electronic **International Journal of Time Use Research**
2008, Vol. 5, No. 1, 26-42 **dx.doi.org/10.13085/eIJTUR.5.1.26-42**

Entropy and stability in time use – An empirical investigation based on the German Time Use Survey

Rainer Hufnagel

Rainer Hufnagel
Department of Home Economics
University of Münster
Philippistraße 2
48149 Münster, Germany
e-mail: hufnagel@hw.uni-muenster.de

Abstract

Flexibility is a tool for preserving the stability of a system. In general, we can expect that the more variant its behaviour, the more stable a system will be. The investigation provides an example of this principle within the discipline of home economics. For a sample of single-person households from Germany's national Time Use Survey 2001/2002, it can be shown that the stability of activity sequences is greater, the higher the entropy of time use. For this purpose, a Markov model is derived from heuristic considerations. The Markov matrices are estimated and their eigenvectors and eigenvalues then calculated. It is evident that the entropy of an attractor is higher, the lower the norm of the second eigenvalue of the corresponding Markov matrix. The main components of this relationship, namely diversity and stability in time use, turn out to be only weakly associated with the usual socio-economic regressors. Hence, new empirically and theoretically relevant dimensions for socio-economic research emerge.

JEL-Codes: A21, C32, C51, D13

Keywords: Time use survey, Markov process, system stability, entropy, teaching

The author is grateful to Dr. Brian Bloch for his comprehensive editing of the manuscript.

1 Introduction

Although this paper may seem merely to protocol econometric results drawn from the German Time Use Survey 2002/2003 (GTUS), it makes a contribution through reporting and analysing progress in basic research with respect to a dynamic theory of the household. GTUS is based on time diaries, hence the concept of Scientific Use File (SUF) of GTUS documents, the sequence of activities in a special computer file. The dynamic features of this file have not been used much in research so far. Most published research simply argues on the basis of the overall frequencies of different activities. Use has rarely been made of the fact that data are also available that document not only the frequency, but also the sequence of activities.[1] Intuitively, there should surely be measures of the diversity, monotony, stability, variation, complexity, determination and randomness of these dynamics, but there is no consensus on the definitions through which these pre-concepts could be formalized. There are also many ways of defining the relevant concepts. One critical condition is that a definition not be self-contradictory. A “good” definition, however, becomes meaningful in the sense that the entity defined is also the object of a theorem.

Take, for example, the field of physics. We have a pre-theory intuitive concept of heat, of the space filled by gas, the elastic resistance that it exerts against compression. The definitions of “pressure” p, “Volume” V and “Temperature” T are difficult to formulate. However, it turns out that it was worthwhile to select just these definitions of many possible related ones, because scientists were then able to formulate the theorem $p \cdot V = T$ (Boyle-Mariotte-Gay-Lussac). Finding these definitions seems easy in retrospect, but a glance at the history of science shows that it took centuries to differentiate meaningfully between impulse, force and kinetic energy, or, to give another example, between temperature and the quantity of heat.[2]

Accordingly, if we wish to research the dynamics of time use, we should not be content with arbitrary definitions of global properties of such a dynamic, but establish those which can be linked to theorems with empirical content. This means theorems that are not true by definition or a priori, but prove to be true by observation or a posteriori.

This paper proposes a first step in this direction. It suggests formalizing time use sequences through a Markov process. The diversity of time use can therefore be measured by the entropy of its attractor, and its stability by the norm of its second eigenvalue. Diversity and stability thus defined prove to have an empirical relationship, at least for the investigated data.

It is found that the higher the entropy of time use, the higher its stability.

---

1 Exceptions known to the author are Grossmann (2007), Hufnagel (2004), Merz and Burgert (2004), Gille and Marbach (2004, 103 ff.), Wilson (1999).

2 See Falk (1990, 125 ff.).

It is clear that such a result can only be the starting point for building a genuine theoretical framework for the investigation of time use data. However, it should constitute a viable foundation.

