Explanatory analysis of weekday and weekend daily physical-virtual activity and travel patterns of individuals in Cairo, Egypt

Ahmed Ibrahem MOSA, Assistant Professor, ITS program, Nile University

Adel Sayed Abd El Maksoud, Associate Professor, Egyptian National Institute of Transportation

Ali Salam Heikal, Professor, Ain-Shams University, Faculty of Engineering, Civil Engineering Department

ABSTRACT

Activity-based travel demand forecasting is increasingly becoming a mainstream approach to travel demand modelling and has received substantial attention in the past three decades. At the core of activity analysis there are issues related to time allocation by individuals and the role time play in organizing activities within specific periods. In the past, this stream of work attempted to summarize human behaviour in terms of representative behaviour of several relatively homogeneous groups using pattern recognition techniques. Understanding activity engagement and time use patterns, however, may require us to examining the effects of correlates on time allocation behaviour, habit persistence, history dependence, and the impact of changes in the correlates. While most of the studies have focused on weekday activities, out-of-home activities, and physical activity and travel patterns of individuals in developed countries, the aim of this paper is to contribute to the understanding of daily activity and time allocation patterns of individuals in developing country. In this paper, we explore the complex relationships among virtual-physical activity participation and travel patterns of individuals in developing country. Using two-day activity and telecommunication diary survey conducted in Cairo Region, in Egypt, individuals' daily time allocation for virtual-physical activities and travel are analyzed. Essentially, we have adopted two approaches; the first approach is to identify relatively homogeneous behavioural groups using cluster analysis. This is done to reduce the great diversity in individuals' behaviour into a few reprehensive patterns of behaviour using activity and travel indicators. In addition, cross-sectional homogeneity of time allocation between different social groups is analyzed. The second approach proposes a mixed multinomial logit model. The model aims to investigate the influential factors that may govern the selection mechanism of daily activity patterns. The analysis showed strong dependence between weekday and weekend day for both activity and travel patterns.

Keywords: Activity-based modelling, time use, virtual and physical activity participation, mixed logit models.

INTRODUCTION

At the core of activity analysis are issues related to time allocation by individuals and the role time plays in organizing activities within specific periods. In the past, this stream of work attempted to summarize human behaviour in terms of representative behaviour of several relatively homogeneous groups using statistical-mathematical pattern recognition techniques (Pas, 1982, 1983; Reeker and McNally, 1986 a,b).

A subsequent step along this line was to study the temporal state dependence in activity engagement and time use patterns. There have been several studies examining the role of state dependence in activity and travel behaviour from temporal perspectives. Early applications include (Pas, 1987, 1988) a state dependent micro-simulation model system in Goulias and Kitamura (1997), and Goulias and Kim (2001), who studied the role of history dependency in explaining activity-travel patterns of travellers.

These studies confirm that the choice of activity and the amount of time allocated to an activity is dependent on the preceding activity engagement and time allocation pattern, and the extent of the dependency is likely to reduce as one proceeds back in time within a one- or two-day time frame. Ma and Goulias (1997a, b) develop a typology for studies in the area of stationary and dynamics analysis of personal activity-travel behaviour patterns. From behaviour stand point, he identifies changes in person's activity and travel behaviour patterns as attributed to two factors; cyclical behaviour (schedule that repeats itself over longer periods) or random variation. In terms of time horizon, Ma classifies changes of activity participation and related mobility as being very short- term (within a day), short-term (from one day to another), medium- term (from one week to another), long term (from one year to another), and very long- term (across multi-year intervals).

Understanding activity engagement and time use patterns, however, requires us to examine the effects of correlates on time allocation behaviour, habit persistence and history dependence, and the impact of changes in the correlates. One very important correlate to time allocation behaviour is the growing Information and Communications Technology (ICT). The impacts of ICT adoption and use are likely to be far-reaching, with the potential to fundamentally alter the life styles of businesses and individuals (see, for example, Droege, 1997, Graham and Marvin, 1996, Boden, 1999, French et al., 1999 and Zimmerman et al., 2001, for broad discussions of the potential influence of ICTs on urban planning, medical care and services, education, and family responsibilities). An interesting aspect of ICT use from a transportation perspective is its impact on personal activity-travel behaviour. Three significant effects of ICT's on activity and travel patterns have been postulated in prior research: substitution, generation, and modification (Mokhtarian, 1990; Salomon, 1998, 2000). Substitution and modification of activity and travel patterns can have a significant impact on transportation system performance. For instance, recent advances in information and communication technologies (ICTs) make it possible to conduct activities virtually, thus obviating the need for physical travel, at least for some types of activities. Further, ICT use may contribute towards reducing urban congestion and air-quality problems (by replacing travel with virtual activities); on the other hand they may also generate significant additional and induced travel due to increased connectivity and access to resources.

While most of the studies have focused on weekday activities, out-of-home activities, and physical activity and travel patterns using data from Anglo-Saxon countries. The aim of this paper is to contribute to the understanding of daily activity participation and time allocation patterns of individuals in Arabic countries by considering the following objectives:

1) To examine the daily time allocation to physical as well as virtual activities and travel,
2) To examine the daily in-home and out-of-home time allocation to maintenance and recreational activities,
3) To investigate the factors that may govern the daily activity pattern choice.

