Appendix C. Technical Appendix

To perform the analysis, we needed to obtain reliable measures of household income, rental and ownership housing costs by income, household transportation costs by income, jobs and employment locations, and other socioeconomic measures of households by income and by place. In this section we explain how we derived or gathered each measure at the census tract level.

Household Incomes

To begin, we first had to identify specific incomes in multiple income bins at the census tract level that would roughly approximate to the standard HUD measures of income, e.g. 30%, 50%, 80%, and 100% of a region's Area Median Income (AMI). The census provides a count of each household by income at the tract level in 16 income bins and uses these bins for several other indicators, such as the percentage of income on housing by income, tenure by income, age of head of householder by income, etc. Therefore, at the tract level, we present the incomes by a nominal vale in six bins rather than as a percentage of AMI since AMI is not available for the 29,628 tracts. A translation table between dollar values and percent AMI for each region is in Table AX in the Appendix.

The census category of income at the tract level was not specific enough for our calculations. The income bin grouping at the tract level leaves two large bins at the bottom and the top that could have wide variation. The bottom bin is "less than $10,000" and the upper income bin is "$200,000 or more". The middle bins are in $5,000 to $10,000 increments. At the same time, there are more groupings than we needed for this analysis. Therefore, we both consolidated the bins from 16 bins to 6 bins, and then within each bin, calculated an average income for the households within each cohort (e.g. $17,982 rather than $15,000 to $19,999) in order to have a specific point rather than a range. The table below shows the income distribution results available at the tract level from the Census for a tract in California. We use both the Family and Non-Family Income fields (P76 and P79) to obtain the count of all households by income. Households in Group Quarters are excluded.

[1] In 2000, there were 105,480,101 households in the U.S. according to the 2000 Census, SF1.
[2] Some households were excluded from the sample if they were in census tracts with fewer than 100 households, or if they lived in group quarters, such as dormitories.
[3] We compared tracts in 1990 and 2000 that had the same boundaries each decennial census for eight regions; Portland, Los Angeles, San Francisco, Dallas, Chicago, Denver, Pittsburgh, and Atlanta.
[4] Other studies have noted this…

Table 4. Income Distribution by Census Tract

Tract 402.02, Riverside County, CA	P. 76 Family Income	P. 79 Non-Family Income	Total
Total:	543	234	777
Less than $10,000	21	13	34
$10,000 to $14,999	40	37	77
$15,000 to $19,999	30	41	71
$20,000 to $24,999	21	12	33
$25,000 to $29,999	33	15	48
$30,000 to $34,999	49	32	81
$35,000 to $39,999	38	13	51
$40,000 to $44,999	19	0	19
$45,000 to $49,999	9	0	9
$50,000 to $59,999	59	11	70
$60,000 to $74,999	84	30	114
$75,000 to $99,999	97	21	118
$100,000 to $124,999	28	9	37
$125,000 to $149,999	6	0	6
$150,000 to $199,999	9	0	9
$200,000 or more	0	0	0

To create six income bins for analysis, we collapsed the above income bins into the six bins in Table 5. Within each bin, we used the PUMS 5% census data to calculate the average income for the households in each bin in a 5% PUMA. The average of each bin from the PUMA was applied to the corresponding bin for each census tract within a PUMA. Table 5 shows the average and range of incomes by income bin calculated using the PUMAs in the 28 metro areas.

Table 5. Average Income by 5% PUMS in Each Income Bracket for 28 Metros

Census Income Bin	Weighted Average	Min	Max	N
<$ 20,000	$10,252	$7,204	$13,560	5,742,029
$20,000 to <35,000	$26,997	$25,340	$29,848	4,181,936
$35,000 to <50,000	$41,531	$40,024	$43,655	3,048,739
$50,000 to <$75,000	$60,146	$56,686	$62,807	2,834,351
$75,000 to <$99,000	$84,992	$78,666	$89,317	1,144,763
$100,000 to <$250,000	$131,036	$110,137	$176,710	971,172
Total 5% PUMAs				963

Housing Costs as a Percentage of Income (H)

For each income bin we used the same methodology. We first gathered the expenditures on housing by income at the census tract level for both renters and owners. However, again, this field as reported at the tract level has large bins at the top and bottom, less than 20% at the bottom and greater than 35% at the top. See Table 5 below. This table shows the results for the same census tract in California used above.

