Transportation (2011) 38:511–524
DOI 10.1007/s11116-010-9311-z

Generating twins to build weekly time use data from multiple single day OD surveys

Marcela Munizaga · Sergio Jara-Díaz · Javiera Olguín · Jorge Rivera

Published online: 10 November 2010

Abstract Origin-Destination surveys, which are regularly conducted in many cities (to calibrate transport models), contain indirect information on individual time use that can be recovered through the declared trip purpose. Although this data source is very rich, it has two limitations for the calibration of time use models: the level of disaggregation regarding time use is constrained by the definition of trip purposes, and the information gathered on different time periods is usually obtained from different individuals. In this paper we propose a new method to overcome the second limitation, transforming the original daily observations into individual-weeks. For every working day observation we build Saturday and Sunday "twins" as a convex combination of observed weekend individuals such that the distance between the attributes of the working day individual and the synthetic twin is minimized. We applied this procedure to the Santiago OD survey, and generated a database of weekly observations particularly rich for model calibration and segmentation.

Keywords Time use · Data processing · OD survey

Introduction

Time use and travel are closely related. This is explicitly acknowledged in the vast literature on Activity Based travel demand models (see Kitamura 1988; Axhausen and Gärling 1992; and Jones et al. 1990) where observations usually include 1 day only. However,

M. Munizaga (✉) · S. Jara-Díaz · J. Olguín
Departamento de Ingeniería Civil, Universidad de Chile, Casilla 228-3, Santiago, Chile
e-mail: mamuniza@ing.uchile.cl

S. Jara-Díaz
e-mail: jaradiaz@ing.uchile.cl

J. Olguín
e-mail: javieraolguin@gmail.com

J. Rivera
Departamento de Economía, Universidad de Chile, Santiago, Chile
e-mail: jrivera@econ.uchile.cl

decisions regarding time use involve periods that encompass both work time (mostly on weekdays) and leisure (mostly on weekends), as acknowledged by Minnen (2010), among others. This has been explicitly captured in the recent microeconomic literature on time allocation, where Jara-Díaz et al. (2008) have developed a time-consumption-travel model whose calibration requires observations along a complete work-leisure cycle, i.e., a minimum of a week (Munizaga et al. 2008). The Activity Based literature has also acknowledged this need (e.g., Spissu et al. 2009). However, weekly observations on time use are not easily available, which is indeed an obstacle for relevant comparison and analysis across different population segments.[1] On the other hand, time related data is mandatory for travel demand analysis, which is why transport is one of the few areas where one can find good and systematically collected information involving time use. Moreover, through the declared trip purpose the information collected to describe in detail the travel pattern using travel diaries allows recovering, with some level of aggregation, a description of time use of the individuals observed.

The objective of this paper is to describe a method to convert a large OD survey containing daily data into a weekly database with the necessary information to calibrate a time use system in a complete work-leisure cycle. This is done for the Santiago 2001 OD survey (DICTUC 2003) that has more than 12,000 households sampled randomly from a six million population. Each household is observed during a whole day using travel diaries where every trip and its purpose are mandatorily reported (avoiding the underreporting of trips mentioned by Hubert et al. 2008). This database is very rich in terms of the information collected, the sample size and the reliability of the data. Although daily observations are perfectly suited to generate an aggregate description of the sample and the segments regarding time use, they cannot be used to model and understand decisions on time use, which, as discussed above, should consider a whole work-leisure cycle. Moreover, the (marginal) values of leisure and work would be erroneously captured if daily models were attempted because the distribution of time use among shopping, errands, entertainment, work, rest and so on is done on a weekly basis. The main challenge, then, is how to mix observations of different people observed on different days to create weekly observations, maintaining the richness of the original data. The rest of the paper is organized as follows: in Sect. 2 we describe the method to transform the information contained in an OD survey into time use information, and to build weekly observations from single day data. In Sect. 3 we present the application to the Santiago OD survey and in Sect. 4 we describe the resulting database. Finally, in Sect. 5 we summarize and conclude.

Data transformation method

From travel information to time use

In a travel diary OD survey, the individuals answer a questionnaire where they are asked to declare the trip purpose, as well as the times when the trip began and ended. To transform that type of travel information into time use data, we propose to use the trip purpose information. The definition of trip purposes limits the number of activities that can be

[1] There are a few travel surveys containing weekly data, as the panels reported by Axhausen et al. (2002) and Löchl et al. (2005) where a reduced number of individuals are observed during several weeks, and the early data collected in Upsala (Hanson and Burnett, 1981; Hanson and Hanson, 1981). Most of the large datasets regarding time use contain daily information only.

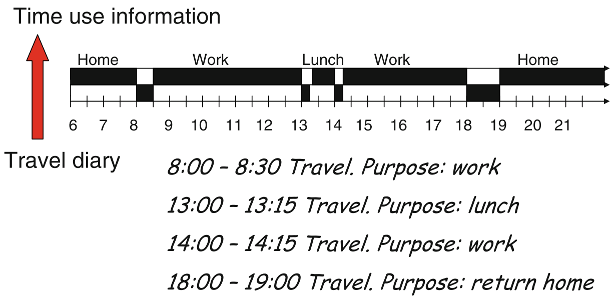

Fig. 1 Turning travel purpose into activity time

identified, as further disaggregation is not possible (although two or more activities can be aggregated). As the travel diary is focused on travelling, activities conducted inside the house have to remain as a single category, while activities conducted outside the house can be identified with more detail, but always constrained to the definition of trip purposes. For example, if only one purpose is recorded per trip, then it is not possible to obtain information about time assignment to more than one activity at the same time and location. However, it is possible to identify time assigned to activities at different locations. The time interval between two consecutive trips is assigned to the activity associated to the purpose declared for the first trip. Figure 1 represents this simple procedure. This data transformation process assumes that the cycle begins when the individual is at home. This assumption can be verified by checking whether the OD zone registered in the survey as the origin of the first trip of the day includes the address of the user's house.