Essentially, we have adopted two approaches; the first approach is to identify relatively homogeneous behavioural groups using cluster analysis. This is done to reduce the great diversity in individuals' behaviour into a few reprehensive patterns of behaviour using activity indicators. In the second approach, this paper proposes a mixed multinomial logit model that accommodates heterogeneity across individuals and incorporates heteroscedasticity among daily activity patterns. The model aims to investigate the influential factors that may govern the selection mechanism of daily activity patterns. Investigation along these dimensions will highlight the way individual's allocates their time with, which in turn enables us to understand better the causal factors affects individuals' daily activity and travel participation.

The rest of this paper is organized as follows. First, the data used are described. Then cluster analysis is used as a pattern recognition technique. This is followed by describing of the econometric model structure and the empirical results. Finally, a later section of the paper presents a summary of the research effort and identifies the major conclusions.

DATA SOURCES

The primary data source used for this analysis is the activity-travel and telecommunication diary survey spanned the period from December 2005 to January 2006. The survey was administrated in three academic and research institutions in Cairo, Egypt. Particularly; Egyptian national institute of transportation, Ain-shams University, and the Information and Technology Institute (ITI). Respondents from the three locations were contacted first by face to face interview to solicit their participations. Respondents who agreed to participate received the relevant activity diary sheets. In addition, a comprehensive explanations of the diary sheets and how can they enter the diary were done Further, they were requested to arrange that all of their household members above 12 years old to log their activities for two days. The response rate was 75.2% (of the 270 households to which questionnaires were distributed).

The activity diary instrument consisted of 4 questionnaires parts, and a core diaries part. The first two parts concerned with various individual and household characteristics, the third part of the survey instrument aimed to collect general data about the physical activities of the respondents, the fourth part of the instrument concerned with daily and non-daily Internet use and availability. The core part of the instrument consists of two diaries; the physical and Internet activities diary; and the telecommunication diary. The physical and Internet activities diary, aimed to collect data about participants' all in-home and out-of-home activities. For each successive activity, respondents were asked to provide information about the type of activity (based on 42 pre-coded scheme of activities), the start and end times of activities (beginning with 3 a.m. on the first diary sheet and ending with 3 a.m. following the second diary sheet), location of participation (respondents were asked to record the location where the activity took place based on 12 pre-coded activity locations, i.e., in-home, shopping centres, work place, school, restaurant, relative/friends house, etc.) and exact geographical location for out-of-home activities. Further, for each successive activity, respondents were asked to report "with whom" they were doing this activity and "for whom" they did it. Respondents were provided with eleven pre-coded social contacts categories. The category reflects the most important social networks of immediate family, relatives and friendships.

The information gathered on travel episodes included information on travel mode used, transfer location, and travel time.

Respondents were asked to complete the diaries for two consecutive days (starting either on a Saturday or a Thursday, so that data would be obtained for one weekend day and one weekday)[1]. After data cleaning and verification the final sample from the survey consisted of 459 respondents belong to 150 households, and contain a total of 15935 weekday and weekend activity and travel episodes. To our knowledge, this is the first attempt to conduct an activity diary and telecommunication survey in Cairo, Egypt.

A secondary data source used in the analysis is a zonal-level land-use, transport level of service and demographic data obtained from the Greater Cairo Transportation Master Plan Study. The secondary data provides the following information for each Traffic Analysis Zones (TAZ): (a) total employment and employment levels disaggregated by sector, (b) zonal population, income and age distribution of the population, (c) the area type of the zone (CBD zone, an urban zone, a suburban zone, or a rural zone), and (d) transportation system level of service in terms of average travel, waiting, and access time for bus and transit services. This information was used to study the impact of the characteristics of the residence zone on household interactions in daily travel patterns.

CLASSIFICATIONS OF DAILY ACTIVITY AND TRAVEL PATTERNS

Pattern Recognition with Cluster Analysis

The objective of pattern recognition is to search for homogeneous groups of observations (persons in this study) with respect to some variables of interest using cluster analysis. In general, cluster analysis seeks to identify a set of groups or cases on the bases of their "nearness". The measures most commonly used are distance and similarity. Distance measures how far apart two observations are, while similarity measures closeness of two observations.

The clustering technique used here is two step clustering algorithm. This algorithm has two steps 1) Pre-cluster the cases into many small sub-clusters which are treated as single cases; 2) Cluster the sub-clusters resulting from pre-cluster step into the desired number of clusters. The pre-cluster step uses a sequential clustering approach. It scans the data records one by one and decides if the current record should be merged with the previously formed clusters or starts a new cluster based on the distance criterion. As the clustering process going on, cases (persons in this study) are classified based on nearest centroid sorting algorithm, i.e., a case is assigned to the cluster for which the distance between the case and the center of cluster (centroid) is smallest. The estimation is based on a strategy, which chooses cases that have large distances between them and uses their values as initial estimates of the cluster centers. As subsequent cases are processed, a case replaces a center if its smallest distance to center is greater than the distance between the two closest centers. The distance here is commonly used log-likelihood distance. The log-likelihood distance is a probability based distance. The distance between two clusters is related to the decrease in log-likelihood as they are combined into one cluster.

[1] The weekend in Egypt is Friday and Saturday. Sunday is an ordinary workday.

