Using GIS to Evaluate a New Source of Transportation Census Data: The American Community Survey

Wende A. O'Neill, Ph. D. & Daniel Baldwin Hess

Introduction

In a rapidly changing society, many people have come to view the decennial (ten year) census as slow and antiquated. Data taken from a census captured only once in a ten-year period is not released until a few years after the decennial year and, for many applications, the data is stale before it is time for a recount. To improve the situation, the U.S. Bureau of the Census is implementing plans as the 21st century approaches for the American Community Survey (ACS). Using refreshed multi-year accumulations, the ACS will provide timely data by sampling households nationwide every month of the year to collect housing, social, and economic data.

Transportation planners typically spend a great amount of effort adjusting decennial Census data to fit their needs. The ACS, however, provides data users with timely, comparative housing, social, and economic data throughout the decade about communities and population groups. Since continuous measurement is an untested process in the United States, transportation planners have expressed concern about the different data structure under which information will be released. This study explores the problems of geographical and data-structuring differences by comparing ACS estimates to locally-derived transportation data that planners are accustomed to using.

Importance of Census Data to Transportation Planning

State and metropolitan planning organizations rely on data from the decennial census for a broad array of applications. Data from the long-form questionnaire, which includes questions covering place of work; mode of transportation to work; carpooling; travel time and time of departure to work;vehicles available; and mobility limitations, are used for planning and modeling travel behavior. These data have been summarized for some state and local areas in the Census Transportation Planning Package (CTTP). Travel demand forecasting models and other transportation planning activities apply these data to estimate the impact of transportation and land use policy on the transportation system. The Transportation Efficiency Act for the 21st Century (TEA-21), the Clean Air Act Amendments of 1990 (CAA 1990), and the Americans with Disabilities Act all increase transportation planning requirements and related data requirements of states and metropolitan planning organizations (MPOs) (*Implications of Continuous Measurement* 1996).

Census transportation data can be used alone to provide descriptive analysis of work trip patterns and, when compared with previous census-year surveys, of how those work-trip patterns are changing over time. In combination with other data, journey-to-work data supports a region's travel demand forecasting package providing input on workplace attractions, work-trip origin-destination distribution patterns, work-trip departure times, work-trip length distribution, and travel-time distribution. To smaller metropolitan areas with limited resources, the survey becomes a foundation on which to build a model system, and in larger metropolitan areas the survey is a useful data base on observed travel behavior with which to calibrate or validate the model system.

The Census Bureau anticipates greater accuracy with continuous measurement in coding and interviewing due to permanent staffing, instead of the temporary staffing associated with taking the census every ten years (*Implications of Continuous Measurement* 1996). Because of the nature of data collection for the American Community Survey, there is high sampling error, so the data is released with caveats. The high sampling error is countered by moving averages, in which multi-year samples are averaged. Because census data would be collected over five years, instead of on the traditional "census day," the bureau will report "rolling averages" instead of numbers fixed in time.

This study extends the previous BTS study by focusing on the impact of the ACS on *data* for transportation planning and policy analysis. An analytical approach is taken to document the *validity* of estimates from the first year of the ACS demonstration project. This is accomplished by comparing 1996 ACS estimates of household characteristics and transportation variables for Multnomah County, Oregon to data developed locally by the Portland Metropolitan Service District (Portland Metro).

Description of the Data and Development of a GIS Application

Because geographical analysis is a framework for much of the transportation planning activity of Portland Metro and is used to link various data sets, GIS has been selected as the primary methodology for this study. Zoning systems are the principal way in which spatial structure is incorporated into transportation planning models. A zone system is a set of contiguous polygons (areas defined by at least three sides). As this study is focused on evaluating estimates from the ACS, census block groups are considered to be fundamental reporting zones. Because census geography is defined independently of its attribute data, census zones are considered " artificial statistical areas," as the geography is unrelated to zone characteristics (Goodchild *et al* 1993). In this sense, census data models are different from those commonly used for transportation data, where the spatial basis is dependent on unique land use and transportation characteristics.

