Lifestyle clusters for labor force participation, occupation and housing use in integrated land use and transportation modeling

Bayarma Alexander
Urban Land Use and Transportation Center
University of California
One Shields Avenue
Davis, CA 95616
U.S.A
Tel: (530) 752 1331
E-mail: baleksandr@ucdavis.edu

Dimantha De Silva
HBA Specto Incorporated
101-6th Avenue SW
Calgary, AB, T2P 3P4
Canada
Tel: (+403) 2321060
E-mail: dds@hbaspecto.com

John E ABRAHAM
Vice-president
HBA Specto Incorporated
101-6th Avenue SW
Calgary, AB, T2P 3P4
Canada
Tel: (+403) 2321060
E-mail: jea@hbaspecto.com

John Douglas Hunt
Professor
Department of Civil Engineering
University of Calgary
2500 University Drive NW
Calgary, AB, CANADA T2N 1N4
Tel: +1 (403) 220 8793
E-mail: jdhunt@ucalgary.ca

Shengyi Gao
Urban Land Use and Transportation Center
University of California
One Shields Avenue
Davis, CA 95616
U.S.A
Tel: (530) 7526303
E-mail: sgao@ucdavis.edu

Abstract: The California statewide integrated land use and transportation modeling has being developed using the PECAS framework. This paper focus on two goals: a) presenting a framework for identifying household lifestyle clusters, and b) applying them to a spatial economic model system. A household's lifestyle is defined by two sets of dimensions: *space use* and *household wage*. Using a two-step clustering algorithm, different household lifestyle clusters were identified. The applicability of these lifestyle clusters for a integrated land use and transportation model was explored.

Key words: Integrated land use and transportation modeling, lifestyle clusters, PECAS, PUMS, California

1. INTRODUCTION

The concept of the lifestyle has been extensively used in market, economic, transportation and studies (Lazer, 1963, Duchin 2003, Bagley & Mokhtarian 1999, Krizek & Waddell 2003 and Walker and Li 2007). It was initially proposed by Lazer (1963) and has been identified as a useful framework for describing clusters of household choices of residential location, labor force activity and auto-ownership that predispose or condition patterns of daily activity and travel behavior. In transportation studies, for instance, Krizek & Waddell (2003) defined four types of "lifestyle" dimensions (travel behavior, activity participation, vehicle ownership and residential location) and Bagley & Mokhtarian (1999) defined eleven lifestyle dimensions to investigate the importance of lifestyle variables to residential neighborhood choice. Walker and Li (2007) developed latent class choice models to segment households according to their preferences in residential location choices. They identified three latent lifestyle segments: suburban dwellers, urban dwellers and transit riders.

In this study, we present a framework for identifying household lifestyle clusters and applying them to PECAS, a spatial economic model system (Hunt and Abraham, 2005). The California statewide integrated land use and transportation modeling is being developed using the PECAS framework. PECAS stands for Production Exchange Consumption Allocation System and it considers '*activities*' located in a set of geographic zones covering the area under considerations. These activities include industrial sectors, households, business establishments, and other institutions. They produce '*commodities*' (goods, services labor, capital, and floor space) and then transport and sell these commodities. They also consume commodities during this production according to their technology. The *Activity Allocation Module* (AA) of the PECAS system allocates the study area-wide quantity of each activity in a three level additive logit model framework based on random utility theory (Abraham and Hunt, 2007). The current paper focus on the *household activities* of the California PECAS model. Households act as consumers in the real estate, goods and services, and transportation markets; as suppliers of labor; and as investors in the operational model system.

The aim of this paper is to increase our understanding of broad decisions of households regarding their labor force participation, housing type, and housing choice. We apply a clustering method to identify a number of classes of households with strong internal similarities in their labor force participation and housing use. The different household lifestyles are formed by a combination of two sets of dimensions (space use and household wage) and they are used to classify households as lifestyle clusters. These dimensions are supported by the data and support a link of the lifestyle clusters with the PECAS spatial economic model system.

The outline of this article is as follows. Section 2 presents a review of past studies on household lifestyles. Section 3 describes the research design and applied methodology of this study. The empirical results are presented in sections 4 and 5. In section 6, we put forward our conclusions and suggest avenues for further research.

2. HOUSEHOLD LIFESTYLES: housing and labor force participation

Ben-Akiva et al. (1996) define lifestyle decisions as “where to live, how many hours to work, where to work, how many hours to allot for social and recreational activities, how many automobiles to own, and which travel mode to use for work trips.”