Pas (1982) provide a guideline to choose the optimal number of clusters. Using cluster analysis to classify daily travel-activity behaviour according to out-of-home activities and five periods in a day, he argued that, when the number of clusters is less than five, the amount of information explained is substantially increased by increasing the number of clusters. However, when the number of clusters is more than 12, each additional cluster accounts for only small amount of information. His findings provide some guidance of the number of clusters to use in our study.

There are thus tradeoffs to consider in selecting the number of clusters. the smaller the number of clusters, the greater the heterogeneity within the cluster-that is, more dissimilar types are being combined into a group whose average may be relatively central but conceals a lot of variation. On the other hand, the larger the number of clusters, the greater the homogeneity within clusters and the more “types” that can be identified. However, comparing characteristics across a large number of relatively small clusters is both more difficult cognitively and less reliable statistically. Thus it is desirable to review a range of cluster solutions on either side of a potential optimum, in order to identify that optimum number of clusters.

In this study the first stage in the cluster analysis was to obtain quick cluster results for sets of 5-10 clusters. These sets are then analyzed to determine the preferred solution, that is, how many clusters are best. The decision is made based on how similar clusters are to each other, how many respondents are within each cluster, and whether the clusters identify interesting subsets of the sample. In addition to the previous step, one way analysis of variance (ANOVA) is used to evaluate the result of clustering and the strength of variables in discriminating cases among clusters.

Classifications of Virtual-Physical Activities and Travel Patterns

In deriving the first daily activity and travel patterns the variables used are:

1) Total amount of time allocated to travel in weekday and weekend (T_dur1,T_dur2),
2) Total amount of time allocated to physical maintenance activity in weekday and weekend (T_phy_main_dur1,T_phy_main_dur2),
3) Total amount of time allocated to virtual maintenance activity in weekday and weekend (T_virtu_main_dur1,T_Vitu_main_dur2),
4) Total amount of time allocated to physical recreational activity in weekday and weekend (T_phy_rec_dur1,T_phy_rec_dur2),
5) Total amount of time allocated to virtual recreational activity in weekday and weekend (T_virtu_rec_dur1, T_virtu_rec_dur2).

A 5-cluster solution is obtained for each of the two days as shown in Table 1. The F ratio of ANOVA indicate that all the variables are significant discriminators for the five groups and the identified groups significantly differ from one another, implying an effective clustering of duration in differentiating among individuals and classifying them into clusters. In the same way, the Bayes information criterion (BIC), Akaike information criterion (AIC) for each number of clusters are computed and used to find the estimate of the number of clusters. The lower the BIC or AIC values, the better the cluster we estimate.

As shown in Table 1 we have five type of activity-travel behaviour with wide variability in time allocation among clusters as desired to identify different behaviours and there correlates.

Table 1- Average Profile of Daily Virtual-Physical Activities and Travel Patterns

Variables used to create clusters		Cluster 1	Cluster 2	Cluster 3	Cluster 4	Cluster 5
Cluster size (%)		31.699	23.638	18.736	18.736	7.190
Weekday	T_dur1	129.072	59.740	98.144	23.269	94.737
	T_phy_m_dur1 (min.)	64.710	410.677	90.464	163.462	148.947
	T_virtu_m_dur1 (min.)	0.475	0.000	0.000	0.000	100.263
	T_phy_rec_dur1 (min.)	307.661	409.844	252.113	684.038	341.842
	T_virtu_rec_dur1 (min.)	34.253	15.625	202.680	29.038	126.053
Weekend	T_dur2	145.929	66.488	48.200	21.712	81.277
	T_phy_m_dur2 (min.)	104.929	346.570	106.200	124.212	188.723
	T_virtu_m_dur2 (min.)	0.429	0.000	0.533	0.205	123.511
	T_phy_rec_dur2 (min.)	458.000	459.884	383.493	623.459	432.511
	T_virtu_rec_dur2 (min.)	35.429	16.364	225.067	43.288	105.957

The first and largest cluster contains 31.699% of the sample has the longest travel times for both two days (about two hours per day) and the shortest time allocated to physical maintenance activities in both days. The persons in this group, however, allocate a good portion of their time on both physical and virtual recreational activities in both days. The second group with substantial size (23.638% of the sample) has the longest physical maintenance activities duration for both days (about six hours per day).Also, it has the second highest physical recreational duration in both two days. Virtual activity duration is also so low. The third cluster with 18.736% of the sample has the longest time allocated to virtual recreational activities and the lowest time allocated to physical recreational activity in both two days. In contrast, cluster 4 of equal size (18.736% of sample) is characterized by the longest time allocated for physical recreational activity and shortest travel duration in both days. The last cluster (7.190% of the sample) has the longest time allocated to virtual maintenance activity. Persons with this pattern are also likely to allocate highest amount of time to virtual recreational activities. In this group physical activity durations are of “average” daily duration. This indicates that people in this group participate in many short duration activities in a day.

Activity-Travel Cluster Composition

The results from the previous analyses suggest typical patterns in time allocation related to various activity and travel activities. These analyses, however, do not show whether a certain patterns are more associated to certain type of individuals. In this section the similarities and differences between the five cluster groups, presented previously are analyzed in terms of the socio-demographic characteristics of the sample. These included household socio-demographics and individual socio-demographics.

