

# AllTransit™ Methods

## **AllTransit Database & Web Tool**

AllTransit is a broad and comprehensive transit database designed to comply with the General Transit Feed Specification (GTFS) in which transit data are developed and assembled by CNT. It includes the location of stops and routes, and the type and frequency of service for all scheduled bus, rail, and ferry service in metropolitan regions over 100,000 in population.

Publicly available GTFS data was gathered, all remaining agencies were contacted to enquire about the availability of GTFS. 102 agencies shared the data with us. For the 305 agencies without GTFS data, we compiled route maps and schedules from their website and/or made phone calls to agencies to get the information. These maps and schedules served as reference to recreate the routes on a simple user-friendly online tool (AllTransit Data Builder) that CNT created.

The previous version of the AllTransit dataset was developed to provide a robust measure of transit access for CNT's ground-breaking Housing and Transportation Affordability (H+T®) Index. We updated and expanded the dataset, and built a more robust and user friendly map-based website devoted solely to AllTransit. It reports measures of transit and its impact for all regions with scheduled transit service and a population of over 100,000.

## **Geographic Level and Data Availability**

AllTransit data was constructed at the Census block group level. Currently 371 Metropolitan and Micropolitan Areas in the United States, also known as Core Based Statistical Areas (CBSAs), defined by the Office of Management and Budget in 2013 with more than 100,000 population and scheduled transit service are included. Smaller regions with readily available GTFS data are also included. CBSAs in Puerto Rico are not included due to insufficient data.

## **Data Sources**

1. 2010-2014 American Community Survey 5-year Estimate (2014 ACS) – an ongoing U.S. Census survey that generates data on housing characteristic, transportation use, community demographics, income, and employment.
2. U.S. Census TIGER/Line Files – geographical features such as roads, railroads, and rivers, as well as legal and statistical geographic areas.
3. U.S. Census Longitudinal Employment-Household Dynamics (LEHD) Origin-Destination Employment Statistics (LODES) – detailed spatial distributions of workers' employment and characteristic detail on age, earnings, industry distributions, and local workforce indicators. LODES data built on 2014 Census data are used here.