A substantial body of studies has used the term ‘*lifestyle*’ as a framework for describing clusters of residential location choice, household's preferences in labor and other activity participations (Ben-Akiva and Bowman 1998, Salomon and Ben-Akiva 1983, Bagley & Mokhtarian 1999, Waddel 2000, Krizek & Waddell 2003, Moekel *et al*. 2003, Cao and Mokhtarian 2005, Walker & Li 2007). For instance, Salomon and Ben-Akiva (1983) used a k-means clustering approach to define five lifestyle segments. They used household structure, labor force participation, and education as input variables for the cluster analysis then estimated mode and destination choice models for each lifestyle segment. Using an attitudinal survey from a San Francisco Bay Area neighborhood, Bagley and Mokhtarian (1999) applied factor analysis to measure ten attitudinal and eleven lifestyle dimensions and applied them with travel and demographic characteristics as explanatory variables in a residential choice model. Krizek and Waddell (2003) referred to lifestyle in the context of the set of household choices: travel behavior, activity participation, vehicle ownership and residential location. They argue that by understanding each lifestyle relative to its socio-demographic dimensions, researchers can gain better understanding of how different phenomena interact and the potential market of households that may respond to various land use and transportation-planning initiatives.

Households operate within multiple markets for housing, labor, and travel. A number of studies have investigated consumer choice in the housing market (Quigley 1976, 1984; Fischer & Aufhauser 1988, Borch-Supan & Pitkin 1988, Kim 1992, Cho 1997). Quigley (1976) estimated the choice behavior of 18 alternatives of residential housing and residential location in Pittsburgh. Quigley found that as family size increases, households are less likely to choose multifamily units and are more likely to choose common-wall units and single detached units; larger family with higher household income choose less dense housing; the increased family size is associated with the larger interior housing size. Using the 1977 Annual Housing Survey (AHS), Borch-Supan and Pitkin explore housing consumption in the Albany-Schenectady-Troy SMSA, New York. They focused on nine housing types and explored the role of demographic and financial variables. Their result indicated that income, relative out-of-pocket costs have a strong influence on housing choices. Similar findings are presented by Kim (1992) and Ahmad (1994) who found that family size, education, and income are the significant determinants of housing choice. Past studies have found that there is a direct relationship between the size of a dwelling and the number of people who live in it (Narrol 1962, Cook and Heizer 1965). Households of higher wealth and/or status live in larger dwellings than those of low wealth/status (Hayden & Cannon 1982, Arnold and Ford 1980, Bawden 1982).

Households make labor supply and labor force participation decisions for each of their members. Household decisions regarding labor force participation (number of working hours) may feed back on decisions regarding housing consumption, and choices of housing consumption may influence employment decisions. For example, workers participating in higher paid jobs are able to afford more housing services; the increased work hours support the annual consumption of housing services. The labor-housing simultaneous decisions have received relatively little research attention in land use and transportation studies. This may be because of the complication of units. The unit of analysis for the labor market is generally the individual, whereas the housing market analysis is at household level (Waddel 2000). There are several economic studies that have examined the household-level housing and labor force decisions. For example, Kohlhase (1986) found that household decisions regarding labor supply and housing demand are not separable but feedback upon each other. He modelled the labor-housing decisions joint for seven demographic groups. His results have shown that the housing and labor market behavior of one-earner households significantly differs from that of two-earner households.

Tokle and Huffman (1991) studied the effects of geographical differences in local economic conditions on wage labor demand and labor force participation decisions of U.S. farm and rural nonfarm couples. They have found that labor force participation decisions of households are also affected by changes in anticipated local economic conditions. For farm households, the probability of wage work increases when expected farm output prices decline and decreases when local labor demand grows. Given the general trend toward larger farms, the results show that existing agriculture is significantly easier for farm households that reside in localities where nonfarm employment is expected to grow rapidly.

3. RESEARCH DESIGN AND METHODOLOGY

3.1 Data and variables

The lifestyle concept has been widely used in different fields of studies including land use and transportation. Each field and each study employs its own specific definition of lifestyle and uses a different number of dimensions depending on the research purpose and data availability. In the current study, we define household lifestyles by two sets of dimensions: *space use* (floor space by 14 space types, estimated square feet of house) and *household wage* (household wage income by 19 occupation types). Table 1 shows the description of these dimensions. Floor space type is a categorical variable with 14 categories of residential space, and the estimated square feet and wages are continues variables.