In the following analysis, a detailed derivation of the central result of this paper is provided. Possible applications of these initial results are discussed in the concluding section. The data used for the investigation are presented in Section 2. Section 3 reports previous research, preliminary heuristic findings and looks for interdisciplinary elements which could be considered in further research. The analysis also explains and justifies why the author decided to use a Markov model. The assumptions and implications of the model are described in considerable detail. The fourth section presents the empirical results yielded by the Markov framework. Section 5 contains a discussion of the findings and the implications for further research.

2 Data used for the investigation

Our investigation is based on the Time Use Survey 2001/2002 of the Statistisches Bundesamt Deutschland (German National Office of Statistics)[3]. We use the 95% Scientific Use File which has been available from the Statistisches Bundesamt since 2005. The sample size is approximately 5,500 households, comprising about 14,500 people. Data was collected by means of the following methods:

- household questionnaires
  (one per household)
- personal questionnaires
  (to be completed by all household members older than 10 years)
- time diaries
  (also to completed by all household members older than 10 years).

The usual socio-economic and socio-demographic data were collected for households and their members. In particular, data is available for household composition, income, living conditions, profession and education, health and satisfaction (with work and leisure-time activities). Based on the time diaries, a computer file was constructed that indicates how much time was used for which activity per day. The list of activities comprises 272 items. The activities are classified hierarchically by subject. For example, we have Activity Number 312 “baking” that belongs to the Activity Group Number 31 “preparing meals”. Again, Activity Group Number 31 is subsumed under Activity Field 3 “housekeeping”.

3 See, for example, Kahle (2004) on the design of the 2001/2002 Time Use Survey. Some initial research on this data by a pioneering group of social scientists was published by the Statistisches Bundesamt (2004).

Time use is documented in a file which shows, for each individual older than 10 years, which activities had been conducted in 10-minute intervals. Normally, such information is available for three days per person, and usually includes two working days.

This following investigation by the author is essentially explorative in nature. Trying to use all the information included the SUF at once, would have been prohibitively complex. Therefore, a sub-sample was selected, based on the following criteria.

Only single-person households are used for the investigation. This is to avoid investigating influences from the activities of several household members.

It is necessary to document time use for two subsequent working days. This allows us to regress the activities performed during a certain hour of one day on the activities of the same hour the day before.

Single people ("singles") who regularly perform paid work are excluded from the study. Considering the answers of paid workers, one can see that they mostly fill in an 8-hour block of time with only one activity: 11= "paid work". In comparison to non-employed individuals, their time use becomes very monolithic or indeed monotonous. Because we are interested in non-trivial reasons for "monotony" and the diversity of time use, regularly employed people are excluded from the study.

Applying these criteria, we obtain a sample of 426 single people. For these individuals, we have information on what they did on two subsequent working days during each ten-minute interval of each day. Accordingly, we have 2×144×426= 122,688 observations of time use. Each of these observations can assume one of 272 values, the number of activities documented by the SUF. Some of these activities rarely emerge, and others do so very frequently. Because we have to estimate these frequencies later in the investigation, we must ensure that there are sufficient observations of each activity to secure a reasonable level of significance of the estimated frequencies. Therefore, the activities are aggregated into 18 activity groups which are listed in Table 1.

3 A Markov model for time-use sequences

By means of the data described in the above section, we can pose the heuristic question of whether it is possible to predict the activities that a person will perform within a certain time interval. In order to deal with this issue, the author has proposed a Boolean grid model (Hufnagel, 2000).

