Relative Accessibility and the Choice of Modes

Andrew Owen[*] Paul Anderson[†] David Levinson[‡]

August 1, 2012

Abstract

The factors influencing commute mode choice are a subject of ongoing research and policy. Existing literature explores a wide range of factors which may influence mode choice; many of these focus on demographic factors as well as user preferences and perception, thereby highlighting the unique characteristics of each mode. This analysis hypothesizes that mode share at the origin, an aggregate expression of individuals' mode choices, is determined in large part by more fundamental properties of transportation systems. Accessibility, a measurement of the ease of reaching destinations, is used as a tool for comparing modes which focuses on their properties as abstract transportation systems. It explores the potential to predict the relative commute shares of non-auto and auto modes from the relative accessibility provided by each. Using public data sources and methods selected for their simplicity and ease of interpretation, a model is estimated which accounts for 41% of the variation in commute mode share at the origin block group level in the Minneapolis–Saint Paul, MN metropolitan area.

1 Introduction

It is increasingly common for urban transportation planning agencies to establish goals of increasing transit ridership. For example, in 2010 the Metropolitan Council of the Twin Cities adopted the goal of doubling transit ridership in its region by 2030 [7]. In seeking that and similar goals, the Council and other agencies will be guided in part by the answer to the question: what makes a traveler choose to make a trip by transit rather than some other mode?

In almost all cases, that "other mode" is driving: according to estimates from the American Community Survey for the years 2006–2010, 90.1% of commute trips in the United States were made by car. Driving is a very different experience than using transit: driving is on-demand while transit is schedule-based; a train passenger can read a book while a driver ought not; a bus passenger can talk to other passengers while a solo driver is relatively isolated; a driver may pay to store her vehicle at the

---

[*] University of Minnesota, Department of Civil Engineering, 500 Pillsbury Drive SE, Minneapolis, MN 55455 USA, owenx148@umn.edu

[†] University of Minnesota, Department of Civil Engineering, ande9474@umn.edu

[‡] University of Minnesota, Department of Civil Engineering, dlevinson@umn.edu `http://nexus.umn.edu`

end of a trip while a streetcar passenger pays to board at the beginning of his. Travelers undoubtedly consider all these differences and more when choosing a mode.

When researchers and planners describe transit and driving as separate *modes* they are focusing on these differences. However, transit systems and road networks share a fundamental purpose: as transportation *systems*, they both are created with the intent that people will use them to reach destinations by paying some cost (in time and/or money). When viewed from this standpoint, it becomes apparent that travel by transit and travel by auto, regardless of their myriad differences, can be compared along two fundamental dimensions: the set of destinations to which they provide access, and the cost of reaching those destinations. Destinations and their cost of access are integrated by measures of accessibility; this study investigates the value of considering the relative accessibility provided by transit and auto in predicting mode share.

2 Background

2.1 Accessibility

Accessibility marries the simpler concept of mobility with an understanding that travel is driven by a desire to reach destinations. It is important to distinguish between *individual accessibility* and *locational accessibility*: the former seeks to characterize the ease with which travelers might reach their destinations, subject to constraints of ability, budget, and other barriers; the latter examines accessibility as a spatial phenomenon by considering the costs and benefits of the potential trips offered by transportation systems between origins and destinations of interest. Horner [5] explores this distinction in the literature and notes that individual accessibility measures are generally poor at "producing…generalized assessments of intraurban structure," while locational accessibility measures are more useful for "understanding relationships between transportation and land use."

Locational accessibility can be a particularly useful tool for transportation planners because it provides a way to evaluate the properties of transportation systems at a level that is aggregate enough to avoid the vagaries of individual users' preferences and constraints, but still detailed enough to provide guidance for planning at the city and regional level. It can be especially useful for multi-modal transportation planning because it is able to provide a level playing field for evaluating modes relative to one another [1]. This is achieved by setting aside the many particular differences between transportation *modes* and considering their relative merits as transportation *systems*.

Many different implementations of locational accessibility measurement are possible. El-Geneidy and Levinson [3] provide a practical overview of historical approaches. Most contemporary implementations can be traced at least back to Hansen [4], who proposes a measure where potential destinations are weighted by a gravity-based function of their access cost and then summed:

$$A_i = \sum_j O_j f\left(C_{ij}\right) \tag{1}$$

$$\begin{aligned} A_i &= \text{accessibility for zone } i \\ O_j &= \text{number of opportunities in zone } j \\ C_{ij} &= \text{time cost of travel from } i \text{ to } j \\ f\left(C_{ij}\right) &= \text{weighting function} \end{aligned}$$

The specific weighting function $f\left(C_{ij}\right)$ used has a tremendous impact on the resulting accessibility measurements, and the best-performing functions and parameters are generally estimated independently in each study or study area. This makes comparisons between modes, times, and study areas challenging. Levine et al. [6] discuss these challenges in depth during an inter-metropolitan comparison of accessibility; they find it necessary to estimate weighting parameters separately for each metropolitan area and then implement a second model to estimate a single shared parameter from the populations of each.