The objective of the method is to convert an OD survey into a database with the necessary information to calibrate a workers time use model considering a whole work-leisure cycle; therefore, those individuals observed on a working day who do not go to work must be omitted from the sample. If during a working day a person stays at home all day, either because s/he works at home or because s/he is sick (or any other reason), the work-leisure cycle cannot be captured and, therefore, that observation shall be omitted from the sample. Evidently this is not the case if a person stays all day at home during a weekend day, as s/he is likely to be conducting leisure activities.

Building a weekly cycle

The more difficult task is how to build a weekly cycle (in our case the relevant work-leisure cycle) from one-day observations, as the activity patterns of the different days of the week come from different individuals. This problem can initially be seen as an imputation problem. Imputation methods are aimed at solving the problem of missing partial data by constructing the missing values using the information available in the database (Pfeffermann and Rao 2009). Roughly speaking, there are two types of methodologies, the hot deck and the regression imputation methods. Hot deck imputation completes the missing values for an individual from similar individuals whose record is complete, using different approaches to define similarity. Regression imputation is based on completing the missing values of a variable by means of a regression between this variable and the others. Both methods are designed to solve the problem of incomplete data in sample surveys (Ford 1983; Deville and Särndal 1994). In both cases it is required that part of the data is complete, which is not the case we are studying. Therefore, this type of method can not be used to solve our problem. There is another type of regression imputation, which, in our case would mean to calibrate a regression model to explain time assigned in a weekend day based on socioeconomic characteristics, and then—for each individual observed in a weekday—replace his/her socioeconomic variables in the estimated model in order to obtain time assigned in a weekend. However, this does not guarantee that the prediction is feasible, i.e., that the summation of all times assigned will add to 24 h a day. Furthermore, for activities with small time assignments nothing prevents to obtain negative values.

Another approach that might be useful is that of the impact evaluation program methods (Heckman and Vytlacil 2006; Imbens 2004). Here, the idea is to consider disjoint samples of individuals, one labelled as treated and the other as untreated. A treated individual is one that receives a *treatment*, which could be for example a monetary incentive, a subsidy or a certain benefit. As the amount of the treatment is not relevant in this type of analysis, it can be represented with a binary variable. The objective of the method is to determine treatment effect, that is, the incidence of the treatment on some relevant individual's characteristic previously defined. The method is based on a concept of similarity (matching) among individuals, defined through observable characteristics to compare their performance in the desired variable. Usually, the matching among treated and untreated individuals is defined by means of the propensity score indicator (propensity score matching; Rosenbaum and Rubin 1983) or based on the observable characteristics of the full sample (matching in characteristics; Imbens 2004).

Although these methods were developed for a different problem, they might be used for the imputation we require in our problem. We need to find, within the Saturday and Sunday samples, individuals who can replicate the behaviour of those observed during working days. For each individual observed in a weekday, we have to find a person who is so much alike that would replicate what the observed person would do on a weekend. The implicit assumption is that people who are similar in socioeconomic (SE) characteristics have similar behaviour, a link that has been established in many ways in the time use and travel literature. As examples, this is present in the so-called structural equations approach or similar comprehensive models as the ones by Lu and Pas (1999) and Bhat and Koppelman (1994), where gender, age and family structure (number of children) are identified as exogenous variables influencing time assignment to activities. These same three variables appear within the line of models inspired by Kitamura (1984)—where time use is directly explained through SE variables—along with others like ethnic origin, marital status and work location (Kitamura 1984; Bhat and Misra 1999; Chen and Mohktarian 2006; Munshi 1993; Meloni et al. 2004; Kitamura et al. 1996; Hirsh et al. 1986; Srinivasan and Bhat 2006; Kato and Matsumoto 2009; Meloni et al. 2007). Finally, these same types of variables have been used for segmentation when analyzing different aspects of time assigned to travel, as for example in Hanson and Hanson (1981) or Mohktarian and Chen (2004).

Next we have to decide which matching procedure to use. Within available matching methods, propensity score have the inconvenience that they reduce the dimensionality of the problem, merging all the variables in a real valued propensity score function. Matching in characteristics is more appropriate for our problem, because it considers all relevant characteristics of the individuals independently. The usual methods to do matching in characteristics are the nearest neighbour or nearest neighbours methods (Abadie and Imbens 2006), where a fixed number of individuals (usually one) is selected according to a distance measure. That would match individuals observed during weekdays to similar individuals observed during the weekend. However, we propose to use a new matching procedure developed by Krell et al. (2008) that does not choose any particular individual, but a combination of some of them that would replicate better the original one. Actually, nearest neighbour is a particular solution of the method, but might not be the best, as illustrated next.