Table 4. Household Income in 1999 by Rent and Selected Monthly Owner Mortgage Costs as a Percentage of Household Income		
Tract 402.02, Riverside County, CA	**H.73 Renter Costs**	**H.97 Mortgage costs**
Total:	225	375
Less than $10,000:	17	0
Less than 20 percent	0	0
20 to 24 percent	0	0
25 to 29 percent	0	0
30 to 34 percent	7	0
35 percent or more	10	0
Not computed	0	0
$10,000 to $19,999:	49	43
Less than 20 percent	0	0
20 to 24 percent	0	0
25 to 29 percent	0	0
30 to 34 percent	6	0
35 percent or more	35	43
Not computed	8	0
$20,000 to $34,999:	48	53
Less than 20 percent	14	29
20 to 24 percent	0	0
25 to 29 percent	14	0
30 to 34 percent	0	0
35 percent or more	20	24
Not computed	0	0
$35,000 to $49,999:	27	40
Less than 20 percent	21	9
20 to 24 percent	0	12
25 to 29 percent	0	0
30 to 34 percent	0	8
35 percent or more	6	11
Not computed	0	0
$50,000 to $74,999:	23	143
Less than 20 percent	11	42
20 to 24 percent	12	56
25 to 29 percent	0	38
30 to 34 percent	0	7
35 percent or more	0	0
Not computed	0	0
$75,000 to $99,999:	49	55
Less than 20 percent	49	33
20 to 24 percent	0	14
25 to 29 percent	0	0
30 to 34 percent	0	0
35 percent or more	0	8
Not computed	0	0
$100,000 or more:	12	32
Less than 20 percent	0	32
20 to 24 percent	0	0
25 to 29 percent	0	0
30 to 34 percent	0	0
35 percent or more	0	0
Not computed	12	0
$150,000 or more:	Not computed	9
Less than 20 percent	Not computed	9
20 to 24 percent	Not computed	0
25 to 29 percent	Not computed	0
30 to 34 percent	Not computed	0
35 percent or more	Not computed	0
Not computed	Not computed	0

Summarizing the 28 metros by renters, owners, and all households, we found 31% of renters are paying more than 35% of their income on housing compared to 18% of owners. Overall, 23% of households are paying more than 35%.

Percent of Households Paying 35% or more of Income by Income in 28 Metros (Census 2000, SF3, H.97, H.73)

Income	Rent	Own	All
Less than $10,000	65%	70%	66%
$10,000 to $19,999	70%	54%	65%
$20,000 to $34,999	31%	39%	34%
$35,000 to $49,999	8%	25%	17%
$50,000 to $74,999	3%	12%	9%
$75,000 to $99,999	1%	5%	4%
$100,000 or more	0%	2%	2%
TOTAL	***31%***	***18%***	***23%***

Again we used the PUMS 5% sample to cross tab the six income bins by the average percentage of income households in each bin were spending on housing. These results were then applied to each specific "percent of income" bin for each income bin for each tract within a PUMA. The summary results at the regional level are displayed below.

On average, households earning less than $35,000 were spending between 31% and 58% of their income on housing.

Note: For the highest income bin, we limited our analysis to households earning <$250,000. This eliminated 5,386,480 household records and reduced total households in our analysis to 41,761,305. Extremely high incomes above $250,000 would have greatly skewed the analysis for this income bin.