To further pursue the goal of parsimony and to ease the process of inter-modal accessibility comparisons, this study adopts a simpler, binary weighting function and implements a *cumulative opportunities* measure of accessibility:

$$f\left(C_{ij}\right) = \begin{cases} 1 & \text{if } C_{ij} \leq t \\ 0 & \text{if } C_{ij} > t \end{cases} \tag{2}$$

$$t = \text{travel time threshold}$$

Accessibility is calculated for specific time thresholds and the result is a simple count of destinations that are reachable within that threshold. This approach involves both advantages and disadvantages. Both calculation and interpretation of the accessibility measure are dramatically simplified. But accessibility must be reported separately for each time threshold of interest, and the model cannot be finely calibrated to account for varying user preferences, values of time, etc.

2.2 Mode Share

Mode share within an analysis zone summarizes the result of local individuals' mode choices. Research into mode share and mode choice has explored a wide range of factors which potentially contribute to individuals' mode choice process. Taylor et al. [11] provide a review of prior research into transit mode share which identifies two major avenues of investigation: *descriptive* analyses which focus on "traveller attitudes and perceptions," and *causal* analyses which examine "environmental, system, and behavioral characteristics." Within each, individual factors are identified as *external* if they are generally outside the direct control of transit planners and managers (*e.g.* population, income), or *internal* when they are endogenous to a specific transit system implementation (*e.g.* fares, vehicle design).

Using this classification, accessibility-based investigations of transit mode share can be described as causal, because they specifically avoid reliance on demographics or traveller preferences. They include both internal and external factors: travel times are the direct result of routing and scheduling decisions made by the operator, while the spatial distribution of opportunities is not.

3 Data

Any analysis of accessibility requires data describing the locations of origins and destinations, the cost of traveling between them, and the opportunities available at each.

Automobiles travel across the network of public roads and highways. Calculating travel times through this network requires two types of information: data describing the structure of the network, and data describing the cost of travel along individual links in the networks. The Metropolitan Council, the metropolitan planning organization for the Twin Cities region, maintains a model of the regional road network. It provides a network topography for freeway, arterial, and collector roadways in the region, designed to model travel using transportation analysis zones (TAZs) as origins and destinations. The most recent version of this network was updated in 2009, and it provides an adequately accurate representation of the state of the regional road network in 2010, with model links conflated to match actual geography. It does not model local roads, but instead provides "dummy links" which connect the centroid of each TAZ to adjacent arterial links. These are coded so that they may only be used for direct access to or from a TAZ centroid; they may not be used for travel between zones.

By itself, this model only describes the structure of the road network; per-link speed information is needed in order to generate travel times. The Twin Cities' regional freeway network is very well-instrumented, and data recorded by loop detectors throughout the system are archived by the Minnesota Traffic Observatory, operated and hosted by the University of Minnesota. Archived loop detector data for every weekday in 2010 provided the basis for average freeway link speeds.

Freeway speeds are derived from direct traffic observations made by embedded loop detectors which record the observed traffic volume and detector field occupancy at 30-second intervals. MnDOT provides an estimated average effective vehicle length for each detector, which allows the calculation of speed from volume and occupancy. These 30-second speed measurements are aggregated by averaging to 5-minute time slots. Finally, the 5-minute speeds measurements taken between 7 and 9 AM on weekdays during the year 2010 are averaged for each detector to produce representation of AM peak period speeds. We assign the resulting detector speeds to links in the model network based on location.

In contrast to freeways, local arterials and collectors are only sparsely instrumented. Average arterial and collector link speeds are estimated from speed measurements made during a regional GPS-based travel survey in 2008 [14]. This data represents a very accurate measurement of traffic speeds at specific locations and specific times. Speeds for unobserved links are estimated from the samples collected on similar links.

Transit users interact with a different type of network than automobile drivers. Instead of navigating physical infrastructure, transit users move through a more abstract network of bus and rail routes provided by the transit operator. Metro Transit, the primary transit operator in the Twin Cities region, provides a publicly-available general transit feed specification (GTFS) dataset. This includes bus and rail routes operated by all six fixed-route transit providers in the region.

Data describing the distribution of labor and employment in the region are drawn from the U.S. Census Bureau's Longitudinal Employer-Household Dynamics program (LEHD). The workplace area characteristic dataset for 2010 provides Census block-level counts of employee work locations.

In general, LEHD is a very useful data source for accessibility evaluation because it is updated yearly and is drawn from actual payroll records collected at the state level—in this case, by the Minnesota Department of Employment and Economic Development. However, it is important to address the fact that LEHD data is *synthetic*: while it is based on actual payroll records, the published results are created by an algorithm designed to produce data which are statistically similar to the underlying data, and which converge to the same distribution when aggregated. To facilitate integration with travel times calculated between TAZs, block-level LEHD employment data are aggregated to TAZs.

The American Community Survey (ACS) collects data describing commute mode choice during its annual national survey of households and individuals. For each of the 2,085 block groups in the study area, the ACS estimates the total number of workers, which includes the workers who report commuting by each mode as well as workers who report working from home. Because the opportunity and the decision to work from home is influenced by a wide range of factors and only partially related to mode choice [10], the number of *commuters*—workers who report some mode of travel to work (including "other")—is used as the total population. In order to facilitate the logarithm-based model described below, block groups where the non-auto mode share is zero are excluded Young and Young [12]; this affects 159 block groups (7.5%), leaving 1,926 included in the analysis.

	n	**Mean**	**Median**	**SD**
Mode share ratio	1926	0.16	0.08	0.27
Accessibility ratio (30 min)	1926	0.13	0.09	0.13

Table 1: Distribution statistics of block group mode share ratios (non-auto commuters / auto commuters) and accessibility ratios (jobs with 30 minutes by transit / jobs within 30 minutes by auto) for the study area.