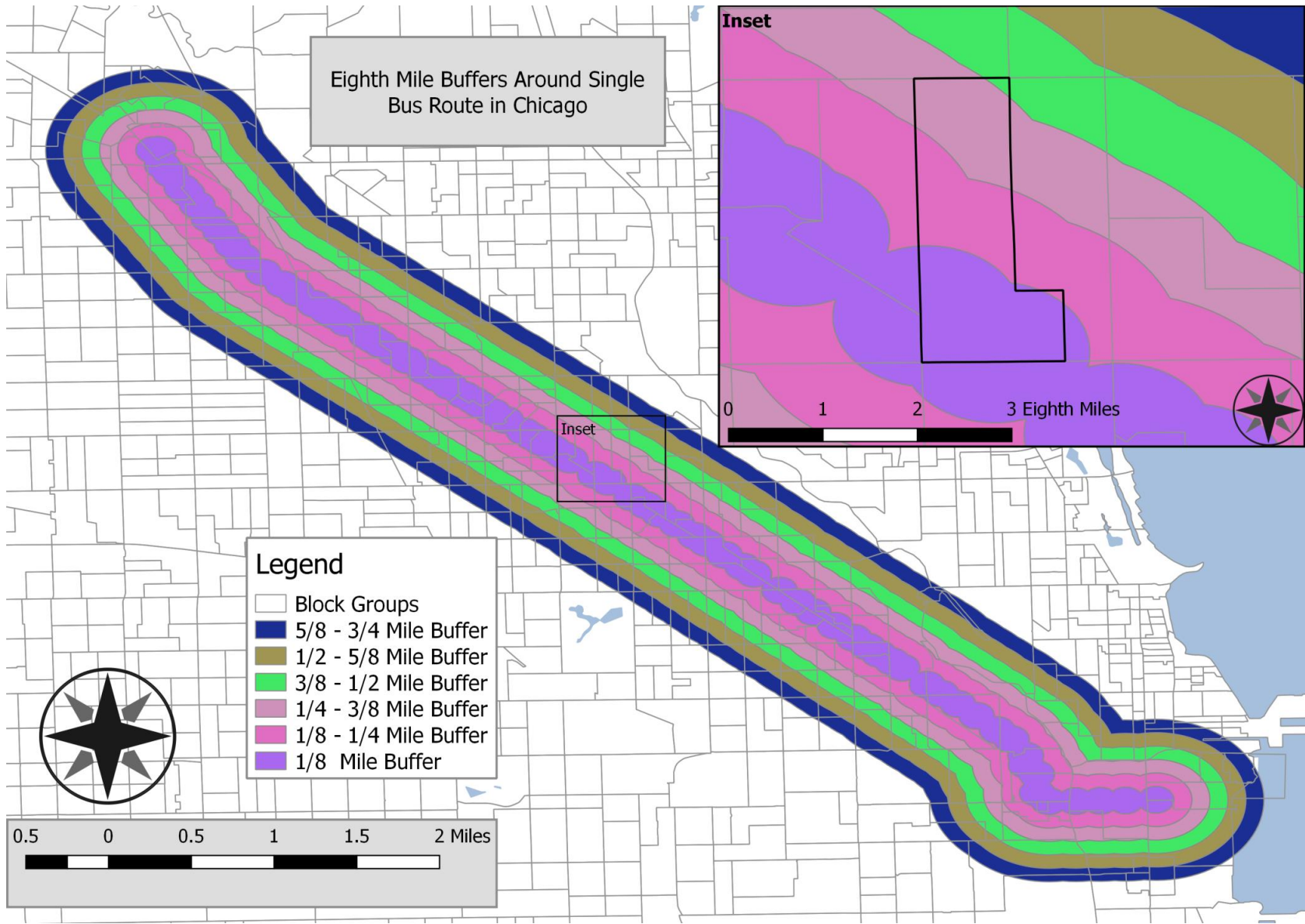
4. AllTransit – a 2015 database of General Transit Feed Specification (GTFS) data developed by the Center for Neighborhood Technology, including bus, rail, and ferry service for both transit agencies that report their GTFS data publicly and those derived by CNT staff for agencies that do not.
5. Low-Income Housing Tax Credit Program – a database maintained by US Department of Housing and Urban Development with information on affordable housing locations.
6. U.S. Department of Agriculture – National Farmers Market Directory

### **Calculations and Definitions**

1. **Transit Stop** – A transit stop is defined from the AllTransit GTFS data. An individual stop is derived from the GTFS stop file ([defined here](#)) and has a few nuances that are worth pointing out.
  - a. If a transit route has stops on either side a street, for example an east bound bus on one side and a west bound bus on the other, that counts as two stops because transit service in both directions provides value.
  - b. Some transit agencies define unique stops for each route using the same location. In this case we would count each unique stop, for example where two bus lines intersect and share a stop this may be coded as one stop for some agencies, but two for others.
  - c. The distinction between a station and a stop is ignored for the purposes of this website.
2. **Transit Route** – A transit route is defined as a set of all stops that are used on a single directional route and is derived from the GTFS route file ([defined here](#)). Each transit agency codes things slightly differently, and the direction is not always available, for example if an elevated rail line runs into a loop (as in the Brown Line in Chicago); where the route begins and ends is not well defined, so there may not be a direction defined for such a route. However, the most important use for the Transit Route in the website is for calculating frequency of service and the lack of consistency of the direction of a route does not affect the overall total trips per week, since we sum that over all directions.
3. **Near Transit** – A ½ mile buffer around each transit stop is the geography of calculation for all “near transit” metrics. Jobs, workers, commuters, households, and population characteristics within all the buffers is summed to calculate the near transit values.
4. **Near High Frequency Transit** – A ½ mile buffer around stops on routes that meet our criteria for “High Frequency Transit” is calculated and data is aggregated to the sum of all such buffers. Household and population characteristics are summed to obtain these calculations.
5. **High Frequency Transit** – High frequency transit is defined three ways:
  - a. Around the Clock – Transit service that operates 7 days per week throughout daytime, evening and overnight hours, and maintains average headways of 15 minutes or less.
  - b. Full Day – Transit service that operates 7 days per week and maintains average headways of 15 minutes or less between the hours of 7AM and 10PM.
  - c. Rush Hour – Transit service that operates Monday through Friday and maintains average headways of 15 minutes or less between peak commuting hours of 7AM – 9AM and 4PM – 6PM.