In contrast to census geography, the design of transportation zoning systems is viewed as an important geographical component of spatial analysis. Transportation Analysis Zone (TAZ) structures endeavor to create homogeneous areas of socioeconomic characteristics and tripmaking tendencies of the population. Portland Metro maintains a set of TAZs for use in planning studies. The zones in the TAZ system constitute the objects, or geographical individuals, that are the basic units for observation and measurement of spatial phenomena.. The current Portland Metro travel demand model forecasting system uses 1260 zones for the four-county area, 481 of which are Multnomah County TAZs and examined in detail for this study.

Figure 1 shows the census block group and TAZ coverages for Multnomah County. Since TAZs and census block groups have noncoterminous boundaries, this study develops spatial basis change methods to transfer socioeconomic data (developed locally at the MPO level) from its reporting zones (TAZs) to an independent set of reporting zones (census block groups). The decision to convert TAZ data to block group data was based on the ACS providing confidence intervals for its estimates. Unfortunately, the Portland Metro data set does not include this information. Only when the spatial basis of household and transportation data has been changed can a one-to-one direct comparative analysis be performed between locally derived data and census data.

Selection of Variables for Comparison

Residential trip generation rates are usually classified by household structure, household size, and auto ownership. These characteristics are important determinants of travel tendencies and are typically considered in trip rate analyses, trip generation studies, and transportation demand models. Unlike household activity surveys, which record information about trips for *all* travel purposes, the census only reports information about the journey to work. But the journey to work is significant because it represents about 25% of all daily trips. More work trips occur during peak travel periods and the average trip length is greater than that for other trip purposes. Destinations for the work trip are less diverse, but origin/destination data are more readily available than the origins and destinations for other trip types (Racca 1998).

This study compares the following variables that are commonly used in travel demand forecasting (Ortúzar and Willumsen 1990):

Household Characteristics	Transportation Characteristics
Number of Households	Means of Transportation to Work
Distribution of Household by Age of Householder	Number of Vehicles Available
Number of Persons in Household	Average Travel Time to Work
Household Income	Time Leaving Home to go to Work

The selected variables' definition and structure are different for ACS and Portland Metro data. Table 1 lists the eight variables chosen for comparison [in Column (1)], each variable's strata in Portland Metro data [in Column (2)], and each variable's strata in ACS data [in Column (3)].

The chief source of Portland Metro data is the 1994/1995 Household Activity Survey. Both temporal and categorical adjustments are necessary to compare the two data sets. First, all Portland Metro data is adjusted to its 1996 equivalent using linear interpolation based on growth factors from other sources. Next, the data are manipulated so that the strata for each variable are identical. Converting some of the variables, such as age of householder and household size requires fairly straightforward aggregation of categories. However, other variables, such as income and means of travel to work (which is the number of work trips by mode) require a great deal of manipulation to achieve similar categories. Some categories cannot be readily converted so there is also information loss in the conversion process. For example, there is no direct conversion of transit trips walk access and park and ride trips to public transportation because we do not know how many of the park and ride trips result in a carpool, vanpool, or public transportation. In this model, we assumed that anyone who carpooled identified their mode as shared ride trips and park and ride trips belong to public transportation. We could not use the variable, time leaving home to go to work because of the aggregate nature of the Portland Metro data. For the travel time to work variable, Census reports the number of households in each time interval. An average value was calculated for each block group by multiplying the mid range value by the number of households, summing then dividing by the number of households. The minimum range value was used for the last interval and the worked at home households were assigned a travel time of 0 minutes. After the adjustments have been made, spatial methods are used to interpolate Portland Metro block group level between the two data sets.

Spatial Methods

A literature review performed for this study found regional and geographic science literature dealing with the theory of spatial basis change (see Batty and Sammons 1978, Batty and Sikdar 1982, Bracken and Martin 1995, Fisher and Langford 1995, Flowerdew and Openshaw 1987, Flowerdew *et al* 1991, Goodchild *et al* 1993, Openshaw and Taylor 1981) and planning and engineering literature dealing with applications of spatial basis change (see Flood and Siaurusaitis 1997, Gan 1994, Hsiao *et al* 1997, Keck 1997, O'Neill *et al* 1992, Peng and Dueker 1993, Zhao 1999). Each allocation method has different data needs and results in errors and different levels of accuracy. The literature review and previous studies point the way toward a methodology for this study.