Table 1
Activity groups used in this investigation

Group	Includes activity No.*)	Group	Includes activity No.	Group	Includes activity No.
1. Sleeping	010-012	*7. Textile care*	330-339	*13. Recreation*	531-532
2. Eating	020-021	*8. Craftmanship and gardening*	340-359	*14. Sports*	610-651
3. Personal care	030-039	*9. Shopping*	360-389	*15. Hobby*	700-729 840-849
4. Paid work and education	100-249	*10. Voluntary work*	400-449	*16. Reading*	800-819
5. Preparing meals	310-319	*11. Social contacts*	510-519	*17. Watching TV*	820-822
6. Cleaning	320-329	*12. Cultural activities*	520-529	*18. Travelling time*	900-993

* For a detailed description of these activities, see the handbook from the SUF of the Statistisches Bundesamt.
Source: Own illustration.

In this model, the activity conducted during a certain time interval depends on:

1. the activities carried out in the time intervals prior to the observed one
2. the activities of other household members
3. natural, social and economic restrictions
4. socio-economic and socio-demographic characteristics, attitudes and norms of the household members.

When the author used Boolean regressions to concretize his Boolean grid model, he found that the number of observations is too small to obtain significant estimates.

A less demanding approach was subsequently attempted, estimating Probit regressions. At first glance, the results seemed promising: hit rates and ψ-R^2s often greater than 90% could be obtained. However, these results turned out to be useless for the purpose of predicting and simulating household behaviour. This will be explained in more detail, because this failure has important implications.

Assume a 0/1-dummy-variable y_t^s. Imagine, for example, that y_t^s indicates whether person number s ($s \in \{1,\dots,S\}$) is sleeping during time interval t or not. y_t^s should depend on the activities performed immediately before (0/1 dummies as well), the length of time these activities took and the squares of the variables, the length of time which the activities had not been conducted in the past, and the squares of the variables (all metric). Furthermore, y_t^s should depend on the time of day and on the social, economic and psychological characteristics of the investigated person.

Because y_t^s is dichotomous and we assume an interdependency of many qualitative, ordinal and metric regressors x_i (i=1,…,n), it is best to work with a Probit [4] model:

(1) $\tilde{y}_t^h = a_0 + a_1 \cdot x_1^s + ... + a_n \cdot x_n^s + \varepsilon^h$

(2) $y_t^s = 1$ *for* $\tilde{y}_t^s \geq 0$ *and* $y_t^s = 0$ *for* $\tilde{y}_t^s < 0$

(3) $\varepsilon \sim N(0,1)$.

Now, suppose the coefficients a_i (i=0,…,n) have been estimated.[5] Let $x_1^s,..., x_n^s$ be the observed values for a certain member of the sample. We can then use the Probit-model for two different forms of prediction (e.g. Woolridge, 2006, 582 ff.):

A) We predict $y_t^s = 1$ *if* $a_0 + a_1 \cdot x_1^s + ... + a_n^s \geq 0$ *and* $y_t = 0$ *otherwise*.

B) We predict the probability that $y_t^s = 1$, to be $\Phi(a_0 + a_1 \cdot x_1^s + ... + a_n \cdot x_n^s)$.[6]

It was found that only the second interpretation of the Probit-estimates was useful for dealing with our time budget data. This is clear if one examines Figure 1, which shows the result of a Probit-estimation for a sample of singles. The ordinate indicates the predicted probabilities $\Phi(a_0 + a_1x_1 + …..+a_nx_n)$. Among the regressors, x_i are a dummy that indicates whether the person had been sleeping during the previous time interval or not, a variable that measures how long the person has been sleeping without interruption in the past 24 hours and the square of this variable. Finally, the time of day is among the regressors and this variable is indicated on the abscissa. At the bottom of Figure 1, we see a cluster of observations that behave as might be expected. The probability of sleeping is low during the day and higher at night. At the top of the plot, there is a cluster of observations with very high probabilities for sleeping, even during the day. The reason for this cluster is that the Probit-procedure indicates that it is very likely that someone will continue sleeping when he has slept during the last ten minutes, even in bright daylight.