6. **Transit Shed** – The Transit Access Shed (TAS) is defined as a geographic area accessible within 30 minutes by public transportation. For each transit stop, all stops that can be reached within 30 minutes were identified. One transfer within 1/4 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 (a quarter mile) of the block group are merged into one grand shed. This area is then assigned as the block group's Transit Access Shed. Jobs, workers, and households data within the each block group's TAS is proportionally summed to it.
7. **Transit Trips Per Week** – Using the transit routes within a ½ mile of a block group the total number is aggregated by summing the total number of trips per week for every such route. Note that by using the route rather than the stops we avoid double and triple counting stops on the same route that serve the block group.
8. **Transit Connectivity Index (TCI)** – The TCI is an index from 0-100 that is scaled by the number of transit trips the average household in a block group can access by walking each week. The TCI is a measure of how connected the average household member is to the availability of a transit ride. This is a placed-based measure that is derived by examining the proximity of all transit routes, and the area covered by each route at different distances. This is accomplished by using a series of six one eighth mile buffers around each route, and calculating the fraction of land area covered by each buffer for a given block group. The map below shows this for a bus route in Chicago. The highlighted block group in the inset has approximately 1/3 coverage from the first eighth mile buffer, approximately 1/3 coverage from the 1/8-1/4 mile buffer, and a little less from the 1/4-3/8 mile buffer, with a very small fraction covered by the 3/8-1/2 mile buffer. Then we add up all the trips per week for each buffer for each route and obtain the frequency of service for the average location in the block group, for a given distance. We take a weighted sum of these six distances to create an overall statistic on how well the average household in the given block group is served each week. This is done separately for bus, and then again for all other transit modes (streetcar, light rail, commuter rail, ferry, etc.). The weights for this weighted sum are determined by using an ordinary least square fit to the percent of people using transit for their journey to work, controlling for the households (using people per household, household income, and commuters per household) and other locational variables (average block size, fraction of renter dwelling units, fraction of single family homes, and transit access shed).

Figure 1: Eighth Mile buffers Around the Stops on the CTA 56 Bus Inbound Route



Once this weighted sum is made for every block group in the USA we create an index that ranges from zero to 100, where the value of zero is assigned to block groups that have no connectivity to transit and 100 is given to the best block group in the country. Therefore, this index is a relative rating of how well connected to transit, or “can I walk out the front door and get a bus or train,” for every block group. The graphs below show the value of the frequency distribution of TCI for all block groups in the USA, Figure 3 shows this magnified for only the highest values of TCI, and shows just how rare it is in the USA to have these very high values of transit availability.