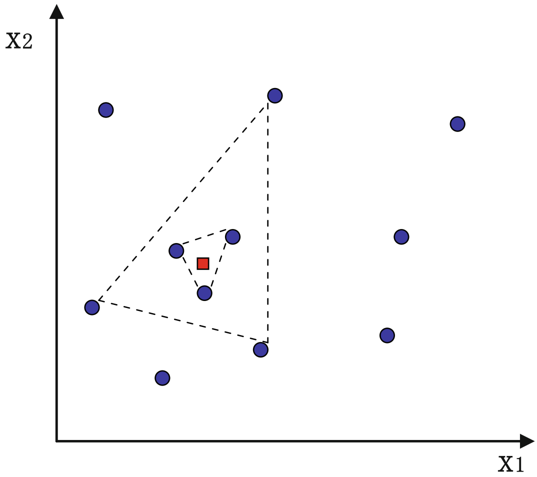

Fig. 2 Illustration of the matching procedure by means of convex combinations

In Fig. 2, the problem is represented in two dimensions where X_1 and X_2 are socio-economic characteristics, the square is the individual to be matched and the dots are all the candidates to build the twins. The nearest neighbour is evidently one of the three dots in the small triangle surrounding the individual. However, an appropriate linear combination of the same three dots (i.e., a convex combination) matches the individual better, in fact exactly in the example. A perfect match could also be obtained with a convex combination of the three dots that form the largest triangle, but using observations that are evidently further from the target.

The method by Krell et al. (2008) is based on the preceding idea, generating the missing information from an optimization problem that minimizes the distance between the attributes of the individual observed during the week and a convex combination (e.g., the triangles) of the data from individuals that are observed in the missing day (Saturday or Sunday), such that the extreme points of the optimal convex combination are as close as possible to the one observed during the week. In essence, this method pairs individuals whose characteristics are "similar" according to a criteria that depends on a vector of individual's attributes, not employing propensity score (Rosenbaum and Rubin 1983) or any related concepts. This way the dimensionality of the matching problem is not reduced. Then we build a Saturday and a Sunday "twin" for each individual observed during a working day applying those convex combination coefficients to the time assignment. Therefore, the Saturday and Sunday twins are not a particular person observed (like in the nearest neighbour method), but a synthetic person built as a combination of real persons observed on those days.

Formally, let us denote by L the set of individuals observed during a working day and by S all individuals observed on Saturday (similar for Sunday), potential candidates to build the twin; denote by $X_i^L,\ X_i^S \in R^K$ the vector of the K socioeconomic characteristics for individual i (belonging to L and S, respectively). For any $i \in L$ let us consider the following optimization problem

$$\min_{(\mu_{ij})} \left\| X_i^L - \sum_{j\in S} \mu_{ij} X_j^S \right\| + \sum_{j\in S} \mu_{ij} \cdot \left\| X_i^L - X_j^S \right\| \tag{1}$$

$s.t$

$$\sum_{j \in S} \mu_{ij} = 1$$
$$\mu_{ij} \geq 0 \quad \forall j \in S. \tag{2}$$

The first term of the objective function considers the distance among attributes whereas the second picks the extreme points that are as close as possible to the characteristics of the individual observed during the week. From the optimization problem, for all $i \in L$ we obtain weights μ_{ij}. For a given individual i on a weekday, the set of individuals j on a weekend such that $\mu_{ij} \neq 0$ contains all the persons that contribute to create the twin. Then, using these weights, the weekend (Saturday or Sunday) time assignment to activity k by the twin of individual $i \in L$ is built as

$$T^k_{i,S} = \sum_{j \in S} \mu_{ij} T^k_j \tag{3}$$

where T^k_j is the time assigned to activity k by individual j. Note that the summation over all activities k for the twin adds up to 24 h by construction. The weekly time assigned by individual i to activity k is obtained adding $T^k_{i,S}$ to the time spent during weekdays.

The same weights can be used to estimate the expected number of individuals performing a given activity at every instant during weekends. Thus, an activity pattern for Saturday and Sunday of the sample of synthetic twins can be constructed, although individual time of day cannot be predicted.

Finally, note that for many individuals identical in SE characteristics the distribution of time assigned to an activity on a weekday could be negatively correlated with the distribution of the same activity on the weekend, e.g., domestic work. With daily information, one is unable to capture these weekday-weekend trade-offs. However, the average weekly time assignment obtained for each SE group would be unbiased, although the variance would probably be overestimated.

Application

We applied this method to the Santiago 2001 Origin Destination Survey (ODS), conducted from July 2001 until April 2002 (DICTUC 2003). The study included household surveys, interception surveys, traffic counts and level of service measurements. To generate the time assignment information we used the household survey, which includes personal interviews to all the household members, to recover information about all the trips in a particular day chosen randomly (from 5 am to 5 am of the next day). The final sample contains 12,346 households (47,903 individuals) in the normal season and 3,191 households (11,860 individuals) in the summer season. These observations are evenly distributed along days of the week. The database contains variables that describe socioeconomic characteristics of both households and individuals. Each trip is described by origin and destination, starting and ending time, purpose, transport mode, walking time and distance. On any weekday 10% of the sample makes no trip at all, which falls within the range reported by Madre et al. (2007) for UK and US. Among those that make at least one trip, some 52% makes two trips, 21% makes four trips and 8% makes six trips.