Table 1 : Dimensions of household lifestyle

	Name		Type of variable
	Dimension 1		
1	Space type	14 categories listed in Table 2	Categorical
2	Estimated square feet		continuous
	Dimension 2	**2000 SOC Codes**	
3	Agriculture worker's wage	11-	
4	Assembly worker's wage	23-	
5	Business worker's wage	15-23	
6	Construction worker's wage	25-1000	
7	Entertainment worker's wage	25-2000, 25-3000, 25-4000, 25-9000	
8	Food worker's wage	27-	
9	Health worker's wage	29-31	
10	Maintenance worker's wage	33-, 37-, 39-	
11	Manager worker's wage	35-	
12	Military worker's wage	41-1000, 41-2000	
13	Non-retail worker's wage	41-3000, 41-4000, 41-9000	
14	Office worker's wage	43-	
15	Post secondary education worker's wage	45-	
16	Primary education worker's wage	47-	
17	Professional worker's wage	49-	
18	Retail worker's wage	51-	
19	Service worker's wage	53-	
20	Transport worker's wage	55-, 99-	

Detailed household information was available from the synthetic population data of the California PECAS Model. The PECAS Activity Allocation (AA) and Space Development (SD) modules use year 2000 as base year. The floor space inventory, household and industry activity totals by transportation analysis zone (TAZ) are required as inputs. We used a Population Synthesizer to derive a synthetic population using known target totals for several key characteristics, such as income class, dwelling type, and age. The population synthesizer works by combining a trial population of households drawn from a sample and altering it by switching new possible households in; if the match with the targets improves, the new household is kept. Two elements are needed to run the Population Synthesizer: samples which are the individual household records to be used and targets which are the control totals for geographies in the model system, such as zones.

The Public Use Microdata Sample (PUMS) 5% data from the U.S. Census was used as the basis for samples. These samples consist of housing units and persons, with detailed descriptive data but anonymized by replacing detailed location data with large area (PUMS Area, or “PUMA”) definitions. A composite sample record was created for each PUMS housing unit and the associated person record(s), with the totals for each of the targets. All targets were available at the block group level, and thus, at the zone level. The targets were treated categorically (for example, rather using an average income, the number of households in seven income categories is represented). Because categorical totals are used, all target weights were set to be equal. The target totals used are: household size, dwelling type, household income, persons by age, auto ownership, workers by occupation, students by school type, number of rooms and number of bedrooms. These targets were all derived from Census SF3 tables, provided at the block group and aggregated to the zonal level as needed.

Each PUMS household was categorized based on the type of housing (Table 2) and the type of household (Table 3). For each zone a housing value threshold (VT in Table 2) was defined, to separate high value (“luxury”) housing from lower value (“economic”) housing. A selection of households in buildings with more than 50 units in the building in PUMAs that contained high rise dwellings were classified as high-rise, since the PUMS data does not have information on the height of any large buildings.

Table 2: PECAS residential space type

	Space type name	Definition
1	Rural luxury residential	Detached, >= 10 Acres, >=VT
2	Rural economic residential	Detached, MH or Van, >= 10 Acres, <VT
3	Acreage luxury residential	Detached, >= 1 Acre

		<10 Acres, >=VT
4	Acreage economic residential	Detached, MH or Van, >= 1 Acre, <10 Acres, <VT
5	Single family detached luxury residential	Detached, <1 Acre, >=VT
6	Single family detached economic residential	Detached, <1 Acre, >=VT
7	Joined luxury residential	Attached or Multiunit <=4 units, >=VT
8	Joined economic residential	Attached or Multiunit <=4 units, <VT
9	Low-rise luxury residential	Multiunit >4 units, >=VT
10	Low-rise economic residential	Multiunit >4 units, <VT
11	High-rise luxury residential	Multiunit >50 units, >=VT, *
12	High-rise economic residential	Multiunit >50 units, <VT, *
13	Urban mobile home residential	Mobile Home, Van, < 1 Acre
14	Group Quarter residential	Non institutional group quarters

* see text for high-rise definition

PECAS household types: The household lifestyles are defined for each PECAS household type. The PECAS household types were defined from the PUMS data for the state of California. We distinguish 25 categories of household based on household income and number of people in a household (see Table 3).

Table 3: PECAS household type

Type ID	PECAS Household Type (income x $1,000/year)	N
1	up to 15 income 1 person	62567
2	up to 15 income 2 persons	14673
3	up to 15 income 3 or 4 persons	16773
4	up to 15 income 5 persons	8013
5	up to 15 income all senior	22946
6	15 to 50 income 1 person	40767
7	15 to 50 income 2 persons	47606
8	15 to 50 income 3 or 4 persons	59803
9	15 to 50 income 5 persons	35280
10	15 to 50 income all senior	37827
11	50 to 100 income 1 person	18574
12	50 to 100 income 2 persons	46697
13	50 to 100 income 3 or 4 persons	61310
14	50 to 100 income 5 persons	31429
15	50 to 100 income all seniors	11225
16	100 to 150 income 1 person	3368
17	100 to 150 income 2 persons	17090
18	100 to 150 income 3 or 4 persons	23375
19	100 to 150 income 5 persons	10374
20	100 to 150 income all seniors	3604

21	150 or more income 1 person	2463
22	150 or more income 2 persons	13259
23	150 or more income 3 or 4 persons	16134
24	150 or more income 5 persons	6582
25	150 or more income all seniors	4376

3.2 Methodology

A methodology of defining lifestyle clusters will be discussed in the next section followed by a framework for applying the lifestyle clusters to integrated land use and transportation model..