First, an area ratio basis change is performed as a baseline areal interpolation scenario. The results are a point of departure for comparing other methods. Second, an areal interpolation with an auxiliary variable (street length) is performed. Third, a set of control zones is designated that capture known residential settlement patterns (from land use information). Because of the rich data source from the local planning agency, two roadway networks representing different street types are used. In the third approach, the spatial basis change is enhanced by adding more information about the source zones. This approach represents an improvement over past studies. The three spatial methods used in this study are as follows.

Spatial Data Conversion Method #1: Area Ratio This spatial data conversion method assigns data from TAZs to block groups according to the ratio of land area in the overlap area to land area in the TAZ. This is the common "polygon overlay" routine provided in most GIS.

Spatial Data Conversion Method #2: Roadway Coverage This spatial data conversion method overcomes the assumption of uniform coverage of households across land area by distributing households according to roadway coverage. The built roadway network gives an indication of an area's ability to support population and the roadway network is used an auxiliary variable for distributing households. TAZ data are converted to block group geography according to the ratio of road length in the overlap area to road length in the TAZ.

Spatial Data Conversion Method #3: Roadway Coverage with Control Zones The preceding spatial data conversion method assumes that all households are located along streets. It is possible to enhance the previous method by assigning households based on the type of road and its surrounding land use. The spatial basis change then more closely models the actual distribution of population across geographic areas. This method introduces of a third set of zones, known as control zones, that contain information about residential density and land use.

Table 1: Data Comparability and Estimate Strata, American Community Survey and Portland Metro

Estimate	Portland Metro 1994	ACS 1996
Households	Number of households	Number of households

Age of Householder	Age of head: less than 25 years Age of head: 25 - 54 years Age of head: 55 - 64 years Age of head: 65 or more years	Age of Householder by Household Type: 15 to 64 yrs Age of Householder by Household Type: 65 yrs or more (summed for various categories of household type)	
Household Size	Household size: 1 person Household size: 2 persons Household size: 3 persons Household size: 4 + Persons	Persons in Household: 4 Persons 1 Person 5 Persons 2 Persons 6 Persons 3 Persons 7 + Persons	
Vehicles Available to Household	Average number of vehicles available per household.	Owner 1 vehicle 2 vehicles 3 vehicles 4 vehicles 5 or more vehicles	Renter 1 vehicle 2 vehicles 3 vehicles 4 vehicles 5 or more vehicles
Household Income	Household Income 1994$ less than $ 17,500 $ 17,500 - $ 28,999 $ 29,000 - $ 40,499 $ 40,500 or more	Adjusted Household Income 1996$ less than $5,000 $37,500 - $39,999 $5,000 - $9,999 $40,000 - $42,499 $10,000 - $12,499 $42,500 - $44,999 $12,500 - $14,999 $45,000 - $47,499 $15,000 - $17,499 $47,500 - $49,999 $17,500 - $19,999 $50,000 - $54,999 $20,000 - $22,499 $55,000 - $59,999 $22,500 - $24,999 $60,000 - $74,999 $25,000 - $27,499 $75,000 - $99,999 $27,500 - $29,999 $100,000 - $124,999 $30,000 - $32,499 $125,000 - $149,999 $32,500 - $34,999 $150,000 or more $35,000 - $37,499	
Means of travel to work	Total home-based work trips Drive alone trips Shared ride trips	Car, truck or van: drove alone Bicycle Car, truck or van: carpooled Walked Motorcycle Other Means	