The reported travel purposes were classified in 13 categories. If the purpose of the trip is "return to the house", the purpose of the forward trip is also recorded. As said before, information about individual activities is contained only in this variable. Those observations where the first trip of the day was originated in a zone different from the zone of the house address, were separated into those were the purpose of the first trip was "Return home" (1.11%)—that were included in the sample, but taking the next trip as the initial trip of the day—and the rest (0.54%) that were excluded from the final sample.

As evident, time assigned to the "Travel" activity can be obtained directly. The travel purposes defined as "work" and "work related", were consolidated into a "Work" activity; "study" trip purpose is associated to "Study" activity; the purposes "visit friends and relatives", "out to eat or have a drink" and "recreation" were added into an "Entertainment" activity; the purposes "drop-off/pick up someone or something", "shopping" and "errands" were classified into the activity "Shopping and errands"; "return to the house" is associated to the activity "Home" and, finally, the purposes "health" and "other" were associated to the activity "Other". The three aggregated activities (entertainment, shopping and errands and other) were built from original activities that received little time assignment in many cases. As evident, all the activities but "staying at home" are conducted outside the house. When a person is at home it is impossible to know weather s/he is studying or conducting a recreation activity, or doing something else; therefore, it was simply coded as "Home". With the start and finish trip times, we obtain the time assignment for the activities between trips.

As explained in the previous section, this data transformation process assumes that the cycle begins when the individual is at home, except when the OD zone registered in the survey as the origin of the first trip of the day does not include the address of the user's house. However, in this application there are some first trips in the morning whose purpose is recorded as "return to the house", in which case the previous activity is also recorded. This association cannot be done if the purpose of the first trip is other than "return to the house" and the individual is located outside his/her zone of residence. In those few cases the observation was eliminated.

Given the focus of this research and the procedure and assumptions previously described, only workers observed during the normal season were included, excepting those who do not report travel during a working day (around 10% in the sample, evenly distributed along working days), and incomplete or inconsistent observations. The resulting sample of workers has 16,887 individuals; 11,863 observed in a working day, 2,306 on a Saturday (including 408 who do not travel on that day) and 2,718 on a Sunday (675 do not make any trip). Using this information, we analyzed the time assignment and made comparisons among the five working days and between working and weekend days. This was done by means of activity patterns and average duration data. An activity pattern shows the percentage of individuals in the sample that is conducting a particular activity at every particular time of day. The average duration of activities is calculated including only the observations with positive time assignment.

The activity patterns for the whole sample are shown in Fig. 3 for the weekday and the 2 weekend days. In a working day, the three main activities are: staying at home, work and travel, while a very small percentage of individuals assign time to other activities. The percentage of people who make some trip with work purpose is 43.5 on Saturday and 19.1 on Sunday. The maximum percentage of people working is observed at noon (36% on Saturday and 15% on Sunday). The activity patterns of the weekend days show that out-of-home recreation activities reach a maximum of 13% on Saturday and 19% on Sunday, in both cases around 6 pm. Individuals who do not travel at all in weekend days are also included, as staying at home was their only activity.

The visual analysis of the data indeed suggests that time assignment during working days (Monday to Friday) is the same, exhibiting important differences with Saturdays and