The individual socio-demographics explored included gender (male, female), age (young, middle, old), and license holding to drive. The household socio-demographic characteristics considered included household income, and number of cars owned by the household.

Individual socio-demographics

Table 2 shows the results of pattern compositions in terms of males and females by cluster for both weekday and weekend. The value of the contingency coefficients for both days (C1= 0.445, C2=0.401) indicates a strong and significant dependency between daily patterns and gender. As expected, males are found in high proportion in the cluster 1 (Traveller group), while cluster 2 (with the longest physical maintenance activities duration for both days) is mostly populated by females, which may be partly due to the greater degree of task allocation and responsibility for child care by female members inside the household. This is also a manifestation of traditional gender role. Interestingly, a substantial amount of females appear to belong to cluster 3 and allocate longer time pursuing virtual recreation activities. On the other hand, males show a clear majority in cluster4 with longer time allocated to physical recreational activity, which again is most likely motivated by the gender role issues.

Table 2: Gender by Daily Virtual-Physical Activities and Travel Patterns

Day	Cluster		Gender		Grand total
			Male	Female	
Weekday	Cluster 1	Count	154	67	221
		%	67.54	29.00	48.15
	Cluster 2	Count	5	91	96
		%	2.19	39.39	20.92
	Cluster 3	Count	46	51	97
		%	20.18	22.08	21.13
	Cluster 4	Count	11	15	26
		%	4.82	6.49	5.66
	Cluster 5	Count	12	7	19
		%	5.26	3.03	4.14
	Total Count		228	231	459
	Total %		100.00	100.00	100.00
Chi-square ($\chi 2$)		113.46			
Contingency coefficient (*C1*)		0.445			
Weekend	Cluster 1	Count	60	10	70
		%	26.32	4.33	15.25
	Cluster 2	Count	22	99	121
		%	9.65	42.86	26.36
	Cluster 3	Count	40	35	75
		%	17.54	15.15	16.34
	Cluster 4	Count	77	69	146
		%	33.77	29.87	31.81
	Cluster 5	Count	29	18	47
		%	12.72	7.79	10.24
	Total Count		228	231	459
	Total %		100.00	100.00	100.00
Chi-square ($\chi 2$)		88.044			
Contingency coefficient (*C2*)		0.401			

Table 3 shows the daily pattern compositions by age. The value of the contingency coefficients for both days (C1= 0.354, C2=0.318) indicates a strong and significant dependency between daily patterns and age for both days. Interestingly, the age groups young and elderly show clearer majority in cluster 1 (about 2 hours allocate to travel per day). This could be attributed to job location inflexibility with starting career, or not being able to afford housing near the job locations, in the case of young individuals, and relatively more free time in the case of older individuals. A substantial amount of young group appears to belong to cluster 3, with higher amount of time spent in virtual recreational activities (4-5 hours per day), which can be expected due to familiarity with ICT services. Also, it indicates that an individual who has an active lifestyle may be more likely to pursue virtual activity, as well as be more likely to generate travel. Among different age groups, middle age individual are the most likely to belong to cluster2 (with more than five hours allocated to maintenance activities per day), perhaps because greater household responsibilities for middle age individuals.

Table 3- Age by Daily Virtual-Physical Activities and Travel Patterns

Day	Cluster		Age (years)			Grand total
			Young (12-29)	Middle (30-49)	Old (50 -60 or older)	
Weekday	Cluster 1	Count	117	39	65	221
		%	48.35	44.83	50.39	48.25
	Cluster 2	Count	33	28	34	95
		%	14	32	26	20.74
	Cluster 3	Count	78	13	6	97
		%	32	15	5	21.18
	Cluster 4	Count	10	1	15	26
		%	4	1	12	5.68
	Cluster 5	Count	4	6	9	19
		%	2	7	7	4.15
Total Count			242	87	129	458
Total %			100.00	100.00	100.00	100.00
Chi-square ($\chi 2$)			65.781			
Contingency coefficient (*C1*)			0.354			
Weekend	Cluster 1	Count	43	17	10	70
		%	17.70	19.54	7.75	15.25
	Cluster 2	Count	52	29	40	121
		%	21.40	33.33	31.01	26.36
	Cluster 3	Count	61	9	5	75
		%	25.10	10.34	3.88	16.34
	Cluster 4	Count	60	25	61	146
		%	24.69	28.74	47.29	31.81
	Cluster 5	Count	27	7	13	47
		%	11.11	8.05	10.08	10.24
Total Count			243	87	129	459
Total %			100.00	100.00	100.00	100.00
Chi-square ($\chi 2$)			51.724			
Contingency coefficient (*C2*)			0.318			

The daily pattern composition by driving license holding is interesting. As shown in Table 4 a strong relationship is observed between license holding and daily activity patterns of individuals with contingency coefficients (C1=0.304 , C2=0.318). Both clusters 3 and 5 are populated with individuals hold a driving license. This reflects a strong relation between mobility and connectivity (use internet).