	Transit trips walk access Park and ride trips Walk and bike trips	Taxicab Worked at home Public transportation (bus, trolley, subway, or railroad)
Travel time to work	Aggregate average travel time to work in minutes for each TAZ.	5 min 25 to 29 min 45 to 59 min 5 to 9 min 30 to 34 min 60 to 89 min 10 to 14 min 35 to 39 min 90 or more min 15 to 19 min 40 to 44 min Worked at home 20 to 24 min
Time leaving home to go to work	Time leaving home to go to work (for a sample of households) for 4 superdistricts within Multnomah County.	12:00 -4:59 am 7:00 - 7:29 am 10:00 - 10:59 am 5:00 - 5:29 am 7:30 - 7:59 am 11:00 - 11:59 am 5:30 - 5:59 am 8:00 - 8:29 am 12:00 - 3:59 pm 6:00 - 6:29 am 8:30 - 8:59 am 4:00 - 11:59 pm 6:30 - 6:59 am 9:00 - 9:59 am Worked at home

Results

ACS estimates are released in the form of a direct estimate as well as a lower and upper bound representing the 90 percent confidence interval. Portland Metro interpolated estimates at the block group level are compared to the ACS 90 percent confidence interval to see whether they lie below, within, or above the confidence interval. If the interpolated estimate falls within the ACS 90 percent confidence interval, the interpolated estimate is considered to be a good fit. The results of a block-group level analysis are summed for the entire county and shown in Table 2. The columns containing the percentage of interpolated estimates that lie within the ACS 90% confidence interval are highlighted.

Some differences in the ACS estimates and Portland Metro interpolated estimates can be explained by differences in source data. Table 3 gives the aggregate counts for Multnomah County variables from the ACS (for block groups) and Portland Metro (for TAZs) prior to any spatial modeling. Both data sets encompass *all* of Multnomah County, so the aggregate county sums for household characteristics and transportation variables can be compared to each other. Income ranges 1 and 4, and carpool, transit, and walk/bike means of transportation to work have large aggregate differences. This is partly attributable to temporal and structural changes made to the Portland Metro data to derive comparable data sets.

Both the area ratio and roadway coverage spatial basis change methods give different estimates for household characteristics and transportation variables. Error analysis (RMSE and MAE) of the results indicate that the control zone method outperforms the roadway coverage method and the roadway coverage method outperforms the area ratio method. The control zone method produces more block group-level estimates that fall within the ACS 90 percent confidence interval, and the roadway coverage estimates deviate less from the ACS direct estimate.

With regard to spatial analysis of the findings, in general there appears to be greater differences between the estimates of the two data sets on the periphery of Multnomah County, particularly the far northwestern and far eastern sections of the county. This is consistent with general findings of spatial interpolation, which performs poorest when the street pattern is irregular and/or the two zone systems are not of similarly sized zones.

Analysis of Variance

The three methods used to perform a spatial basis change from TAZ to Block Group and the ACS direct estimates are compared using Analysis of Variance. ANOVA tests were performed for eighteen variables to compare the means of the data generated from the three spatial change methods and the ACS generated data to each other. The null hypothesis tested was:

$H_0: \mu_{acs} \mu_1 = \mu_2 = \mu_3$

Where μ_{acs} = mean of data from the 1996 ACS

μ_1 = mean of data determined from the area ratio spatial basis change method

μ_2 = mean of data determined from the roadway coverage spatial basis change method, and

μ_3 = mean of data determined from the control zone spatial basis change method.

Table 2: Comparison of American Community Survey (ACS) Estimates with Portland Metro Spatially Interpolated Estimates for Household Characteristics and Transportation Data Universe: All Block Groups in Multnomah County

