An Analysis of Accessibility, Social Interaction, and Activity-Travel Fragmentation in California

August 2019

A Research Report from the Pacific Southwest Region University Transportation Center

Elizabeth C. McBride, University of California Santa Barbara

Adam W. Davis, University of California Davis

Konstadinos G. Goulias, University of California Santa Barbara

Pacific Southwest Region UTC
University Transportation Center

UC SANTA BARBARA

TECHNICAL REPORT DOCUMENTATION PAGE

1. Report No. PSR-18-09	**2. Government Accession No.** N/A	**3. Recipient's Catalog No.** N/A
4. Title and Subtitle An Analysis of Accessibility, Social Interaction, and Activity-Travel Fragmentation in California		**5. Report Date** August 31, 2019 **6. Performing Organization Code** N/A
7. Authors Elizabeth C. McBride, UCSB Adam W. Davis, UC Davis Kostadinos G. Goulias, UCSB https://orcid.org/0000-0002-7436-9447		**8. Performing Organization Report No.** PSR-18-09
9. Performing Organization Name and Address University of Southern California 650 Childs Way, RGL 216 Los Angeles, CA 90089		**10. Work Unit No.** N/A **11. Contract or Grant No.** USDOT Grant 69A3551747109
12. Sponsoring Agency Name and Address U.S. Department of Transportation Office of the Assistant Secretary for Research and Technology 1200 New Jersey Avenue, SE, Washington, DC 20590		**13. Type of Report and Period Covered** Final report (Feb. 2019–Mar. 2020) **14. Sponsoring Agency Code** USDOT OST-R

15. Supplementary Notes
Report URL: https://www.metrans.org/research/an-analysis-of-accessibility-social-interaction-and-activity-travel-fragmentation-in-california

16. Abstract

Fragmentation of activities and travel is defined here as the multiple sequencing of many relatively short activities and trips that happen in a person's daily schedule. These are combined with much longer activities and travel to form a complete schedule of activities and travel by each observed individual. Fragmentation of activity-travel schedules may lead to increased transport demand because many activities, enabled by mobile communication technologies and other societal innovations, are no longer bound to specific times and specific places. Our main objective in this research is to close the research gap in understanding how and why individuals engage in activity-travel fragmentation. Studying the correlation of activity and travel fragmentation with social interaction and accessibility offered by the environment in which people live and close this research gap. Closing this gap will provide policy recommendations in the context of SB 375 on land use and travel. This will enable distinguishing between people that face social exclusion and the dichotomy between women spending more time in the private sphere, and less in the public one - and vice versa for men (Kwan, 1999; Kwan 2000; Turner, 1997). A secondary objective is to develop robust statistical methods for fine grained spatio-temporal data to improve travel demand forecasting models.

17. Key Words Traffic; Transportation demand		**18. Distribution Statement** No restrictions.	
19. Security Classif. (of this report) Unclassified	**20. Security Classif. (of this page)** Unclassified	**21. No. of Pages** 57	**22. Price** N/A

Form DOT F 1700.7 (8-72) Reproduction of completed page authorized

TABLE OF CONTENTS

Contents

About the Pacific Southwest Region University Transportation Center

The Pacific Southwest Region University Transportation Center (UTC) is the Region 9 University Transportation Center funded under the US Department of Transportation's University Transportation Centers Program. Established in 2016, the Pacific Southwest Region UTC (PSR) is led by the University of Southern California and includes seven partners: Long Beach State University; University of California, Davis; University of California, Irvine; University of California, Los Angeles; University of Hawaii; Northern Arizona University; Pima Community College.

The Pacific Southwest Region UTC conducts an integrated, multidisciplinary program of research, education and technology transfer aimed at *improving the mobility of people and goods throughout the region.* Our program is organized around four themes: 1) technology to address transportation problems and improve mobility; 2) improving mobility for vulnerable populations; 3) Improving resilience and protecting the environment; and 4) managing mobility in high growth areas.

U.S. Department of Transportation (USDOT) Disclaimer

Disclosure

Principal Investigator, Co-Principal Investigators, others, conducted this research titled, "An Analysis of Accessibility, Social Interaction, and Activity-Travel Fragmentation in California" at University of California, Santa Barbara. The research took place from 8/15/2018 to 8/14/2019 and was funded by a grant from the PSR UTC in the amount of $100,000 and UCSB in the amount of $10,000. The research was conducted as part of the Pacific Southwest Region University Transportation Center research program.

Acknowledgements

Funding for this research was provided by the Pacific Southwest Region 9 University Transportation Center and the GeoTrans Laboratory at the Department of Geography, University of California Santa Barbara.

Abstract

Sequence analysis is used in this project to measure fragmentation in activity participation and travel. Studying sequences of daily episodes (each activity at a place and each trip) is preferable over other techniques of studying activity-travel behavior because sequences include the entire trajectory of a person's activity during a day while jointly considering the number of activities and trips, their ordering, and their durations. We first identify places visited and duration at each place on a minute-by-minute basis, then we derive representative daily behavior patterns using hierarchical clustering. Our study shows there are at least nine distinct daily patterns with different sequencing of activities and travel as well as travel time ratios and modal split. As expected, day of the week plays a major role in the type of daily activity-travel patterns. Travel time ratios are also examined for each daily pattern and we find differences in the role played within each pattern between central city, suburban, exurban, and rural dwellers. In a comparison of couples, we find systematically higher fragmentation in households that have children and their parents are employed with women showing higher fragmentation in the activity-travel patterns.

An Analysis of Accessibility, Social Interaction, and Activity-Travel Fragmentation in California

Executive Summary

Fragmentation of activities and travel is defined here as the multiple sequencing of many relatively short activities and trips that happen in a person's daily schedule. These are combined with much longer activities and travel to form a complete schedule of activities and travel by each observed individual. Fragmentation of activity-travel schedules may lead to increased transport demand because many activities, enabled by mobile communication technologies and other societal innovations, are no longer bound to specific times and specific places. Our main objective in this research is to close the research gap in understanding how and why individuals engage in activity-travel fragmentation. Studying the correlation of activity and travel fragmentation with social interaction and accessibility offered by the environment in which people live can close this research gap. Closing this gap provides policy recommendations in the context of SB 375 on land use and travel. This will enable distinguishing between people that face social exclusion and the dichotomy between women spending more time in the private sphere, and less in the public one – and vice versa for men. A secondary objective is to develop robust statistical methods for fine grained spatio-temporal data to improve travel demand forecasting models.

To study fragmentation a new method of sequence analysis in activity participation is used in this project. To test the feasibility of the new methods and identify the best indicators we use the residents in the Central Coast of California (San Louis Obispo and Santa Barbara Counties) that participated in the California Household Travel Survey in 2012-13 (CHTS). The method explores sequences of daily activity and travel employing techniques from the sequencing of events in the life course of individuals. Studying sequences of daily episodes (each activity and each trip) are preferable over other techniques of studying activity-travel behavior because sequences include the entire trajectory of a person's activity during a day while at the same time considers the number of activities, order of activities in a day, and their durations jointly. We find substantial fragmentation in activity participation among persons with children and in specific age groups (25 to 65) amplified by the presence of children in the household. We also find poverty plays an important inhibiting role. Examinations of the days of the week shows significant and substantial differences among the different days of a week with both Sundays and Saturdays being different but also substantial differences among the weekdays.

We repeat this analysis with refinements using a statewide sample of 12,704 person that also participated in CHTS and we obtain nine distinct daily patterns. These include patterns of

people staying at home for long periods in a day, people that follow typical daily weekday working and school schedules. We also find people that travel for an entire day and people that stay at home in the morning but then travel for the rest of their interview day. We also have two patterns of running errands with very different time of day rhythms. The ninth pattern is by people that spent most of the time in a day at locations that are not home, work, or school and travel for very short time in a day. Each pattern also has different memberships in terms of gender, age, and day of the week (in addition to the working and/or student status as expected). Each of the days of the weekend has different mix of daily patterns challenging not only the typical day used in regional plans but also shows the Saturdays and Sundays are different in terms of activity and travel behavior. In addition to systematic differences among workers and students we also find systematic differences in time of day patterns between males and females and age groups. We also find evidence of higher fragmentation by center city dwellers but this is different across the daily patterns we derived in this analysis. Moreover, rural and exurban residents tend to need to spend longer times travelling to enjoy the same amount of time in activities than their counterpart suburban and center city dwellers.

The effect of children on schedule fragmentation is substantial with parents having by far higher fragmentation in their schedules than other adults employed or otherwise. Women even when they are not employed but they are in households with children also have fragmented schedule and employed women in households with children have even more fragmented schedules. All this conforms the household responsibility hypothesis.

Introduction

Fragmentation of activities and travel is defined in this paper as the sequencing of many short activities and trips that happen in a personal daily schedule. These are combined with other activities and travel that are much longer to form a complete string of episodes and durations of each episode by each individual we observe. Fragmentation is of interest to travel behavior researchers due to concerns raised by Couclelis's (*2000*, *2004*) hypothesis that economic, societal and political developments increase the individual's flexibility in scheduling daily activities. The more recent emergence of disruptive transportation services (e.g., Uber/Lyft) and automation (e.g., self-driving cars) also has the potential of added flexibility to reach places and therefore increased fragmentation.

Fragmentation in a schedule that is made of a sequence of activities means multiple switching between different activities in a day, e.g., the sequence of:

> escorting children to schools—go to work—eat meal with colleagues—run errand—go back to work—go to a social event—go back to work—pick up children from schools—go shopping—return home—escort a child to soccer practice—do some work using mobile technologies—escort child back home—work at home

Patterns like this lead to increased transport demand because many activities are no longer bound to specific times and specific places, different people need to be escorted in different activity locations, and work can often be done ubiquitously. This is further enabled by Information and (tele)Communication Technologies (ICT) that release spatial and temporal constraints. The usual mode enabling fragmentation and flexibility in scheduling is the private car. This, however, may change dramatically due to the rapid emergence of ride-hailing services and other shared mobility services that release even more spatio-temporal constraints. These services are labeled as disruptive transportation and—under specific circumstances and for specific demographic segments—using disruptive transportation increases trip making. This theme was explored further in the conference IATBR2018 (see IATBR2018.org). In this report, we demonstrate a relatively new method of travel behavior analysis to examine daily patterns in a holistic way. We set the foundation to understand the potential impact of disruptive transportation by identifying ho*w and why individuals engage in activity-travel fragmentation.*

Sequence analysis is used here to describe fragmentation and daily patterns. In travel behavior, Wilson (*1998a*, *1998b*) uses biology-inspired sequence alignment methods to study the sequences of activities, Joh et al. (*2001*) explores different techniques to introduce space in the sequence analysis, Goulias (*1999*) uses Mixed Markov Latent class (MMLC) models to study year-to-year and day-to-day variation and transitions in activity-travel patterns, and McBride et

al. (*2016*) look at lead and lag effects of land use on travel behavior using MMLC to analyze car ownership and travel as a function of demographic changes and land use.

Sequences of activities and the daily transitioning from one activity to another as well as the amount of time spent in each activity has been an important direction of travel behavior analysis (Auld, Rashidi, Javanmardi, & Mohammadian, 2011; Bhat & Pinjari, 2000; Ettema, 1995; Přibyl & Goulias, 2007). Examples include the Dutch diary analysis in Leszczyk and Timmermans (2002), in which gender and age are important determinants of moves from one activity type to another, and the formulation of methods to create daily models of activity participation and travel as in Auld et al. (2011). Analysis of sequences of activities and travel is of paramount importance in formulating econometric models embedded in activity-based daily simulations of household activity-travel patterns for large-scale travel demand analysis (Bhat et al., 2013).

In the following analyses, we move one step deeper in understanding and explaining activity sequencing during a day at specific locations, activity duration by type, and their correlation with spatial opportunities as well as social and demographic characteristics. The *sequence analysis* here examines places visited by a person during a day jointly with the duration of activities at each place. It also examines the travel episodes and time spent to reach these places. Entire daily sequences of activities and travel are described by three indicators called *Entropy* (depicting variety in daily schedules), *Turbulence* (depicting complexity in daily schedules), and *Complexity* (based on entropy). These are summary indicators that complement each other to capture daily activity-travel patterns for each individual in a succinct mathematical way. In parallel, we also derive representative patterns of daily place-time allocation and show their relationship with these indicators and determinants of travel. All this uses samples of persons in the California Household Travel Survey (CHTS) (NUSTATS, 2013).

The key questions we answer are:

1. Is there a clear taxonomy of daily sequences?
2. Are these types of sequences different in their fragmentation?
3. What are the characteristics of the people that use different types of sequences?
4. How different are the behaviors within these patterns?

Data and Data Processing

The data used in this analysis comes from the 2012 California Household Travel Survey (NUSTATS, 2013). This was a comprehensive travel survey conducted over all of 2012 and some of 2013, including information collected at the household and person level, vehicle ownership information, a one-day travel diary for every respondent, and other information not used here. In total, there were 114,639 respondents in 45,362 households.

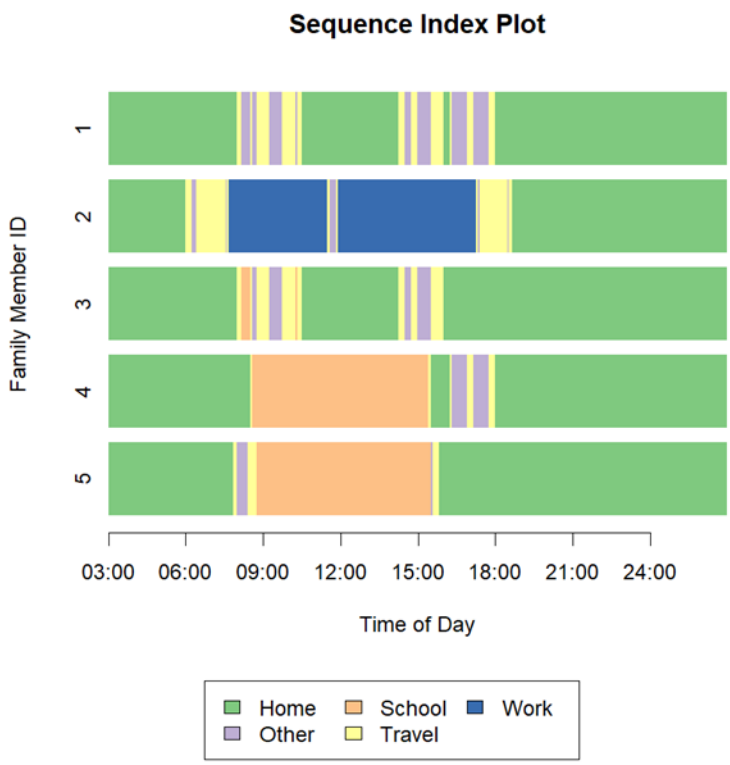
As a test area for the first sequence analysis method (pilot analysis), we use only respondents residing in San Luis Obispo (SLO) and Santa Barbara (SB) counties in the Central Coast region of California. These counties were chosen for the test area because the researchers' familiarity with the region allows for confirmation of geographic accuracy. In total for this analysis, we use data from 2,942 persons. The Santa Barbara and San Luis Obispo regions comprise an area with a variety of land use types—ranging from high density/urban to low density/rural. We use person and household characteristics and one-day travel diaries for the respondents in Santa Barbara and San Luis Obispo. The one-day travel diary takes place across the entire year (including holidays, weekends, and weekdays). It covers from 03:00 on the survey day until 02:59 on the following day. Respondents report every place they go, their travel mode, the top three activities performed at each place, and several other characteristics of the places visited and their travel. Land use surrounding each residence is depicted by indicators that are based on a detailed establishments inventory following techniques reported in Chen et al. (Chen et al., 2011) and McBride et al. (2017). We classify activities here by the place where people where during the interview that are activity at Home, Work, School, and all other.

The second analysis is a statewide random sample of 5,000 households. These are 12,704 persons living in many different places in California. We classify activities in the interview day of these persons in the same way we did for the Central Coast sample (i.e., activity at Home, Work, School, all other, and travel). Then, we repeat the same analysis as for the Central Coast. We also compare male and female fragmentation within the same households using the entire statewide database. The California Household Travel Survey is available at the Transportation Secure Data Center of the National Renewable Energy Laboratory (https://www.nrel.gov/transportation/secure-transportation-data/index.html)

Methods for Sequences

A sequence is a series of time points at which a subject can move from one discrete "state" to another. In this report, these states are based on the types of places people visit and stay during their diary day: Home, Work, School, and Other. Travel between these places is also considered a "state." We consider people who go through many states in their day to have fragmented schedules. In this report, we use sequence analysis to statistically analyze the fragmentation of respondents' days using a minute-by-minute time series. Every minute of the day contains a specific state for each person in the study. Figure 1 shows an example of the sequences identified from each person's diary in one family in the study area.