Defining lifestyle clusters: We applied two-step clustering method to identify a number of classes of household (henceforth referred to as representative lifestyle clusters) with strong internal similarities in their labor force participation and housing use. There are many clustering methods available, general classes are: hierarchical and non-hierarchical clustering. Hierarchical clustering proceeds successively by either merging smaller clusters into larger ones or by splitting larger clusters. Non- hierarchical clustering, on the other hand, attempts to directly decompose the data set into a set of discrete clusters. The criterion function that the clustering algorithm tries to minimize may emphasize the local structure of the data by assigning clusters to peaks in the probability density function or the global structure. Typically the global criteria involve minimizing some measure of dissimilarity in the samples within each cluster, while maximizing the dissimilarity of different clusters. A commonly used non-hierarchical clustering method is k-means clustering.

We tested some of methods to classify households including k-means clustering, hierarchical clustering, fuzzy clustering and two-step clustering using SPSS and S-PLUS software. The k-means, EM clustering and other traditional clustering methods were effective and accurate on small data sets, but did not scale up to the large data sets. We chose the two-step clustering method because it had several advantages over the other methods. For instance, the two-step clustering algorithm handles very large datasets. In addition, it is capable of handling both continuous and categorical variables and requires only one data pass in the procedure. The two-step clustering algorithm has two steps: (i) pre-cluster the cases into many small sub-clusters (ii) cluster the sub-clusters resulting from pre-cluster step into a desired number of clusters.

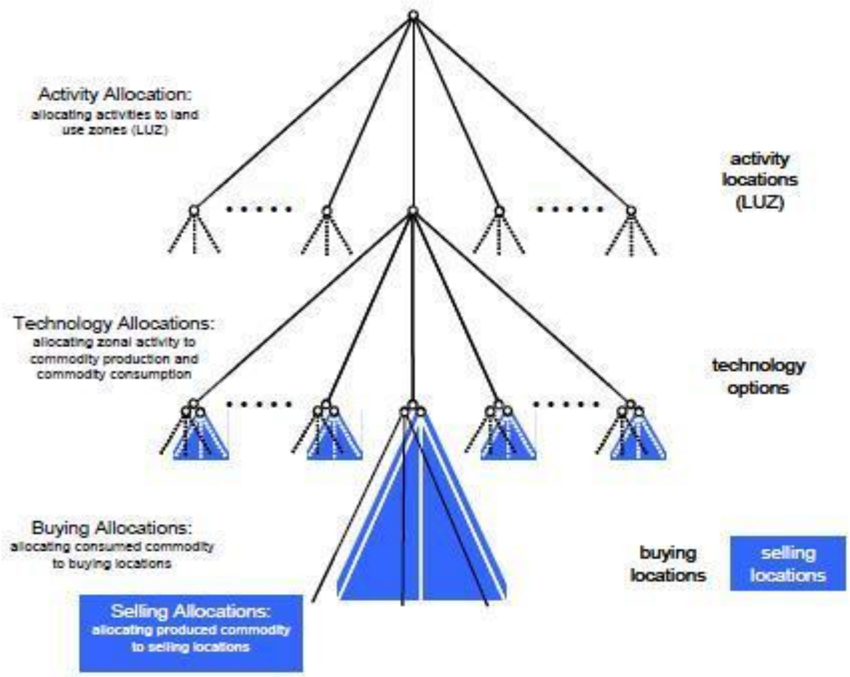

Figure 1 Three Level Nesting Structure Used in AA (Source_ PECAS Theoretical Formulation)

Framework for applying lifestyle clusters to a spatial economic model: The PECAS framework includes two basic modules: the Activity Allocation (AA) and Space Development (SD) modules (Hunt and Abraham, 2005). The allocation process in the AA module uses a three-level additive/nested logit model (Abraham and Hunt, 2007) with a nesting structure as shown in Figure 1. The AA module allocates production and consumption activities across land use zones (LUZ). At the highest level of the nesting structure, the locations of activities are represented by LUZs, which are aggregations of TAZs. At the middle level, the quantity of each activity in each LUZ is allocated among available technology options. At its lowest level, the total quantity of each commodity produced at the intermediate level was allocated among the exchange locations.