Table 6. Percent of Income on Housing by 5% PUMA for 28 Metros

MSA	<$ 20,000	$20,000 to <35,000	$35,000 to <50,000	$50,000 to <$75,000	$75,000 to <$99,000	$100,000 to <$250,000	N
Anchorage	65%	35%	26%	22%	18%	14%	55
Atlanta	59%	33%	25%	20%	16%	14%	660
Baltimore	58%	33%	26%	21%	17%	14%	1070
Boston--Worcester-Lawrnece	56%	33%	25%	21%	18%	14%	1219
Chicago--Gary--Kenosha, IL	59%	31%	24%	20%	18%	14%	2055
Cincinnati	51%	26%	21%	18%	15%	12%	476
Cleveland--Akron	52%	27%	21%	18%	15%	12%	872
Dallas-Fort Worth	57%	29%	22%	18%	16%	13%	1050
Denver-Boulder	59%	33%	25%	21%	18%	14%	614
Detroit	55%	27%	21%	18%	15%	12%	1567
Honolulu	61%	35%	27%	22%	20%	16%	210
Houston-Galveston	56%	27%	21%	17%	15%	12%	878
Kansas City, MO	51%	26%	20%	17%	14%	12%	493
Los Angeles--Riverside	63%	36%	27%	23%	20%	16%	3356
Miami--Fort Lauderdale	63%	35%	27%	21%	18%	14%	623
Milwaukee--Racine	54%	28%	21%	18%	16%	13%	453
Minneapolis--St. Paul	54%	30%	23%	19%	16%	13%	741
New York--North	64%	36%	27%	22%	19%	15%	5072
Philadelphia--Wilmington	57%	31%	24%	19%	17%	13%	1568
Phoenix--Mesa	58%	31%	23%	19%	16%	13%	692
Pittsburgh, PA	47%	24%	18%	16%	14%	11%	702
Portland--Salem	59%	32%	25%	20%	17%	14%	484
San Diego, CA	63%	35%	27%	23%	20%	16%	602
San Franciso--Oakland, CA	65%	39%	30%	25%	21%	17%	1455
Seattle--Tacoma, WA	60%	34%	26%	22%	19%	15%	769
St. Louis, MO	51%	25%	19%	16%	14%	12%	524
Tampa--St. Petersburg--Cle	53%	28%	21%	17%	15%	12%	546
Washington	61%	35%	27%	22%	18%	14%	1025
Average	58%	31%	24%	20%	17%	14%	1065

Transportation Cost as a Percentage of Income (T)

The premise for this study is the examination of the fraction of income a household spends on transportation and housing. Housing cost is relatively easy to assess, since the US Census and many other sources gather it. However, the amount of money a household has to spend on transportation, especially for a specific location, is not as readily available. To do this study, we have applied a model developed to calculate the average household transportation costs using a regression analysis based on the analysis and theory of the Location Efficient Mortgage® (LEM), which was peer reviewed and developed by a group of researchers including the Center for Neighborhood Technology[5]. For this study we use a model that was developed under the Urban Markets Initiative of the Brookings Institution's Metropolitan Policy Program by the Center for Neighborhood Technology with the Center for Transit-Oriented Development[6]. This model calibrated with data from the Minneapolis/St Paul metropolitan area provides output that give the

5 John Holtzclaw, Robert Clear, Hank Dittmar, David Goldstein, and Peter Haas, "Location Efficiency: Neighborhood and Socio-Economic Characteristics Determine Auto Ownership and Use—Studies in Chicago, Los Angeles, and San Francisco," Transportation Planning and Technology 25(1) (2002): 1-27, available online at www.tandf.co.uk/journals/online/0308-1060.html.
6 See http://www.brook.edu/metro/umi.htm and http://www.cnt.org/publications/Affordability-Index-White-Paper-Draft-0805.pdf for more detailed discussion.

average household transportation cost within a census tract given the household's income and size.