Figure 1. One family's sequence in a day (reproduced from McBride et al., 2019)



There are many techniques in the travel behavior field that can be used to measure the duration of episodes within a single state and the transition rates from one state to another (Auld et al., 2011; Bhat & Pinjari, 2000; Ettema, 1995; Kroesen, 2014; Přibyl & Goulias, 2007). These define the state of the art in longitudinal data analysis (Kroesen & Goulias, 2016). They can be useful for measuring fragmentation in a person's day but are cumbersome or infeasible when the number of transitions is very high.

We use Entropy, Turbulence, and Complexity that can handle very long sequences. The explanation follows McBride *et al.* (2019) closely, with some important additions.

Entropy is a measurement of "prediction uncertainty" (Gabadinho, Ritschard, Studer, & Muller, 2010).

$$h(x) = h(\pi_i, \ldots \pi_s) = -\sum_{i=1}^{s} \pi_i \log(\pi_i) \qquad \text{(Eq. 1)}$$

Where x is the sequence, s is number of potential states and π_i is proportion of occurrences of the *i*th state in the considered sequence. The proportion of minutes allocated to each state over the course of the entire day and the number of distinct states that a person visits drive the value of Entropy. For this measure, the number of state changes and the contiguity of states do

not matter. It simply uses the proportion of total time spent in each state, regardless of the number of different episodes that time is spread over.

If a person has no state changes during the entire day, for instance if they spend all day at home, their Entropy would be zero. In contrast, someone who moves around a lot will have "high" Entropy. The range of entropy values depends on the number of distinct states. Sequences with more unique states have higher potential maximum Entropy values, and Entropy is at its highest when people spend equal amounts of time in each state. In our study with five distinct states (Home, Travel, Work, School, and Other), the maximum Entropy is 1.61.

The second measure – Turbulence – is a bit more complicated than Entropy in terms of what it uses for its calculations.

$$T(x) = \log_2\left(\phi(x)\frac{s^2_{t,max}(x)+1}{s^2_t(x)+1}\right) \qquad \text{(Eq. 2)}$$

- x represents the sequence of activities and travel episodes in one person's diary;
- $\phi(x)$ is the number of distinct subsequences in sequence x;
- t_i is duration in each distinct state and is used to compute the mean consecutive time and variance below (i=1,…, number of distinct episodes);
- s^2_t is variance of the state-duration for the x sequence;
- $s^2_{t,\,max}$ is the maximum value that the variance can take given the total duration of the sequence x

  $$s_t, \max = (n-1)\left(1-\bar{t}\right)^2 \qquad \text{(Eq. 3)}$$
- n is length of distinct state sequence
- $\bar{t}$ is mean consecutive time spent in the distinct states

Turbulence uses the number of distinct subsequences in a given sequence and the number of consecutive time points spent in a given state (Elzinga & Liefbroer, 2007; Gabadinho et al., 2010). Consider a person with a daily sequence H-T-W-T-H meaning the person was at home (H) in the morning, traveled (T) to work (W), and after work traveled (T) back home (H). This sequence would contain the following subsequences: an empty sequence; the full sequence itself; subsequences of the type T-W-T-H, W-T-H, and T-W-T; discontinuous subsequences like T-T-H (which skips the work activity); and single activities H, T, and W. Enumerating all these subsequences yields ($\phi(x) = 27$) possible combinations that respect the precedence of activities in the H-T-W-T-H sequence. For a given sequence of activities (x), the measure of Turbulence is a measure of variability in terms of distinct activities, the order of these activities, and the variance of the durations of these activities in a day. All this makes Turbulence a measure of schedule complexity.

Gabadinho et al (2010) define another indicator they call Complexity that is based on Entropy and the transitions within a sequence (s).

$$C(s) = \sqrt{\frac{nt(s)}{(l(s)-1)} \frac{h(s)}{h_{max}}} \qquad \text{(Eq. 4)}$$

This is a function of the Entropy and the number of transitions (nt(s) = l(s) -1) in a sequence s, normalized by the maximum theoretical entropy (h_{max}) and the length of the sequence (l(s)). This indicator will have a value between 0 and 1, with zero corresponding to Entropy zero and no transitions.

Table 1 provides a few examples of sequences with state duration for each activity, counts of subsequences, Entropy, Turbulence, and Complexity. Person 1 in Table 1 stays at home all day and has the sequence of (H, 1440), Entropy zero, and Turbulence 1. The number of distinct subsequences is 2 (i.e., the empty sequence and the sequence itself). Persons 2 and 3 have activity patterns with 5 episodes that are 2 activities at home, 2 trips, and one at work (Persons 2) or some other place (Person 3). Both persons have 27 subsequences and they are all different in the Entropy of their sequences because the number of minutes allocated to each episode are different between them. Similarly, for the Turbulence, the variance across the durations of activities is different between the two subjects. The Complexity indicator combines the advantages of Entropy (variety of time use) and Turbulence (reflecting the possibility of many subsequences), but in a simpler form than Turbulence replacing subsequences with the length of the sequence and the number of transitions. Persons 4 and 5 show how the number of subsequences increases dramatically when more activity episodes are added and how this is reflected in the three summary indicators we use. We will use these indicators in the summary of findings.

Table 1. Examples of Sequences

	(Activity/Place, Duration in minutes)	Pattern	Number of Subsequences	Entropy h(x)	Turbulence T(x)	Complexity C(x)
Person 1	(H,1440)	H	2	0.000	1.00	0.00
Person 2	(H,830)-(T,10)-(W,320)-(T,10)-(H,270)	H-T-W-T-H	27	0.372	6.63	0.0322
Person 3	(H,255)-(T,45)-(O,120)-(T,30)-(H,990)	H-T-O-T-H	27	0.302	6.10	0.0290
Person 4	(H,600)-(T,15)-(O,60)-(T,10)-(O,20)-(T,10)-(H,725)	H-T-O-T-O-T-H	79	0.203	7.87	0.0291
Person 5	(H,485)-(T,5)-(W,169)-(T,2)-(H,10)-(T,14)-(O,70)-(T,10)-(O,25)-(T,15)-(H,125)-(T,15)-(W,321)-(T,14)-(H,160)	H-T-W-T-H-T-O-T-O-T-H-T-W-T-H	9632	0.641	16.01	0.0790

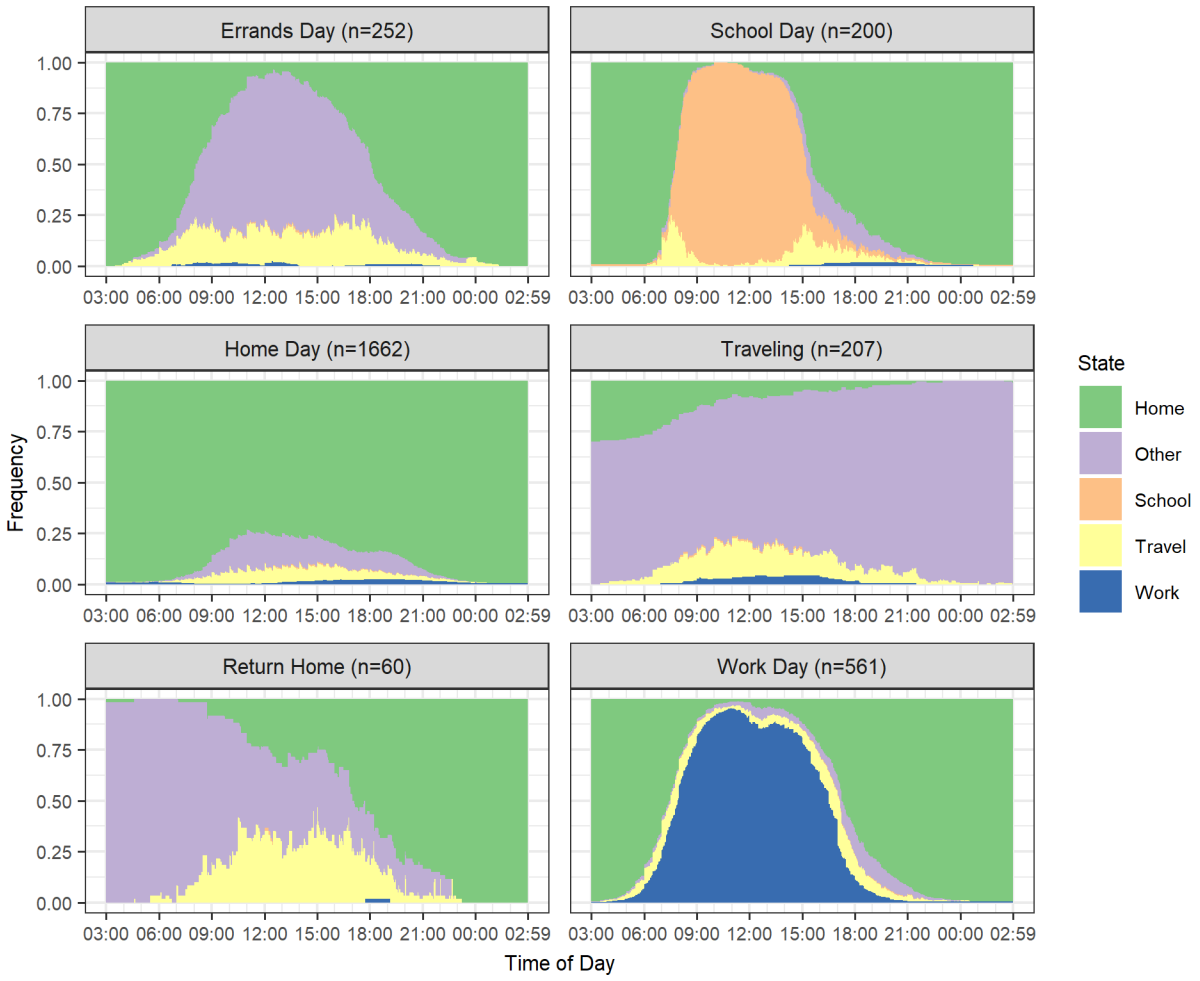
Pilot Test in San Louis Obispo and Santa Barbara

Our first objective is to find groups of sequences that are similar using a small sample as a pilot test of the methods in this project. For each of the 2942 sequences, there is a series of 1440 bins—one for each minute of the survey day starting at 3:00 AM and ending the next day at 2:59 AM. Each bin is colored by a letter (H, T, W, S, and O).

To identify similar sequences among the 2942 person-sequences, we need to have a rule of comparison. For example, we can perform different operations to reproduce one sequence departing from another and assign penalties to each operation (Wilson, 1998a, 1998b). Measuring the difference between two sequences depends on the number of operations and sum of penalties accumulated in the comparison. The operations applied to this comparison are replacement, insertion, and deletion (indel). In the sequence alignment literature, the measurement of dissimilarity and the number of operations needed to make two sequences exactly the same is called a "distance". The distance between two sequences is the minimum combination of replacements and indel (Abbott & Tsay, 2000). For ease of interpretation, in this chapter we will call this the *dissimilarity score* between two sequences. The output of an algorithm that does these operations among all the sequences is a matrix of dissimilarity scores. In the analysis here, this is a matrix of 2942 by 2942 (=8655364) cells containing the dissimilarity scores among sequences for each person we have in our sample.

This matrix can then be analyzed using clustering techniques to identify a small number of groups of sequences that represent similar time of day activities and travel patterns in our sample. To do this, we use the agglomerative nesting clustering method. We start with 2942 sequences and group pairs of sequences based on their dissimilarity scores. Then, we compare all the cluster dissimilarity score averages to each other and lump together clusters with smaller dissimilarity cores. We proceed until all observations are in one cluster (Kaufman & Rousseeuw, 1990). This process can be thought of as a tree that starts with every sequence as an individual "leaf" and ends with one cluster as the "trunk." After inspecting the overall time of day patterns, we selected the six-cluster solution because it shows clear representations of time of day time allocation patterns to places/activities (Figure 2).

Figure 2. Six clusters of daily sequences of places and travel

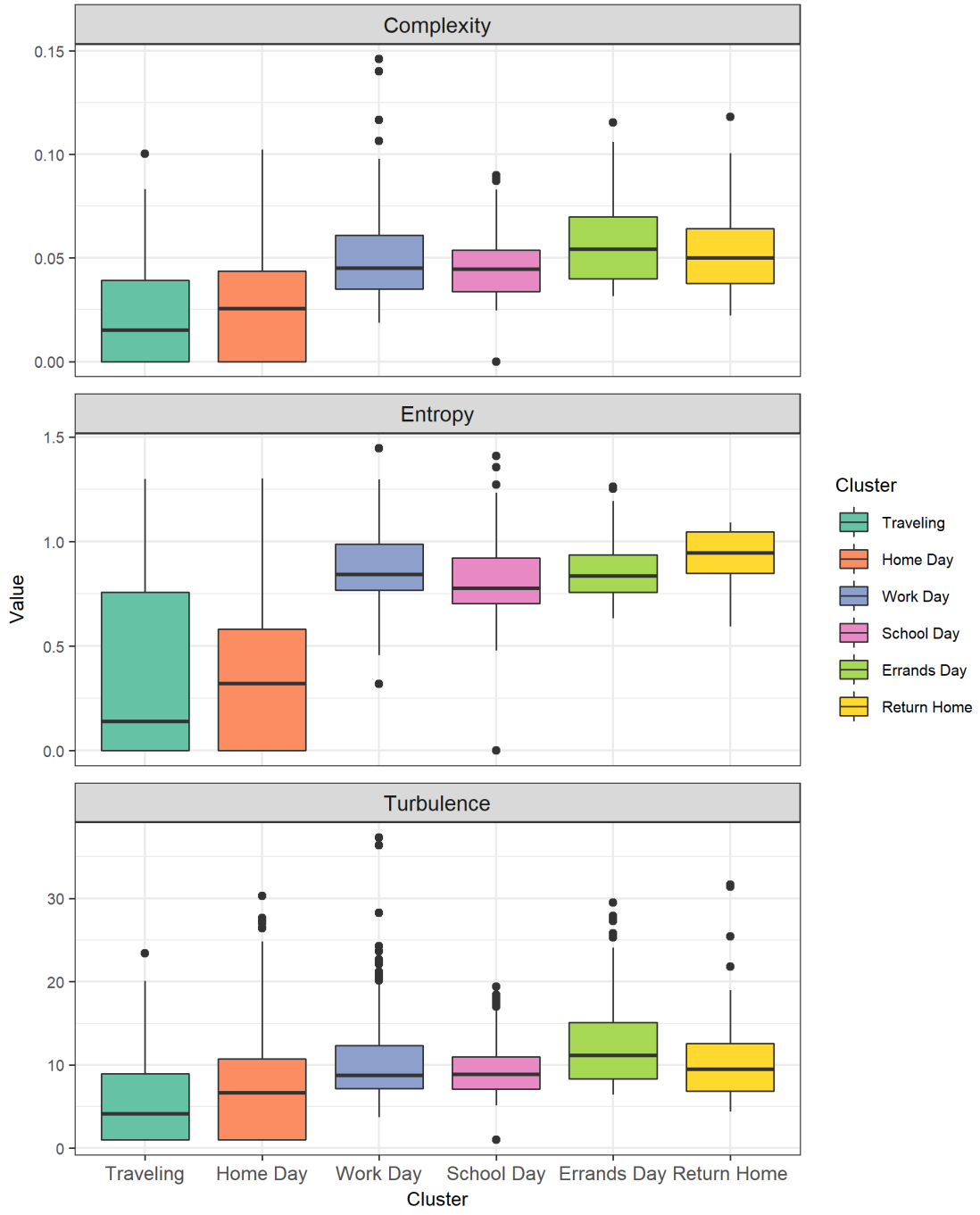


The *Traveling* cluster in Figure 2 is the 207 persons that were outside the region during the survey or left during the day of the interview and travelled outside the region of SLO and SB. In the *Traveling* cluster, people are mostly out of home and we set accessibility to zero and distance from home when missing to 350km. The *Home Day* cluster consists of 1662 persons that spent most of their time at home with very short trips to places that are classified as Other. It has the smallest median kilometers from home at 5.1 km. This is indicative of the fact that people in the *Home Day* cluster do not venture far from home when they do leave. The *Work Day* cluster, with 561 persons, is the usual working day with travel before and after work and some visits to other places in the after-work hours. The *School Day* cluster is the usual school day for 200 survey participants with substantial travel before and after school and some work activities after school (presumably high school and college students). This cluster has a median kms from home of 7.6, which is the second smallest behind *Home Day* (5.1 km). This shows that students go to school relatively close to home. The *Errands Day* cluster, from 252 participants, shows a typical errands day, where respondents start at home, go to "Other" locations all day

and return home in the evening. The *Errands Day* respondents travel throughout the day with no noticeable fluctuations by time of day, indicating that they move from destination to destination all day in many small trips. People in this cluster drive frequently, but around home. The *Return Home* cluster is a group of 60 respondents who are returning home from locations outside the region. They travel more than all other groups. Not all workers are in the *Work Day* cluster, indicating that people with irregular work timing and/or short durations at work locations are in every other type of cluster. In contrast, the place "School" does not appear to be a preeminent sequence state in any other cluster except *School Day*. Locations categorized as "Other" appear in all six patterns, pointing out the need to explore this category further. As these patterns only look at respondents' travel on a single day, it would be expected that depending on the day surveyed, respondents' cluster membership would vary. The composition of the clusters is further examined later with the multinomial logit model.