Now imagine that we attempt to simulate a single's behaviour on the basis of such a Probit-estimation. We start with a person who is sleeping at four in the morning. Then, the cross applying to this person will fall in the upper cluster in Figure 1, Φ will be about 0.95. Consequently, according to the above Rule A, we would predict "sleeping" for the next time interval. Again, the observation would be in the upper cluster and, continuing this process, we would predict sleeping for the entire day. To conclude, trying to simulate a single's behaviour on the basis of Probit-estimates and the above prediction Rule A, would lead us into a quite unreal world.

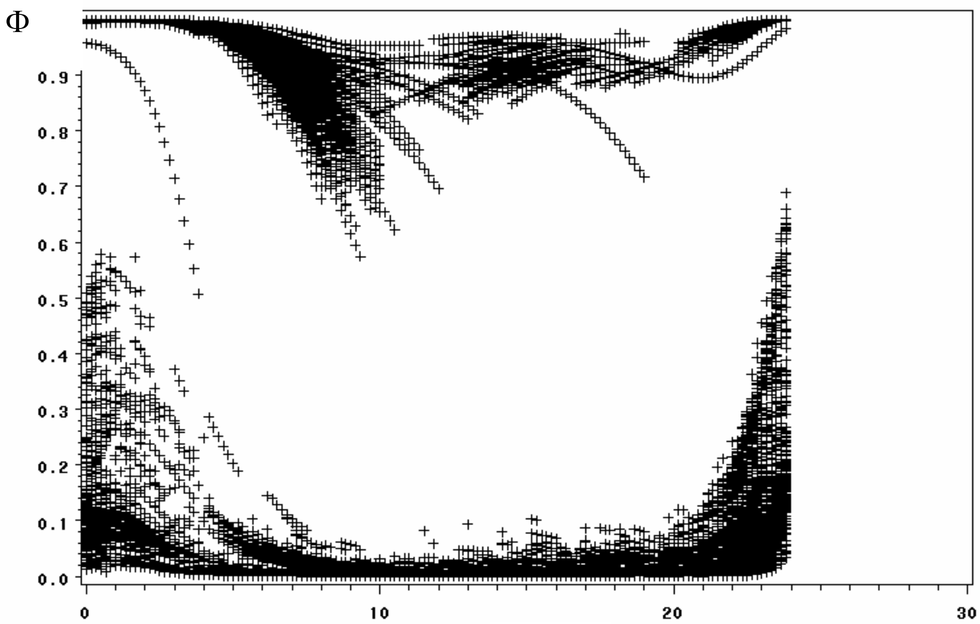
---

[4] Clearly, a Logit or anything similar would do as well in the context of this paper.

[5] Assume that non-significant regressors have been deleted.

[6] Φ denotes the cumulative distribution function of the standard normal.

Figure 1
Predicted probabilities for the activity of sleeping



Φ=Probability for the activity of sleeping; Hour= time of day.
Source: Own illustration, based on German Time Use Survey 2001/2002.

One way out of this dilemma, however, is to use the above prediction Rule B. Let us start with a person sleeping at 8 o'clock in the morning. We predict that she will sleep during the next time interval with a probability of about 90%. If we continue this process for the next 4 hours, we will obtain a probability of $0.90^{4\times 6}$ = 8% that this person will still be sleeping at noon. This seems a much more realistic result than that yielded with predictions through Rule A.

To summarize: there are not enough observations in the SUF to estimate a complex model of time use behaviour. We can successfully reduce complexity by using Probit instead of Boolean regressions. However, we must then accept that we cannot predict which activity will be conducted, but only the probabilities with which they will be conducted. Hence, with the data and methods on hand, the central object of our analysis must be vectors

(4) $$x^{\langle t\rangle} = \left(x_1^{\langle t\rangle}, \ldots, x_n^{\langle t\rangle}\right)^T$$

where $x_i^{\langle t\rangle}$ yields the probability that Activity Number i will be performed in time interval Number t. From heuristic investigations based on Probit estimations, the insight was gained that the probabilities $x_i^{\langle t\rangle}$ generally and consistently depend on the following regressors:

a) the activity conducted in the time interval immediately before the current one
b) the time of day
c) some socioeconomic and socio-demographic characteristics.