The household transportation costs consist of a combination of auto ownership auto use and transit use and therefore the model estimates each cost separately. This allows each to be estimated separately based on the neighborhood and the household size and income. These three components are the dependent variables in the model and are affected by the combination of seven independent variables about the built environment and two independent household variables, household size and income. Together, these nine variables represent the independent place-based neighborhood characteristics and the socioeconomic characteristics that predict household transportation costs at the census tract level, a geography that approximates a neighborhood. It is important to model these costs at a neighborhood level, given that many of the independent variables can vary block by block.

To develop the regression formula, we tested each of the independent variables separately against the dependent variables, and then in combination to determine their relationship. The analysis showed that the independent variables co-vary and are interdependent of one another. Thus, no one variable, such as transit accessibility or household income, by itself completely determines transportation costs. Rather, it is the combination of these variables that determines how many autos a household owns, how many miles members drive each vehicle, and how much transit they use. Because transportation is an integral part of our daily routines, it makes sense that it is the combination of how a household's workers commute to work, the distance to services such as a grocery store, how children get to school or other activities, and how much a family earns that determines total household transportation costs.

It's important to note, while many of the findings by Housing/Transportation trade-off are directly related to the variables in the model that predict the transportation costs, we focused on the other characteristics in the neighborhood or region that are related to transportation and housing costs but are not directly related to the variables in the model, such as auto ownership, which is predicted by the model but also reported by the census. In the findings, we are able to compare modeled auto ownership to reported auto ownership and use reported auto ownership to cite a finding rather than just the modeled cost.

In this analysis we used the model described above to assess the household transportation cost for households within each of the six income cohorts described above. As model inputs we used the average income in that income bin, and used the average size of households in the census tract. We then took the weighted average of these costs to determine the overall average household transportation cost. We then divided that cost by the cohort's income to obtain our transportation cost burden (T). The following table shows this burden for each metropolitan regions by income bin, and for the average Household in each region.

	> $20,00	$20,000 - $30,000	$30,000 - $50,000	$50,000 - $75,000	$75,000 - $100,000	>= $100,000	All Households
Anchorage	53%	34%	25%	19%	14%	10%	18%
Atlanta	62%	37%	27%	20%	15%	10%	21%
Baltimore	52%	33%	24%	19%	14%	10%	19%
Boston	56%	34%	26%	19%	15%	10%	19%
Chicago	51%	31%	23%	18%	14%	9%	18%
Cincinnati	59%	37%	27%	20%	15%	10%	23%
Cleveland	55%	35%	26%	19%	15%	10%	22%
Dallas	59%	35%	26%	19%	15%	10%	21%
Denver	52%	32%	24%	18%	14%	9%	19%
Detroit	58%	36%	26%	19%	15%	10%	21%
Honolulu	47%	28%	21%	16%	13%	8%	16%
Houston	60%	36%	26%	20%	15%	10%	22%
Kansas Cit	60%	38%	28%	21%	15%	10%	23%
Los Angeles	50%	31%	23%	17%	13%	9%	19%
Miami	52%	32%	23%	18%	13%	9%	21%
Milwaukee	54%	34%	26%	20%	15%	10%	22%
Mn-St Paul	54%	34%	26%	19%	15%	10%	19%
New York	45%	27%	20%	16%	13%	8%	16%
Philadelphia	52%	33%	24%	19%	14%	10%	20%
Phoenix	55%	34%	25%	19%	14%	9%	21%
Pittsburgh	59%	37%	27%	20%	15%	10%	25%
Portland	58%	36%	26%	20%	15%	10%	22%
San Diego	51%	32%	24%	18%	13%	9%	19%
San Francisco Bay Area	51%	31%	23%	17%	13%	8%	15%
Seattle	55%	34%	25%	19%	14%	9%	19%
St. Louis	58%	36%	27%	20%	15%	10%	23%
Tampa	60%	37%	27%	20%	15%	10%	25%
Washington	55%	33%	24%	18%	14%	9%	17%
Total	53%	33%	24%	18%	14%	9%	19%

Job Locations

To locate and define the size of the employment centers for a region, we use the Census Transportation Planning Package 2000 that provides the total number of employees per census tract.
This analysis used a simple clustering analysis to determine where the centers of employment are within the region and the size of each employment center based on the number of employees within its boundaries.