Figure 3 shows the distribution of Entropy, Turbulence and Complexity values for each of the six clusters. The *Traveling* cluster has the lowest median values for all three indicators due to the way data are recorded for people that are out of town during the interview or leave from home on that day and go far. *Home Day* has the second lowest median and overall distribution for all three indicators among the clusters but shows substantial variation, meaning people in this cluster have a wide variety of sequences. The other four types have similar complexity and variety with *Work Day*. The *Work Day* cluster shows a few outliers in the values of Turbulence and Complexity (recall these two include sequence and subsequence lengths). Figure 3 demonstrates that we cannot study sequences with only clusters or only indicators of Entropy, Complexity and Turbulence. Instead, we need to explore sequences using multiple methods.

Figure 3. Box plots of within-cluster entropy, complexity, and turbulence



Correlation of Sequences with Person Characteristics

In a parallel analysis (McBride et al., 2019), we use regression methods to study the correlation between Entropy and Turbulence for the entire 2942 sample of people with social and demographic characteristics of this sample. In summary, people age 25 to 65 have the most fragmented schedules (particularly as measured by Turbulence), especially when they have

children over age 4 in the household. Escorting and joint participation in activities with children is a clear motivation for this. We also find significant differences among people of different incomes, with poverty inhibiting activity fragmentation for a person. Ethnicity and nativity also play a role in distinguishing among sequences. Hispanic people have sequences that are simpler than the US native group, but still more complex than other groups. Gender also emerges as a major covariate for Entropy, but not for Turbulence. People that live in urban and suburban environments, however, tend to have more fragmented schedules most likely due to the mixing of short and long activities in their schedule. Another major factor of fragmentation is the day of the week. Each day of the week appears to have a different composition of activities and durations. Sunday is the day with the least fragmented schedules and Friday shows the highest fragmentation. In addition, older children in the household motivate a more fragmented daily schedules of activities.

Multinomial Logit Model

To better understand the composition of each of these clusters, we estimate a multinomial logit model with categories for the six clusters for each person and identify variables that explain their cluster membership. Table 2 shows the results of this multinomial logit model.

Table 2. Multinomial logit of cluster membership

	Cluster Type				
	Home Day	Work Day	School Day	Errands Day	Return Home
Constant	1.938	-15.706	-2.681	-0.234	-2.123
	*t = 8.888****	*t = -127.915****	*t = -6.231****	*t = -0.804*	*t = -4.596****
Respondent is Worker	-0.275	17.520	-0.910	0.139	1.076
	*t = -1.779**	*t = 142.692****	*t = -2.560***	*t = 0.700*	*t = 3.188****
Respondent is Student	-0.844	-2.760	3.766	-0.200	0.514
	*t = -3.659****	*t = -4.313****	*t = 10.468****	*t = -0.710*	*t = 1.261*
Respondent is Female	0.091	-0.066	-0.213	-0.218	0.609
	t = 0.610	*t = -0.368*	*t = -0.854*	*t = -1.151*	*t = 1.984***
Number of Children Under 16 in Household	0.374	0.498	0.576	0.516	0.471
	*t = 3.551****	*t = 4.079****	*t = 4.347****	*t = 4.395****	*t = 2.882****
Survey Day: Tuesday	0.232	0.529	0.845	0.324	-0.686
	t = 0.814	*t = 1.655**	*t = 2.009***	*t = 0.876*	*t = -1.167*
Survey Day: Wednesday	0.353	0.650	1.365	0.373	-0.998
	t = 1.300	*t = 2.111***	*t = 3.371****	*t = 1.056*	*t = -1.597*
Survey Day: Thursday	0.290	0.592	0.907	0.347	-2.158
	t = 0.996	*t = 1.795**	*t = 2.052***	*t = 0.917*	*t = -2.011***
Survey Day: Friday	0.466	0.637	0.856	0.924	-0.913
	t = 1.512	*t = 1.858**	*t = 1.893**	*t = 2.450***	*t = -1.301*
Survey Day: Saturday	-0.245	-2.367	-3.743	-0.049	-0.010
	t = -0.935	*t = -5.721****	*t = -3.510****	*t = -0.141*	*t = -0.023*
Survey Day: Sunday	0.156	-2.338	-3.399	0.043	-0.251
	t = 0.639	*t = -6.306****	*t = -4.305****	*t = 0.132*	*t = -0.583*
Akaike Inf. Crit.	5,634.98	5,634.98	5,634.98	5,634.98	5,634.98
Note:	*p<0.1; **p<0.05; ***p<0.01				
Loglikelihood at convergence = -2762.488 (degrees of freedom = 55) Loglikelihood with constants only = -3818.690 ; McFadden pseudo R^2 = 0.277					

The model tests which sequence cluster people belong to, given the independent variables in the model (worker, student, male, number of children under 16 in the household, and day of survey). All comparisons are made to a reference group that is excluded from the model. In this case, the reference group is the *Traveling* cluster. Across all clusters, workers are more likely to be in the *Work Day* cluster and the *Return Home* cluster. If a respondent is a student, they are more likely to be in the *School Day* cluster. We also control for the day of the week of the travel diary in this model and use Monday as the reference day. CHTS respondents assigned to a Saturday or Sunday have low propensity to be in the *Work Day* and *School Day* clusters. The exact opposite is true for the weekdays. The number of children under 16 in the household has positive and significant coefficients for all five groups in Table 2 (also confirmed by the descriptive statistics of Table 3). This reflects the fact that the reference group (*Traveling*) has

the lowest number of children in the household (Table 4). Within clusters, we find different coefficients for the days of the week. For the *Home Day* cluster, none of the coefficients of the days of the week are significantly different than zero, meaning *Home Day* as a pattern of sequences is spread almost uniformly throughout all days of the week for this sample. A similar trend is shown for *Errands Day*, except for Friday.

Table 3. Within-cluster sample characteristics of categorical variables

Variable	Traveling	Home Day	Work Day	School Day	Errands Day	Return Home	Overall
N females	108	916	266	82	117	39	1528
	52.17%	55.11%	47.42%	41.00%	46.43%	65.00%	51.94%
N people with disabilities	12	149	16	2	11	0	190
	5.80%	8.97%	2.85%	1.00%	4.37%	0.00%	6.46%
N ppl in hhs with kids age 00 to 03	16	164	47	42	29	8	306
	7.73%	9.87%	8.38%	21.00%	11.51%	13.33%	10.40%
N ppl in hhs with kids age 04 to 15	38	388	125	157	83	21	812
	18.36%	23.35%	22.28%	78.50%	32.94%	35.00%	27.60%
N ppl in hhs with kids age 16 to 18	36	163	55	63	31	9	357
	17.39%	9.81%	9.80%	31.50%	12.30%	15.00%	12.13%
N students	33	177	3	185	47	15	460
	15.94%	10.65%	0.53%	92.50%	18.65%	25.00%	15.64%
N weekend responders	81	536	26	3	81	33	760
	39.13%	32.25%	4.63%	1.50%	32.14%	55.00%	25.83%
N workers	92	636	561	16	113	38	1456
	44.44%	38.27%	100.00 %	8.00%	44.84%	63.33%	49.49%
Household income at or below poverty line	13	118	25	32	26	2	216
	6.28%	7.10%	4.46%	16.00%	10.32%	3.33%	7.30%

Table 4. Within-cluster sample characteristics of continuous and count variables

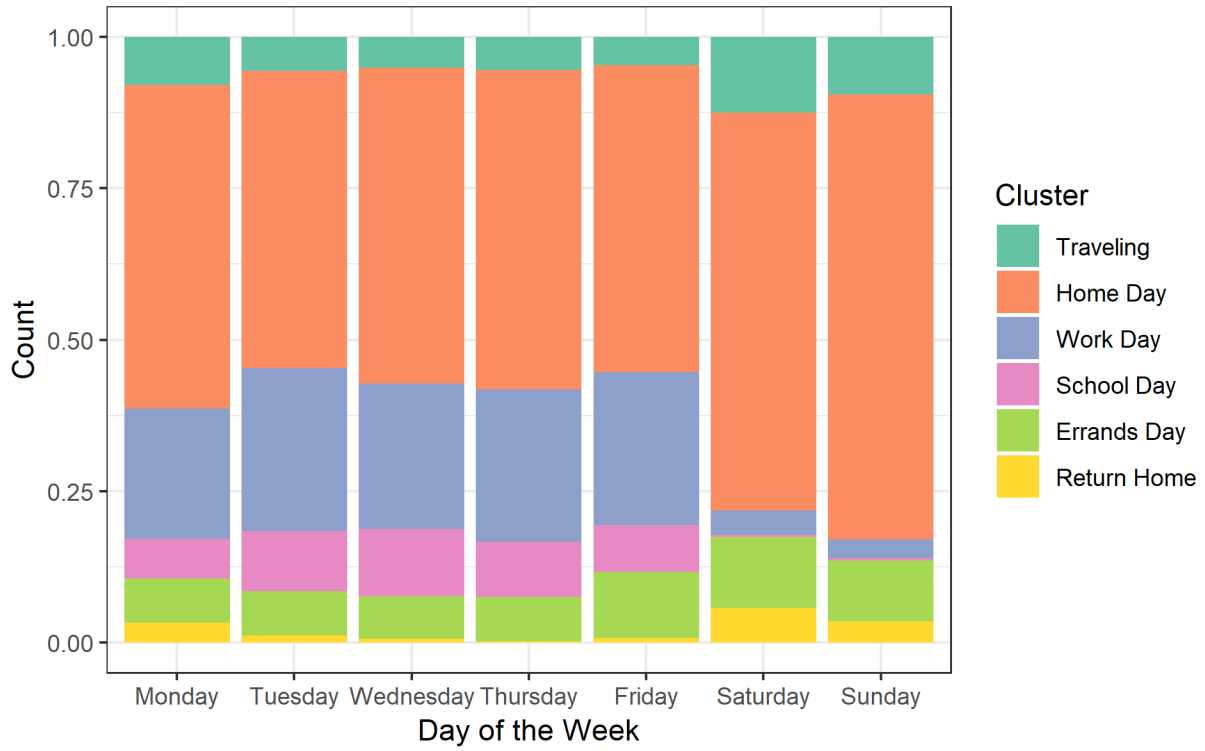
Variable	Cluster	Mean	Minimum	Median	Maximum	St. Deviation
Complexity (C(s), Eq. 4)	Errands Day	0.057	0.032	0.054	0.115	0.020
	Home Day	0.026	0.000	0.026	0.102	0.023
	Return Home	0.053	0.022	0.050	0.118	0.021
	School Day	0.046	0.000	0.045	0.090	0.015
	Traveling	0.021	0.000	0.015	0.100	0.024
	Work Day	0.050	0.019	0.045	0.146	0.018
	Total	**0.035**	**0.000**	**0.035**	**0.146**	**0.025**
Customer Service Establishments Near Homes	Errands Day	8.262	0.000	6.490	24.550	6.072
	Home Day	7.754	0.000	6.285	25.240	5.708
	Return Home	8.010	0.000	6.620	24.320	5.729
	School Day	8.504	0.000	7.225	25.240	6.235
	Traveling	2.495	0.000	0.000	24.440	5.193
	Work Day	8.256	0.000	6.500	25.240	6.125
	Total	**7.579**	**0.000**	**6.220**	**25.240**	**5.991**
Maximum Distance from Home (kms)	Errands Day	66.681	0.337	22.215	350.000	105.821
	Home Day	15.283	0.000	3.555	350.000	46.253
	Return Home	197.782	0.116	278.819	350.000	157.157
	School Day	10.972	0.000	5.417	350.000	26.990
	Traveling	302.914	6.915	350.000	350.000	113.791
	Work Day	28.360	0.297	14.592	350.000	52.360
	Total	**45.846**	**0.000**	**8.398**	**350.000**	**99.830**

Table 5 (continued). Within-cluster sample characteristics of continuous and count variables

Variable	Cluster	Mean	Minimum	Median	Maximum	St. Deviation
Number of Children in Household Between 0 and 3	Errands Day	0.123	0.000	0.000	2.000	0.352
	Home Day	0.131	0.000	0.000	2.000	0.423
	Return Home	0.133	0.000	0.000	1.000	0.343
	School Day	0.250	0.000	0.000	2.000	0.519
	Traveling	0.092	0.000	0.000	2.000	0.336
	Work Day	0.102	0.000	0.000	2.000	0.357
	Total	**0.130**	**0.000**	**0.000**	**2.000**	**0.407**
Number of Children in Household Between 4 and 15	Errands Day	0.635	0.000	0.000	5.000	1.038
	Home Day	0.406	0.000	0.000	5.000	0.842
	Return Home	0.583	0.000	0.000	3.000	0.889
	School Day	1.445	0.000	1.000	5.000	1.078
	Traveling	0.242	0.000	0.000	4.000	0.599
	Work Day	0.348	0.000	0.000	4.000	0.731
	Total	**0.477**	**0.000**	**0.000**	**5.000**	**0.889**
Number of Children in Household Between 16 and 18	Errands Day	0.143	0.000	0.000	2.000	0.403
	Home Day	0.113	0.000	0.000	3.000	0.365
	Return Home	0.150	0.000	0.000	1.000	0.360
	School Day	0.425	0.000	0.000	3.000	0.705
	Traveling	0.193	0.000	0.000	2.000	0.442
	Work Day	0.107	0.000	0.000	2.000	0.337
	Total	**0.142**	**0.000**	**0.000**	**3.000**	**0.408**
Number of Household Vehicles	Errands Day	2.079	0.000	2.000	6.000	0.983
	Home Day	2.089	0.000	2.000	8.000	1.108
	Return Home	2.333	1.000	2.000	5.000	0.877
	School Day	2.215	0.000	2.000	6.000	1.060
	Traveling	2.222	0.000	2.000	6.000	0.924
	Work Day	2.160	0.000	2.000	7.000	1.028
	Total	**2.125**	**0.000**	**2.000**	**8.000**	**1.064**

Figure 4 shows the observed relationship between Day of the Week and cluster types. From this analysis, we do not detect any major gender differences except for the Return Home cluster in which there are 39 females (65% of this group). Investigation into this revealed that the median of maximum distance traveled from home in the *Return Home* cluster is much lower for females than it is for non-females (80 km vs the imputed maximum of 350km). As it turns out, many more females start from a closer location to home than non-females. They stay in town and run errands before returning to their own homes.

Figure 4. Observed cluster membership by day of the week (DOW)



Linear Regression Models of Within-Cluster Complexity

The main objective of this chapter is to explore fragmentation of daily schedules. In this section, we analyze within-cluster complexity to understand the relationship between fragmentation, person characteristics, household structure, accessibility, and distance travelled to reach places. To study the propensity of persons to fragment their daily activity-travel pattern, we estimate six regression models. For each of the six patterns described above we use the complexity indicator (C(s) in Eq. 4) computed for each individual in the cluster group as the dependent variable. As explanatory variables, we use person and household characteristics. Table 5 shows all six regression models.