Figure 2: TCI Frequency Distribution for All Block Groups

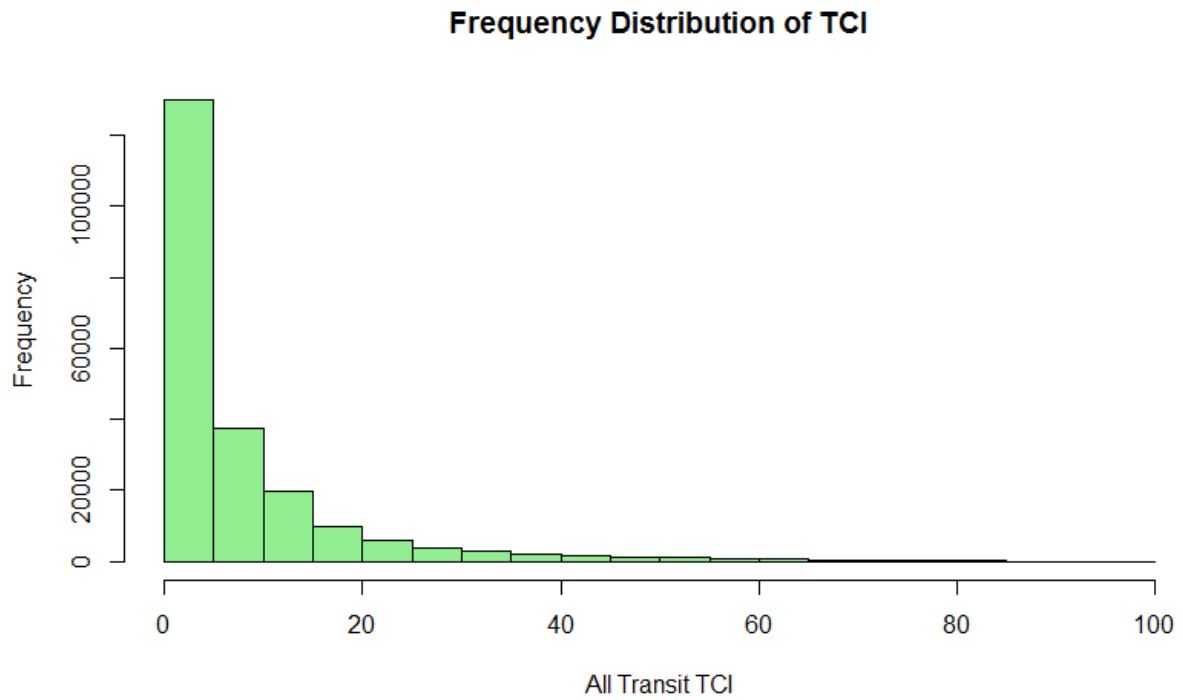
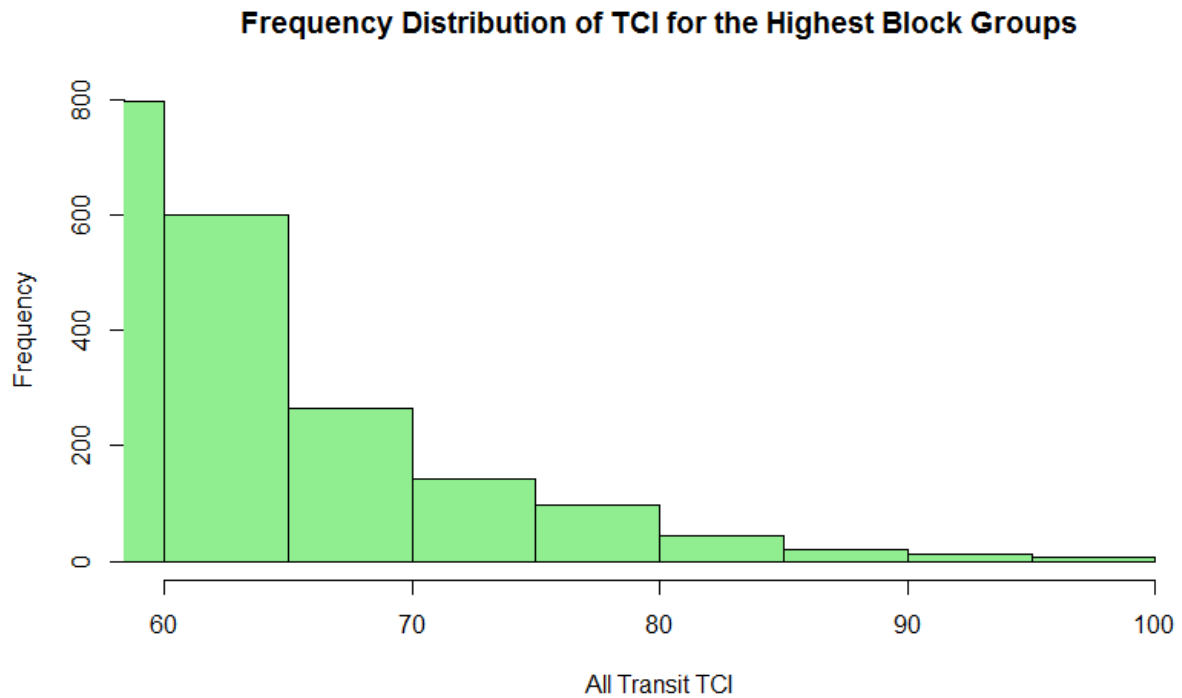