Characteristic	County Total ACS Estimate	Spatial Conversion Method #1: Area Ratio[1]					Spatial Conversion Method #2: Roadway Coverage[1]					Spatial Conversion Method #3: Control Zones[1]				
		County Total		Census Block Groups[2]			County Total		Census Block Groups[2]			County Total		Census Block Groups[2]		
		Estimate	Difference	Within	Greater Than	Less Than	Estimate	Difference	Within	Greater Than	Less Than	Estimate	Difference	Within	Greater Than	Less Than
Number of Households	260,800	259,600	- 0.46 %	44 %	26 %	30 %	261,200	+ 0.15 %	45 %	28 %	27 %	258,300	- 0.96 %	46 %	29 %	25 %
HOUSEHOLD INCOME																
Income Range 1	64,900	79,800	+22.96 %	82 %	17 %	1 %	79,900	+ 23.11 %	82 %	17 %	1 %	79,700	+ 22.80 %	81 %	18 %	1 %
Income Range 2	53,000	57,200	+7.92 %	83 %	14 %	3 %	57,600	+ 8.68 %	91 %	8 %	1 %	57,300	+ 8.11 %	91 %	8 %	1 %
Income Range 3	46,200	43,900	-4.98 %	87 %	10 %	3 %	44,400	- 3.90 %	94 %	5 %	1 %	44,000	- 4.76 %	95 %	4 %	1 %
Income Range 4	96,700	77,900	-19.44 %	87 %	7 %	5 %	79,300	- 17.99 %	97 %	2 %	1 %	77,400	- 19.96 %	97 %	2 %	1 %
AGE OF HOUSEHOLDER																
64 years or less	206,600	203,300	- 1.60 %	89 %	7 %	4 %	204,700	+ 0.92 %	77 %	15 %	8 %	202,000	- 2.23 %	90 %	6 %	4 %
65 years or more	54,200	56,300	+ 3.87 %	82 %	9 %	9 %	56,500	+ 4.24 %	83 %	8 %	9 %	56,000	+ 3.32 %	91 %	5 %	4 %
PERSONS IN HOUSEHOLD																
1 person	87,300	82,000	- 6.07 %	50 %	27 %	23 %	82,300	- 5.73 %	49 %	29 %	22 %	81,600	- 6.53 %	59 %	23 %	18 %
2 persons	88,600	85,800	- 3.16 %	65 %	19 %	16 %	86,400	- 2.48 %	56 %	25 %	19 %	85,300	- 3.72 %	67 %	19 %	14 %
3 persons	36,700	39,000	+ 6.27 %	78 %	19 %	3 %	39,200	+ 6.81 %	66 %	26 %	8 %	38,800	+ 5.72 %	77 %	20 %	2 %
4 or more persons	48,200	52,800	+ 9.54 %	90 %	9 %	1 %	53,200	+ 10.37 %	90 %	8 %	2 %	52,600	+ 9.13 %	90 %	8 %	2 %
MEANS OF TRANSPORTATION TO WORK																
All Work Trips	287,100	267,900	- 6.69 %	81 %	8 %	11 %	249,200	- 13.20 %	83 %	6 %	11 %	270,700	- 5.71 %	86 %	6 %	8 %
Drive Alone	200,300	184,800	- 7.74 %	60 %	17 %	23 %	172,100	- 14.08 %	61 %	12 %	28 %	185,800	- 7.24 %	67 %	14 %	19 %
Carpool	36,400	40,500	+ 11.26 %	66 %	28 %	6 %	37,700	0.00 %	69 %	25 %	6 %	41,000	+ 12.64 %	66 %	29 %	5 %
Transit	32,000	22,400	- 30.00 %	66 %	14 %	20 %	20,800	- 35.00 %	67 %	12 %	21 %	23,000	- 28.13 %	66 %	15 %	19 %

Walk/Bike	16,100	20,100	+ 24.84 %	75 %	24 %	1 %	18,500	+ 14.91 %	79 %	19 %	2 %	20,300	+ 26.09 %	74 %	25 %	1 %
Other	2,400	3,400	+ 41.67 %	-	-	-	3,100	+29.17 %	-	-	-	3,400	+ 41.67 %	-	-	-
Vehicles Available (per household)	1.53 (avg)	1.50 (avg)	- 1.96 %				1.50 (avg)	- 1.96 %				1.54 (avg)	+ 0.65 %			
Travel Time to Work (minutes)	23.25 (avg)	24.71 (avg)	+ 6.28 %				27.20 (avg)	+ 16.99 %				25.27 (avg)	+ 8.69 %			

[1] - Portland Metro Interpolated Estimate

[2] - Interpolated Estimate Relationship to ACS 90% Confidence Interval

Results of the ANOVA are shown in Table 4. Generally, the spatial basis change methods have the most impact on the transportation related variables. The null hypothesis is rejected for all transportation variables except vehicles available per household and two of the household income variables. For these variables, we conclude that the data resulting from the spatial conversion algorithms and the ACS data are not the same when the level of significance is 0.05.