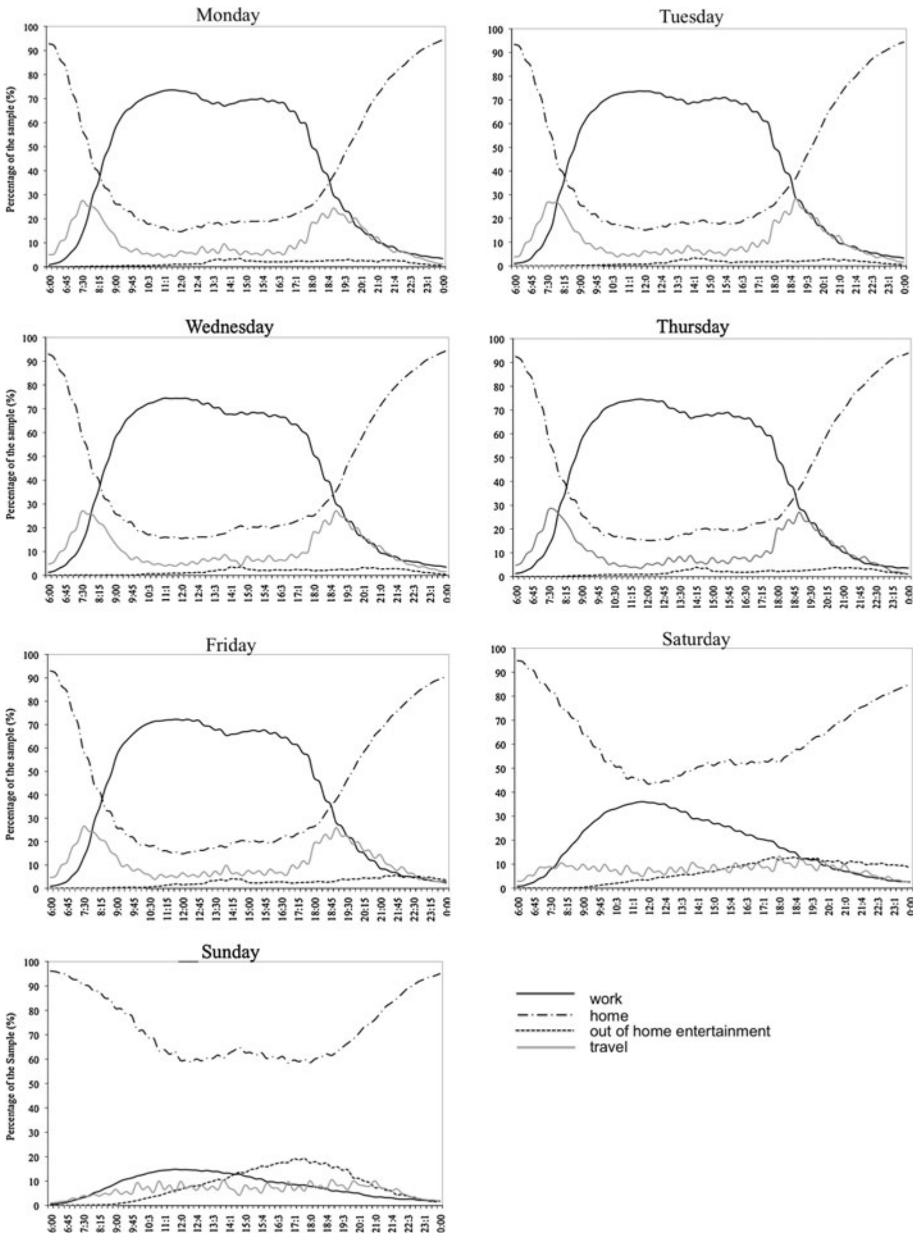

Fig. 3 Activity patterns by day

Sundays, different between themselves as well. We applied the Kruskal–Wallis test (Kruskal and Wallis 1952) and concluded that time assignments are not significantly different between weekdays. Therefore, all Monday to Friday observations were considered "weekday".

Therefore, to transform the original daily observations into individual-weeks, every working day individual time assignment should be multiplied times five adding the time assignment of the weekend twins. Previously, individuals in extreme or peculiar cases were discarded: those with 1 h or less assigned to work exhibiting a high salary (who presumably work at home, impossible to detect); those with more than 15 daily hours at work; those with exceptionally large travel time or travel expenses; and those that reported no income or whose income did not add up to the family income. Eliminating these observations reduced the weekday sample to 9,464 individuals to which we input weekend information obtained from the 2,306 individuals observed on a Saturday and the 2,718 observed on a Sunday in order to obtain weekly observations.

In order to create the twins, the socioeconomic characteristics used to search for the optimal convex combination of individuals are: sex, age, location, education, link with the house head, number of vehicles, income, driving license, internet and cable TV. The norm induced by the inverse of the co-variance matrix of the data (Mahalanobis norm) is employed to measure distance in the objective function.

In order to verify that socioeconomic variables respond adequately to the method, we performed an experiment to predict daily time use for a subsample of travellers on a weekday using the characteristics of the rest of the individuals to build the twins. We selected randomly 10% of the 1970 travellers on Tuesday and applied the matching procedure by Krell et al. (2008) using the socioeconomic variables listed above. The procedure assigned from one to seven individuals with an average of 2.22 to form the convex combination. The resulting time prediction was quite accurate, with a correlation of 0.94 between predicted and observed values and a sum of squared residuals 50% lower than what was obtained using the nearest neighbour method.

Results

After applying the method by Krell et al. (2008) the number of weekend individuals necessary to create a "twin" happens to vary from 1 to 11 (2.5 in average); in 60% of the cases only one or two weekend individuals are needed and almost all the weekend individuals (98%) are used at least once. It is important to mention that the time assignments obtained are different from those obtained using the (simpler) nearest neighbour method.

Note that the original daily observations could be used to provide aggregate descriptions of the samples and the segments regarding time use for every day of the week for comparative purposes. This is something that does not need the creation of weekly observations, as daily samples are (presumably) representative for each day. The twins could be used to provide an aggregate description of time use during weekend days as well, and these aggregate values could be compared against those obtained with the original data. However, the observations created for Saturday and Sunday for each and every weekday individual are built using the original observations, and all originals are used at least once for each weekend day. Therefore, the daily averages for each activity using the twins are practically equal by construction to those of the original weekend samples. Nevertheless, we did verify that this is the case.

To see how the weekly data constructed with our method looks like, we present an aggregate analysis that will illustrate the differences between segments (which will prove useful in modelling). In Fig. 4 we show the average weekly time assigned to the different activities by men and women living in the East side of the city (the richest) and those from

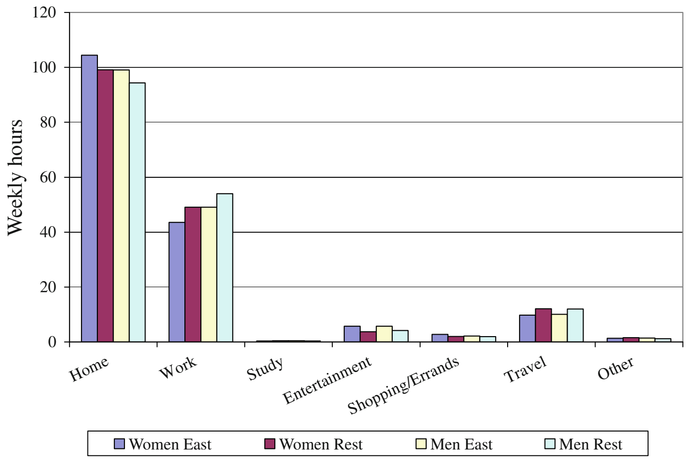

Fig. 4 Average weekly time assignment by zone and gender (25–64 age group)