Figure 3: TCI Frequency Distribution for Block Groups with High Value of TCI



**All Transit Performance Score** – This scores the overall quality of transit as it pertains to actual use of transit. It combines with the TCI, described above, that measures “can I get transit” with the Jobs Accessible in 30 Minute Transit Ride, which measures “what can I get to once I am on transit,” and combines them in a way that is reflective of what fraction of people use transit for a given type of trip – their journey to work – answering the question “how should I get to work.” In order to determine how to combine these two measures, like the TCI, we use an ordinary least square fit using these two variables as well as the same control variables, for the households (people per household, household income, and commuters per household) and other locational variables (average block size, fraction of renter dwelling units, fraction of single family homes, and transit access shed size). Once this weighted sum is made for every block group in the USA we create an index that ranges from zero to 100, where the value of zero is assigned to block groups that have no connectivity to transit and 100 is given to the highest value block group in the country. Figure 4 shows the frequency distribution for the percent of commuters using transit for their journey to work. Figure 5 shows the frequency distribution of this intermediate index for all block groups in the USA, Figure 6 shows this magnified for only the highest values of the intermediate index, and again shows how rare it is in the USA to have these very high values of transit quality.

Figure 4: Percent Journey to Work by Transit Frequency Distribution for All Block Groups

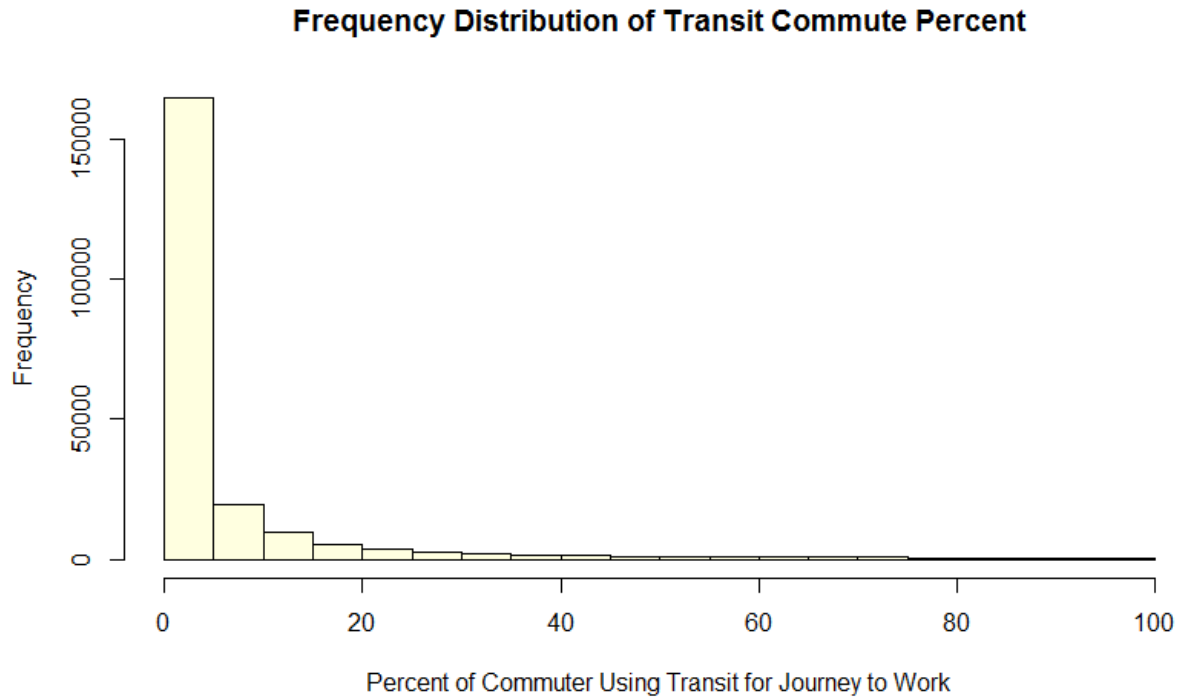


Figure 5: Intermediate Index TCI Frequency Distribution for All Block Groups

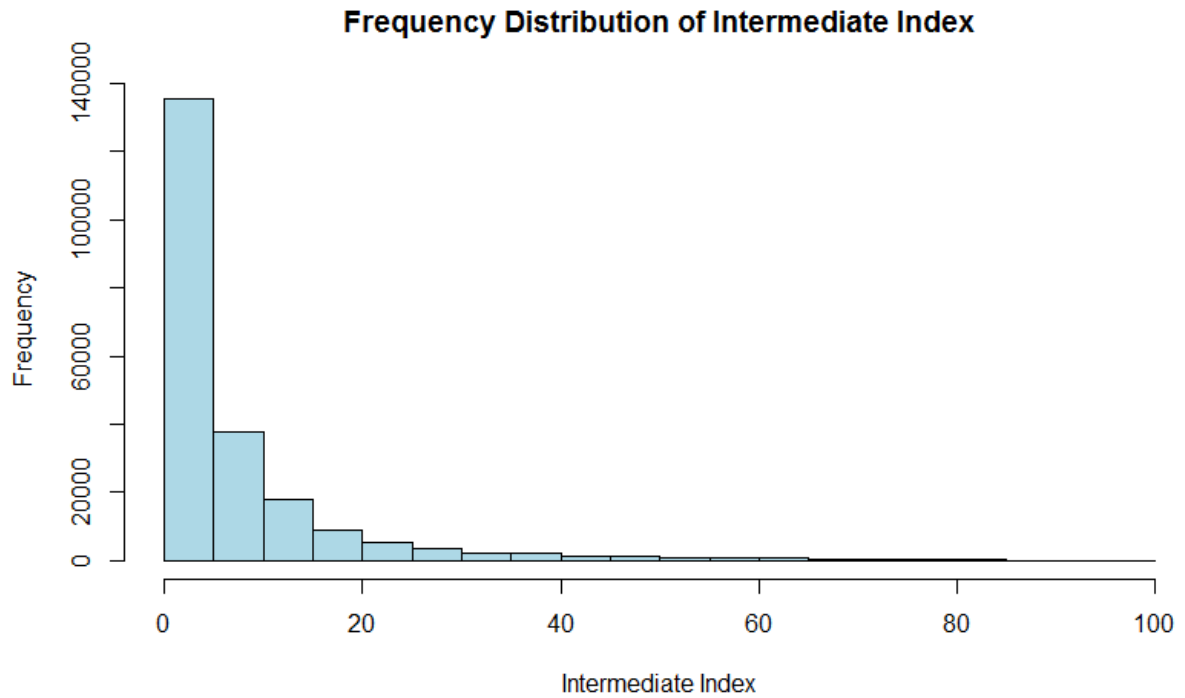
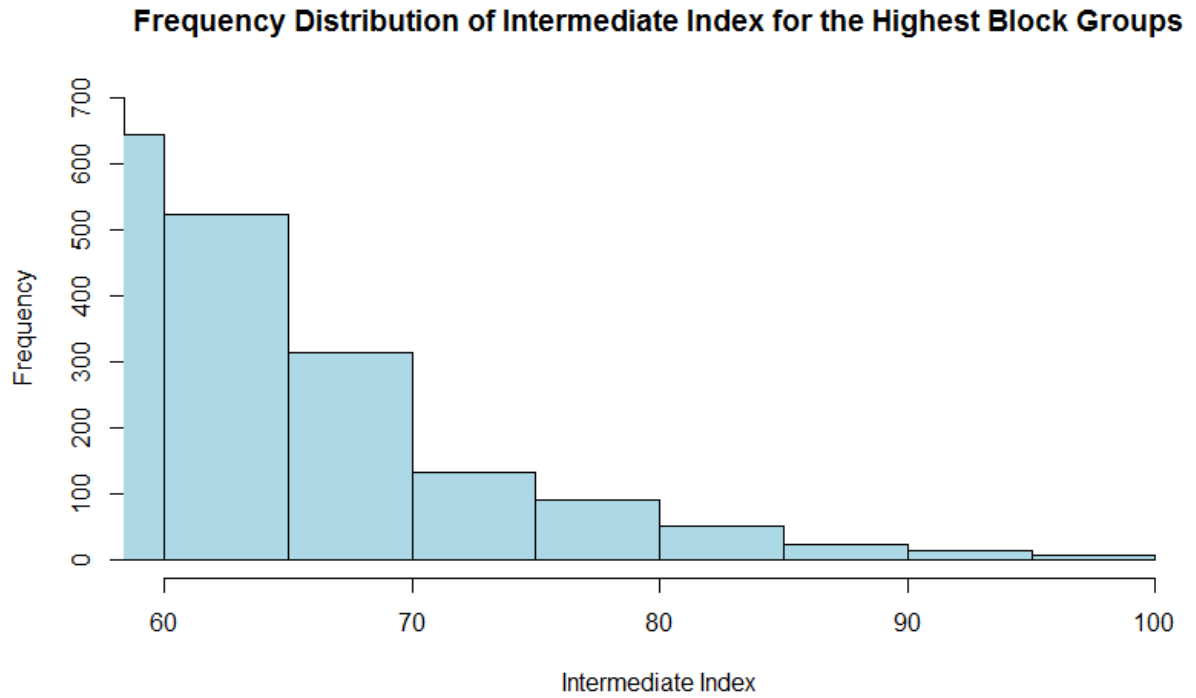