Table 6. By-cluster complexity linear models

	Dependent Variable is Complexity C(s) (Eq. 4)					
	Cluster Type					
	Traveling	Home Day	Work Day	School Day	Errands Day	Return Home
Constant	0.025	0.030	0.048	0.045	0.054	0.039
	*t = 3.192****	*t = 13.473****	*t = 16.145****	*t = 7.741****	*t = 9.372****	*t = 2.803****
Disability	0.003	-0.006	0.001	-0.015	0.006	
	t = 0.443	*t = -2.741****	*t = 0.310*	*t = -1.281*	*t = 0.874*	
Household Income Near or Below the Poverty Line	-0.011	-0.011	-0.007	-0.009	-0.019	0.018
	t = -1.581	*t = -4.313****	*t = -1.833**	*t = -2.305***	*t = -3.291****	*t = 1.180*
Low to Medium Household Income	0.001	-0.006	-0.003	-0.002	-0.006	-0.004
	t = 0.113	*t = -3.796****	*t = -1.241*	*t = -0.581*	*t = -1.629*	*t = -0.417*
Medium to High Household Income	-0.006	-0.003	0.00001	-0.001	-0.006	0.011
	t = -1.552	*t = -2.141***	*t = 0.006*	*t = -0.532*	*t = -1.744**	*t = 1.651*
Weekend	0.002	-0.005	-0.008	0.002	0.001	0.003
	t = 0.672	*t = -3.894****	*t = -1.954**	*t = 0.263*	*t = 0.336*	*t = 0.490*
Number of Children Under 4 in Household	0.005	-0.004	-0.0003	-0.001	0.006	0.029
	t = 1.104	*t = -3.161****	*t = -0.113*	*t = -0.365*	*t = 1.439*	*t = 2.751****
Number of Children Age 4 to 15 in Household	-0.007	0.003	0.002	0.001	0.002	0.0003
	*t = -2.263***	*t = 3.995****	*t = 2.153***	*t = 0.760*	*t = 1.278*	*t = 0.069*
Number of Children Age 16 o 18 in Household	-0.013	0.004	0.002	-0.0005	0.005	0.003
	*t = -3.342****	*t = 2.798****	*t = 0.903*	*t = -0.268*	*t = 1.611*	*t = 0.327*
Female	-0.003	0.001	0.0004	-0.001	0.005	0.009
	t = -0.852	*t = 0.547*	*t = 0.240*	*t = -0.399*	*t = 1.922**	*t = 1.530*
Worker	0.006	0.003		0.009	0.003	-0.010
	*t = 2.014***	*t = 2.691****		*t = 2.195***	*t = 1.189*	*t = -1.468*
Student	0.016	-0.002	0.041	0.006	-0.008	-0.014
	*t = 3.118****	*t = -0.925*	*t = 3.869****	*t = 1.338*	*t = -2.113***	*t = -1.471*
Number of Household Vehicles	0.001	-0.002	-0.0004	-0.001	0.0004	-0.002
	t = 0.301	*t = -4.422****	*t = -0.494*	*t = -0.838*	*t = 0.249*	*t = -0.594*
Mean Customer Service Establishments within 10km of Home	0.002	0.0003	0.0003	-0.0001	0.0003	0.001
	*t = 6.206****	*t = 2.940****	*t = 1.860**	*t = -0.737*	*t = 1.397*	*t = 0.931*
Maximum kilometers traveled from home	-0.00003	0.0001	0.00001	-0.00004	0.00001	0.00004
	*t = -1.697**	*t = 11.809****	*t = 0.700*	*t = -1.078*	*t = 0.933*	*t = 1.931**
Observations	190#	1,543#	526#	195#	240#	58#
R^2	0.328	0.144	0.057	0.090	0.132	0.403
Adjusted R^2	0.274	0.137	0.033	0.019	0.078	0.226
Residual Std. Error	0.020 (df = 175)	0.022 (df = 1528)	0.018 (df = 512)	0.015 (df = 180)	0.019 (df = 225)	0.019 (df = 44)
F Statistic	6.089*** (df = 14; 175)	18.411*** (df = 14; 1528)	2.385*** (df = 13; 512)	1.266 (df = 14; 180)	2.440*** (df = 14; 225)	2.283** (df = 13; 44)

Notes: # sample sizes are lower than the cluster membership due to missing values for some of the explanatory variables used here
*p<0.1; **p<0.05; ***p<0.01

Traveling Cluster

The number of people in a respondent's household between 4 and 15 years old has a significant negative effect on Complexity. More children in the age range of 4 to 15 corresponds to lower Complexity for members of the *Traveling* cluster. More household members in the 16 to 18 age range also significantly corresponds to lower Complexity in the *Traveling* cluster. The number of children between ages 0 and 3 is not a significant determinant of Complexity in this cluster because there are very few persons in this group with children in this age group. Respondents in the *Traveling* cluster that are employed have higher Complexity. Presumably traveling for work and combining their work trip with other activities. We find a similar effect for students.

Accessibility around home measured by average customer serving establishments within 10 km from each home location is highly statistically significant. Considering that 25% of this group started their day at home, it is possible they use access to opportunities around their homes before departing for a long-distance trip. The variable of maximum kilometers a destination was from home captures the interaction between choices in space and timing of trips. Longer distances from home correspond to lower Complexity scores indicating an inhibiting impact on Complexity when people consume longer travel time to far away destinations. However, we are not distinguishing among the locations included in other and this may mask other factors.

Home Day Cluster

It is important to note that the *Home Day* cluster does include some movement. People in the *Home Day* cluster are those who mostly spent the day at home, with little other activity. However, a few persons did not stay home all day. The majority of respondents with disabilities are in the *Home Day* cluster (149 of 190). Respondents in the *Home Day* cluster who have disabilities have significantly lower Complexities than those without disabilities. The linear model seems shows that not only do disabled respondents belong mostly to the group that did not leave the house much, but also, they tend to travel less than non-disabled people within the cluster. Disability appears to limit how people travel and therefore has a significant impact on their access to opportunities. People in the poverty group belong to households with an income near or below the poverty line. For members of the *Home Day* cluster, being in the lowest income group has a strong significant negative impact on Complexity as compared to those that are above the poverty line. This shows how poverty can have a limiting effect on opportunities and movement.

Home Day cluster members who did their travel diary on a weekend day had significantly lower Complexity than those who did it on a weekday. This indicates that, within the *Home Day* cluster, on a weekend day people tend to go to fewer places than they do on weekdays. Having children under 4 in the household significantly reduces Complexity for the *Home Day* cluster. Unsurprisingly, a baby or toddler is a strong anchor to the home. Having more children between 4 and 15 years old has a significant positive effect on Complexity for those in the *Home Day* group. Caretakers who are not working on the diary day are probably responsible for most of

the errands or transportation of children. Even if they are spending most of the day at home, there are more likely to be at places they need to go since they have children that need escorting. Having children in a household between 16 and 18 also has a significant positive effect on Complexity for *Home Day* cluster members. Even though in this age group teenagers can legally obtain a driver's license, this does not mean they drive without their parents when they turn 16. Teens may be more independent, but they are not completely self-reliant. The outcome is parents with higher Complexity in their schedules to serve children with a need to be at different places in a day. Workers in the *Home Day* cluster are likely enjoying a day off. Their Complexity is higher than cluster members who are not workers. On days off, workers will have more errands to run that they could not complete on a work day.

In this group, a higher number of vehicles corresponds to significantly lower Complexity. If there are more vehicles available in a household, individuals can go where they want without needing to combine trips with other people in the household and therefore using simpler activity-travel patterns. In contrast, more vehicles in a household also correspond with higher income that increases complexity and we control for income with the presence of the poverty measure in the model. Higher numbers of customer service establishments around the home correspond to higher Complexity in the *Home Day* cluster. People with more establishments around their homes can still have "Home Days", where they spend most of their time at home while still participating in shorter trips. They can probably run errands closer to home more easily. Opposite of the previous group this group of people trade-off time and distance differently. Traveling farther from home means higher Complexity for *Home Day* cluster members. There were lots of people who did not leave their homes at all on their Home Day, but for those that did, the farther they went from home the more fragmented their day was. The respondents whose maximum distance is farther away from home are likely running errands that day, but not enough errands to place them in the *Errands Day* cluster. However, they will still have more complex schedules than those who stay at home or close to their home.

Work Day Cluster

For people in the *Work Day* cluster, being below or near the poverty line has a negative effect on Complexity. Respondents with household income below/near the poverty line who are working on the diary day have less fragmented schedules than those earning more money. This could be for several reasons.

- Respondents might not take a single occupancy vehicle to work, so they are not as mobile.
- They might not have the monetary freedom to participate in opportunities.
- It could also be that the type of job they hold does not allow for as much schedule flexibility as the higher-income positions.

On weekend days, people in the *Work Day* cluster tend to have lower Complexity. This could just be similar to the effect discussed in the *Home Day* section – that people do not do as much on weekend days. It could also have to do with the types of jobs that are available over weekends. Very few office or manual labor positions are worked on weekends. The jobs would likely be in the service industry. These types of jobs might be more limited in the way they can be performed – they need to be happening in the same place all the time – which reduces opportunities for fragmentation. On work days, Complexity is only significantly impacted by having children age 4 to 15 in the household. A higher number of children in this age range corresponds to an increase in Complexity. Children in this age range need to be driven around to appointments, activities, etc. They also need to be picked up and dropped off from school. If there are more children in a household, the adult worker will escort the children to different places. A student who is in the *Work Day* cluster is a student who also has a job. This situation leads to increased Complexity because the needs to be at different places at different times increases fragmentation.

Work Day cluster members tend to have corresponding higher Complexity with the customer service establishment numbers around their home. They have more access to establishments around their homes, so they can more easily access opportunities without having to travel far.

School Day Cluster

Our search for significant indicators for this pattern did not yield as many variables as for the other patterns. This is mainly due to the membership in this group that are children and teenagers that are students. However, being at/around the poverty line lowers Complexity, even for students on a school day. This is indicative of the problems in equity for students in poverty. K-12 students in poverty may not have access to the same numbers of after-school activities as their peers. This lack of access to opportunities at an early age sets young people up for fewer opportunities later. *School Day* cluster members who are also workers have higher Complexity than those who are not workers. This has already been addressed in the *Work Day* cluster discussion: students who are also workers have more complex days because they have more places to be and things to do than someone who is just a worker or just a student.

Errands Day Cluster

For people in the *Errands Day* cluster, having household income around the poverty line means they have lower Complexity. As in other clusters, this is indicative of the impact of poverty on the mobility of people and access to opportunities. *Errands Day* cluster members in households with higher numbers of children in the 4 to 15 age range have statistically higher Complexity. Adults in this cluster likely spend a lot of their day transporting children, running errands, and performing other general tasks required to manage a household. As will be discussed later, children in this age range are often accompanying parents on errands if it is not a school day. Females in *Errands Day* cluster members have significantly higher Complexity than non-females (males in this sample). This result is in line with past research that has shown that women tend

to take on more of the household tasks and is consistent with the household responsibility hypothesis (Crane, 2007; Turner & Niemeier, 1997), even if both partners work full-time. A worker in the *Errands Day* cluster has higher Complexity than a non-worker in the cluster. This was discussed earlier with the *Home Day* cluster: on a worker's day off, they might need to run the errands that they do not have time for on work days. This would result in more fragmentation on their day off. Students in the *Errands Day* cluster have lower Complexity than non-students. This could be because some cluster members are children who come along for errands with their parents, but not for all of them. Younger students might be participating in other activities that are not errands where they do not travel as much (e.g., playing at a friend's house). This student category also includes university students. For that group, irregular schedules allow for them to spread their errands out more, so in a single free day they would not have as much fragmentation as non-student adults. For the *Errands Day* cluster, higher maximum distance traveled from home means higher Complexity. Destinations with more opportunities, like a mall, might be farther away from the home, and while a person is there they go to several different destinations.

Return Home Cluster

This group is made by 65% females that also show higher Complexity than their counterparts males. The median for maximum kilometers for females and non-females in the *Return Home* differs greatly: for females it is 80.8 km, while it is 350 km (the maximum imputed value) for non-females. The means were also different: 174 for females and 241 for non-females.

So, the *Return Home* cluster in fact contains two separate groups of people: people who travel long distances, and people who stayed around town but did not start at home (and went to places marked "Other" during the day). Women primarily make up the latter group. Presumably, women returning home from a trip far from home when they arrive at home are also running local errands. This may be an indication of the multiple roles played by women in the household responsibility hypothesis discussed above. In this group, having children under 4 in the household increases respondents' Complexity significantly. Since the *Return Home* cluster consists of people's travel days, this shows that traveling with very young children leads to more fragmentation and therefore complexity in patterns (however very few persons have children under 4 years old in this group). Workers who are in the *Return Home* cluster have lower Complexity than non-workers.

All together, these results show that clustering of daily patterns using this type of data and then studying the fragmentation characteristics of the cluster members leads to important behavioral conclusions about task allocation and the role played by income, accessibility, and the trade-offs with distances travelled. In essence, we see heterogeneity in correlation across and within clusters. This has not been analyzed in this depth before.

Statewide Analysis

In this section, we report the findings from sequences that are based on a random sample of 12,704 persons in 5,000 households. We limit the analysis to 12,704 persons because the computation of differences in the *dissimilarity score* between pairs of sequences requires a very large matrix that exceeds our local computational facilities. The analysis here uses a matrix that is 12,704 by 12,704 (=161,391,616) cells, containing the dissimilarity scores among sequences for each person we have in this sample. This matrix is then analyzed using the agglomerative nesting clustering method. We start with 12,704 sequences and group pairs of sequences based on their dissimilarity scores. Then, we compare all the cluster dissimilarity score averages to each other and lump together clusters with smaller dissimilarity cores. We proceed until all observations are in one cluster (Kaufman & Rousseeuw, 1990). This process can be thought of as a tree that starts with every sequence as an individual "leaf" and ends with one cluster as the "trunk." After inspecting the overall time of day patterns, we select the nine-cluster solution because it shows clear representations of time of day time allocation patterns to places/activities. Below, we use the words "cluster" and "daily pattern" interchangeably. Figure 5 shows these nine distinct patterns. The cluster names are based on the daily behavior each cluster represents. For each of these nine patterns, we study the membership in terms of sociodemographic characteristics of respondents using descriptive statistics (Table 6 and Table 7). We also expect these patterns to be correlated with the day of the week assigned to each respondent, and Table 8 shows this correlation.

Figure 5. Nine clusters of daily sequences of places and travel statewide

Table 7. Cluster/pattern membership in daily patterns by person characteristics

Daily Pattern	**Females**	**Workers**	**Students**	**Disabled**	**Weekend**
Home Day	50.55%	36.04%	15.75%	10.72%	36.49%
School Day	46.96%	2.60%	92.96%	2.51%	2.80%
Typical Work Day	45.05%	99.55%	1.78%	1.83%	7.71%
Errands Type 1	49.73%	46.29%	19.17%	5.24%	44.85%
Mostly Out of Home	49.17%	50.09%	18.37%	5.01%	31.54%
Errands Type 2	43.58%	52.20%	14.70%	7.60%	29.73%
Non-typical Work Day	32.97%	100.00%	9.19%	2.70%	19.46%
Leave Home	47.60%	48.50%	22.75%	6.89%	33.23%
Traveling	50.27%	45.60%	32.69%	4.67%	27.20%

Table 8. Cluster/pattern membership in daily patterns by age group of respondent

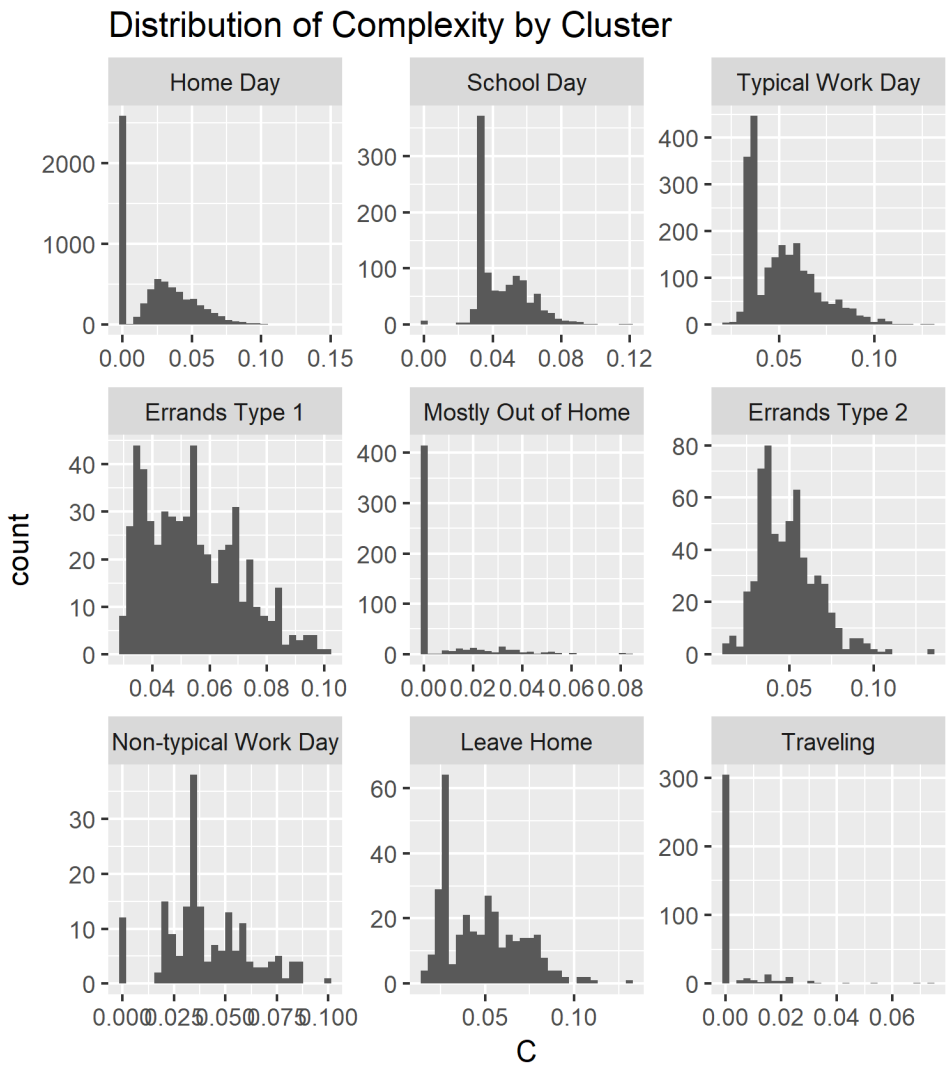
	Age00-03	*Age04-15*	*Age16-18*	*Age19-24*	*Age25-34*	*Age35-44*	*Age45-54*	*Age55-65*	*Age65+*	*Did not Tell*
Home Day	257	766	212	305	463	654	943	1412	1614	230
School Day	35	740	173	32	19	7	5	2	2	22
Typical Work Day	0	0	9	102	272	408	611	618	141	83
Errands Type 1	15	70	19	29	31	64	87	121	93	24
Mostly Out of Home	16	57	21	31	41	53	80	115	95	30
Errands Type 2	24	68	13	25	37	71	120	136	83	15
Non-typical Work Day	0	0	10	33	26	38	32	34	6	6
Leave Home	6	47	20	32	20	35	56	59	52	7
Traveling	16	87	21	14	31	42	50	52	33	18
Total	369	1835	498	603	940	1372	1984	2549	2119	435