Table 4- License Holding by Daily Virtual-Physical Activities and Travel Patterns

Day	Cluster		License holding		Grand total
			Yes	No	
Weekday	Cluster 1	Count	73	148	221
		%	54.48	45.54	48.15
	Cluster 2	Count	7	89	96
		%	5.22	27.38	20.92
	Cluster 3	Count	38	59	97
		%	28.36	18.15	21.13
	Cluster 4	Count	3	23	26
		%	2.24	7.08	5.66
	Cluster 5	Count	13	6	19
		%	9.70	1.85	4.14
Total Count			134	325	459
Total %			100.00	100.00	100.00
Chi-square (*χ2*)			46.593		
Contingency coefficient (*C1*)			0.304		
Weekend	Cluster 1	Count	19.000	51	70
		%	14.179	15.69	15.25
	Cluster 2	Count	34	87	121
		%	25.37	26.77	26.36
	Cluster 3	Count	21	54	75
		%	15.67	16.62	16.34
	Cluster 4	Count	26	120	146
		%	19.40	36.92	31.81
	Cluster 5	Count	34	13	47
		%	25.37	4.00	10.24
Total Count			134	325	459
Total %			100.00	100.00	100.00
Chi-square (*χ2*)			51.748		
Contingency coefficient (*C2*)			0.318		

Household socio-demographics

Table 5 shows the daily pattern compositions by household income. The value of the contingency coefficients for both days (C1=0.296, C2=0.378) indicates a strong and significant dependency between daily patterns and household income for both days. The traveller cluster (cluster 1) has the largest percentages of individuals from a household with higher or medium incomes than individuals from low income category. This suggests that individuals in higher income households can afford to do the travel due to the greater availability of disposal income and mobility resources. The low income group, on the other hand, has the highest percentages in both clusters 2 and 4 (the physical activity clusters), which may be attributable to their insufficient resources and mobility. As expected, both cluster 3 and 4 (the virtual activity clusters) are populated with individuals from higher income households, which could also be attributed to the availability of resources.

The daily activity and travel patterns compositions by household car ownership are shown in Table 6. The value of the contingency coefficients for both days (C1=0.280, C2=0.394) indicates a significant dependency between daily patterns and household car ownership especially for weekend. As expected, both cluster 3 and 4 (the virtual activity clusters) are populated with individuals from households with more than one car owned, which is similar to the findings reported previously.

Classification of Daily in-home and Out-of-home Activity Patterns

In the previous section, using the total time allocated to physical, virtual and travel activities, daily activity and travel patterns of individuals are developed. This however, is unsatisfying since it necessarily ignores the behaviour mechanism in selecting the daily activity and travel patterns. In this section we turn our focus on describing and predicting the propensity that individuals will choose a particular daily activity pattern (i.e., clusters derived from previous section). Furthermore, we hypothesis that an individual's select a specific daily activity pattern based on the time he allocates for in-home as well as out-of home activities. In other words, individuals allocate their time on different virtual, physical, and travel activities based on the time they allocate on in and out-of-homes activities.

In this section, therefore, we turn our focus on describing and predicting the propensity that individuals will choose a particular daily in-home and out-of-home activity pattern. Using the same methodology described previously, we derived new patterns considering the total time individuals allocate to maintenance and recreational activities, for in-home as out-of home activities. In deriving the clusters the variables used are:

1) Total in-home maintenance activity duration in weekday and weekend (H_m_dur1,H_m_dur2)
2) Total out-home maintenance activity duration in weekday and weekend (Out_m_dur1,Out_m_dur2)
3) Total in-home recreational activity duration in weekday and weekend (H_r_dur1,H_r_dur2)
4) Total out-home recreational activity duration in weekday and weekend (Out_r_dur1, Out_r_dur2).

Table 7 describes the four clusters in terms of the person's behaviour with each cluster. The first cluster behaviour and the largest (42.4% of sample) is characterized by the shortest maintenance out-of-home activity duration for both weekday and weekend. The persons in this group, however, allocate a good portion of their time on in-home recreational activities. The second group with substantial size (22.2 % of the sample) has the longest recreational activities duration at home for both two days and shows a small amount of time that is allocated on out-home recreational activities in both two days. The third cluster, containing 17.86% of the sample, has the longest out-home recreational duration for both days. Persons with this pattern are also likely to allocate highest amount of time in out-of-home maintenance activities during the weekday. The fourth cluster with 17.5% of the sample is characterized by their home bond activity behaviour in both days. Persons in this group allocate the majority of time on in-home maintenance and recreational activities (longest in-home maintenance activities duration in both days) and a very short time on out-of-home maintenance and recreational activities.

Table 5- Household Income by Daily Virtual-Physical Activities and Travel Patterns

Day	**Cluster**		**Household Income**			**Grand total**
			Low income	**Middle income**	**High income**	
Weekday	Cluster 1	Count	43	119	59	221
		%	44.33	50.42	46.83	48.15
	Cluster 2	Count	36	47	13	96
		%	37.11	19.92	10.32	20.92
	Cluster 3	Count	11	45	41	97
		%	11.34	19.07	32.54	21.13
	Cluster 4	Count	7	16	3	26
		%	7.22	6.78	2.38	5.66
	Cluster 5	Count	0	9	10	19
		%	0.00	3.81	7.94	4.14
Total Count			97	236	126	459
Total %			100.00	100.00	100.00	100.00
Chi-square (*χ2*)			44.070			
Contingency coefficient (*C1*)			0.296			
Weekend	Cluster 1	Count	14	38	18	70
		%	14.43	16.10	14.29	15.25
	Cluster 2	Count	40	47	34	121
		%	41.24	19.92	26.98	26.36
	Cluster 3	Count	10	43	22	75
		%	10.31	18.22	17.46	16.34
	Cluster 4	Count	33	94	19	146
		%	34.02	39.83	15.08	31.81
	Cluster 5	Count	0	14	33	47
		%	0.00	5.93	26.19	10.24
Total Count			97	236	126	459
Total %			100.00	100.00	100.00	100.00
Chi-square (*χ2*)			76.405			
Contingency coefficient (*C2*)			0.378			