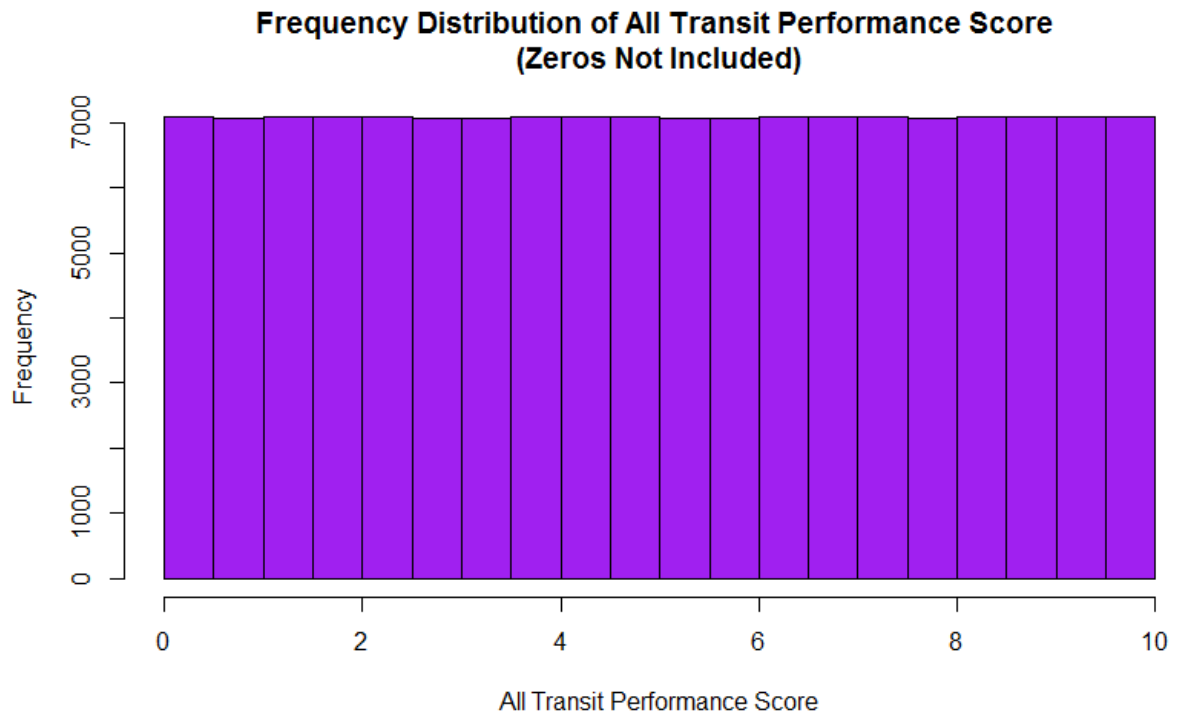
Figure 6: Intermediate Index TCI Frequency Distribution for the Highest Value Block Groups



The All Transit Performance Score is calculated by ranking for all block groups from highest to lowest, using the this intermediate index, then rescaling this rank value so that the highest ranking block group is given 10.0 and the lowest are given 0 (there are many with value = 0), making the score a number between 0 and 10 which essentially is the percentile for the block group (divided by 10). Figure 7 shows the frequency distribution for the All Transit Performance Score.