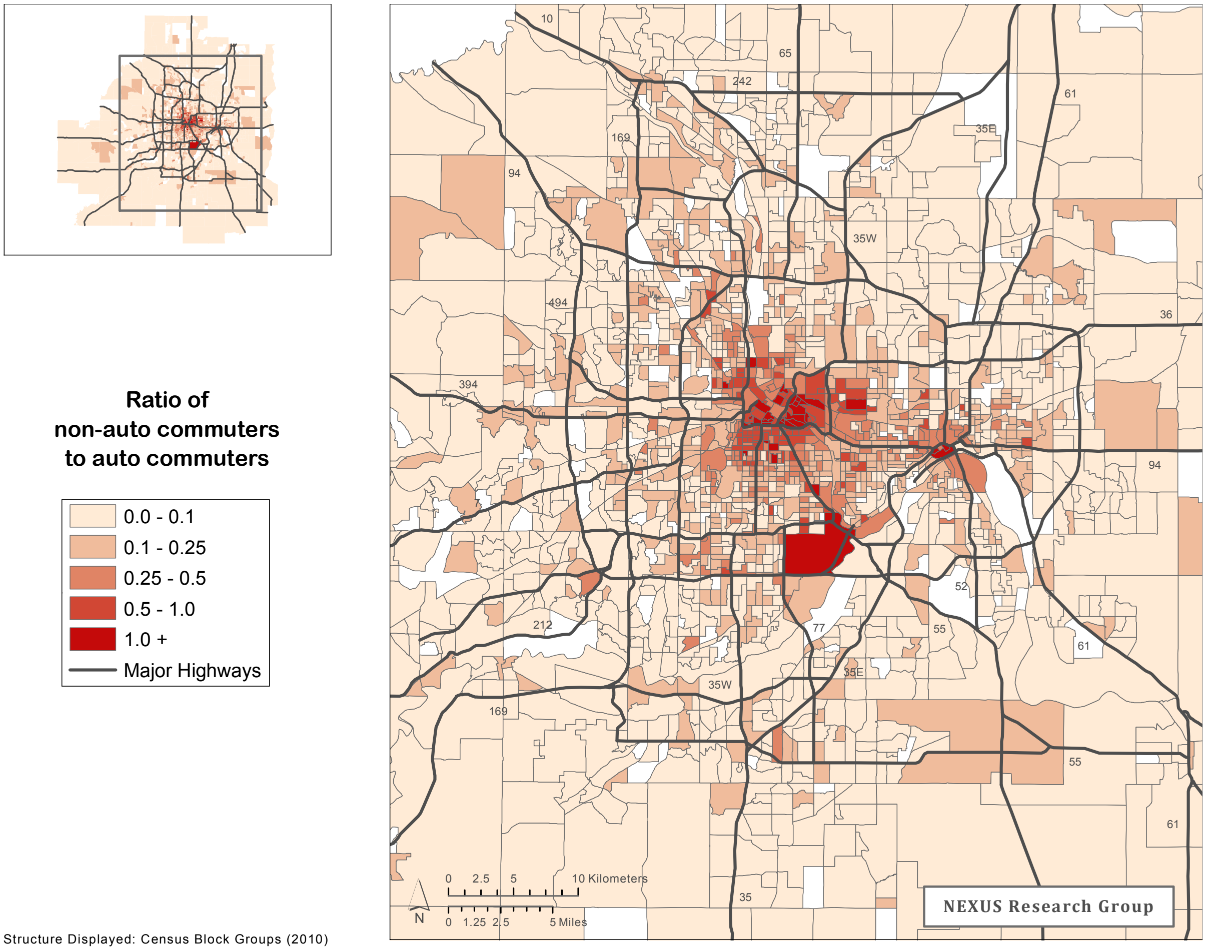
The dependent variable in this analysis is the ratio of non-auto mode share to auto mode share, which, because they are shares from the same population of commuters, is equivalent to the ratio of the number of non-auto commuters to the number of auto commuters. Table 1 describes the distribution of this ratio for 1,926 block groups in the study area, and Figure 1 provides a map of the same information.

4 Calculating Travel Times

Once average AM peak period speeds are assigned to each link in the model network, each link's traversal time cost is calculated from its speed and length. At this point the network is effectively a weighted graph, which is used to calculate a full shortest-path matrix of AM peak period travel times between each pair of TAZ centroids.

Transit travel times consist of four components: time spent accessing a stop or station from an origin, time spent on board transit vehicles, time spend waiting for and making transfers, and time spent accessing a destination from the final stop or station. This investigation depends on the assumptions that origin and destination access is achieved by walking at a speed of 5 km/h, that users make a maximum of one transfer, that users will spend a maximum of 15 minutes waiting for a transfer, and that transfers are possible between stop within 400m of each other. On-vehicle travel times are based on the schedules reported in the GTFS dataset.