Table 9. Cluster/pattern membership in daily patterns by day of the week

Daily Pattern	Monday	Tuesday	Wednesday	Thursday	Friday	Saturday	Sunday	Total
Home Day	13.19%	12.57%	13.14%	12.43%	12.18%	17.97%	18.52%	100.00%
School Day	16.49%	19.67%	17.36%	22.37%	21.31%	1.45%	1.35%	100.00%
Typical Work Day	16.22%	20.45%	21.30%	19.39%	14.93%	4.28%	3.43%	100.00%
Errands Type 1	11.39%	10.67%	9.76%	10.67%	12.66%	24.95%	19.89%	100.00%
Mostly Out of Home	10.58%	14.47%	13.54%	15.03%	14.84%	14.47%	17.07%	100.00%
Errands Type 2	16.39%	13.85%	11.99%	14.36%	13.68%	10.81%	18.92%	100.00%
Non-typical Work Day	17.30%	14.05%	15.68%	15.14%	18.38%	14.05%	5.41%	100.00%
Leave Home	11.38%	12.28%	12.87%	14.37%	15.87%	19.46%	13.77%	100.00%
Traveling	11.26%	14.56%	15.93%	19.23%	11.81%	17.03%	10.16%	100.00%

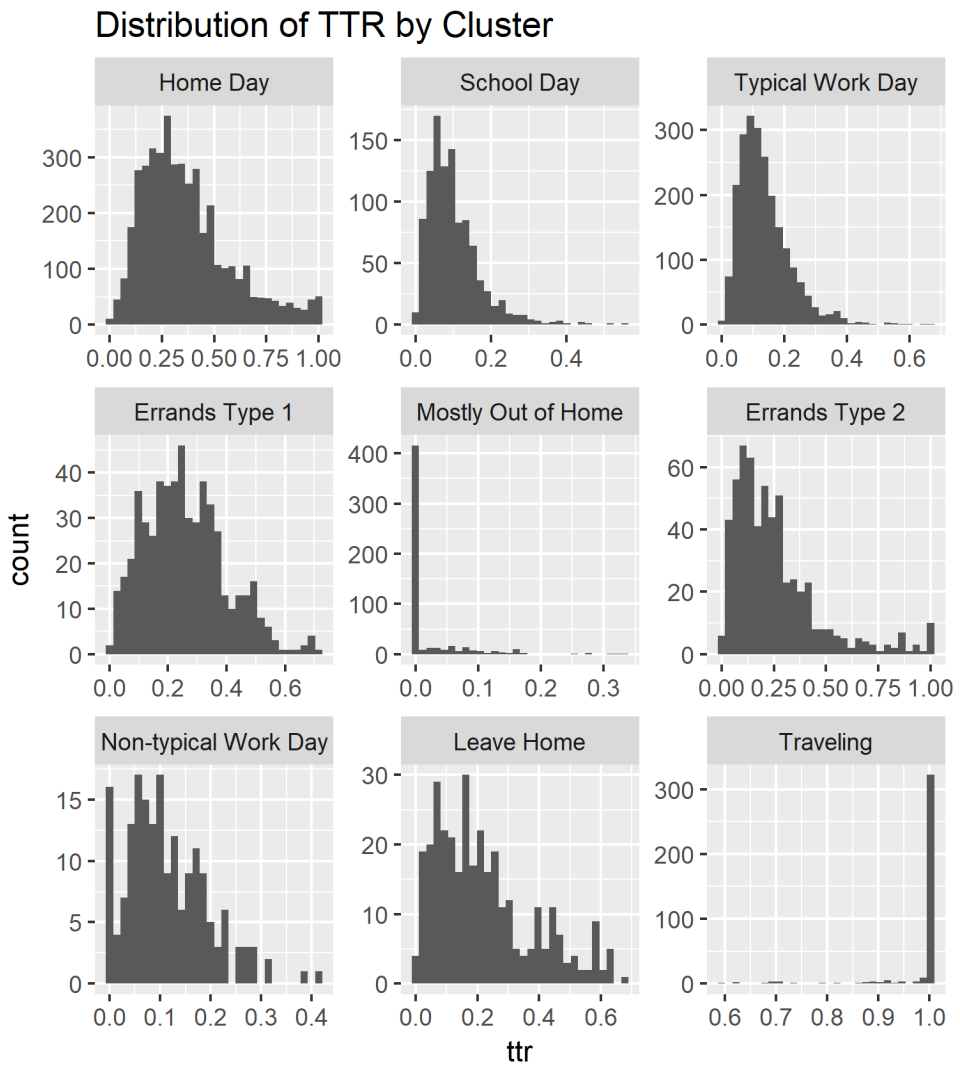
One of the key objectives in this project is to explore place-activity-travel fragmentation. As shown in the pilot analysis using the Central Coast data, the indicator named Complexity (C(s) in Equation 4) is sufficient as an indicator of sequence fragmentation. Figure 6 shows the histograms of the C(s) values for each of the nine daily patterns.

Figure 6. Complexity C(s) histograms by daily pattern



In this analysis, we also use the travel time ratio (Dijst and Vidakovic, 2000, Schwanen and Dijst, M., 2002, Susilo and Dijst, 2009 and 2010, Milakis et al., 2015, Dharmowijoyo et al., 2016, Milakis and van Wee, 2018). Travel time ratio (TTR) is a compact indicator to represent trade-offs of people between travel and activity time. In this report, TTR is defined as the total travel time in a day divided by the sum of the total time outside the home plus the total travel time in a day. In this way, we can study the percent of time for travel over the total length of Hagerstrand's time-space prism, which is the time elapsed between the first departure from home and arrival at home at the end of the day. We modify this to fit patterns with no home stay and use total amount of time that is not home in the denominator of TTR. Figure 7 shows the histograms of the TTR for each of the nine daily patterns derived here.

Figure 7. Travel Time Ratio (TTR) by daily pattern



Similar to Table 5 we did for the Central Coast, we estimate linear regression models for indicator C (Table A.1) and TTR (Table A.2) for each of the nine patterns to explain within-cluster variation in fragmentation and TTR. Table 9 contains descriptive statistics of the Complexity and TTR used as dependent variables in the models of Appendix A.

Table 10 shows the number of trips in each cluster and the modal splits within each cluster. The total number of trips for this sample is 41,175, corresponding to 3.24 trips per person per day.

Home Day Cluster

The most populous cluster is the *Home Day*, with 54% (6856 persons) of the sample selecting this pattern. These persons spent most of their time at home and a few of them (4,265) travel to other places. This is also the second most popular pattern for weekends (Table 6). Notable is the slightly more than half of the persons in this cluster being women, and 11% disabled persons (compared to 7% of total sample being disabled), reflecting movement restrictions for this group. Figure 6 shows the composition of this cluster clearly, with a substantial number having zero complexity because they stayed at home all day. This contributes to the average complexity (Table 9) being the third lowest among the nine patterns. This pattern also has the second highest travel time ratio indicating that for persons who left home, 36.8% of the time spent outside home was for travel. The average number of trips per person day in this pattern is 2.76 (lower than the overall average), and based on Table 10, 83.41% of these trips are by private motorized means with 37.38% driving alone.

In Appendix A, the regression models for complexity (Table A.1) show that the presence of children in the age group 4 to 15 years old contributes positively to higher fragmentation. In contrast, living in the suburbs, exurbs, or rural areas is an inhibitor to fragmentation when compared to the center of the city that has higher density of opportunities. Females also have higher fragmentation than males in this pattern.

The travel time ratio regression (Table A.2) shows that from among the persons that have out of home activities in this cluster, rural residents have a 0.06 higher ratio than center city dwellers. Exactly the opposite happens when the respondent is a child below 15 years old, and students have 5% lower TTR that non-students. Table A2 also shows clearly the impact of children of any age in the household making the TTR between 0.02 and 0.04 higher than persons with no children in the household.

Typical Work Day Cluster

The second most populous pattern is the *Typical Work Day* pattern (2,244 people representing 17.7% of the sample) that shows usual morning and afternoon peaking of work with a noon break for lunch. As expected, 99.55% of respondents in this pattern are employed persons (Table 6), and no children display this pattern (Table 7). Weekdays make up the majority of the days of the week this pattern occurs on, with small percentages on Saturday and Sunday (Table 8). This pattern has the second highest fragmentation (Table 9), and 13.6% of the out of home time is travel time. Table A1 in Appendix A shows the presence of children in the age groups 4 to 15 and 16 to 18 years old are correlated with higher fragmentation, but higher car ownership is negatively correlated with fragmentation. This shows the impact of decreased constraints for persons in households with more cars. Senior residents and person in the very low levels of poverty are also more likely to fragment their place-travel less than other groups. The presence of children in the age group 4 to 15 is positively correlated with the travel time ratio presumably needing rides to different places. Being a student is also positively correlated with

TTR, indicating the need for students to travel to work and other non-work activities and therefore having a higher TTR by 0.04 than non-students on a typical work day. People in this cluster make on average 4.24 trips per day, of which 61.44% are driving alone, second only to the non-typical work day discussed later.

Table 10. By-cluster complexity and travel time ratio (TTR)

	Mean C(s)	Std. Dev C	Mean TTR	Std. Dev TTR
Home Day	0.024	0.024	0.368	0.213
School Day	0.045	0.015	0.100	0.070
Typical Work Day	0.052	0.017	0.136	0.081
Errands Type 1	0.054	0.016	0.257	0.140
Mostly Out of Home	0.007	0.014	0.019	0.048
Errands Type 2	0.049	0.018	0.253	0.208
Non-typical Work Day	0.041	0.020	0.113	0.080
Leave Home	0.049	0.022	0.220	0.159
Traveling	0.003	0.009	0.984	0.060

Table 11. By-cluster number of trips and modal split

	vehicle driving alone_ratio	vehicle driving others_ratio	vehicle passenger_ratio	other motorized_ratio	Total	Trips/Person
Home Day	37.38%	21.87%	23.45%	0.71%	83.41%	2.76
School Day	7.26%	2.44%	61.00%	2.22%	72.91%	3.52
Typical Work Day	61.44%	13.39%	6.26%	1.83%	82.92%	4.24
Errands Type 1	26.83%	24.69%	33.28%	0.72%	85.52%	5.24
Mostly Out of Home	20.67%	29.92%	30.51%	4.53%	85.63%	1.52
Errands Type 2	39.33%	17.17%	23.84%	1.34%	81.68%	4.41
Non-typical Work Day	64.59%	7.94%	9.33%	1.21%	83.07%	3.13
Leave Home	30.83%	22.28%	31.88%	1.95%	86.95%	3.99
Traveling	27.48%	17.90%	31.04%	2.46%	78.88%	3.24
	bike_ratio	walk_ratio	transit_ratio	other non-motorized_ratio	other_ratio	Total
Home Day	1.63%	12.11%	2.36%	0.50%	0.00%	16.59%
School Day	2.33%	16.43%	8.03%	0.30%	0.00%	27.09%
Typical Work Day	2.07%	11.05%	3.88%	0.04%	0.03%	17.08%
Errands Type 1	1.07%	9.52%	3.83%	0.07%	0.00%	14.48%
Mostly Out of Home	1.38%	11.22%	0.59%	0.39%	0.79%	14.37%
Errands Type 2	1.15%	11.46%	4.91%	0.15%	0.65%	18.32%
Non-typical Work Day	2.42%	10.71%	3.80%	0.00%	0.00%	16.93%
Leave Home	1.20%	7.80%	2.93%	0.00%	1.13%	13.05%
Traveling	2.97%	13.06%	3.82%	0.17%	1.10%	21.12%

School Day Cluster

The third most populous pattern is the *School Day* (1,037 persons and 8.2% of the sample). Table 6 shows that 92.56% of the persons in this cluster are persons classified in the survey as full-time students. Table 7 shows the majority of these persons are age 4 to 15 (740 persons) and age 16 to 18 (173 persons) and this pattern is typical of weekdays (Table 8) with a very small percentage on Saturday (1.45%) and Sunday (1.35%). Figure 6 shows that we have two groups of people in this cluster: (a) a group that has the same low complexity in their schedule and (b) another with high variety. Table A.1 shows that the presence of children in ages 4 to 15 and 16 to 18 increases the need to fragment schedules (presumably, children accompanying

other children of the household in different activity locations). The travel time ratio, however, is only 10%, and for children in the age group below 15 years old is even lower. This indicates that both the school location and the other activity locations are most likely in close proximity. Typical of the persons in this cluster is their modal split with 61.00% riding cars as passengers, 8.03% as transit passengers (the highest among all clusters), and 16.43% walking (also the highest among all clusters).

Errands Day Clusters

The next two daily patterns are by persons that visit places classified as *other,* and we name them *Errands Type 1* (553 person and 4.4% of the sample) and *Errands Type 2* (592 persons and 4.7% of the sample). Both have patterns reaching a peak of visiting places other than home, work, or school, and they both have substantial amounts of traveling. The major difference between the two clusters is the time of day the peak is reached. Both patterns have more men than women with *Errands Type 1* having a higher number of women, but still lower than men. *Errands Type 1* is also the preferred pattern for weekends (44.85% in Table 6), with Saturday getting almost a quarter of the persons in this pattern (Table 8). Both daily patterns have high fragmentation and high within-cluster average TTR of 25.7% and 25.3% (Table 9). The values of fragmentation for both clusters are spread substantially (Figure 6). The same is true for their TTR (Figure 7). The *Errands Type 1* daily pattern has the highest average number of trips per person (5.24 in Table 10), and an almost even spread in the use of private cars, but still 85.52% of trips are made by private motorized modes. In contrast, *Errands Type 2* has a much lower number of trips per person (4.41 in Table 10), a higher driving alone ratio (39.33% in Table 10), but still high private car modal split.

For *Errands Type 1*, complexity is lower when this pattern happens on weekends, by senior residents, when the household has children 16 to 18 years old, and by students (Table A.1). In contrast, fragmentation is higher for households that have children age 4 to 15 and live in the exurbs. For *Errands Type 2* we also see lower fragmentation in weekends, higher fragmentation for households that have children age 4 to 15 and higher fragmentation when the respondent is female or a worker (Table A.1). The TTR is higher in *Errands Type 1* for households that have children age 4 to 15, and substantially higher for residents in exurbs and rural areas when compared to their counterparts in suburbs and center of a city (Table A.2). The *Errands Type 2* TTR variation is not explained by any of the variables we tested (Table A.2).

Mostly Out of Home Cluster

The next daily pattern is *Mostly Out of Home*, with 539 persons (4.2% of the sample). This reflects the definition of places as *other,* which includes second homes, hotels, camping grounds, etc. that could not be assigned as the primary home location. It is also the third most popular pattern for weekends (31.54% in Table 6). This pattern has the second lowest complexity and lowest TTR (Table 9). Tables A1. and A.2 shows the inhibiting role of very young children in fragmentation for this pattern and the even lower TTR for senior residents in this

pattern. The histograms in Figures 5.2 and 5.3 also reflect this lack of variation and therefore inability of models to find significant factors affecting fragmentation and TTR. This cluster has the lowest number of trips among all patterns and the highest driving others percentage of trips (Table 10).