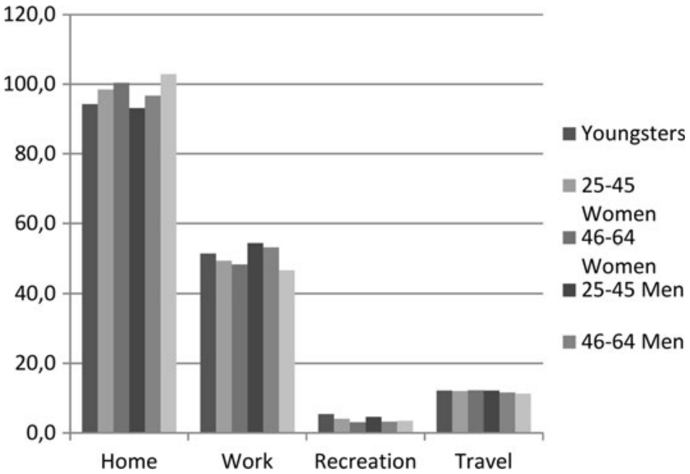

Fig. 5 Average weekly time assignment by age and gender

the rest, for individuals in the 25–64 age range. The most evident differences are in the Home and Work activities, particularly for Women East and Men Rest.

In Fig. 5 we show the average time assignment to the four most important activities for six different groups segmented by age and gender from the whole sample. It can be seen that elders and women spend more time at home, while men and youngsters assign more time to work. The segment with the largest assignment to out of home entertainment time is that of people under 25.

Finally, in Fig. 6 we show the complete activity pattern for the generated database following the procedure explained by the end of Sect. 2, with Saturday and Sunday proportions calculated as expected values. This aggregate picture shows how work is clearly the most important out of home activity during working days and that there is a significant amount of work also on weekends.

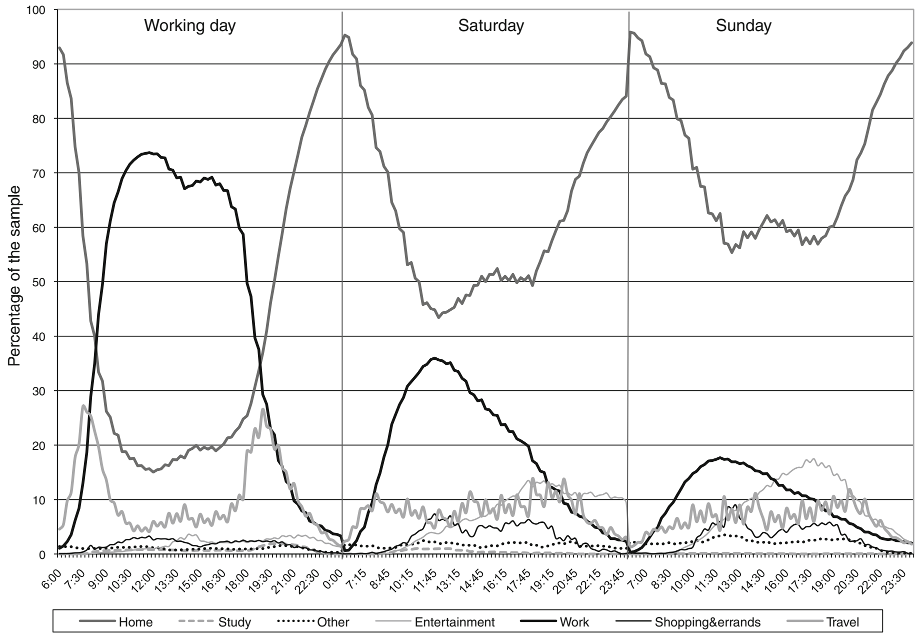

Fig. 6 Complete activity pattern

Time use data generated with this procedure proved very useful, as explanations of the differences in individual behaviour (decisions) across segments regarding time use cannot be built unless weekly observations are used. In particular, estimates of the (marginal) values of leisure and work would be erroneously captured if daily models were attempted because the distribution of time use among shopping, errands, entertainment, work, rest and so on is done on a weekly basis. If modelled daily, the marginal utility of activities that are assigned more time in a particular day would (erroneously) seem larger. In fact, the new time use database permitted the estimation and comparison of time use values for interesting segments in Santiago, revealing differences between genders, zones of residence, age and family structure (Jara-Díaz et al. 2010). So building a weekly time use data base from the OD survey has revealed differences in the choices made by individuals belonging to different segments of the population in Santiago.

Synthesis and conclusions

We have presented a method to transform information from a travel diary based OD survey, where each individual is surveyed a single day, into weekly based time use data usable for modelling purposes. The method is based upon the hypothesis that individuals of similar socioeconomic characteristics have similar behaviour, and a new matching approach is used. Substitution between time assigned to a particular activity during weekdays and on weekends cannot be explicitly introduced in the model given the information available. This method can be applied to represent aggregate behaviour of any database of similar characteristics, allowing comparisons between different cultures or between different types of cities.

The 2001 Santiago Origin Destination Survey is a very large database, containing detailed, reliable and representative information of worker-travellers for every day of the week. The methodology applied generates a time use database equally rich. However, the aggregation level is determined by the definition of trip purposes. Therefore, activities conducted inside the house have to be aggregated into a single category. Fortunately, the activities conducted outside the house are clearly identified through the trip purpose, which allowed us to identify seven different (aggregated) activities: Travel, Home, Work, Shopping and Errands, Recreation, Study, Other.