The following describes how the automated algorithm locates and defines each employment center within a region using GIS software, Census TIGER/Line® Files, and CTPP 2000 job data.

Automated Algorithm to Identify Employment Centers with GIS

1. Calculate the land area for each polygon (excluding water) within the region
2. Calculate the job density (jobs per acre of land) for each polygon
3. Use 7 jobs/acre as the minimum job density threshold
4. Sort the polygons by the total number of jobs within each polygon
5. Assign the tract with the most jobs a number starting at one. This will be the first employment center cluster.
6. A single census tract is most likely only a portion of an entire employment center cluster. Therefore, the additional neighboring tracts that are also part of the cluster must be identified and assigned to this cluster. To do so, the neighboring tracts are scanned and those where the

density is higher than or equal to the chosen minimum density threshold are assigned to the cluster. The area of the first employment center cluster is now defined.

7. Continue adding polygons in step 6 until there are no new adjacent polygons to add to the cluster.

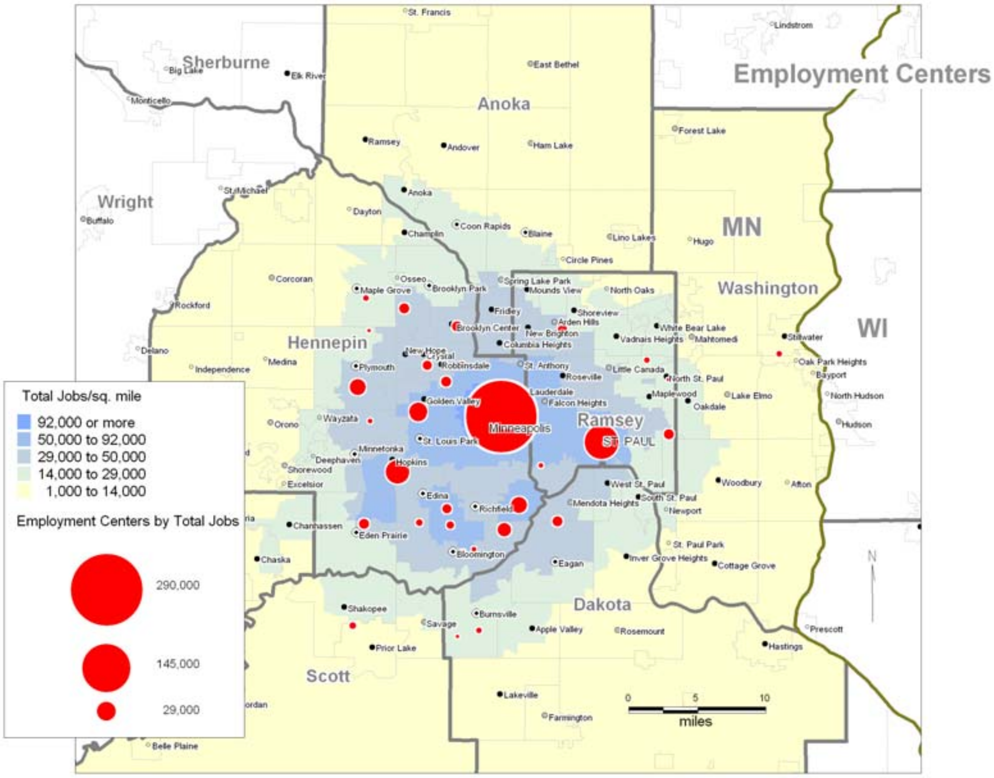
8. To identify the remaining employment center clusters, remove the polygons that have been assigned to an employment center cluster from the list and repeat steps 4 through 7 until there are no more polygons that have a job density above or equal to the minimum density threshold.

9. We chose a weighted center to find the employment geographical center so we can definne a center point from which to measure distance..