To produce a matrix of transit travel times between TAZs, it is necessary to associate transit stops with TAZs and estimate origin and destination access times. Using the coordinates of transit stops



Zone Structure Displayed: Census Block Groups (2010)
Primary Data Source: US Census Bureau

Figure 1: Ratio of non-auto commuters to auto commuters by block group. Non-auto commuters includes all workers who report travelling to work by any mode other than auto or motorcycle. Workers who report working from home are excluded.

as reported in the GTFS dataset, each Census block centroid is linked to the two closest stops, and then any remaining stops are linked to the closest block centroid in order to ensure that all stops are considered. Finally, each stop is linked to the TAZs which contain its linked block centroids. Walking access times between each stop and its linked TAZs are calculated based on the distance between the stop and the TAZ centroid, multiplied by a circuity factor of 1.2 to account for a typical level of indirectness on local streets [8, 9].

GTFS trip schedule data is used to calculate the travel time between each pair of stops, subject to the restrictions that the trip must both depart and arrive between 7 and 9 AM, and that it honors the maximum transfers, transfer waiting times, and transfer distances described above. For each TAZ linked to the origin stop and each TAZ linked to the destination stop, the corresponding access times are added to the stop-to-stop travel time to represent the total travel time between each pair of linked

TAZs. The shortest discovered travel time between each pair of TAZs are retained as the final transit travel time for that pair.

For some origin zones the estimated walk time to a nearby zone is less that the time required to access a transit stop or station. Because essentially every transit trip includes a walking component, it is reasonable to include walking in a calculation of transit accessibility. Therefore, the final travel time table is updated to reflect the walking time between zones where that time is lower that the lowest time provided by transit, or where travel by transit is not possible. This occurs very infrequently in the central urban part of the study area; in outlying areas it allows for pedestrian travel between zones where no transit service exists.

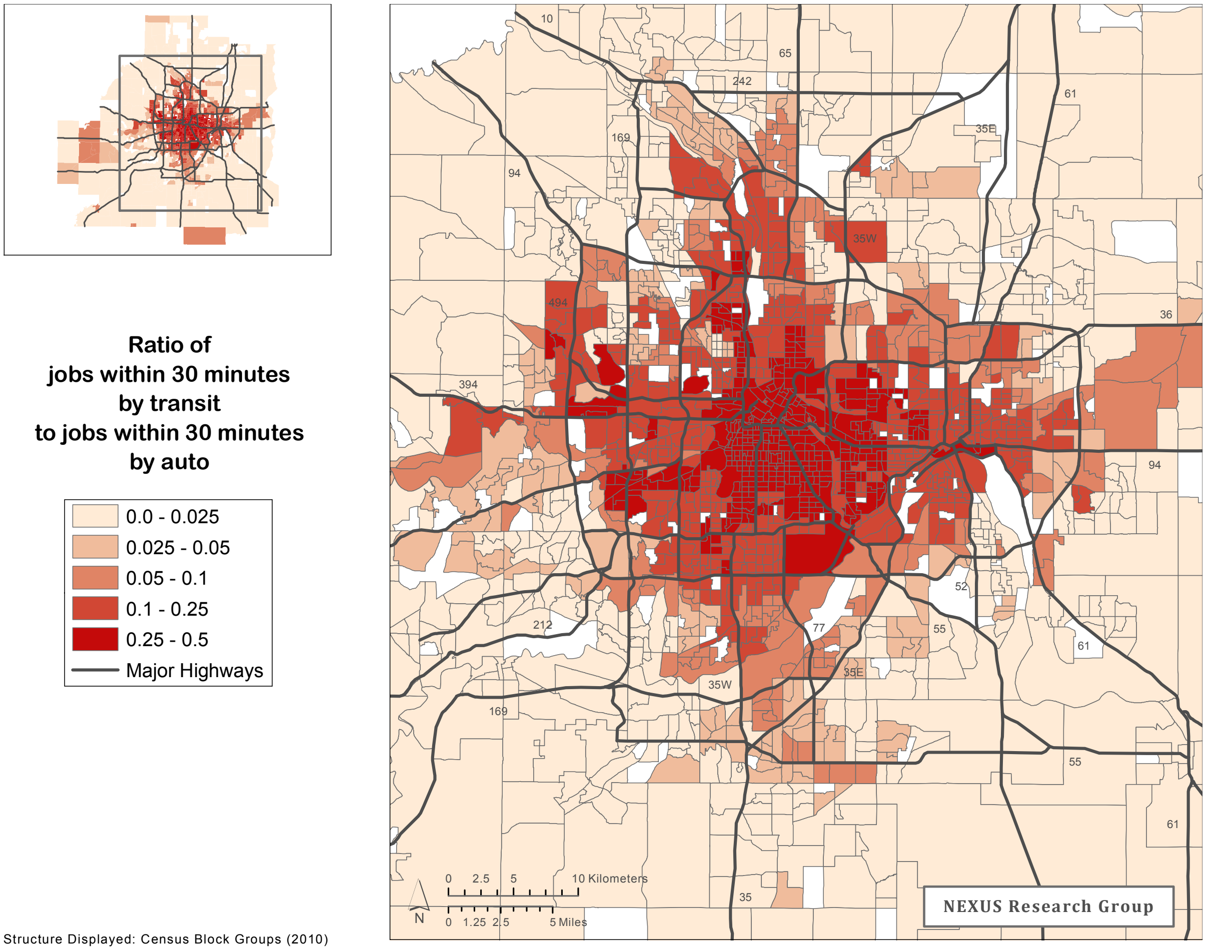
5 Calculating Accessibility

Once the appropriate land use and travel time datasets are assembled, calculating accessibility is comparatively straightforward. A cumulative opportunities measure of accessibility is implemented at the TAZ level using the method described above. First, for each origin TAZ, all destination TAZs reachable within a specified time threshold by a specified mode are identified. Next, the number of opportunities is summed across those reachable destinations.