Figure 7: All Transit Performance Score Frequency Distribution for All Block Groups



### Metrics – Calculation & Data Source Overview

Metric	Calculation	Data Source
<b><i>Jobs</i></b>		
Jobs accessible in 30 minute transit ride	Jobs data aggregated to the Transit Access Shed, weighted by households	Longitudinal Employer-Household Dynamics, 2014
Workers accessible in 30 minute transit ride	Workers data aggregated to the Transit Access Shed, weighted by households	Longitudinal Employer-Household Dynamics, 2014
Jobs near transit	Jobs data aggregated to ½ mile around transit	Longitudinal Employer-Household Dynamics, 2014; AllTransit Data
Workers near transit	Workers data aggregated to ½ mile around transit	Longitudinal Employer-Household Dynamics, 2014; AllTransit Data
Commuters near transit	Commuters data aggregated to ½ mile around transit	American Community Survey 5-year Estimate (2014 ACS); AllTransit Data
Total Commuters	Sum of commuters in selected geography	American Community Survey 5-year Estimate (2014 ACS)
<b><i>Economy</i></b>		
Customer Households Accessible in 30 minute transit ride	Households data aggregated to the Transit Access Shed	American Community Survey 5-year Estimate (2014 ACS)
Transportation Costs	Transportation costs as a percent of income for regional typical household for block groups within ½ mile of transit - calculated from the H+T Index	CNT's Housing + Transportation Affordability Index, American Community Survey 5-year Estimate (2014 ACS) ; AllTransit Data
<b><i>Health</i></b>		
Walkable neighborhoods	Census block sizes aggregated to ½ mile around transit	U.S. Census TIGER/Line Files, 2014
Commute by bicycle	Bicycle commuters aggregated to ½ mile around transit	American Community Survey 5-year Estimate (2014 ACS) ; AllTransit Data
Commute by walking	Commuters by walking data aggregated to ½ mile around transit	American Community Survey 5-year Estimate (2014 ACS) ; AllTransit Data

Farmers markets near transit	Farmers market locations aggregated to ½ mile around transit	U.S. Department of Agriculture: National Farmers Market Directory, 2015 ; AllTransit Data
Farmers markets accessible in 30 minute transit ride	The count of all farmers markets within the Transit Access Shed	U.S. Department of Agriculture: National Farmers Market Directory, 2015 ; AllTransit Data
<b>Equity</b>		
Population near transit	Population data aggregated to ½ mile around transit	American Community Survey 5-year Estimate (2014 ACS) ; AllTransit Data
Households near transit	Household data aggregated to ½ mile around transit	American Community Survey 5-year Estimate (2014 ACS) ; AllTransit Data
Population near high frequency transit	Population data aggregated to ½ mile around high frequency transit	American Community Survey 5-year Estimate (2014 ACS) ; AllTransit Data
Households near high frequency transit	Household data aggregated to ½ mile around high frequency transit	American Community Survey 5-year Estimate (2014 ACS) ; AllTransit Data
LIHTC buildings near transit	Total low-income housing building locations aggregated to ½ mile around transit	US Department of Housing and Urban Development - LIHTC Database, 2013; AllTransit Data
LIHTC units near transit	Total low-income housing unit locations aggregated to ½ mile around transit	US Department of Housing and Urban Development - LIHTC Database, 2013; AllTransit Data
<b>Transit Quality</b>		
AllTransit Performance Score	Overall transit score for a neighborhood that includes measures of transit quality, the jobs that can be reached in a 30 minute transit ride, compared to the number of people using transit for commuting	AllTransit Data
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	AllTransit Data

Transit Trips per week	Total trips possible within the block group and ½ mile of its border	AllTransit Data
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	AllTransit Data
<b>Mobility</b>		
Transit Routes within ½ mile	Total transit routes available within the block group and ½ mile of its border	AllTransit Data
High Frequency Transit Routes within ½ mile	Total high frequency transit routes available within the block group and ½ mile of its border	AllTransit Data
Transit Stops within ½ mile	Total transit stops available within the block group and ½ mile of its border	AllTransit Data
Bikeshare locations near transit	Total bikeshare locations aggregated to ½ mile around transit	AllTransit Data
Carshare locations near transit	Total carshare locations aggregated to ½ mile around transit	AllTransit Data
Carshare vehicles near transit	Total carshare vehicles aggregated to ½ mile around transit	AllTransit Data