The analysis of the data shows that time assignment during working days (Monday to Friday) is very similar; the corresponding observations were considered as equivalent, exhibiting important differences with Saturdays and Sundays, different between themselves as well. The original daily observations had to be transformed into individual-weeks. To do this, we take every working day observation as an individual whose time assignment repeats from Monday to Friday. The time assignment of each of these individuals during Saturday and Sunday was generated creating twins using a new method developed by Krell et al. (2008). With the sample created of 9,464 individuals with weekly time assignment information, we are now able to look at the activity patterns of different groups of individuals, analyze differences between groups or segments, calibrate activity models, and calculate and analyze their values of leisure and work.

For synthesis, the method explained and applied here helps converting massive travel data, usually obtained daily, into weekly time use observations, which ultimately contributes to the analysis of the revealed preferences regarding time use. Understanding activity patterns behind travel patterns lies in the foundations of the activity-based approach, which is presently undergoing various revisions and improvements. Among them is the need to consider weekly data and the advantages of the joint estimation of time use and travel models, possibly including various types of discrete choices. This is important not only from a modelling perspective but also for appraisal and prediction (Munizaga et al. 2008).

Regarding further research we envision at least three lines related with the limitations and reliability of the method. One is the gross aggregation imposed by the trip purpose definition, which does not allow the recovery of time use within a given location e.g., home. This difficulty could be softened with the help of additional information regarding the distribution of home activities from subsamples of travellers. A second line is the application of the convex combination matching procedure to a large and reliable weekly time use database also available on a daily basis, such that the observed weekend time assignment of a subsample could be compared with the time use predicted using the twins procedure. Finally, with this same type of information, we could also analyze substitution of time assigned to certain activities (e.g., shopping, domestic work) between weekday and weekend, in order to test whether the method could be improved to help to create weekly observations using daily samples.

Acknowledgments This research was partially funded by Fondecyt grants 1090204, 1080140 and 1109518 and the Institute Complex Engineering Systems (ICM P-05-004-F, CONICYT FBO16). SECTRA provided the Santiago OD survey data. We want to thank the comments of four unknown referees for their useful suggestions, including the (weak) test of the method described at the end of Sect. 3. Remaining errors are of course ours.

References

Abadie, A., Imbens, G.: Large sample properties of matching estimators for average treatment effects. Econometrica **74**(1), 235–267 (2006)

Axhausen, K., Zimmermann, A., Schönfelder, S., Rindsfüser, G., Haupt, T.: Observing the rhythms of daily life: a six-week travel diary. Transportation **29**(2), 95–124 (2002)
Axhausen, K., Gärling, T.: Activity-based approaches to travel analysis: conceptual frameworks, models, and research problems. Transp. Rev. **12**(4), 323–341 (1992)
Bhat, C., Koppelman, F.: A structural and empirical model of subsistence activity behavior and income. Transportation **21**(1), 71–89 (1994)
Bhat, C., Misra, R.: Discretionary activity time allocation of individuals between in-home and out-of-home and between weekdays and weekends. Transportation **26**(2), 193–209 (1999)
Chen, C., Mohktarian, P.: Tradeoffs between time allocations to maintenance activities/travel and discretionary activities/travel. Transportation **33**(3), 223–240 (2006)
Deville, J.C., Särndal, C.E.: Variance estimation for the regression imputed Horvitz-Thompson estimator. J. Off. Stat. **10**(4), 381–394 (1994)
DICTUC: Actualización de encuestas Origen Destino de viajes, V Etapa. Informe Final a Sectra, Santiago (2003)
Ford, B.L.: An overview of hot-deck procedures, dans incomplete data in sample surveys. In: Madow, W.G., Olkin, I., Rubin, D.B. (eds.) Vol II: theory and bibliographies, pp. 185–206. Academic Press, New York (1983)
Hanson, S., Burnett, K.O.: Understanding complex travel behavior: measurement issues. In: Stopher, P.R., Meyburg, A.H., Brög, W. (eds.) New horizons in travel-behaviour research, pp. 207–230. Croom Helm, New York (1981)
Hanson, S., Hanson, P.: The impact of married women's employment on household travel patterns: a Swedish example. Transportation **10**(2), 165–183 (1981)
Heckman, J., Vytlacil, E.J.: Econometric evaluation of social programs 6. In: Heckman, J., Leamer, E. (eds.) Handbook of econometrics. North Holland, Amsterdam (2006)
Hirsh, M., Prashkea, J., Ben-Akiva, M.: Dynamic model of weekly activity pattern. Transp. Sci. **20**(1), 24–36 (1986)
Hubert, J.P., Armoogum, J., Axhausen, K.W., Madre, J.L.: Immobility and mobility seen through trip based versus time use surveys. Transp. Rev. **28**(5), 641–658 (2008)
Imbens, G.: Nonparametric estimation of average treatment effects under exogeneity: a review. Rev. Econ. Stat. **86**(1), 4–29 (2004)
Jara-Díaz, S.R., Munizaga, M., Greeven, P., Guerra, R., Axhausen, K.: Estimating the value of leisure from a time allocation model. Transp. Res. B **42**(10), 946–957 (2008)
Jara-Díaz, S.R., Munizaga, M., Olguín, J.: (2010) Understanding the value of work behind time allocation patterns: the case of Chilean workers. Working paper, División Ingeniería de Transporte, Departamento de Ingeniería Civil, Universidad de Chile
Jones, P., Koppelman, F., Orfeuil, J.: Activity analysis: state-of-the-art and future directions. In: Jones, P. (ed.) Development in dynamic and activity-based approaches to travel analysis, pp. 34–55. Gower Publishing Company, Aldershot (1990)
Kato, H., Matsumoto, M.: Intra-household interaction in a nuclear family: a utility-maximizing approach. Transp. Res. B **43**(2), 191–203 (2009)
Kitamura, R.: A model of daily allocation to out-of-home activities and trips. Transp. Res. B **18**(3), 255–266 (1984)
Kitamura, R.: An evaluation of activity-based travel analysis. Transportation **15**(1–2), 9–34 (1988)
Kitamura, R., Yamamoto, T., Fuji, S., Sampath, S.: A discrete-continuos analysis of time allocation to two types of discretionary activities which accounts for unobserved heterogeneity. In: Lesort, J.B. (ed.) Transportation and traffic theory, pp. 431–453. Pergamin Press, Oxford (1996)
Krell, R., Rau, T., Rivera, J.: An optimization-based matching procedure. Working Paper No 279, Departamento de Economía, Universidad de Chile (2008)
Kruskal, W.H., Wallis, W.A.: Use of ranks in one-criterion variance analysis. J. Am. Stat. Assoc. **47**(260), 583–621 (1952)
Löchl, M., Axhausen, K.W., Schönfelder, S.: Analysing Swiss longitudinal travel data. In: Paper Presented at the Fifth Swiss Transport Research Conference, Ascona, March 2005
Lu, X., Pas, E.: Socio-demographics, activity participation and travel behavior. Transp. Res. A **33**(1), 1–18 (1999)
Madre, J.L., Axhausen, K.W., Brög, W.: Immobility in travel diary surveys. Transportation **34**(1), 107–128 (2007)
Meloni, I., Guala, L., Loddo, A.: Time allocation to discretionary in home, out of home activities and to trips. Transportation **31**(1), 69–96 (2004)
Meloni, I., Spissu, E., Bez, M.: A model of the dynamic process of time allocation to discretionary activities. Transp. Sci. **41**(1), 15–25 (2007)