This method is used to produce, for each TAZ, a set of accessibility measures representing the number of jobs reachable by auto within 10, 20, 30, 40, 50, and 60 minutes, and by transit within the same thresholds. To facilitate comparison with mode share ratios at the block group level, accessibility values are assigned to block groups from the TAZ which contains each block group's centroid, and then the ratio of transit accessibility to auto accessibility is calculated for each. Table 1 describes the distribution of this ratio for 1,926 block groups in the study area, and Figure 2 maps this information.

Figure 3 illustrates the direct relationship between relative accessibility and relative mode share in the study area. In this visualization, areas where a one-to-one relationship overpredicts non-auto mode share are blue, while areas in red indicate underprediction of non-auto mode share. Neutral colors show where the relationship is approximately one-to-one. There is a high degree of local variation, but some trends are apparent. Most apparently, areas outside the outer highway beltway (formed by I-494 and I-694) generally have mode share ratios many times higher than their accessibility ratios. Within the highway ring, mode share ratios tend to be roughly equal to or lower than accessibility ratios, with local exceptions in downtown Minneapolis, near the University of Minnesota, and to a lesser degree in downtown Saint Paul.

The strength of this inside/outside trend relative to the beltway deserves comment. A likely contributing factor is that the transit travel time calculation method only allows for transit stop/station access by walking. Metro Transit and other regional transit operators provide park-and-ride facilities at many locations near and outside the beltway; the transit service at these locations generally focuses on peak-period express trips to and from major employment centers. By allowing transit stop/station access by auto these facilities provide greater accessibility to jobs than is captured by the current accessibility calculation method.



Zone Structure Displayed: Census Block Groups (2010)
Primary Data Source: US Census Bureau

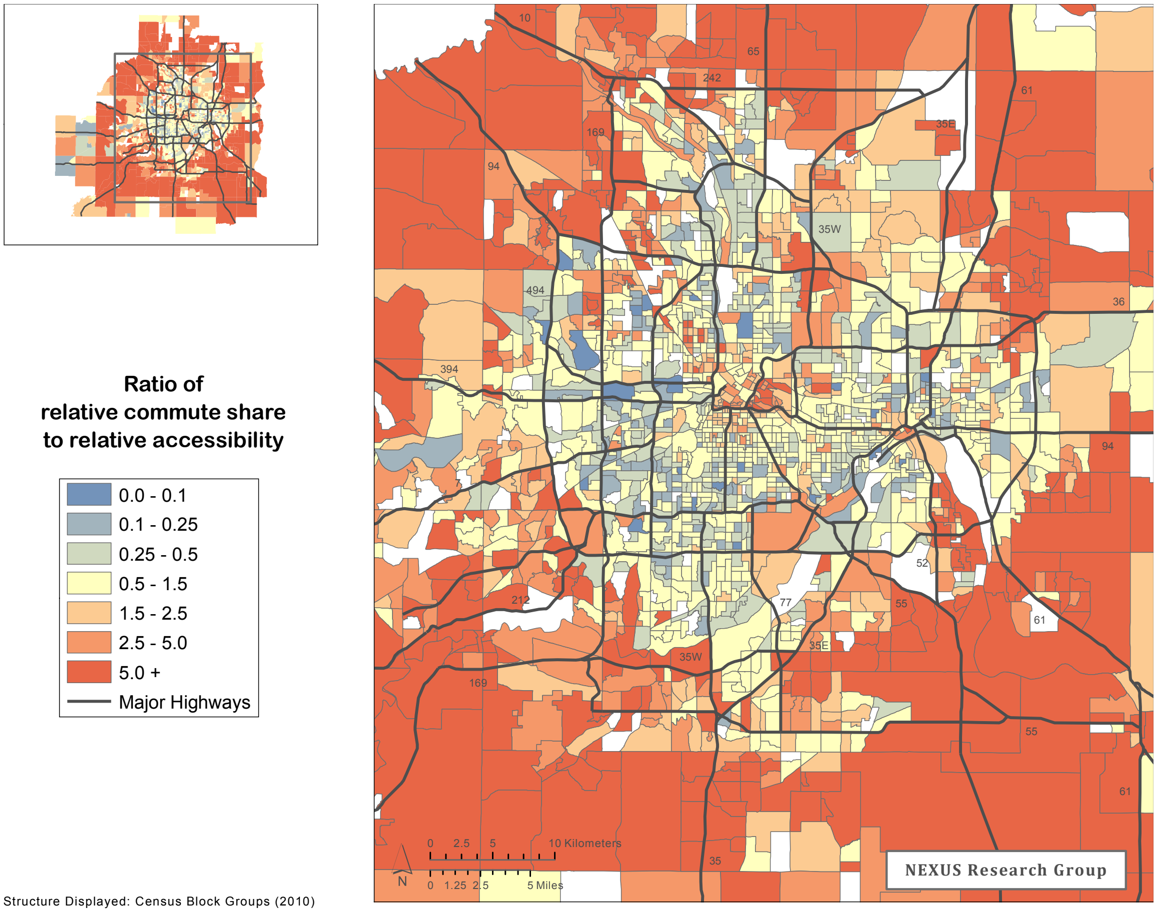
Figure 2: Ratio of jobs within 30 minutes by transit to jobs within 30 minutes by auto for each block group in the study area