Table 6- Number of Cars Owned by Daily Virtual-Physical Activities and Travel Patterns

Day	Cluster		Number of Cars Owned				Grand total
			0-car	1-car	2-car	3 or more cars	
Weekday	Cluster 1	Count	115	71	23	12	221
		%	52.04	39.23	63.89	57.14	48.15
	Cluster 2	Count	59	33	3	1	96
		%	26.70	18.23	8.33	4.76	20.92
	Cluster 3	Count	33	50	8	6	97
		%	14.93	27.62	22.22	28.57	21.13
	Cluster 4	Count	12	13	1	0	26
		%	5.43	7.18	2.78	0.00	5.66
	Cluster 5	Count	2	14	1	2	19
		%	0.90	7.73	2.78	9.52	4.14
Total Count			221	181	36	21	459
Total %			100.00	100.00	100.00	100.00	100.00
Chi-square ($\chi 2$)			38.938				
Contingency coefficient (C)			0.280				
Weekend	Cluster 1	Count	36	22	9	3	70
		%	16.29	12.15	25.00	14.29	15.25
	Cluster 2	Count	61	44	10	6	121
		%	27.60	24.31	27.78	28.57	26.36
	Cluster 3	Count	32	37	3	3	75
		%	14.48	20.44	8.33	14.29	16.34
	Cluster 4	Count	92	49	5	0	146
		%	41.63	27.07	13.89	0.00	31.81
	Cluster 5	Count	0	29	9	9	47
		%	0.00	16.02	25.00	42.86	10.24
Total Count			221	181	36	21	459
Total %			100.00	100.00	100.00	100.00	100.00
Chi-square ($\chi 2$)			84.144				
Contingency coefficient (C)			0.394				

Table 7- Average Profile of Daily In-home and Out-home Activity Patterns

Variables used to create clusters		Cluster 1	Cluster 2	Cluster 3	Cluster 4
Cluster size (%)		42.37	22.22	17.86	17.54
Weekday	H_m_dur1 (min.)	54.51	121.44	90.34	408.10
	Out_m_dur1 (min.)	13.27	90.25	92.07	22.53
	H_r_dur1 (min.)	328.95	595.51	290.48	419.62
	Out_r_dur1 (min.)	26.87	12.37	193.62	12.66
Weekend	H_m_dur2 (min.)	88.87	103.34	102.04	398.23
	Out_m_dur2 (min.)	11.03	47.41	80.78	15.49
	H_r_dur2 (min.)	413.27	636.14	272.26	426.34
	Out_r_dur2 (min.)	95.77	29.52	287.71	48.88

A MIXED LOGIT MODEL OF DAILY IN-HOME AND OUT-OF-HOME ACTIVITY PATTERNS CHOICE

Model Structure

In this section the focus is on describing the propensity that individuals will choice a particular daily in-home or out-of-home activity pattern. For this purpose a mixed multinomial logit model is applied (Train, 1998; Bhat, 1998). The formulation accommodates heterogeneity (*i.e.*, taste variation) across individuals due to both observed and unobserved individual attributes. Correlation in common unobserved factors influencing the choice of daily activity pattern are also considered. Based on Train (1998) the formal model can be described as follows. Given that an individual *q* chooses among *i* possible daily activity patterns on choice occasion *t*. The utility *Uqit* that an individual *q* (*q*=1,2,…,*Q*) associates with an alternative *i* (*i*=1,2,…,I) on choice occasion *t* (*t=1,2,…,Tq*) may be written as:

$$U_{qit} = \beta' q x_{qit} + \varepsilon_{qit}$$

Where x_{qit} is a vector of observed variables affecting individual *q* for alternative i at *t*th choice occasion, β_q is a corresponding column vector of coefficients that is unobserved for each *q* and varies randomly over the individuals representing each individuals' preference, and ε_{qit} is an unobserved random term that is distributed iid extreme value, independent of β_q and x_{qit} . The coefficients β_q are allowed to vary across individuals and this variance induces correlation in utility over daily patterns and choice occasions. Each random parameter β_q is defined as the average preference in the population, b' and an individual deviation, η'_q which represent the individuals' preference relative to the average preference for a particular daily pattern. The utility is:

$$U_{qit} = b' x_{qit} + \eta'_q \, x_{qit} + \varepsilon_{qit}$$

The unobserved part of the utility is $\eta' q x_{qit} + \varepsilon_{qit}$, and this term is correlated over daily patterns and choice occasions. If β_q were known to take the value of β , the probability that individual *q* chooses daily pattern *i* at choice occasion *t* would be standard logit:

$$L_{qit} = \frac{e^{\beta' x_{qit}}}{\sum_{i=1}^{I} \beta' x_{ijt}}$$

However, the preferences of the individuals are not known. Assume that the preferences vary in the population with the density *f (β/θ*)*, where *θ** are the parameters of the distribution. The probability for the individual is the integral over all possible values of β , which depends on the parameters of the distribution of β . the actual probability for individuals' choice of daily pattern is:

$$Q_{qit}(\theta_{qit}(\beta) f(\beta / \theta^*)) d\beta$$

Let *i(q,t)* indicate the daily pattern that individual *q* chose at choice occasion *t*, and if $\beta_q = \beta$ then the probability of individual *q's* observed sequence of choices is:

$$Sq(\beta) = \prod_t Lqi(q,t)t(\beta)$$

The actual probability since β_q is not known, is the integral of the above equation over all values of β:

$$P_q(\theta^*) = \int S_q(\beta) f(\beta/\theta^*)d\beta$$

The log likelihood function is:

$$LL(\theta) = \sum_q \ln P_q(\theta)$$

The integral in the above equation has no closed form and therefore, exact maximum likelihood estimation is not possible. However, it can be approximated by simulation and the result is taken as the approximate choice probability:

$$SPq(\theta) = (1/R) \sum_{r=1,\dots R} Sq(\beta^{r/\theta})$$

Where R is the number of draws in the simulation, and $SPq(\theta)$ is the simulated probability of individual *q*'s sequence of choices.. The simulated log-likelihood function is:

$$SLL(\theta) = \sum_i \ln(SPq(\theta))$$

The estimated parameters are those that maximize SLL. The bias in SLL decreases as the number of draws increase.

Variable Specification and Overall Measures of Fit

Simulated maximum likely-hood estimation, using Halton draws, was used to estimate the parameter of the model. The estimations in the paper were carried out using Limdep Econometric Software version 8, Hensher et al. (2005). The data for estimation were prepared as follows. The dependent variable was the choice of daily pattern membership (cluster number). For each individual eight pattern types (four for each day as presented previously) were considered to constitute the choice set. The best fitting model is shown in Table 8 with cluster 1 served as the reference group. Its log- likelihood is -1049.522 (with 27 degree of freedom) against a value of -1271.232 for the market share model. Additionally, one can observe consistency in sign of coefficients across the model. All parameters of the model are statistically significant at 90% confidence level or better. The signs of all utility parameters seem to be correct and unambiguous.

Model Results and Discussion

The results provide significant evidence on the effect of different variables on propensity to choice a particular daily activity pattern. Household socio-demographics contribute greatly to determining individuals' activity patterns in various ways. As expected, individuals from a household with low or middle income are more likely to belong to cluster 4. Household head, couple household, and individuals from large households are unlikely to belong to cluster 4 compared to other household types. On the other hand, the results indicate that rented apartment dwellers are more likely to belong to cluster 4 than individuals in other dwelling units. Regarding the individual socio-demographic and employment characteristics, the results indicate that middle age individual are more likely to belong to cluster 4. The availability of driving license has a positive effect on belonging to cluster 3. Employment or students with more than 8 hours working or study per day are less likely to belong to cluster 4. The results suggest a higher propensity to belong to cluster 3 for households located in zones with more diverse in economical activities. Individuals residing in poor transit network zones are more likely to belong to cluster 4.

There is also evidence of a substantial amount of heterogeneity in the propensity to participate in a particular daily activity pattern, specifically with regard to the cluster 2, and cluster 4. Perhaps most interesting of the model is the estimated unobserved covariance between daily patterns, indicating complementary and/or substitution mechanism between daily patterns. Table 8 shows that these covariance between cluster 4 and cluster 3 are positive, suggesting a complementary effect between these daily activity patterns (*i.e.*, the more individual allocates time for in-home activities; the more he is likely to participate in out-home recreational activities).On the other hand, Table 8 shows that the covariance between cluster 4 and cluster 2 is negative, indicating substitution between the two daily patterns.

Table 8- Daily In-home and Out-of-Home Activity Pattern Choices

Explanatory Variable	**Cluster 2**		**Cluster 3**		**Cluster 4**	
	Parameter	**t-statistic**	**Parameter**	**t-statistic**	**Parameter**	**t-statistic**
Constant	-1.22	-5.50	-0.60	-2.15	-1.75	-3.93
Household Socio-demographics						
Low income household (Less than 500 L.E /Month)	1.38	5.24			0.52	1.90
Middle income household (501 to 2000 L.E/Month)			-0.57	-1.66	1.01	3.11
Household head					-2.74	-5.06
Couple					-2.16	-1.80
Household size					1.11	2.12
Total Number of high-education student in the household	-0.20	-1.79	-0.49	-4.12	-0.38	-2.34
Rented apartment					2.60	4.23
Individual Socio-demographic and Employment Characteristics						
Middle age "individual between 30 and 49 years of age"					0.99	2.44
Access to a private car			0.51	2.12		
Work/school duration more than 8 hours/day					-1.37	-2.24
Zone Socio-demographic and Transportation characteristics						
Employment diversity "economic activity diversity per zone"	-1.63	-2.57	0.85	1.94	-3.25	-2.48
Average zone Transit system waiting time “Min."					0.04	1.75
Standard deviations of unobserved individual heterogeneity						
Cluster 2	0.74	2.37				
Cluster 3			0.37	0.80		
Cluster 4					1.42	1.99
Unobserved covariance between						
Cluster 3	0.37	0.84				
Cluster 4	-0.87	-1.44	1.12	2.25		