In the second level of logit choice of the AA module, an activity chooses to produce and consume goods, services, labor, and space at a fixed rate or elastic rates. This choice is defined as *technology option choice*. A household in the AA module is treated as an activity which produces labor and consumes space at elastic rates while consuming goods and services at a fixed rate. In this paper, we define households' technology options as choices of lifestyles. Figure 2 presents a sample *Technology Options* table (household technology options are shown) of the PECAS. The Technology Options table is one of the input tables of the AA module:

	A	B	C	D	E	F	G	H	I	J	K
1		Activity	Option Name	Option size	Commodities						
2					occup_1	occup_2	occup_3	…….	spaceID	spaceII	spaceID3
3	Producing sectors	100 to 150 income 3 or 4 persons	Cluster1	25	73.33	0	3511.11	0	0	0	0
4		100 to 150 income 3 or 4 persons	Cluster2	20	0	1749.28	2492.75	0	0	0	0
5		100 to 150 income 3 or 4 persons	Cluster3	62	0	768.69	0	0	0	0	0
6		100 to 150 income 3 or 4 persons	Cluster4	99	69.93	362.94	0	0	0	0	0
7		100 to 150 income 3 or 4 persons	Cluster5	143	0	1650	6351.33	0	0	0	0
8		15 to 50 income 1 person	Cluster 65	264	0	1710.2	0	0	0	0	0
9		15 to 50 income 1 person	Cluster 66	257	0	0	2844.06	0	0	0	0
10		15 to 50 income 1 person	Cluster 67	325	0	47653.78	0	0	0	0	0
11		15 to 50 income 1 person	Cluster 68	252	0	41019.94	0	0	0	0	0
12		15 to 50 income 1 person	Cluster 69	336	0	0	35847.27	0	0	0	0
13		15 to 50 income 1 person	Cluster 70	436	37517.58	0	0	0	0	0	0
14											
15	Consuming sectors	100 to 150 income 3 or 4 persons	Cluster1	25	0	0	0	0	-1463.3	0	0
16		100 to 150 income 3 or 4 persons	Cluster2	20	0	0	0	0	0	-640.9	0
17		100 to 150 income 3 or 4 persons	Cluster3	62	0	0	0	0	0	0	0
18		100 to 150 income 3 or 4 persons	Cluster4	99	0	0	0	0	-672.4	0	0
19		100 to 150 income 3 or 4 persons	Cluster5	143	0	0	0	0	0	-693.2	0
20		15 to 50 income 1 person	Cluster 65	264	0	0	0	0	0	0	-1256.97
21		15 to 50 income 1 person	Cluster 66	257	0	0	0	0	0	-717	0
22		15 to 50 income 1 person	Cluster 67	325	0	0	0	0	0	-650.7	0
23		15 to 50 income 1 person	Cluster 68	252	0	0	0	0	-1275.4		

Figure 2 Lifestyle clusters in AA input table

-Column A is the technology types: production and consumption
-Column B is Activity (PECAS household activities: 25)
-Column C is Option Name
-Column D is Option weight, the a-priori likelihood of a household choosing the option
-Columns E to G are production of labor by households
-Columns I to K are consumption of space by households (negative numbers indicated consumption rather than production)

The average use and make rates of household technology options were calculated for each lifestyle cluster and entered in the TechnologyOptionI table.

4. EMPRICAL RESULTS

4.1 Representative lifestyle clusters

The two-step clustering analysis was carried out with 20 variables (Table 1) to identify the representative lifestyle clusters with internal similarities in their space use, occupation, and wages in SPSS (Version 18.0). One categorical variable and 19 continues variables were included. SPSS estimated the scale parameters for all 19 continuous variables. We applied the cluster analysis for each PECAS household type (Table 3). The first step calculates Bayesian information criterion (BIC) for each number of clusters within a specified range and uses it to find the initial estimate for the number of clusters. The second step refines the initial estimate by finding the greatest change in distance between the two closest clusters in each hierarchical clustering stage. The log-likelihood measure was used to calculate the distance between clusters.

The number of lifestyle clusters range from 12 to 231. The mean number of clusters is 108.2. Three or four person households with an income of $50,000 to $100.000 have the highest number of lifestyle clusters (N=231). One person households with an income of up to $15,000 have the fewest clusters (N=12). Figure 3 shows the lifestyle cluster number by each PECAS household type. Large households with high household income have more lifestyle clusters than small and low-income households, since large households with high income often have many employed people, leading to more variations in the labor market than small, low-income households. A large high income household with many members of working age has more options in both the labor market and the housing market than a low income household with 1 or even 0 workers.