6 Results and Discussion

The ratio of non-auto commuters to auto commuters in each block group is modeled as a exponential function of the ratio of jobs accessible by transit to jobs accessible by auto, testing the hypothesis that relative accessibility explains relative mode shares. Parameters are estimated separately for 10, 20, 30, 40, and 50 minute travel time thresholds:

$$\left(\frac{N_{\text{non-auto}}}{N_{\text{auto}}}\right) = \beta_0 e^{\beta_1\left(\frac{A_{\text{transit}}}{A_{\text{auto}}}\right)} \tag{3}$$

$N =$ number of commuters
$A =$ accessibility



Zone Structure Displayed: Census Block Groups (2010)
Primary Data Source: US Census Bureau

Figure 3: Ratio of relative mode share (non-auto commuters / auto commuters) to relative accessibility (jobs within 30 minutes by transit / jobs within 30 minutes by auto) for each block group in the study area

Table 2 summarizes the results of regression analyses against all six travel time thresholds. The most significant finding is that the transit/auto 30-minute jobs accessibility ratio is able to predict over 40% of the variation in the non-auto/auto mode share ratio ($r^2 = 0.412$); this relationship has a high degree of statistical significance ($p < 2.0 \times 10^{-16}$). The model accurately captures the overall trend in the relationship between relative accessibility and relative mode share, but a good deal of variation remains uncaptured (Figure 4).

Because the regression is based on ratios this result requires careful interpretation: the non-auto / auto accessibility ratio will change when either or both of its components changes. Therefore it is not possible to expect simply that non-auto mode share will increase with increasing transit accessibility, because if auto accessibility increases as well the ratio may remain constant or decline. Instead, it is more useful to consider the effects of changing the accessibility offered by one mode while holding the other constant. For example, if an area's transit accessibility increases while auto accessibility remains constant, these results predict that a greater share of local commuters will choose a mode other than

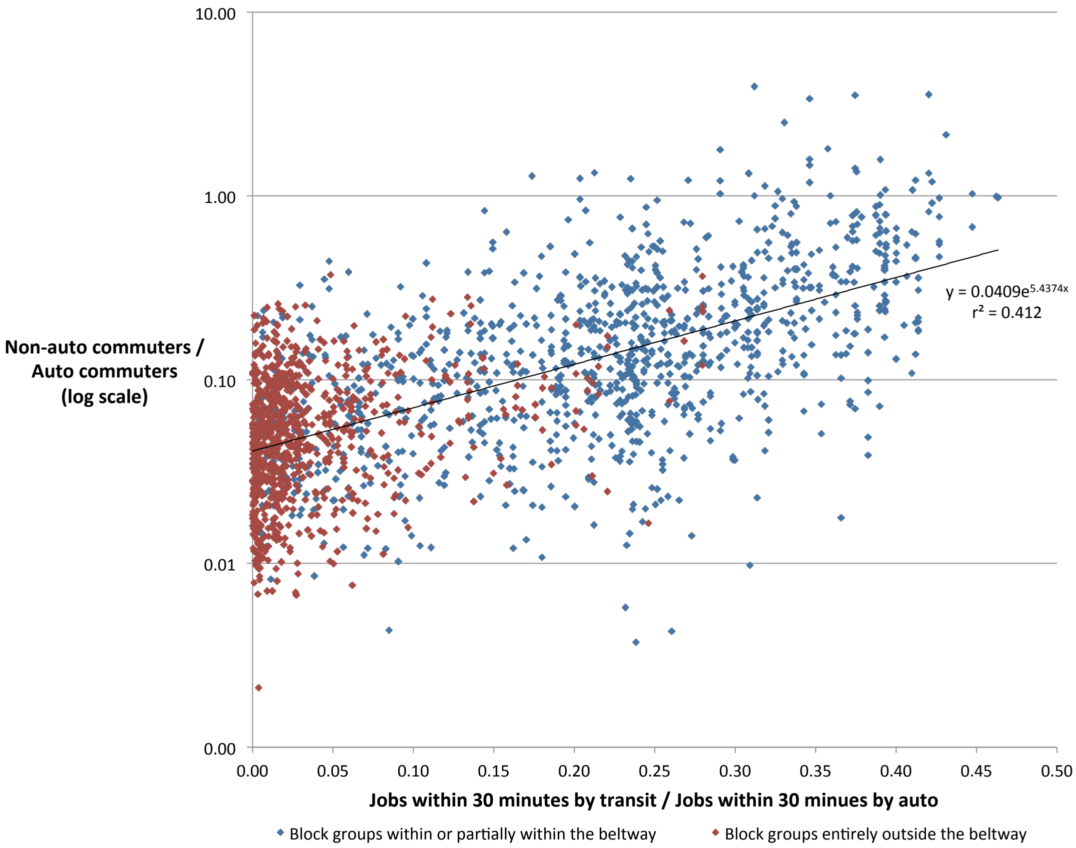


Figure 4: Scatter plot of data points for each block group compared with model results using parameters estimated for 30-minute accessibility ratio. Block groups inside and outside the beltway are color-coded separately, but all are included in the regression

driving; the same result is predicted if transit accessibility remains constant while auto accessibility declines. These results are intuitive, consistent with existing research, and are directly applicable to transportation planning issues at the city and regional level.