10. The final step is to choose only those employment centers that have at least 5000 jobs associated with them.

The total number of jobs is a measure of employment used in the transportation model, and in our classification of job access. We use the gravity model to measure the employment density in the area of each tract. That is, for a given tracts we the sum of all the number of jobs in every other tract in the region divided by the square of the distance, included in that sum is also the number of jobs in the census tract itself. Note that although this measure has units of "jobs/square mile" and therefore an job density measure, it should only be interpreted as a relative measure of job access.

The following map shows the employment centers with a background of the job density measure.



Source: The Census Transportation Planning Package (CTPP) 2000

Figure 7. Employment Center clusters in Minneapolis/St. Paul region.

Worker Commuting Characteristics

In order to define commuter characteristics and congestion, we looked at four different but related statistics. These are the mode of commute, the time of commute, the distance of commute and the average speed of commute.

The first of these is very straight forward. The mode of the journey to work is part of the long form in Census 2000. The following table shows how this breaks out for the census tract in California:

Census Tract 403.02, Riverside County,	Workers
Total:	2,940
Car, truck, or van:	2,792
Drove alone	2,412
Carpooled	380
Public transportation:	37
Bus or trolley bus	9
Streetcar or trolley car (publico in Puerto Rico)	0
Subway or elevated	0
Railroad	28
Ferryboat	0
Taxicab	0
Motorcycle	8
Bicycle	15
Walked	9
Other means	9
Worked at home	70

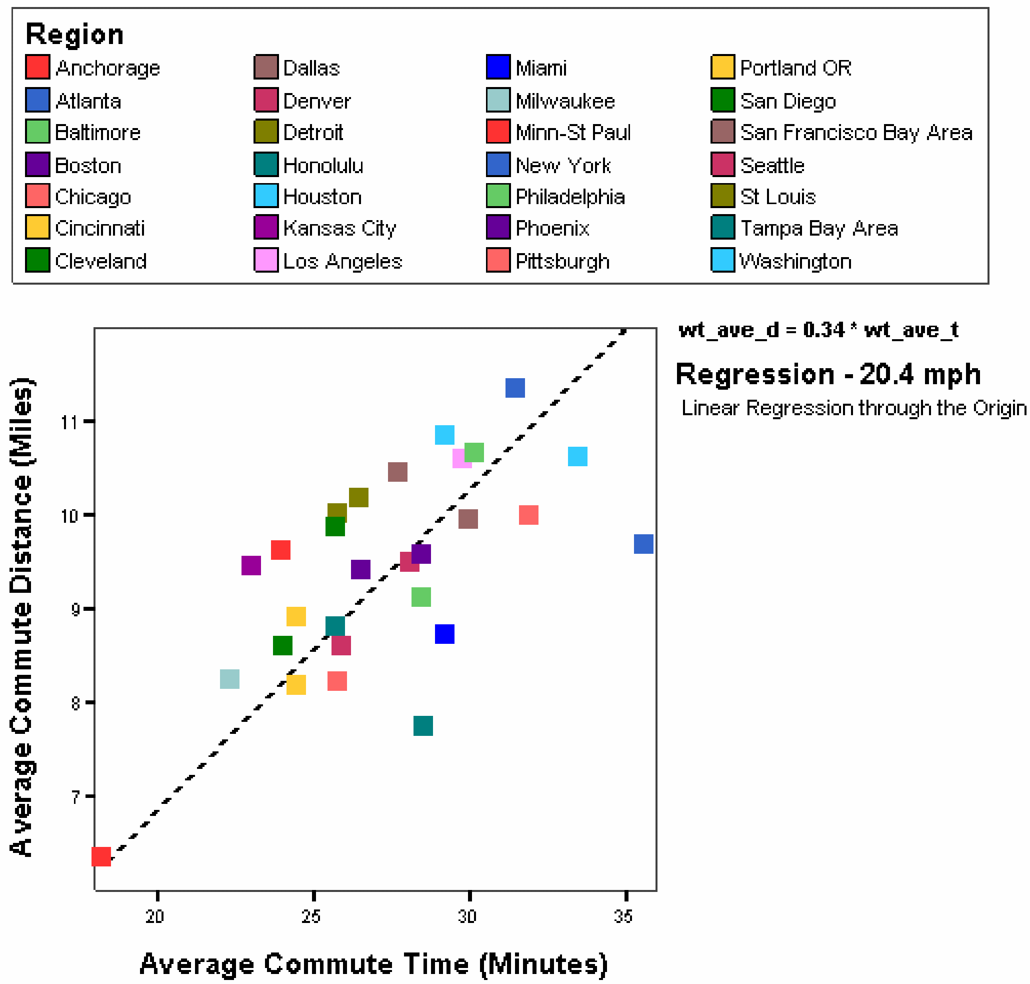
In order to measure the time, distance and average speed to get to work we have used the CTPP data once again. Here we used the part 3 portion of the CTPP, This gives for every tract the tracts that those workers live in, and the mode and time they use to get there. We exported these tract pairs to a GIS and calculated the distance from the center of each tract to get an "as the crow flies" distance of each commute. For the workers that live and work in the same tract, we use the use the average radius of the tract ($r=\sqrt{area/\pi}$). We then group the modes into by auto, public transportation and other. We can then calculate the weighted average of the time and distance of each commute to obtain the time, distance and speed by mode and overall.