Non-Typical Work Day Cluster

The second working day pattern is the *Non-Typical Work Day*, and it is the least populous with 185 respondents (1.5% of the sample). This is an interesting daily pattern because it is entirely made up of workers that show starting and ending times of work that span the entire 24 hour interview period. It is more populated by men (about 67% in Table 6), spread throughout the age groups over 15 years old, and as shown in Table 8, a substantial portion of the cluster responded on Fridays and Saturdays (unlike the other typical work day daily pattern that has very few people on Saturday). This pattern shows substantial fragmentation, but low TTR (Table 9), presumably due to workers living close to the workplace and/or spending longer hours at work. Differences in fragmentation within this pattern are only due to the presence of very young children in the household and females having higher fragmentation than males. The TTR ratio is lower for suburban and exurban residents when compared to the center city and rural dwellers. Higher car ownership also decreases the TTR. This pattern shows lower than average number of trips per person at 3.13 trips and is the pattern with the most driving alone trips at 64.59%.

Leave Home Cluster

The next pattern, *Leave Home* (334 persons and 2.6% of the sample), is characteristic of persons that stayed at a location classified as *other* with substantial traveling. 48.5% in this group are workers (in fact this pattern contains some activity at workplaces) and 22.75% are students. This pattern also shows substantial fragmentation and substantial TTR reflecting the traveling far from home component of the pattern. Figure 6 shows that we most likely have two groups of people in this pattern: a) one that leaves home and does not do a lot where they arrive; and b) a group that participates in multiple visits to places. Figure 7 shows a wide spread of TTRs within this group. The regression models in Appendix A show that disability, poverty, and residing in the exurbs inhibit fragmentation (Table A.1). The TTR ratio regression in Table A.2 show substantial differences due to the presence of children 4 to 18 years old increasing TTR by 6%, exurban and rural living increasing TTR by 6% and 9% respectively. In contrast, students have a TTR 11% lower than non-students. Table 10 shows this pattern has higher than average number of trips per day and a substantial portion of them by car as a driver or passenger.

Traveling Cluster

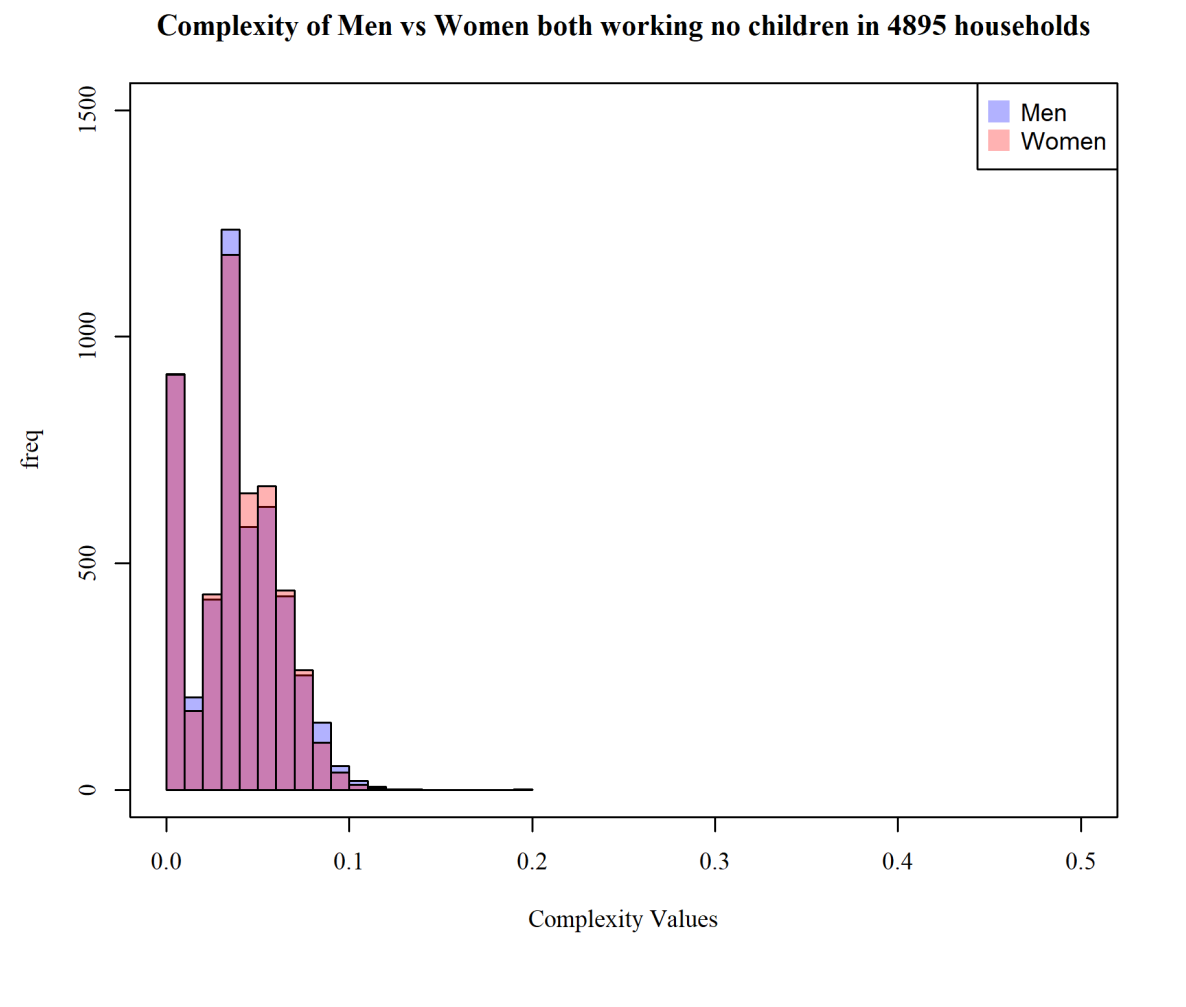
This is a pattern characterized by mostly travelers of all ages and 50.27% females (364 persons and 2.9% of the sample). The pattern is spread throughout the days of the week with the highest percentage on Thursday and lowest on Sunday (Table 8). Reflective of this pattern is

the lowest fragmentation of 0.003 and highest TTR 98.4% (Table 9 and histogram of Figure 6 and Figure 7). The only variable that increases complexity is if the respondent is a child younger than 15 years old. There is a decrease for persons in households with children 4 to 15 years old. TTR for this pattern is higher on the weekend days and for persons in households with children younger than 15 years old. In contrast, people in poverty and child respondents younger than 15 years old have lower TTR. The number of trips in this cluster is exactly at the overall average number of trips with the highest proportion of trips as passengers in a car and the highest bike share (2.97% in Table 10).

Fragmentation within Households

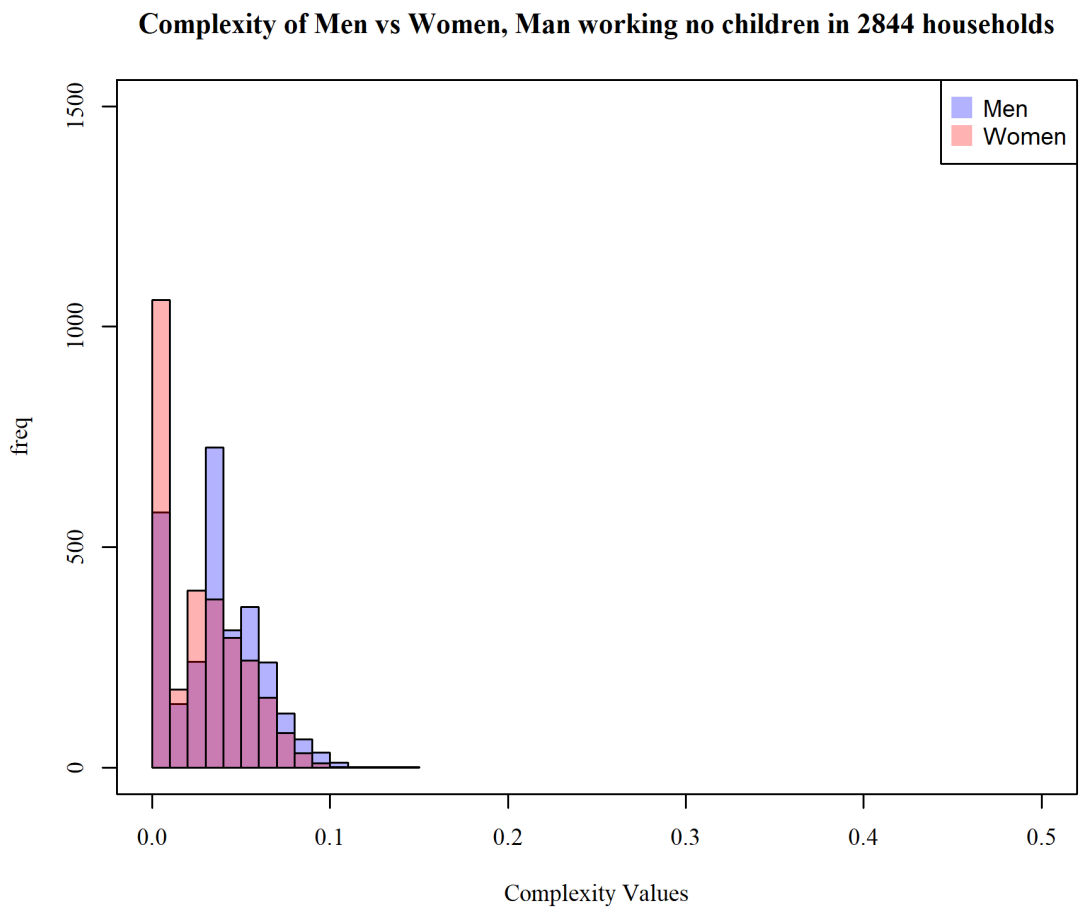
In this section, we explore fragmentation of daily schedules of households that are made by an adult man and an adult woman with and without children. Although we found in sections 4 and 5 that men and women have different time of day activities and travel for some patterns, we need to verify if men and women within the same households have different fragmentation in their daily patterns. This will show if the women are different in their daily schedule than men in the same household when they are employed and when they have children. We are using the full CHTS with observations that were complete enough to build place-travel sequences. We have a total 114,639 persons in 45,362 households. Of these households, 4,895 are adult couples with both spouses working, 2,844 adult couples in which only the man is a worker, and 2,142 adult couples in which only the woman works. Figures 6.1, 6.2, and 6.3 show the three histograms of men versus women for couples with no children. From these figures and values of the C(s), we see that working men and women have high schedule complexity and when we examine couples in which the man works but not the woman, the woman has lower schedule complexity (fragmentation). Exactly the opposite happens when the woman works and the man does (Figure 10), displaying a reversal of roles (at least in terms of fragmentation).

Figure 8. Couples with both working



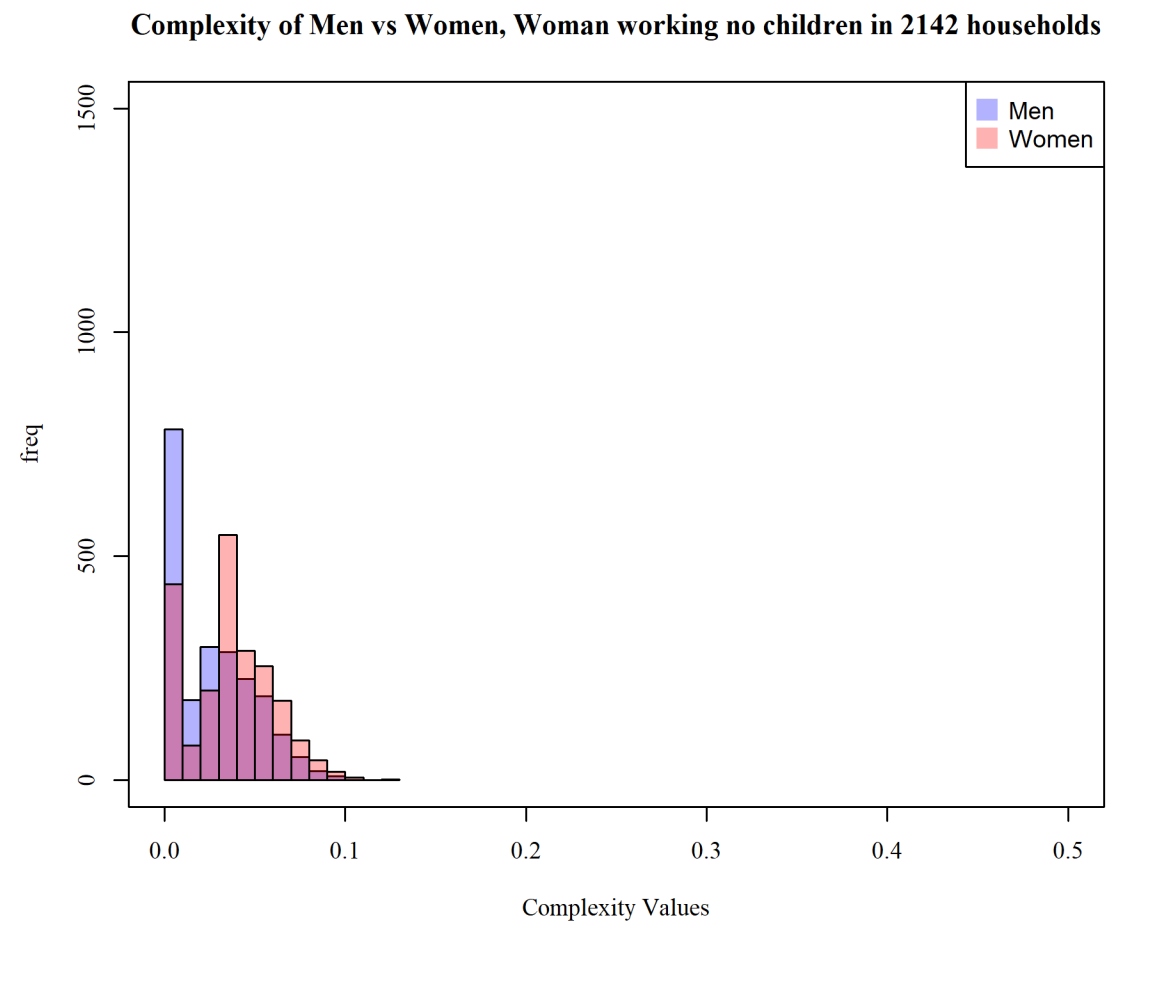
Average C(s) for men (standard deviation) = 0.04 (0.02)
Average C(s) for women (standard deviation) = 0.04 (0.02)

Figure 9. Couples with only the man working



Average C(s) for men (standard deviation) = 0.04 (0.02)
Average C(s) for women (standard deviation) = 0.03 (0.02)

Figure 10. Couples with only the woman working



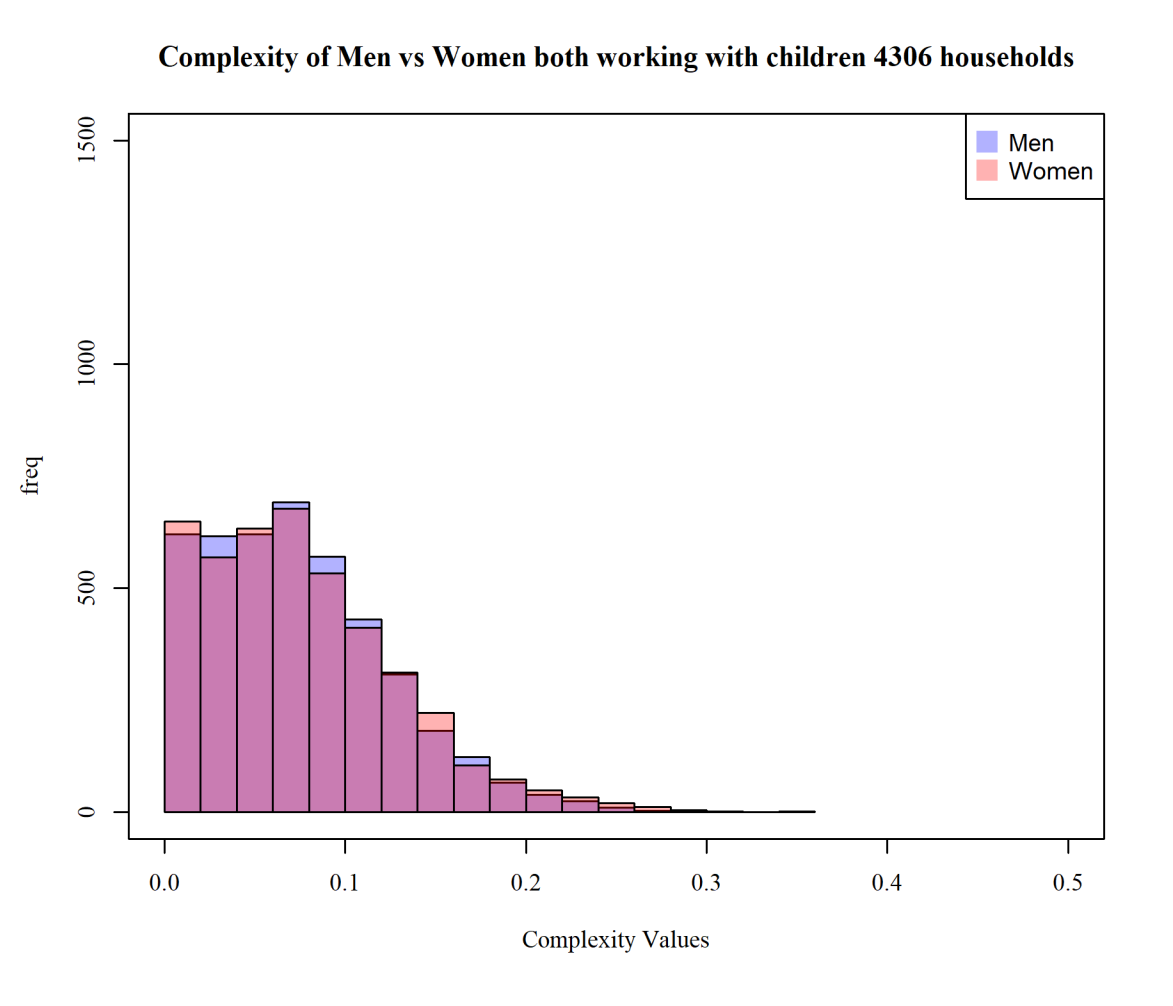
Average C(s) for men (standard deviation) = 0.03 (0.02)
Average C(s) for women (standard deviation) = 0.04 (0.02)

We turn now to households containing two adults of different gender with children, which we are identifying as families, in which both adults work (Figure 11), families in which only the adult man works (Figure 12) and families in which only the adult woman works (Figure 13).