Minnen, J., Glorieux, I.: Two days a week? A comparison of the quality of time-use data from 2-day diaries, 7-day diaries and a weekly work grid. Presented at the Second Time Use Observatory Workshop, Jahuel, Chile, March 2010
Mohktarian, P., Chen, C.: TTB or not TTB, that is the question: a review and analysis of the empirical literature on travel time (and money) budgets. Transp. Res. A **38**(9–10), 643–675 (2004)
Munizaga, M., Jara-Díaz, S.R., Greeven, P., Bhat, C.: Econometric calibration of the joint time assignment-mode choice model. Transp. Sci. **42**(2), 208–219 (2008)
Munshi, K.: Urban passenger travel demand estimation: a household activity approach. Transp. Res. A **27**(6), 423–432 (1993)
Pfeffermann, D., Rao, C.R.: Handbook of Statistics, Sample Surveys: Design, Methods and Applications, Chapter 10, vol. 29A. Elsevier, Amsterdam (2009)
Rosenbaum, P.R., Rubin, D.R.: The central role of the propensity score in observational studies for causal effects. Biometrika **70**(1), 41–55 (1983)
Spissu, E., Pinjari, A.R., Bhat, C.R., Pendyala, R.M., Axhausen, K.W.: An analysis of weekly out-of-home discretionary activity participation and time-use behavior. Transportation **36**(5), 483–510 (2009)
Srinivasan, S., Bhat, C.R.: A multiple discrete-continuous model for independent- and joint-discretionary-activity participation decision. Transportation **33**(5), 497–515 (2006)

Author Biographies

Marcela Munizaga is Assistant Professor at Universidad de Chile. She holds a PhD on Engineering Sciences from the P. Universidad Católica de Chile. Her research focuses on modelling behaviour of travellers and processing transport and time use data. She is Associate Researcher in the Complex Engineering Systems Institute and Vice-President of the Chilean Transport Engineering Society (SOCHITRAN).

Sergio Jara-Díaz is Professor of Transport Economics at Universidad de Chile. Holds a PhD and MSc from MIT, where he has taught during various terms. Author of Transport Economic Theory (Elsevier, 2007) and more than 80 articles in journals and books, on the microeconomics of transport demand, multioutput analysis in transport industries, public transport modelling and pricing, and time allocation. He is Associate Editor of Transportation Science and serves in the Editorial Board of various other scientific journals; is one of the Vice Chairs of WCTRS. He teaches frequently in Spain. Resides in Ñuñoa, Santiago, with his only wife.

Javiera Olguín born in Santiago de Chile, is Civil Engineer and holds an MSc in Transport Engineering from Universidad de Chile. Between 2008 and 2010 she worked at the Chilean Secretariat for Transport Planning (SECTRA). Currently she is studying an MSc in Town Planning at University College London.

Jorge Rivera is currently Associate Professor, Department of Economics, University of Chile, post he has held since 2000. He receives his PhD in Applied Mathematics from University of Paris 1 (2000). His areas of specialization are applied optimization and theoretical microeconomics.