It is important also to consider the regression results for other accessibility time thresholds. The 30-minute accessibility ratio has the greatest predictive power, but the 40-minute accessibility ratio has only slightly less ($r^2 = 0.403$), while the 50- and 60-minute ratios involve modest declines. This group is dramatically different from the 20- and especially the 10-minute accessibility ratios, suggesting a high degree of cross-correlation between accessibility ratios of 30 minutes and above, and little to no correlation between these ratios and the 10- and 20-minute ratios. A direct investigation reveals that this is the case; Table 3 describes the correlation between these accessibility ratios.

The fact that the 30-minute accessibility ratio has the greatest power to predict mode share ratios is consistent with data indicating that average commute length in the Twin Cities metropolitan area

Threshold (minutes)	β_1	β_0	$\Pr < \lvert t \rvert$	r^2
10	0.722	0.0728	<0.0001	0.026
20	4.787	0.0479	<0.0001	0.267
30	5.437	0.0409	<0.0001	0.412
40	4.542	0.0377	<0.0001	0.403
50	3.887	0.0344	<0.0001	0.378
60	3.388	0.0327	<0.0001	0.337

Table 2: Regression results for 10, 20, 30, 40, 50, and 60-minute accessibility ratios. The 30-minute accessibility ratio provides the best prediction of relative mode share.

	20	**30**	**40**	**50**	**60**
10	0.540	0.321	0.270	0.225	0.198
20		0.815	0.776	0.717	0.657
30			0.970	0.928	0.872
40				0.978	0.940
50					0.983

Table 3: Correlation between transit/auto accessibility ratios at selected time thresholds. There are strong correlations between the ratios for threshold of 30 minutes and higher.

is 24.3 minutes (based on 2006–2010 ACS data; the national average commute length is estimated to be 25.2). The distribution of commute times differs between auto and transit commuters, with transit commuters generally having longer commutes (Figure 5). Since thresholds of 30 minutes and above accommodate the majority of commute trips by both auto and transit, it is reasonable that they are the best at predicting the aggregate results of commuters' mode choices. Relatively few commute trips in the Twin Cities (15.8%) are under 20 minutes and far fewer are under 10 minutes, which likely contributes to the low predictive power of those time thresholds.

6.1 Limitations

The model and methodology employed here are intentionally simplistic and some directions for improvement are apparent. Perhaps the greatest opportunity for improvement lies in rationalizing the comparison between auto travel times, which are expressed as averages over time, and transit travel times, which are reported at specific departure times. While this analysis takes the approach of selecting the best transit travel time within a time period of study, this is likely not the most accurate representation: commuters desire to depart or arrive at different times, some of which are not accommodated by the transit schedule. Basing transit travel times on any single trip ignores the travel times provided by other trips, and disregarding the requirement that users wait for a scheduled departure certainly results in overestimation of aggregate accessibility. Comparison of accessibility between on-demand and scheduled transportation systems remains an unresolved challenge.

A more careful consideration of geographical aggregation levels may also improve the performance