Fragmentation of schedules is by far higher for families in which both men and women work (Figure 11), and by far more variable (high standard deviation in addition to a wider spread of the histogram). Women, however, in this case have on average higher values of C(s) than men. This is an indication of household responsibility hypothesis conforming the findings in the previous sections.

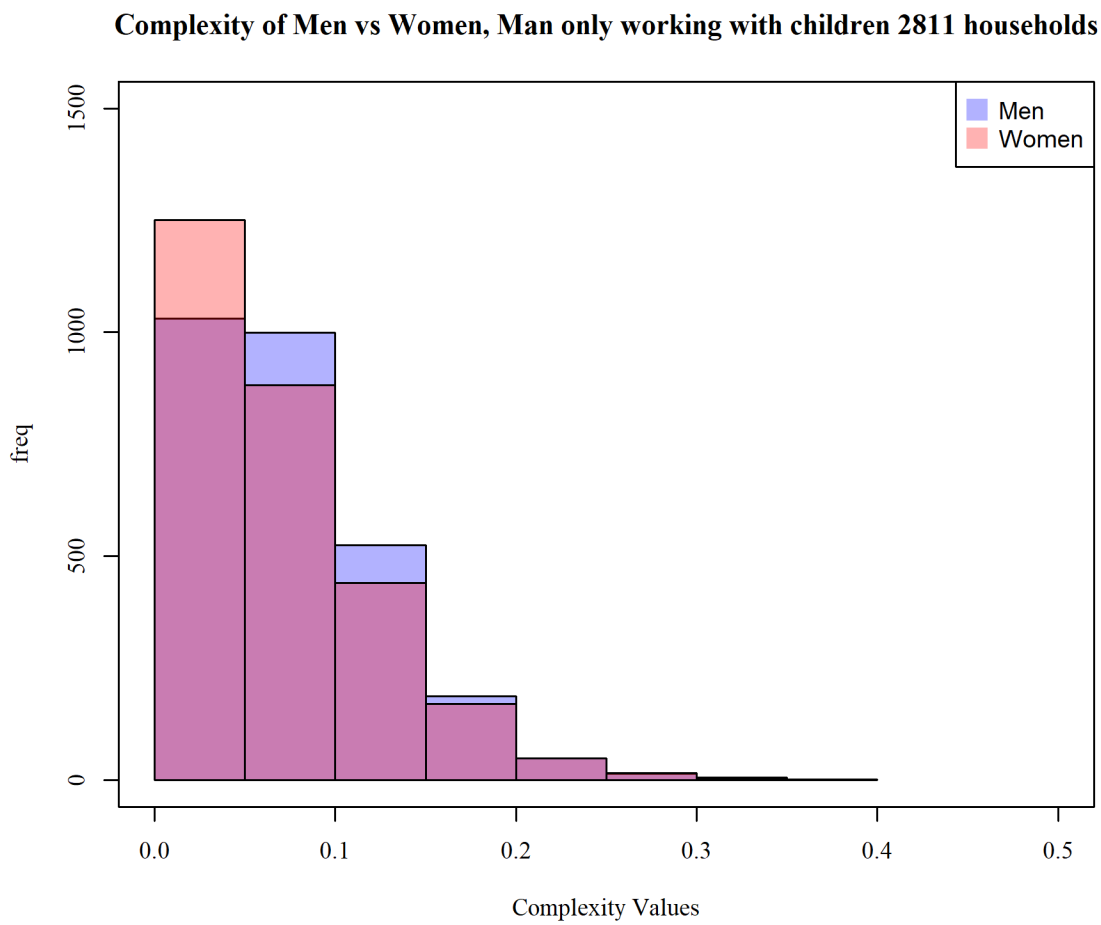
Figure 12 shows C(s) for couples with children in which only the man works and Figure 13 shows the C(s) for couples with children in which only the woman works. Unlike the couples without children, this time there is no reversal in fragmentation, with women having consistently high fragmentation and variability of this fragmentation, and often higher than men independently of their employment status. This result further strengthens the household responsibility hypothesis and the role children play in motivating schedule fragmentation.

Figure 11. Adult couple with children, both adults work



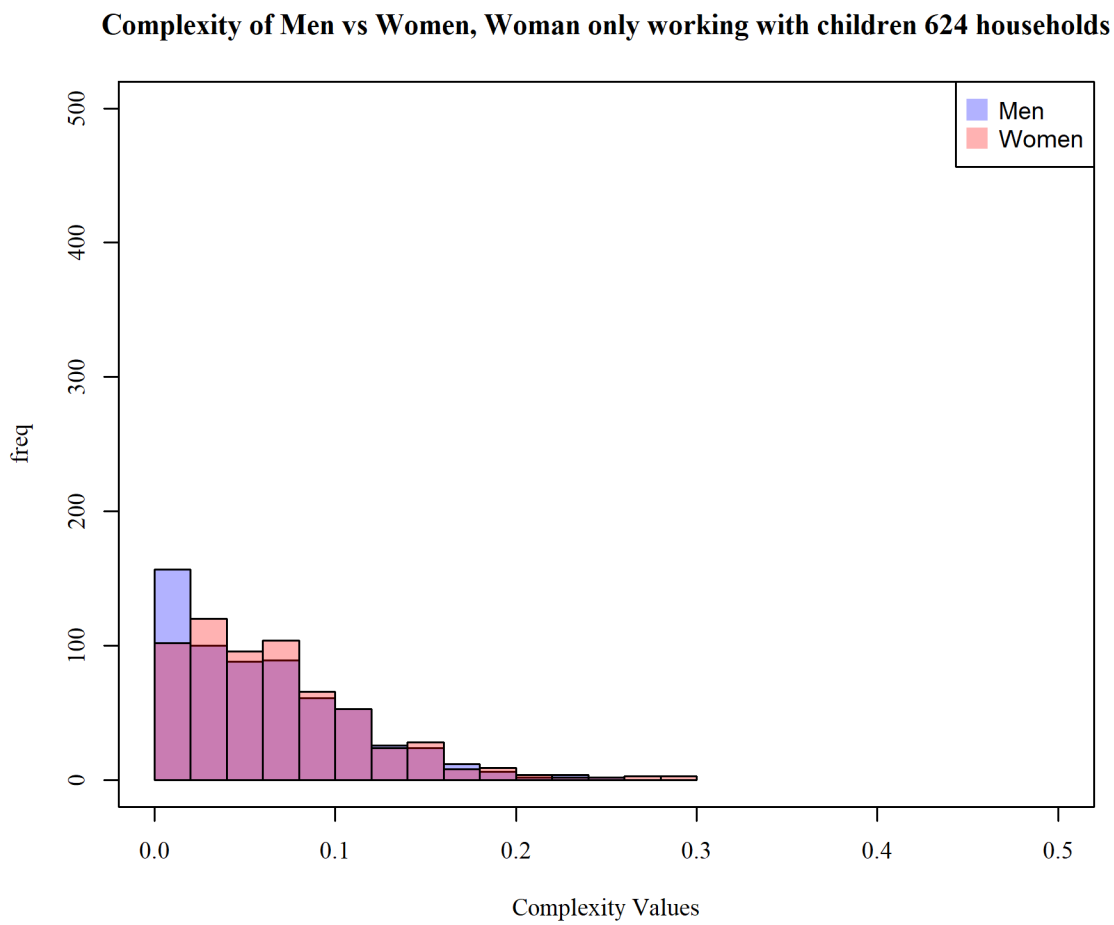
Average C(s) for men (standard deviation) = 0.07 (0.05)
Average C(s) for women (standard deviation) = 0.08 (0.05)

Figure 12. Adult couple with children, only man works



Average C(s) for men (standard deviation) = 0.07 (0.05)
Average C(s) for women (standard deviation) = 0.07 (0.06)

Figure 13. Adult couple with children, only woman works



Average C(s) for men (standard deviation) = 0.06 (0.05)
Average C(s) for women (standard deviation) = 0.07 (0.06)

Summary of Findings and Travel Behavior Research

This project uses a new type of fragmentation in activity-travel behavior. We focus first on place-travel sequences and then report on the activities at each place. The analysis is first done on data from SLO-SB to test the methods and then expanded using a random sample statewide. The findings in the pilot section using data from SLO and SB are encouraging because they confirm findings using other methods to study activities and their durations, but we also find these new techniques provide new insights about scheduling activities. One key finding is the two clusters of daily activities (*Traveling* and *Return Home*) that are, in essence, absent from contemporary activity-based models. These are not visitors to special travel generators such as events and hotels. These are residents of the study area that either were absent from the study area throughout the period of observation, left the area for a long-distance trip, or came back to the study area from a long distance trip. This has implications for synthetic population generation. When synthetic population is used to generate the entirety of residents of a study region, we also need to account for the proportion of this population that will have activity and travel behaviors of *Traveling* and *Return Home* clusters. This exploratory analysis finds that employment and education status are key determinants of daily schedules. It also shows the number of children at different ages play different roles within each of the six clusters of sequences used here. Overall, however, the presence of children in the household increases the complexity of place-travel daily patterns. Poverty emerges as an important determinant of daily patterns and requires further scrutiny together with car ownership, car availability, and public transportation services. This is of particular importance with the emerging mobility services as discussed in other chapters of this book. In addition, the analysis here confirms the household responsibility hypothesis for women.

Using a larger sample of 12,704 persons from 5,000 households spread throughout California yields nine distinct daily patterns. These include patterns of people staying at home for long periods in a day, people that follow typical daily working schedules and typical school schedules. We also find people that travel for an entire day and people that stay at home in the morning but then travel for the rest of their interview day. We also have two patterns of running errands with very different time of day rhythms. The ninth pattern is by people that spent most of their time in a day at locations that are not home, work, or school and travel for very short time. Each pattern also has different memberships in terms of gender, age, and day of the week (in addition to the working and/or student status as expected).

We also make comparisons between men and women that live in the same household and find that in couples with no children, employment status influences fragmentation of activities in such a way that the employed person has a more fragmented schedule. Men and women that are not employed but their partners are appear to be having similar fragmentation values. When we examine adult couples with children, women have consistently more fragmented schedules than men in the same household regardless of employment status. All this further

strengthens the household responsibility hypothesis for women who, in addition to work outside the home, also run a variety of errands, and for this reason need to visit multiple places in a day.

From a land use and transportation viewpoint, if more people moved to dense urban environments and adapted similar lifestyles as the observed data here, we should expect them to have more fragmented schedules during the *Home Day*, but not major differences for all the other patterns. The added flexibility of Mobility as a Service (MaaS) integrating different services may better serve the higher fragmentation patterns we found here (i.e., *Typical Work Day, Errands Type 1, and Errands Type 2*). But, in order to do this, MaaS will need to become a suitable alternative to and compete successfully with the private car that offers the flexibility to give rides to other people.

References

Abbott, A., & Tsay, A. (2000). Sequence analysis and optimal matching methods in sociology: Review and prospect. *Sociological Methods and Research*, *29*(1), 3–33. https://doi.org/10.1177/0049124100029001001

Auld, J., Rashidi, T., Javanmardi, M., & Mohammadian, A. (2011). Dynamic Activity Generation Model Using Competing Hazard Formulation. *Transportation Research Record: Journal of the Transportation Research Board*. https://doi.org/10.3141/2254-04

Bhat, C. R., Goulias, K. G., Pendyala, R. M., Paleti, R., Sidharthan, R., Schmitt, L., & Hu, H. H. (2013). A household-level activity pattern generation model with an application for Southern California. *Transportation*. https://doi.org/10.1007/s11116-013-9452-y

Bhat, C. R., & Pinjari, A. R. (2000). Duration modeling. In *Handbook of transport modelling* (6th ed., pp. 91–111).

Chen, Y., Ravulaparthy, S., Deutsch, K., Dalal, P., Yoon, S., Lei, T., … Hu, H.-H. (2011). Development of Indicators of Opportunity-Based Accessibility. *Transportation Research Record: Journal of the Transportation Research Board*, *2255*, 58–68. https://doi.org/10.3141/2255-07

Couclelis, H. (2000). From Sustainable Transportation to Sustainable Accessibility: Can We Avoid a New Tragedy of the Commons? In *Information, Place, and Cyberspace* (pp. 341–356). Berlin, Heidelberg: Springer.

Couclelis, H. (2004). Pizza over the internet: E-commerce, the fragmentation of activity and the tyranny of the region. *Entrepreneurship and Regional Development*. https://doi.org/10.1080/0898562042000205027

Crane, R. (2007). Is there a quiet revolution in women's travel? Revisiting the gender gap in commuting. *Journal of the American Planning Association*. https://doi.org/10.1080/01944360708977979

Dharmowijoyo, D. B., Susilo, Y. O., & Karlström, A. (2016). Relationships among discretionary activity duration, its travel time spent and activity space indices in the Jakarta Metropolitan Area, Indonesia. *Journal of Transport Geography*, *54*, 148-160.

Dijst, M., & Vidakovic, V. (2000). Travel time ratio: the key factor of spatial reach. *Transportation*, *27*(2), 179-199.

Elzinga, C. H., & Liefbroer, A. C. (2007). De-standardization of Family-Life Trajectories of Young Adults: A Cross-National Comparison Using Sequence Analysis. *European Journal of Population / Revue Européenne de Démographie*, *23*(3–4), 225–250. https://doi.org/10.1007/s10680-007-9133-7

Ettema, D. (1995). Competing risk hazard model of activity choice, timing, sequencing, and duration. *Transportation Research ….*

Gabadinho, A., Ritschard, G., Studer, M., & Muller, N. S. (2010). Mining sequence data in R with the TraMineR package: A user's guide. *Department of Econometrics and Laboratory of Demography, University of Geneva, Switzerland*.

Goulias, K. G. (1999). Longitudinal analysis of activity and travel pattern dynamics using generalized mixed Markov latent class models. *Transportation Research Part B: Methodological*. https://doi.org/10.1016/S0191-2615(99)00005-3

Joh, C. H., Arentze, T. A., & Timmermans, H. J. P. (2001). Multidimensional sequence alignment methods for activity-travel pattern analysis: A comparison of dynamic programming and genetic algorithms. *Geographical Analysis*, *33*(3), 247–270. Retrieved from https://doi.org/10.1111/j.1538-4632.2001.tb00447.x

Kaufman, L., & Rousseeuw, P. J. (1990). *Finding Groups in Data: An Introduction to Cluster Analysis (Wiley Series in Probability and Statistics)*. *eepe.ethz.ch*.