DISCUSSION AND CONCLUSION

In this paper cluster analysis was employed to summarize daily activity and travel behaviour of individuals in developing country. Using two-day activity daily and telecommunication survey conducted in the Cairo region, we examine individuals' time allocation for daily physical and virtual activities and travel. Moreover, clusters compositions in terms of selected socio-demographics variables of the sample are analyzed. The daily activity and travel time allocations analysis reveals that there are five clusters solutions are obtained for each of weekday and week end day. The five types of activity and travel behaviour patterns have a wide variability in time allocation among clusters as desired to identify different behaviour and there correlates. Moreover, a similar exercise has been done to examine time allocation to maintenance and recreational activities for both in-home and out-home.

The second part of this chapter proposed a mixed logit model formulation to investigate influential factors that may govern the selection mechanism of daily activity patterns. The model accommodates heterogeneity across individuals and incorporates heteroscedasticity among daily activity patterns. The results provide significant evidence of substitution and complementary mechanism between different daily activity patterns.

REFERENCES

Bhat, C. R.: Accommodating Variations in Responsiveness to Level-of-Service Variables in Travel Mode Choice Modeling, Transportation Research, Vol. 32, No. 7, 1998, pp. 495-507.

Boden, R.J. (1999). Flexible working hours, family responsibilities, and female self-employment: gender differences in self-employment selection. *American Journal of Econometrics and Sociology*, 58, 71-83.

Droege, P., ed. (1997). *Intelligent Environments*. Elsevier, Oxford.

French, D., *et al.*, eds. (1999). *Internet Based Learning: An Introduction and Framework for Higher Education and Business*. Stylus, sterling. VA.

Graham, S. and S. Marvin (1996). *Telecommunications and the City: Electronic Spaces, Urban Places*. Routledge, London.

Goulias, K. G., and R. Kitamura (1997) Regional travel demand forecasting with dynamic microsimulation models In *Panels for Transportation Planning: Methods and Applications* (eds Golob, Kitamura, Long). Kluwer. Chapter 13, pp. 321-348.

Goulias K. G. and T. Kim (2001) Multilevel analysis of activity and travel patterns accounting for person- and household-specific observed and unobserved effects simultaneously. *Transportation Research Record*, 1752, pp. 23-31.

Greater Cairo Region in the Arab Republic of Egypt (CREATS 2001), *PHASE I FINAL REPORT Volume II: Urban Transport Policy and Strategy*, November 2002 Pacific Consultants International (PCI)

Hensher, D.A., Rose, J. and Greene,W.H.: Applied choice analysis a primer, Cambridge, Cambridge University Press, 2005.

Mokhtarian P. L, 1990. A typology of relationships between telecommunications and transportation. *Transportation Research Part A,* Vol. 24A, No. 3, PP231-242.

Ma, J. and K. G. Goulias (1997a) A dynamic analysis of activity and travel patterns using data from the Puget Sound transportation panel. *Transportation*, 24(1) , pp.1-23.

Ma J. And K. G. Goulias (1997b) An analysis of activity and travel patterns in the Puget Sound transportation panel. Chapter 10 In *Activity-Based Approaches to Travel Analysis* (Editors D.F. Ettema and H. J.P. Timmermans). Pergamon Elsevier Science LTD, The Netherlands. pp. 189-207.

Pas E.I. (1982) Analytically Derived Classifications of Daily Travel-Activity Behavior: Description, Evaluation, and Interpretation. *Transportation Research Record*, 897, pp. 9-15.

Pas E.I. (1983) A Flexible and Integrated Methodology for Analytical Classification of Daily Travel-Activity Behavior. *Transportation Science*, Vol. 17, No. 4, pp. 405-429.

Pas, E. I., and F. S. Koppelman (1987). “An Examination of the Determinants of Day-to-Day Variability in Individuals’ Urban Travel Behavior.” *Transportation,* vol. 14, no. 1, pp. 3-20.

Pas, E. I. (1988). “Weekly Travel-Activity Behavior.” ***Transportation,*** vol. 15, nos. 1-2, pp. 89-109.

Recker, W.; McNally, M.; and Root, G. (1986a) A Model of Complex Travel Behavior: Part 1: Theoretical Developments. *Transportation Research* 20A(4): 307-318.

Recker, W.; McNally, M.; and Root, G. (1986b) A Model of Complex Travel Behavior: Part 2: An Operational Model. *Transportation Research* 20A(4): 319 330.

Salomon, I. (1998). Technological Change and Social Forecasting: The case of telecommuting as a travel substitute. *Transportation Research, Part C*. **6**, 17-45.

Salomon, I. (2000). Can Telecommunications Help Solve Transportation Problems? *Handbook of Transport Modeling*. D. A. Hinsher and J. B. Kenneth, Pergamon**:** 449-462.

Train, K., “Recreation demand models with taste differences over people”, *Land Economics*, vol. 74, pp.185-194, 1998.

Zimmerman, C.A., Campbell, J.L., and C. Cluett (2001). Impact of new information and communication technologies on transportation agencies: a synthesis of highway practice. NCHRP Synthesis 296, National Cooperative Highway Research Program, Transportation Research Board.