Table 3: Percent Difference Between ACS and Portland Metro Estimates Prior to Spatial Basis Change

Characteristic	Year	Multnomah County Total		
		ACS Estimates for Block Groups Number of Households	Portland Metro Estimates for TAZs Number of Households	Percent Difference
Number of Households	1996	260,800	262,500	+ 0.65 %
HOUSEHOLD INCOME				
Income Range 1	1996	64,900	80,200	+ 23.57 %
Income Range 2	1996	53,000	57,800	+ 9.06 %
Income Range 3	1996	46,200	44,600	- 3.46 %
Income Range 4	1996	96,700	79,800	- 17.48 %
AGE OF HOUSEHOLDER				
64 years or less	1996	206,600	205,800	- 0.39 %
65 years or more	1996	54,200	56,700	+ 4.61 %
PERSONS IN HOUSEHOLD				
1 person	1996	87,300	82,600	- 5.38 %
2 persons	1996	88,600	86,900	- 1.92 %

3 persons	1996	36,700	39,500	+ 7.63 %
4 or more persons	1996	48,200	53,500	+ 11.00 %
MEANS OF TRANSPORTATION TO WORK				
All Work Trips	1996	287,100	282,500	- 1.60 %
Drive Alone	1996	200,300	195,200	- 2.55 %
Carpool	1996	36,400	42,700	+ 17.31 %
Transit	1996	32,000	24,000	+ 31.88 %
Walk/Bike	1996	16,100	20,600	+ 16.46 %
Other	1996	2,400	3,400	
Vehicles Available	1994	(average) 1.57 vehicles per hh	(average) 1.56 vehicles per hh	- 0.64 %
Travel Time to Work	1994	(average) 23.38 minutes	N.A.	-

For each variable for which the previous null hypothesis was rejected, a separate ANOVA was run comparing only the spatial basis change methods to each other. In other words, the null hypothesis is:

$H_0 : \mu_1 = :\mu_2 = :\mu_3$

Results from this analysis are in Table 5. The spatial basis change methods are not significantly different for the two income ranges. However, there is a significant difference among the methods for the transportation variables.

Paired T-tests between the ACS data and the data generated using method 1, areal interpolation, on household income ranges 1 and 4 result in rejecting the null hypothesis that the means are equal at a level of significance = 0.05. The same holds true comparing the ACS income data to the results from the other two spatial basis change methods. Consequently, none of the spatial basis change methods are capable of generating data similar to that collected in the ACS for household income ranges 1 and 4. However, part of this conclusion may be a function of the data conversion methods used to achieve comparable categories in each database prior to applying spatial basis change. Particularly, temporal adjustment and linear interpolation between categories may have an effect on the data.

Paired T-tests were also performed for the transportation variables to determine if any of the methods resulted in values that were similar to the ACS. For total number of work trips, the hypothesis, that the difference in the means equals zero, is rejected when comparing each method individually to the ACS. However, when comparing data from the areal interpolation method (1) and the control zone method (3) we fail to reject the hypothesis that the means are equal. Similar conclusions are reached for the drive work trips, carpool work trips, transit work trips, and walk/bike work trips. The exception is average travel time to work. In this case, none of the methods compare favorably with the ACS and the null hypothesis for the test comparing Method 1 and Method 3 is rejected.

Conclusions

From a research perspective, the big question at hand is how reliable are the household and transportation estimates of the ACS? For most household characteristics, local data falls within the ACS 90 percent confidence interval 80 to 90 percent of the time. For most transportation variables, local data falls within the ACS 90 percent confidence interval 75 to 90 percent of the time.