Thus, together with the insight that only the probabilities $x_i^{\langle t \rangle}$ of the activities being performed can constitute the object of analysis, the following approach is suggested

(5) $$x^{\langle t+1 \rangle} = A(t,s) \cdot x^{\langle t \rangle}$$

with a n×n matrix A. The element a_{ij} of the matrix A indicates the probability that the i^{th} activity will be conducted in time interval t+1 given, that the j^{th} activity had been conducted in the time interval t, i.e. in the interval before. These transition probabilities depend on the time of day, and hence finally on t, and possibly on the socioeconomic characteristics described by a vector s. Therefore, in Equation (5), we denote the transition matrix exactly by A(t,s).

Next, it is assumed, for working days, that the matrix A depends on the time of day, but not on the day of the week. Formally, we express this as:

(6) A(t,s) = A(t+144,s)

From Equation (5), we have

(7) $$x^{\langle t+144 \rangle} = M(t,s) \cdot x^{\langle t \rangle}$$

with

(8) $$M(t,s) = \prod_{\tau=0}^{143} A(t+\tau, s) .$$

From Assumption (6), we know that there are 144 distinct matrices M(t,s) (t=1,…,144). We now assume formally that there were only working days and no interruptions in the form of Saturdays and Sundays. Then, each of these matrices M(t,s) would constitute a Markov-process by

(9) $$x^{\langle t+d\cdot 144 \rangle} = M(t,s)^d \cdot x^{\langle t \rangle} \qquad d=1,2,3,\dots .$$

We consider a single Markov-process for a given t, characterized by a matrix M with only non-negative elements. M has n eigenvalues λ_i and eigenvectors u_i.

The eigenvectors can be selected as real and for the norms of the eigenvalues, the following holds

$$1 = \lambda_1 \geq \|\lambda_2\| \geq .... \geq \|\lambda_n\| \geq 0 .$$

If $\|\lambda_2\| < 1$, the eigenvector u_1 is the only attractor of Iteration (9). This means it does not matter which distribution of time-use frequencies a person starts with, because it will finally end with the distribution u_1. In other words, the ultimate result will always be the same frequency distribution. The iteration approaches u_1 more quickly, the lower the norm of λ_2. We can therefore interpret the vector u_1 as an indicator of typical time use and the norm of λ_2 as a measure of the stability of time-use behaviour. Given this interpretation, it is of great interest to estimate the values of u_1 and of λ_2, even if, in reality, a sequence of working days is interrupted by weekend. The attractor shows the typical time use which would be achieved if there were no forced or random deviations, and λ_2 shows how fast such "disarrangements" will be smoothed out in the iteration.

To conclude, we try to estimate the matrices M(t,s). We can then evaluate their eigenvalues and eigenvectors. These allow us to learn something about typical time use and its stability.

Finally, we must deal with some details of the estimation of the matrices M(t,s). For a given t, we have to count the absolute frequencies N_{ij}, indicating in how many cases activity j is followed by activity i exactly 24 hours later. From this information, we derive the absolute frequencies $N_{.j} = \sum_{i=1}^{n} N_{ij}$. We can estimate the element m_{ij} of the matrix M(t,s) by

(10) $$\hat{m}_{ij} = \frac{N_{ij}}{N_{.j}}.$$

Through Bernoulli's law, we have

(11) $$\Pr ob\left(\left|\hat{m}_{ij} - m_{ij}\right| \le \varepsilon\right) \ge 1 - \eta \qquad if \quad N_{.j} \ge \frac{1}{4 \cdot \varepsilon^2 \cdot \eta}.$$

Therefore, even if we request very modest values for ε and η as e.g. ε=0.1 and η=0.1, it should hold that

(12) $N_{.j} > 250$.