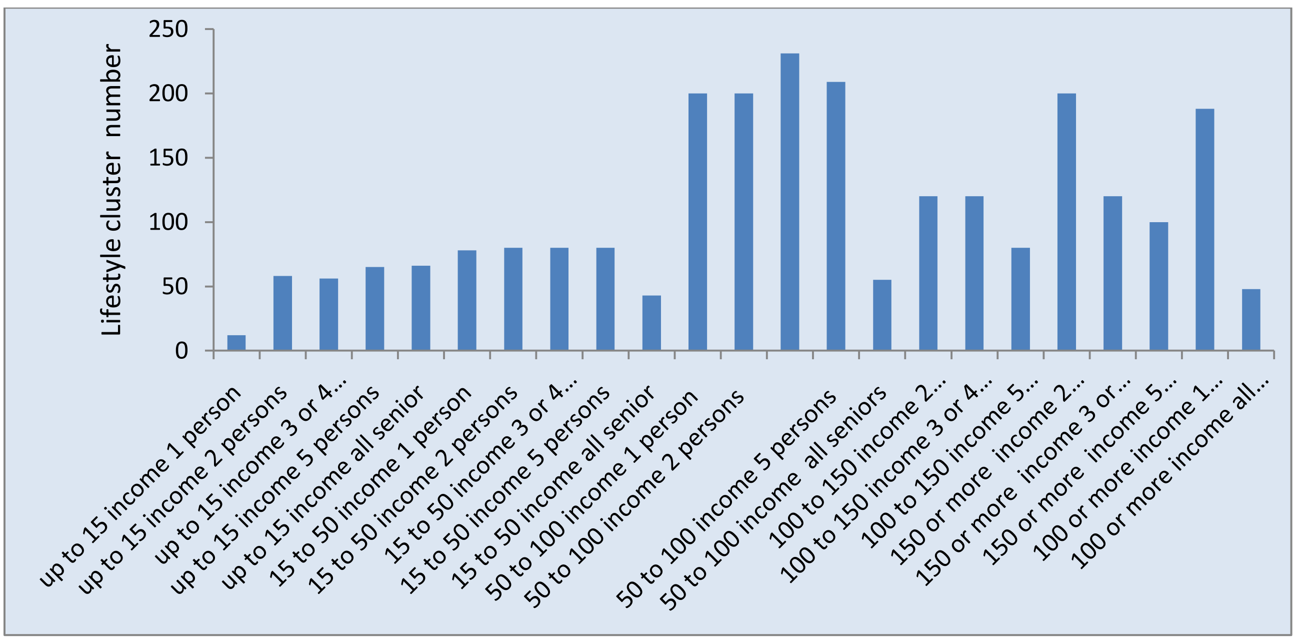

Figure 4 Lifestyle cluster numbers by PECAS household types

4.2 Representative lifestyles in California PECAS

In order to determine more realistic and representative space types for each cluster, we have conducted some preliminary analysis. Figure 5 below shows a snapshot of a cross tabulation between PECAS space types and lifestyle clusters of five-person households with incomes of $100,000 to $150,000. Eighty lifestyle clusters were determined for this household type. Figure 1 shows the distributions of space types for 17 clusters.

The number of clusters was chosen so that most of the lifestyle clusters used one space type. However, some lifestyle clusters consumed two or more types of space. For example, lifestyle cluster 12 in Figure 5 has 70 households and they consume two space types: 34 households use the "rural economy residential" space type and 36 households use the "joined lux residential" space type. In the PECAS model, logit model “size terms” (Daly, 1982) are calculated for each technology option based on the type of space being consumed, so that the *Ceteris paribus* probability of choosing a cluster is based on the quantity of the type of housing in the cluster. Thus all clusters must contain only one space type. We adjusted the number of clusters based on the mixing of space types and labor types in the cluster, we wanted the clusters to represent choices of occupations and housing for the family, and increased the number of clusters until the analysis most clusters (generally all but 3) were clearly dominated by one space type, and 75% or more of the clusters were dominated by one occupation type. The few (generally 3 or less) clusters that remained with multiple space types were than manually split into sub-clusters. In example provided in Figure 5, we split cluster 12 into two sub-clusters (cluster 12a and cluster 12b) in order represent more space types.

	A	B	C	D	E	F	G	H	I	J	K	L	M
1													
2			rural lux resi	rural econ	acreage lux r	acreage ec	SFD lux resi	SFD econ re	joined lux	joined eco	lowrise lux	lowrise ecor	urban MH r
3				Frequency	Frequency	Frequency	Frequency	Frequency	Frequency	Frequency	Frequency	Frequency	Frequency
4	Cluster	1	0	0	0	0	70	0	0	0	0	0	0
5		2		3	10	7	0	0	2	36	1	9	0
6		3	0	0	0	0	99	0	0	0	0	0	0
7		4		0	0	0	189	0	0	0	0	0	0
8		5	0	0	0	0	94	0	0	2	0	0	0
9		6		0	0	0	0	0	0	0	0	131	0
10		7	0	0	0	0	0	211	0	0	0	0	0
11		8		0	0	0	0	140	0	0	0	0	0
12		9	1	0	0	9	0	116	0	0	0	0	0
13		10		0	0	0	0	214	0	0	0	0	0
14		11	0	0	0	0	0	301	0	0	0	0	0
15		12	3	34	4	0	0	0	36	0	0	0	0
16		13	0	0	0	0	0	173	0	0	0	4	0
17		14		0	0	0	0	90	0	0	2	3	0
18		15	2	2	22	8	108	0	2	5	0	0	2
19		16		0	0	0	373	0	0	0	0	0	0
20		17	0	0	0	0	0	0	0	30	12	21	0