Table 4: ANOVA Results

Variable	F	p-value	Fcrit	Conclusion
Number of households	0.043552	0.987917	2.609099	Fail to reject
Age of head 15-64 years old	0.109285	0.954697	2.609099	Fail to reject

Age of head 65+ years old	0.392199	0.758634	2.609099	Fail to reject
Household size – 1 person	0.829145	0.477725	2.609099	Fail to reject
Household size – 2 persons	0.30456	0.822121	2.609099	Fail to reject
Household size – 3 persons	0.893733	0.443613	2.609099	Fail to reject
Household size – 4+ persons	1.656376	0.174417	2.609099	Fail to reject
Household income – Range 1	10.06968	1.38E-06	2.609099	Reject
Household income – Range 2	1.899782	0.127538	2.609099	Fail to reject
Household income – Range 3	0.583593	0.625772	2.609099	Fail to reject
Household income – Range 4	8.125438	2.22E-05	2.609099	Reject
Total Work Trips	129.3125	9.33E-77	2.609212	Reject
Drive alone work trips	93.98567	5.33E-57	2.609234	Reject
Carpool work trips	161.2389	7.45E-94	2.609234	Reject
Transit work trips	111.9429	4.6E-67	2.609298	Reject
Walk/Bike work trips	61.86093	2.63E-38	2.609184	Reject
Vehicles available per HH	1.482552	0.217342	2.609184	Fail to reject
Travel time to work	26.34294	1.01E-16	2.609184	Reject

Table 5: ANOVA Comparison of Methods only

Variable	F	p-value	Fcrit	Conclusion
Household income – Range 1	0.006299	0.993721	3.001389	Fail to reject
Household income – Range 4	0.120231	0.886724	3.001389	Fail to reject
Total Work Trips	161.2782	2.49E-64	3.00156	Reject
Drive alone work trips	118.2347	1.68E-48	3.00156	Reject
Carpool work trips	186.1194	3.88E-73	3.00156	Reject
Transit work trips	193.0855	2.09E-75	3.001674	Reject

Walk/Bike work trips	63.47004	3.16E-27	3.001496	Reject
Travel time to work	15.02155	3.45E-07	3.001489	Reject

Additional analysis reveals that the three spatial basis change methods generate comparable data to the ACS estimates for the socio-demographic data except certain income ranges. The three methods do generate data in these income ranges that are comparable with each other. None of the methods produces data comparable to each other or the ACS data for the transportation variables except average number of vehicles available to a household. However, for the remaining transportation variables, with the exception of average travel time to work, method 1 and method 3 produce comparable results by block group. Consequently areal interpolation yields similar results to the control zone method for all but one variable.

None of the methods are sufficient for reproducing transportation data similar to the ACS estimates. There are a few reasons this may occur. First, the Portland Metro data may be dissimilar to the ACS estimates given the different survey methods used. Additionally, no information on the sampling error associated with the Portland Metro survey is provided. According to Keith Lawton at Portland Metro, the activity-based survey in Portland undercounts the people who are too busy with activities to answer the CATI but there is no published information on the sources and sizes of sampling errors. Second, the transportation variables all required some type of transformation to generate similar categories for comparing ACS data to Portland Metro data. These transformations, including temporal, introduce error into the data that may be significant enough to impact the results from the spatial basis change methods. Third, none of the spatial basis change methods may capture the underlying geographical influences to adequately convert the data. For example, the number of work trips by transit is influenced more by accessibility (proximity) to transit services than the ubiquity of the street network. Also, carpool trips to work are greatly influenced by the proximity of those living near you or in your travel corridor who also work near you or in your travel corridor. None of these spatial relationships are considered by the spatial basis change methods described here.

Further consideration of the impacts on transforming variable types is certainly recommended. Unlike the geography problem, which will disappear when the journey to work coding of the ACS is performed on TAZs, this problem on characterization of data by Census and transportation planners will not change until planning models change. Consequently, planners will have to transform ACS data into the structure with which they are used to working.

References

Batty, M. And R. Sammons. 1978. On Searching For the Most Informative Spatial Pattern. *Environment and Planning A.* Vol. 10, pp. 747-779.

Batty, M. And P.K. Sikdar. 1982. Spatial Aggregation in Gravity Models. *Environment and Planning A.* Vol. 14, pp. 795-882.