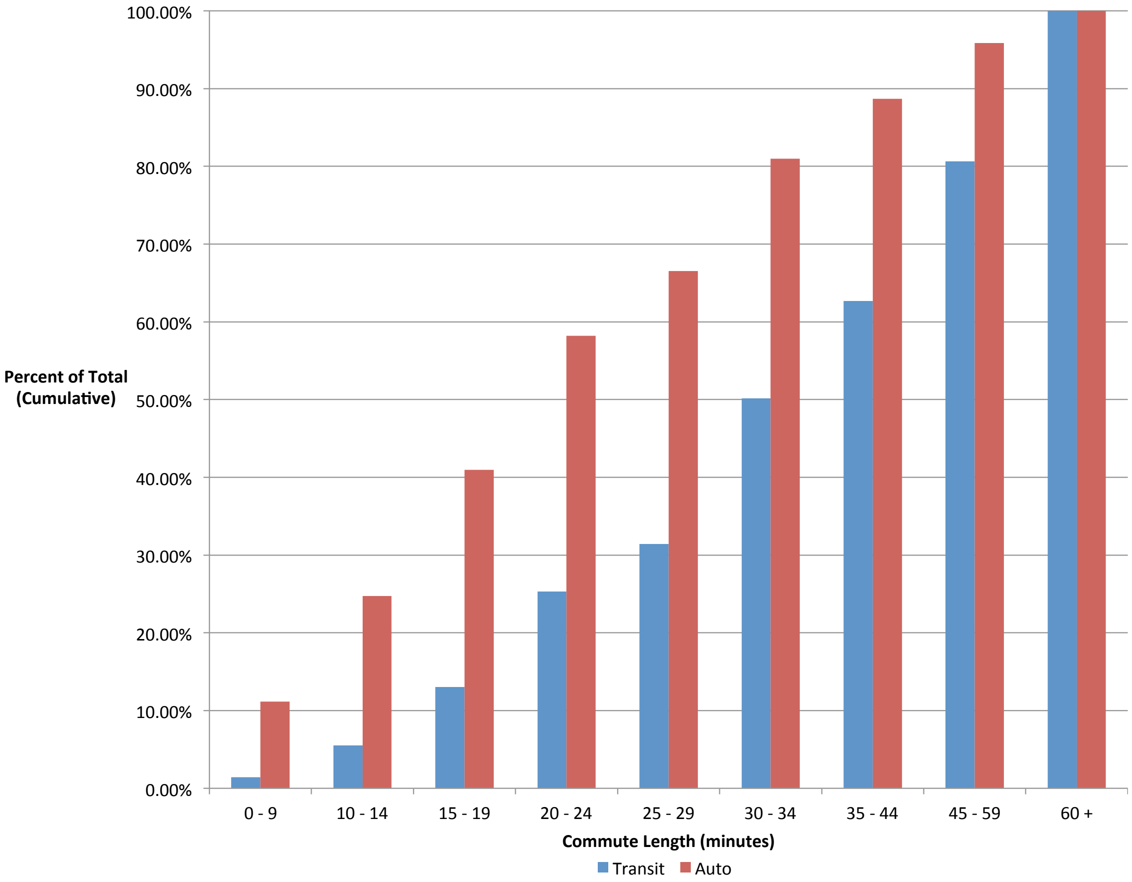


Figure 5: Cumulative distribution of commute times for auto and transit commuters in the Minneapolis-Saint Paul metropolitan area

of the model. Dark and Bram [2] review the extensive literature dealing with the modifiable areal unit problem (MAUP), and Zhang and Kukadia [13] explores its implications for transportation research. While there is no single best solution to the MAUP, it is clear that the choice of geographical units for analysis can influence results. Accessibility research in particular involves a geographical balancing act. In general, at higher levels of aggregation opportunity count estimates become more precise while travel time estimates become less so; at lower levels of aggregation the reverse is the case. In this analysis, geographical units are chosen based solely on data availability and compatibility; a more nuanced selection should provide better results.

Additionally, this analysis considers only one implementation of accessibility measurement. Even within the framework provided by Hansen [4], there exist a wide range of possible implementations and calibrations which may be more suited to predicting relative mode share.

Finally, it seems clear that accessibility ratios are unable to completely describe the variation in relative mode shares (Figure 4). As discussed above, this approach focuses solely on causal factors influencing mode choice while ignoring the descriptive factors that capture the effects of user preferences and perceptions. It also omits demographic causal factors such as income and car ownership. Selective inclusion of additional well-considered variables may improve overall predictive power without adding unnecessary complexity. In particular, the current results suggest that stratification of transit markets by stop/station access mode—walking vs. auto—could significantly improve the performance of the model.

References

[1] Paul Anderson, David M Levinson, and Pavithra Parthasarathi. Using twin cities destinations and their accessibility as a multimodal planning tool. Technical Report MnDOT 2012-05, Minnesota Department of Transportation, 2012.

[2] Shawna J. Dark and Danielle Bram. The modifiable areal unit problem (maup) in physical geography. *Progress in Physical Geography*, 31(5):471–479, 2007. doi: 10.1177/0309133307083294. URL `http://ppg.sagepub.com/content/31/5/471.abstract`.

[3] Ahmed M El-Geneidy and David M Levinson. Access to destinations: Development of accessibility measures. Technical Report MN/RC-2006-16, Minnesota Department of Transportation, 2006.

[4] W.G. Hansen. How accessibility shapes land use. *Journal of the American Institute of Planners*, 25 (2):73–76, 1959.

[5] Mark W. Horner. Spatial dimensions of urban commuting: A review of major issues and their implications for future geographic research. *The Professional Geographer*, 56(2):160–173, 2004. doi: 10.1111/j.0033-0124.2004.05602002.x. URL `http://www.tandfonline.com/doi/abs/10.1111/j.0033-0124.2004.05602002.x`.

[6] J. Levine, J. Grengs, Q. Shen, and Q. Shen. Does accessibility require density or speed? *Journal of the American Planning Association*, 78(2):157–172, 2012.

[7] Metropolitan Council. 2030 transportation policy plan, 2010.

[8] G.F. Newell. *Traffic flow on transportation networks*. MIT Press, 1980.

[9] S. O'Sullivan and J. Morrall. Walking distances to and from light-rail transit stations. *Transportation research record: journal of the transportation research board*, 1538:19–26, 1996.

[10] P.O. Plaut. Non-commuters: the people who walk to work or work at home. *Transportation*, 31 (2):229–255, 2004.

[11] Brian D. Taylor, Douglas Miller, Hiroyuki Iseki, and Camille Fink. Nature and/or nurture? analyzing the determinants of transit ridership across us urbanized areas. *Transportation Research*

Part A: Policy and Practice, 43(1):60 – 77, 2009. ISSN 0965-8564. doi: 10.1016/j.tra.2008.06.007. URL `http://www.sciencedirect.com/science/article/pii/S0965856408001274`.

[12] Kan Hua Young and Lin Ying Young. Estimation of regressions involving logarithmic transformation of zero values in the dependent variable. *The American Statistician*, 29(3):pp. 118–120, 1975. ISSN 00031305. URL `http://www.jstor.org/stable/2683436`.

[13] Ming Zhang and Nishant Kukadia. Metrics of urban form and the modifiable areal unit problem. *Transportation Research Record: Journal of the Transportation Research Board*, 1902:71–79, 01 2005. doi: 10.3141/1902-09. URL `http://dx.doi.org/10.3141/1902-09`.

[14] Shanjiang Zhu. *The Roads Taken: Theory and Evidence on Route Choice in the wake of the I-35W Mississippi River Bridge Collapse and Reconstruction*. PhD thesis, University of Minnesota, 2010.