into; a score of 5.5 for a particular block groups represents that that block group is in the 55th percentile of all block groups. The following equations show this calculation:

$$V_r = \sum_{i=1}^{n} C_i \times f_i(X_i)$$

Equation 9: Calculation of Generic Raw Value V_r

Where:

i is the index or the variables used in this score
n is the total number of variables used for this score
C_i is the fit coefficient from the regression equation for the i^{th} variable
X_i is the value of the i^{th} variable for this block group
$f_i()$ is the linear transformation for the i^{th} variable

This value is then transformed into a number from 0 – 100 by using the same equations used in the Bus Access Index, the Rail Access Index and the Employment Mix Index, shown below:

$$I_r \equiv 100 \times \frac{V_r - V_{min}}{V_{max} - V_{min}}$$

Equation 10: Calculation of Generic Raw Index I_r

Where:

V_{min} is the minimum value for all block groups and
V_{max} is the maximum value for all block groups.

The value of this index is used then to rank all block groups (using a "dense ranking" where two block groups with the exact same value get the same rank, and the next one in gets the next rank) then this rank is turned into a number from 1 to 10 much as above:

$$S_{10} \equiv 10 \times \frac{R_r - R_{min}}{R_{max} - R_{min}}$$

Equation 11: Calculation of Generic Score S_{10}

Where:

R_r is the dense rank of the block group
R_{min} is the minimum dense rank (usually equal to one)
R_{max} is the maximum dense rank

This then gives the score which goes from 0 to 10.

The three scores use different inputs and regression equations, listed in Table 9 below.

Table 9: Neighborhood Characteristic Scores Definitions

Score	List of Independent Variables	Regression Equation
Compact Neighborhood Score	• Gross Household Density • Regional Household Intensity • Fraction of Single Family Detached Housing • Fraction of Rental	Autos per Household

	Housing • Block Density	
Job Access Score	• Employment Gravity • Employment Mix Index	Autos per Household
AllTransit™ Performance Score	• Transit Connectivity Index • TAS Jobs • Average Available Transit Trips per Week	Percent Transit Journey to Work