Bracken, I. and D. Martin. 1995. Linkage of the 1981 and 1991 U.K. Censuses Using Surface Modelling Concepts. Environment and Planning A. vol. 27. pp. 379-390.

Caliper Corporation. 1996. *Caliper TransCAD User's Guide. Version 3.0 for use with MicroSoft Windows.*

Fisher, P.F. and M. Langford. 1995. Modelling the Errors in Areal Interpolation Between Zonal Systems by Monte Carlo Simulation. *Environment and Planning A.* vol. 27. pp. 221-224.

Flood, Michael J. and Victor Siaurusaitis. 1997. Geographic Analysis Techniques for Use in Building or Updating a Travel Demand Model. *Proceedings of the Institute of Transportation Engineers 67th Annual Meeting.* 1997, Boston, Mass.

Flowerdew, R. and S. Openshaw. 1987. *A Review of the Problem of Transferring Data from One Set of Areal Units to Another Incompatible Set.* Northern Regional Research Laboratory, Universities of Lancaster and Newcastle-Upon-Tyne.

Flowerdew, R., M. Green, and E. Kehris. 1991. Using Areal Interpolation Methods in Geographic Information Systems. *Papers in Regional Science.* vol. 70. pp. 303-315.

Gan, Cheng-Tin. 1994. A GIS-Aided Procedure for Converting Census Data for Transportation Planning. *ITE Journal.* November, 1994. pp. 34-40.

Goodchild, M.F., L. Anselin, and U. Deichmann. 1993. A Framework for the Areal Interpretation of Socioeconomic Data. *Environment and Planning A.* vol. 25. pp. 383-397.

Greenshields, Bruce Douglas and Frank Mark Weida. 1978. *Statistics, With Applications to Highway Traffic Analysis.* Westport, Conn.: Eno Foundation for Transportation.

Griffith, Daniel. 1987. *Spatial Autocorrelation: A Primer.* Washington, D.C.: Association of American Geographers.

Hsiao, Shirley, Jian Lu, James Sterling, and Matthew Weatherford. 1997. Use of Geographic Information Systems for Analysis of Transit Pedestrian Access. *Transportation Research Record 1604.* pp. 50-59.

Implication of Continuous Measurement for the Uses of Census Data in Transportation Planning. Washington, D.C.: U.S. Department of Transportation, Bureau of Transportation Statistics.

Keck, J. Kevin. 1997. A GIS Based Census Block Group Level Transit Mode Analysis. *Proceedings of the Institute of Transportation Engineers 67th Annual Meeting.* 1997, Boston, Mass.

Lawton, T. Keith and Eric I. Pas. 1995. Resource Paper for Survey Methodologies Workshop. *Proceedings of a Conference on Household Travel Surveys: New Concepts and Research Needs.* March 12-15, 1995, Irvine, California.

O'Neill, Wende A., R. Douglas Ramsey, and JaChing Chou. 1992. Analysis of Transit Service Areas Using Geographic Information Systems. *Transportation Research Record 1364.* pp. 131-

138.

Openshaw, S. and P.J. Taylor. 1981. The Modifiable Area Unit Problem. in *Quantitative Geography: A British View.* eds. Wrigley, N., R.J. Bennett. London: Routledge & Kegan Paul. pp. 60-70.

Ort_zar, Juan de Dios and Luis G. Willumsen. 1990. *Modelling Transport.* New York, New York: John Wiley & Sons.

Peng, Zhongren and Kenneth J. Dueker. 1993. Error and Accuracy in Spatial Data Allocation. *Proceedings of GIS/LIS Conference.* 1993, Minneapolis, Minn.

Racca, David P. 1998. Using GIS to Identify Markets for Transit in Delaware. *Proceedings of the Geographic Information Systems for Transportation Symposium.* 1998, Salt Lake City, Utah.

Wright, Tommy. 1998. Sampling and Census 2000: The Concepts. *American Scientist.* vol. 86. pp. 245-253.

Zhao, Fang. 1999. *GIS Analysis of the Impact of Community Design on Transit Accessibility.* (Working paper.) Miami, Florida: Florida International University.