Figure 5: Snapshot of cross tabulation of PECAS space type and five persons with incomes 100 to 150 (x1000 $)

Households with low income ($50,000 or less) are more likely to use *Single Family Detached Economy*, *Low Rise Economy, and Urban Mobile Home space types.* On the other hand, high income households ($100,000 or more) mainly use *Single Family Detached Luxury* or *Single Family Detached Economy space types.* Both low and high income households seldom use *High Rise Luxury* and *Economy* space types. Space types vary substantially across different occupations. For example, assembly, construction, service and transport workers generally live in *Low Rise Economy* and *Urban Mobile home;* managers, professional workers generally choose *Single Family Detached Luxury*, *Single Family Detached* and *Low Rise Luxury* residential types. After defining the consumption of space of each lifestyle, the use and make rates of household technology options were calculated and entered in the Technology Options table of the AA module. Fourteen types of use rate (see Table 2) were estimated from the cluster analysis as the mean rate of each lifestyle cluster. The make rates were calculated for eighteen occupation types as the mean rate of each lifestyle cluster.

In applying the clusters in the PECAS framework, it was discovered that the representation of the consumption of high-rise space was too limited – only a few household types consumed high-rise space because the system for separating high-rise multifamily from low-rise multifamily was based partly on household income. It was decided that, from a consumer's (residence) perspective in the PECAS AA module, and from the perspective of the PUMS data, the large-building multifamily options could be combined. The SD module of PECAS, which models developer behavior, will still favor high-rise over low-rise multifamily development when land is scarce.

5.CONCLUSION

This paper developed the framework for describing lifestyle clusters of household choices. The main aim has been to identify different household lifestyles with strong internal similarities in their labor force participation and housing use and apply them to the spatial economic model, PECAS.

The PECAS model represents location choices in two other levels of a nested-additive logit framework – residential location choice in the top level above lifestyle choice, and location for interactions (with regard to jobs, schools, shopping, etc.) in the lower level below lifestyle choice. This allowed us to focus on non-location choices, specifically the type and quantity of residential structure and the participation in different labor markets defined by occupation.

Larger households and higher income households have more variety in their labor force participation and housing consumption. These categories required more clusters to represent their flexibility to earn money and live in housing. Smaller and lower income households were more constrained in their choices, represented with less clusters. In policy analysis with the PECAS model, this representation will tend to show that policies leading to large variety of housing types and access to many jobs of different occupations will provide a large benefit to the larger and higher income household categories, while lower income and smaller household categories will be more affected by the wage and rent variables endogenously generated by the PECAS framework.

Clusters are often used to reduce a complex problem into a simpler representation. The number of resulting clusters in this analysis simplifies the labor and housing market representations in the model, which would otherwise have to consider every possible combination of housing/labor types with extreme endpoints to allow PECAS to interpolate to realistic options. However the number of clusters seems still too large for reporting purposes; analysts seem overwhelmed by the reports of technology choice produced by the PECAS model in operation. The clusters could be further grouped into “meta-clusters” for reporting purposes.

Households in PECAS consume goods, services and other interactions (non-priced items such as park visits or social visits), as well as housing. The PUMS data applied here contains little information on expenditures and travel, which should be interrelated with labor market decisions (more labor income allows higher expenditure) and housing decisions (e.g. larger housing units require more utility expenditures.) Other data sources could be explored to add expenditures to the cluster definition, but expenditure data that has been investigated is rarely linked with household housing and occupation data. A purpose-designed survey may be necessary to fully expand PECAS’s lifestyle representation to goods and services consumption.

The categorical definitions of houses based on income limits the options presented to the clustering algorithm: most (but not all) high income households produce a lot of labor and many consume large households. The option to reduce or eliminate household jobs, and earn and consume less, is not represented in these cluster options because households who have made such a choice would show up in a different household category. PECAS models could be designed so that household categories represent more intrinsic and permanent household variables, such as size and skill sets, instead of variables that represent, to some degree, the outcomes of household choices, such as income. However any categorical system will cause similar types of problems, a more satisfying approach would be to develop a microsimulation version of the PECAS AA module, based on the same utility functions, but treating each household based on its own characteristics instead of its definitional category. A microsimulation of households was once designed and implemented to work with PECAS (Hunt *et al*, 2004), but more recent research to ground the PECAS AA in random utility theory (Abraham and Hunt, 2007) suggests a more direct microsimulation version of PECAS would be a better way to avoid the shackles of categorization.