For example the census tract that we have been using as an example in Riverside County California, there are 170 census tracts where those workers go to work (including the tract it self), the distances vary from 0.8 miles (the tract itself) to 792 mile (4 workers who work in Colorado). The mean time of commute go from 2 minutes to 200 minutes. Because of the nature of self reported surveys like the census, we had to eliminate tracts that did not make sence, for instance the four workers who work in Colorado say they commute by auto and it takes them 20 minutes, this would mean that they would be going 2374 mph (that is over 3 time the speed of sound!) what this probably represents is people who work in a different city and commute there and stay the week, and then use an auto in Colorado. We eliminated such cases from our average by only accepting commutes where the speed is less than or equal to 65 mph.

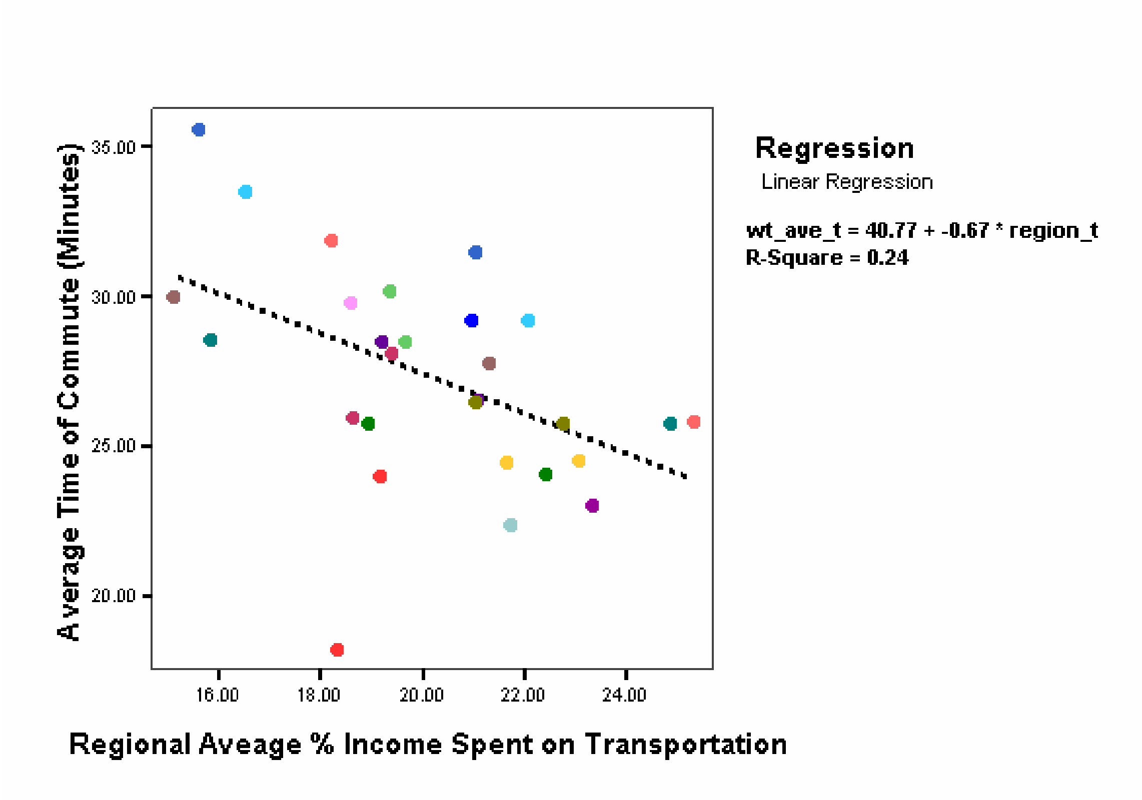
The following table shows the average of these by metro region:

region	Overall Average distance (miles)	Average time (minutes)	Average speed (miles/hour)	Auto Commuters Average speed (miles/hour)	Transit Commuters Average speed (miles/hour)
Anchorage	6.37	18.22	20.16	20.38	8.8
Atlanta	11.36	31.49	21.6	21.97	12.05
Baltimore	10.67	30.16	21.04	21.72	10.74
Boston	9.6	28.46	19.45	20.49	9.95
Chicago	10.01	31.88	18.72	19.53	12.96
Cincinnati	8.94	24.48	21.04	21.36	10.9
Cleveland	8.61	24.06	20.81	21.18	10.71
Dallas	10.47	27.73	22.15	22.34	12.27
Denver	8.62	25.91	19.75	20.14	11.7
Detroit	10.19	26.45	22.49	22.7	10.65
Honolulu	7.77	28.52	16.64	17.34	9.66
Houston	10.87	29.2	21.98	22.28	13.77
Kansas City	9.46	23.02	23.65	23.83	9.98
Los Angeles	10.62	29.8	20.87	21.39	11.13
Miami	8.74	29.21	18.1	18.42	10.44
Milwaukee	8.26	22.36	21.44	21.96	9.07
Minn-St Paul	9.65	23.99	23.18	23.77	11.28
New York	9.69	35.58	17.06	19.77	9.93
Philadelphia	9.14	28.44	19.11	19.99	10.7
Phoenix	9.43	26.53	21.44	21.68	10.64
Pittsburgh	8.24	25.78	18.95	19.65	9.33
Portland OR	8.2	24.44	19.7	20.38	9.46
San Diego	9.89	25.71	22.57	22.98	11.84
San Francisco Bay Area	9.98	29.98	19.51	20.46	11.43
Seattle	9.52	28.09	19.98	20.67	11.43
St Louis	10.03	25.76	22.54	22.82	11.18
Tampa Bay Area	8.82	25.72	20.33	20.46	10.93
Washington	10.64	33.47	18.81	19.93	10.59

The following graphs show these data. Note that the overall fit gives 20.4 mph for the average speed. This speed recall is not the average speed of the vehicle transporting the worker since it is a direct line from the center of the two census tracts but this makes for a good surrogate for congestion as the map on page xx shows.



One more plot like this is very interesting. In the following plot we use the same colors as above, but plot these points as average commute time vs average cost of transportation from our model. The regression shows that as the average time of commute goes down the average total fraction of income spent on transportation is reduced – this would imply that households within regions are optimizing their commute time but not their overall transportation burden.



We believe there is much more that can be done with this measure, but for this study we limit it here.