Kroesen, M. (2014). Modeling the behavioral determinants of travel behavior: An application of latent transition analysis. *Transportation Research Part A: Policy and Practice*. https://doi.org/10.1016/j.tra.2014.04.010

Kroesen, M., & Goulias, K. G. (2016). Modelling activity-travel behaviour dynamics with panel data: The state-of-the-art. *European Journal of Transport and Infrastructure Research*.

Leszczyc, P. T. L. P., & Timmermans, H. (2002). Unconditional and conditional competing risk models of activity duration and activity sequencing decisions: An empirical comparison. *Journal of Geographical Systems*. https://doi.org/10.1007/s101090200083

McBride, E. C., Davis, A. W., & Goulias, K. G. (2019). Fragmentation in Daily Schedule of Activities using Activity Sequences. *Transportation Research Record,* 2673, Issue 4, pp. 844–854. https://doi.org/0361198119837501

McBride, E. C., Davis, A. W., Lee, J. H., & Goulias, K. G. (2017). Incorporating Land Use into Methods of Synthetic Population Generation and of Transfer of Behavioral Data. *Transportation Research Record: Journal of the Transportation Research Board*. https://doi.org/10.3141/2668-02

McBride, E., Lee, J. H., Lundberg, A. M., Davis, A. W., & Goulias, K. G. (2016). Behavioural micro-dynamics of car ownership and travel in the Seattle Metropolitan Region from 1989 to 2002. *European Journal of Transport and Infrastructure Research*, *16*(4).

Milakis, D., Cervero, R., Van Wee, B., & Maat, K. (2015). Do people consider an acceptable travel time? Evidence from Berkeley, CA. *Journal of Transport Geography*, *44*, 76-86.

Milakis, D., & van Wee, B. (2018). "For me it is always like half an hour": Exploring the acceptable travel time concept in the US and European contexts. *Transport Policy*, *64*, 113-122.

NUSTATS. (2013). *2012 California Household Travel Survey Final Report*. Austin, TX. Retrieved from http://www.dot.ca.gov/hq/tpp/offices/omsp/statewide_travel_analysis/Files/CHTS_Final_Report_June_2013.pdf

Přibyl, O., & Goulias, K. (2007). Simulation of Daily Activity Patterns Incorporating Interactions Within Households: Algorithm Overview and Performance. *Transportation Research Record: Journal of the Transportation Research Board*. https://doi.org/10.3141/1926-16

Schwanen, T., & Dijst, M. (2002). Travel-time ratios for visits to the workplace: the relationship

between commuting time and work duration. *Transportation Research Part A: Policy and Practice*, *36*(7), 573-592.

Susilo, Y. O., & Dijst, M. (2009). How Far Is Too Far? Travel Time Ratios for Activity Participation in the Netherlands. *Transportation Research Record*, *2134*(1), 89-98.

Susilo, Y. O., & Dijst, M. (2010). Behavioural decisions of travel-time ratios for work, maintenance and leisure activities in the Netherlands. *Transportation planning and technology*, *33*(1), 19-34.

Turner, T., & Niemeier, D. (1997). Travel to work and household responsibility: New evidence. *Transportation*. https://doi.org/10.1023/A:1004945903696

Wilson, C. (1998a). Activity pattern analysis by means of sequence-alignment methods. *Environment and Planning A*. https://doi.org/10.1068/a301017

Wilson, C. (1998b). Analysis of Travel Behavior Using Sequence Alignment Methods. *Transportation Research Record*. https://doi.org/10.3141/1645-07

Data Management Plan

Products of Research
No new data were collected in this study. The California Household Travel Survey (CHTS) was used and is available at the Transportation Secure Data Center of the National Renewable Energy Laboratory (NREL) (https://www.nrel.gov/transportation/secure-transportation-data/index.html)

Data Format and Content
Data format and the contents of each file are available at the Transportation Secure Data Center of the National Renewable Energy Laboratory (https://www.nrel.gov/transportation/secure-transportation-data/index.html).

Data Access and Sharing
The general public can access the data through the Transportation Secure Data Center of the National Renewable Energy Laboratory (https://www.nrel.gov/transportation/secure-transportation-data/index.html)

Reuse and Redistribution
There are no restrictions on how the data can be reused and redistributed by the general public subject to the permissions mentioned on the website of NREL.

Appendix—Correlation with Sociodemographics of Complexity and Travel Time Ration Statewide

Table A.1 By-cluster complexity statewide

	Cluster Type Dependent Variable								
	Complexity C(s) (Eq. 4)								
	Home Day	School Day	Typical Work Day	Errands Type 1	Mostly Out of Home	Errands Type 2	Non-typical Work Day	Leave Home	Traveling
Constant	0.03 *t = 27.15****	0.04 *t = 15.94****	0.04 *t = 8.16****	0.06 *t = 23.40****	0.01 *t = 4.63****	0.05 *t = 17.88****	0.04 *t = 9.72****	0.05 *t = 13.50****	0.004 *t = 2.28***
Disability	-0.01 *t = -8.80****	-0.004 *t = -1.22*	-0.003 *t = -0.93*	-0.005 *t = -1.50*	-0.001 *t = -0.50*	-0.0003 *t = -0.09*	-0.002 *t = -0.26*	-0.01 *t = -1.82**	-0.002 *t = -0.81*
Household Income Near or Below Poverty Line	-0.002 *t = -2.26***	-0.003 *t = -1.74**	-0.01 *t = -5.10****	-0.002 *t = -0.81*	-0.002 *t = -0.81*	-0.003 *t = -1.07*	-0.01 *t = -1.03*	-0.01 *t = -2.15***	0.002 *t = 1.05*
Weekend	-0.003 *t = -5.66****	0.001 *t = 0.37*	-0.003 *t = -2.43***	-0.004 *t = -3.11****	0.0001 *t = 0.04*	-0.004 *t = -2.39***	0.002 *t = 0.60*	-0.003 *t = -1.25*	-0.001 *t = -0.95*
Respondent is Under 15 Years Old	-0.003 *t = -2.92****	-0.001 *t = -0.69*		-0.002 *t = -0.62*	-0.001 *t = -0.29*	-0.002 *t = -0.67*		-0.003 *t = -0.59*	0.003 *t = 1.71**
Respondent is Over 65 Years Old	-0.002 *t = -3.09****	0.03 *t = 2.87****	-0.003 *t = -1.95**	-0.004 *t = -2.21***	-0.004 *t = -2.34***	-0.001 *t = -0.25*	-0.003 *t = -0.30*	-0.01 *t = -1.55*	-0.001 *t = -0.45*
Presence of Children Under 4	-0.002 *t = -1.99***	-0.001 *t = -0.91*	0 *t = 0.02*	-0.001 *t = -0.45*	-0.005 *t = -2.21***	0 *t = 0.01*	0.01 *t = 1.92**	-0.003 *t = -0.68*	0.0003 *t = 0.24*
Presence of Children Age 4 to 15	0.004 *t = 5.55****	0.002 *t = 0.98*	0.004 *t = 5.34****	0.004 *t = 2.27***	-0.002 *t = -1.23*	0.004 *t = 1.95**	-0.003 *t = -0.81*	-0.001 *t = -0.42*	-0.002 *t = -1.96**
Presence of Children Age 16 to 18	-0.001 *t = -1.15*	0.001 *t = 0.59*	0.003 *t = 2.49***	-0.01 *t = -2.55***	-0.003 *t = -1.43*	0.002 *t = 0.89*	-0.002 *t = -0.47*	-0.01 *t = -1.46*	-0.002 *t = -1.30*
Female	0.002 *t = 3.17****	0.001 *t = 1.39*	0 *t = -0.06*	0.002 *t = 1.64*	0 *t = 0.04*	0.003 *t = 2.02***	0.01 *t = 2.16***	-0.002 *t = -0.72*	-0.0002 *t = -0.25*
Worker	0.004 *t = 6.62****	-0.001 *t = -0.32*	0.01 *t = 1.46*	-0.001 *t = -0.69*	-0.0001 *t = -0.07*	0.004 *t = 2.03***		0.003 *t = 1.17*	0.001 *t = 1.09*
Student	-0.003 *t = -2.98****	0.003 *t = 1.41*	0.005 *t = 1.73**	-0.01 *t = -2.22***	0.0004 *t = 0.17*	-0.003 *t = -0.95*	0.004 *t = 0.76*	0.002 *t = 0.51*	-0.001 *t = -0.92*
Number of Household Vehicles	0.0002 *t = 0.57*	0 *t = 0.05*	-0.001 *t = -2.92****	0.0001 *t = 0.11*	-0.001 *t = -1.09*	-0.001 *t = -1.09*	-0.001 *t = -0.43*	0.0001 *t = 0.08*	-0.0002 *t = -0.37*
Suburban Household	-0.002 *t = -3.55****	-0.0003 *t = -0.25*	0.001 *t = 0.88*	-0.001 *t = -0.80*	-0.001 *t = -0.94*	-0.0001 *t = -0.07*	-0.003 *t = -0.81*	0.004 *t = 1.56*	0.0005 *t = 0.41*
Exurban Household	-0.003 *t = -3.92****	-0.002 *t = -1.35*	-0.001 *t = -0.93*	0.004 *t = 1.83**	0.001 *t = 0.40*	-0.0004 *t = -0.17*	-0.01 *t = -1.22*	-0.01 *t = -2.24***	0.0003 *t = 0.17*
Rural Household	-0.01 *t = -5.92****	-0.005 *t = -2.13***	0.0001 *t = 0.09*	-0.0002 *t = -0.07*	0.003 *t = 1.31*	0.001 *t = 0.22*	-0.01 *t = -1.12*	-0.01 *t = -1.45*	-0.002 *t = -0.77*
Observations	6,853	1,037	2,243	553	537	592	185	334	364
R^2	0.05	0.03	0.04	0.08	0.04	0.04	0.08	0.11	0.04
Adjusted R^2	0.05	0.01	0.03	0.06	0.02	0.02	0.01	0.07	-0.003
Residual Std. Error	0.02 (df = 6837)	0.01 (df = 1021)	0.02 (df = 2228)	0.02 (df = 537)	0.01 (df = 521)	0.02 (df = 576)	0.02 (df = 171)	0.02 (df = 318)	0.01 (df = 348)
F Statistic	24.76*** (df = 15; 6837)	1.77** (df = 15; 1021)	6.44*** (df = 14; 2228)	3.22*** (df = 15; 537)	1.57* (df = 15; 521)	1.63* (df = 15; 576)	1.09 (df = 13; 171)	2.66*** (df = 15; 318)	0.92 (df = 15; 348)

*Note: *p<0.1; **p<0.05; ***p<0.01*

Table A.2 By-cluster TRR statewide

	Cluster Type								
	Dependent Variable								
	Travel Time Ratio (TTR)								
	Home Day	School Day	Typical Work Day	Errands Type 1	Mostly Out of Home	Errands Type 2	Non-typical Work Day	Leave Home	Traveling
Constant	0.42	0.14	0.13	0.28	0.03	0.26	0.17	0.2	1
	*t = 37.59****	*t = 11.98****	*t = 5.02****	*t = 13.08****	*t = 4.10****	*t = 8.27****	*t = 10.08****	*t = 6.96****	*t = 77.39****
Disability	-0.01	0.06	0.02	-0.01	-0.0004	0.05	-0.04	-0.02	0.01
	t = -0.67	*t = 4.65****	*t = 1.40*	*t = -0.47*	*t = -0.04*	*t = 1.34*	*t = -1.00*	*t = -0.70*	*t = 0.77*
Household Income Near or Below Poverty Line	0.01	0.01	-0.01	-0.02	-0.01	-0.03	-0.02	-0.01	-0.02
	t = 1.21	*t = 1.70**	*t = -1.21*	*t = -1.11*	*t = -0.61*	*t = -0.91*	*t = -1.04*	*t = -0.28*	*t = -1.97**
Weekend	-0.04	0.01	0.003	0.01	0.001	0.01	0.02	-0.02	0.01
	*t = -6.44****	*t = 0.82*	*t = 0.53*	*t = 0.50*	*t = 0.27*	*t = 0.65*	*t = 1.48*	*t = -0.84*	*t = 1.99***
Respondent is Under 15 Years Old	-0.06	-0.04		-0.01	-0.01	-0.06		-0.01	-0.03
	*t = -4.40****	*t = -5.42****		*t = -0.29*	*t = -1.18*	*t = -1.60*		*t = -0.18*	*t = -2.55***
Respondent is Over 65 Years Old	0.002	0.09	0.01	-0.01	-0.02	0.03	0.02	0.02	-0.003
	t = 0.24	*t = 1.94**	*t = 1.72**	*t = -0.44*	*t = -2.90****	*t = 1.11*	*t = 0.75*	*t = 0.77*	*t = -0.20*
Presence of Children Under 4	0.04	0.01	0.0004	-0.05	-0.01	-0.01	0.03	0.03	-0.002
	*t = 3.38****	*t = 1.43*	*t = 0.06*	*t = -2.39***	*t = -1.43*	*t = -0.39*	*t = 1.20*	*t = 0.94*	*t = -0.21*
Presence of Children Age 4 to 15	0.02	0.02	0.01	0.04	-0.01	-0.02	-0.01	0.06	0.01
	*t = 2.54***	*t = 2.03***	*t = 3.23****	*t = 2.48***	*t = -1.14*	*t = -1.11*	*t = -1.00*	*t = 2.72****	*t = 1.79**
Presence of Children Age 16 to 18	0.02	-0.01	0.01	-0.06	-0.01	0.03	-0.01	0.06	0.02
	*t = 1.92**	*t = -1.66**	*t = 1.35*	*t = -3.35****	*t = -1.91**	*t = 1.17*	*t = -0.56*	*t = 2.22***	*t = 2.65****
Female	-0.02	0.0003	-0.01	0.02	0.0005	0.01	0.002	-0.03	-0.005
	*t = -3.23****	*t = 0.08*	*t = -2.42***	*t = 1.46*	*t = 0.12*	*t = 0.56*	*t = 0.17*	*t = -1.74**	*t = -0.73*
Worker	-0.01	0.05	0.01	-0.02	-0.001	0.03		0.01	-0.01
	t = -0.89	*t = 3.33****	*t = 0.55*	*t = -1.68**	*t = -0.23*	*t = 1.34*		*t = 0.71*	*t = -0.91*
Student	-0.05	-0.02	0.04	-0.04	0.01	-0.05	0.01	-0.11	0.0002
	*t = -4.06****	*t = -1.80**	*t = 3.04****	*t = -1.91**	*t = 1.41*	*t = -1.57*	*t = 0.42*	*t = -3.69****	*t = 0.02*
Number of Household Vehicles	-0.02	-0.01	-0.004	-0.01	-0.003	0.001	-0.02	0.0001	-0.01
	*t = -4.39****	*t = -2.52***	*t = -2.22***	*t = -1.67**	*t = -1.17*	*t = 0.17*	*t = -2.64****	*t = 0.01*	*t = -2.23***
Suburban Household	-0.001	-0.005	-0.003	0.01	-0.0001	-0.01	-0.03	0.02	-0.001
	t = -0.10	*t = -0.98*	*t = -0.63*	*t = 0.73*	*t = -0.03*	*t = -0.37*	*t = -2.34***	*t = 0.81*	*t = -0.07*
Exurban Household	0.003	0.001	-0.01	0.05	0.0004	-0.03	-0.04	0.06	0.01
	t = 0.34	*t = 0.23*	*t = -2.09***	*t = 2.51***	*t = 0.07*	*t = -1.24*	*t = -1.93**	*t = 2.52***	*t = 1.08*
Rural Household	0.06	0.02	-0.003	0.07	0.001	-0.05	0.005	0.09	0.02
	*t = 4.27****	*t = 2.07***	*t = -0.39*	*t = 2.71****	*t = 0.15*	*t = -1.58*	*t = 0.18*	*t = 2.19***	*t = 1.18*
Observations	4,265	1,037	2,243	553	537	592	185	334	364
R_2	0.04	0.1	0.02	0.08	0.04	0.07	0.15	0.14	0.08
Adjusted R_2	0.04	0.09	0.01	0.05	0.01	0.05	0.08	0.1	0.04
Residual Std. Error	0.21 (df = 4249)	0.07 (df = 1021)	0.08 (df = 2228)	0.14 (df = 537)	0.05 (df = 521)	0.20 (df = 576)	0.08 (df = 171)	0.15 (df = 318)	0.06 (df = 348)
F Statistic	12.53*** (df = 15; 4249)	7.65*** (df = 15; 1021)	3.02*** (df = 14; 2228)	3.09*** (df = 15; 537)	1.28 (df = 15; 521)	3.02*** (df = 15; 576)	2.25*** (df = 13; 171)	3.40*** (df = 15; 318)	1.95** (df = 15; 348)

*Note: *p<0.1; **p<0.05; ***p<0.01*