Acknowledgement

We are grateful for the funding and facilities for this study provided by the California Department of Transportation and the Urban Land Use and Transportation Center at the University of California, Davis. Thoughtful information and valuable work were provided by many colleagues at the Urban Land Use and Transportation Center and HBA Specto Inc. of Calgary, Canada.

REFERENCES:

Abraham, J.E., J.D. Hunt (2007) Random utility location/production/exchange choice, the additive logit model, and spatial choice microsimulations, **Transportation Research Record** 2003:1-6

Ahmed N (1994) A joint model of tenure choice and dem and for housing in the city of Karachi, **Urban Studies** 31. 1691-1706

Arnold JE, Ford A (1980) A statistical examination of settlement patterns at Tikal, Guatemala. **American Antiquity** 45, 713-726

Bagley MN, Mokhtarian PL (1999) The role of lifestyle and attitudinal characteristics in residential neighborhood choice. In: Ceder A (ed) Transportation and traffic theory: **Proceedings of the 14th International Symposium on Transportation and Traffic Theory**. Pergamon Press, Oxford pp 735–758

Bawden G (1982) Community organization reflected by the household: a study of Pre-Columbian social dynamics, **Journal of Field Archaeology** 9, 165-181

Ben-Akiva, M Bowman JL, (1995) Integration of activity-based model system and a residential location model, **Urban Studies** 35 (7), 1131-1153

Ben-Akiva M, Bowman JL, Gopinath D (1996) Travel demand model system for the information era. **Transportation** 23:241–266

Bhat C, Guo J (2006) A comprehensive analysis of built environment characteristics on household residential choice and auto ownership levels. In: Presented at the **85th Annual Meeting of the Transportation Research Board**, Washington DC

Börsch-Supan A, Pitkin J (1988) On discrete choice models of housing demand. ***J. Urban Econ****.* 24. 153-172

Cao X, Mokhtarian P (2005) How do individuals adapt their personal travel? Objective and subjective influences on the consideration of travel-related strategies for San Francisco Bay Area commuters. **Transport Policy** 12:291–302

Cho C (1997) Joint choice of tenure and dwelling type: Multinomial logit analysis for the city of Chongju, **Urban studies** 34 (9) 1459-1473

Daly A (1982) Estimating Choice Models Containing Attraction Variables, **Transportation Research Part B**, 16(1), pp 5-15

Duchin F (1998) Structural Economics: Measuring Change in Technology, **Lifestyles and Environment,** (Washington DC, Island Press)

Fischer MM, Aufhauser E. (1988) Housing choice in a regulated market: a nested multinomial logit analysis, **Geographical Analysis** , 20, pp. 47-69

Hayden B, Cannon A (1983) Where the garbage goes: reuse disposal in the Maya Higlands, **Journal of Anthropogical Archaeology** 2, 117-163

Hunt, J.D. and J.E. Abraham (2005) Design and implementation of PECAS: A generalized system for the allocation of economic production, exchange and consumption quantities. In: **Foundations of Integrated Land-Use And Transportation Models: Assumptions and New Conceptual Frameworks**, Lee Gosselin and Doherty editors, Elsevier

Hunt, J.D., J.E. Abraham and T.J. Weidner (2004) Household Allocation Module of Oregon2. **Transportation Research Record** 1898:98-107

Kim SJ. (1992) A model of rental housing choices in the Korean market, **Urban Studies** 29 1247-1264.

Kohlhase, Janet E (1986) Labor Supply and Housing Demand for One- and Two-Earner Households. **The Review of Economics and Statistics**, Vol. 68, No. 1 (Feb., 1986), 48-57

Krizek K, Waddell P (2003) Analysis of lifestyles choices: neighborhood type, travel patterns, and activity participation. **Transp Res Rec** 1807:119–128

Lazer, W (1963) Lifestyle Concepts and Marketing. **Toward Scientific Marketing.S Greyser**. Chicago, American Marketing Association

Moekel R, Spiekermann K, Schurmann C, Wegener M (2003) Microsimulation of land use. Int J **Urban Sci** 7(1):14–31

Netting, Robert McC (1982) Some home truths on household size and wealth. **American Behavioral Scientist** 25. 641-662

Quigley, JM. (1976) Housing demand in the short run: an analysis of polytomous choice, **Exploratio n in Econom ic Research** , 3. 76-102

Salomon I, Ben-Akiva M(1983) The use of the lifestyle concept in travel demand models. **Environ Plann A** 15:623–638

Waddell P (2000) Towards a behavioral integration of land use and transportation modeling. In: **Proceedings of the 9th International Association for Travel Behavior Research Conference**, Queensland, Australia

Walker J, Li J (2007) Latent lifestyle preferences and household location decisions, J **Geograph Syst** 9 77-101