Considering Request (12), one might consider it necessary to include as many people as possible in the estimation of the matrices M(t,s). However, these matrices might depend on the vector of characteristics s and therefore, the respondents included should be as homogenous as possible with respect to socioeconomic characteristics. Accordingly, the number of respondents used to estimate the matrices M must be kept fairly small. Nevertheless, it is possible to increase the number of observations. For this purpose, we do not consider all 272 activities given in the Time Budget Survey, but aggregate them into those 18 Activity Groups given in Table 1. Further on, we assume that the matrices M(t,s) are the same, if the time intervals t occur at the same hour of the day h. Thus, the task simplifies to an estimation of 24 18×18-matrices M(h,s), h = 1 … 24. It remains a problem that some activity groups occur rather infrequently. Therefore, for some j and some h and h' (h,h'$\in\{1,\ldots24\}$), further assumptions of the form

(13) $m_{ij}(h,s) = m_{ij}(h',s)$

had to be made. For which activity groups j, and hours h, assumptions of the form (13) were made, is depicted in Table 2. The selection was also made by looking at the kind of activity on hand, as well by securing counts of more than about 400 for estimating the transition probabilities m_{ij}.

Table 2
Numbers of counts per hour and activity*

<table>
<tr><th rowspan="2">Activity group No.</th><th colspan="24">Hour of the day</th></tr>
<tr><th>4</th><th>5</th><th>6</th><th>7</th><th>8</th><th>9</th><th>10</th><th>11</th><th>12</th><th>13</th><th>14</th><th>15</th><th>16</th><th>17</th><th>18</th><th>19</th><th>20</th><th>21</th><th>22</th><th>23</th><th>24</th><th>1</th><th>2</th><th>3</th></tr>
<tr><td>1</td><td>2488</td><td>2241</td><td>1611</td><td>771</td><td colspan="2">441</td><td colspan="5">442</td><td colspan="7">343</td><td>904</td><td>1837</td><td>2308</td><td>2460</td><td>2510</td><td>2517</td></tr>
<tr><td>2</td><td colspan="4">502</td><td colspan="3">607</td><td colspan="2">621</td><td colspan="2">411</td><td colspan="3">543</td><td>386</td><td>345</td><td colspan="8">383</td></tr>
<tr><td>3</td><td colspan="3">433</td><td>384</td><td colspan="4">473</td><td colspan="9">558</td><td colspan="7">637</td></tr>
<tr><td>4</td><td colspan="4">502</td><td>660</td><td>799</td><td>894</td><td>907</td><td>694</td><td>802</td><td>874</td><td>805</td><td>621</td><td>418</td><td colspan="3">562</td><td colspan="7">256</td></tr>
<tr><td>5</td><td colspan="6">397</td><td colspan="2">1029</td><td colspan="2">407</td><td colspan="4">311</td><td colspan="10">425</td></tr>
<tr><td>6</td><td colspan="4">127</td><td colspan="2">342</td><td colspan="18">959</td></tr>
<tr><td>7</td><td colspan="12">357</td><td colspan="12">430</td></tr>
<tr><td>8</td><td colspan="12">651</td><td colspan="12">591</td></tr>
<tr><td>9</td><td colspan="4">21</td><td colspan="2">318</td><td colspan="3">716</td><td colspan="3">414</td><td colspan="12">571</td></tr>
<tr><td>10</td><td colspan="9">489</td><td colspan="5">533</td><td colspan="9">340</td><td>310</td></tr>
<tr><td>11</td><td colspan="9">422</td><td colspan="4">478</td><td colspan="2">453</td><td colspan="2">450</td><td colspan="7">342</td></tr>
<tr><td>12</td><td colspan="24">797</td></tr>
<tr><td>13</td><td colspan="11">386</td><td colspan="13">392</td></tr>
<tr><td>14</td><td colspan="11">456</td><td colspan="13">626</td></tr>
<tr><td>15</td><td colspan="14">794</td><td>635</td><td colspan="9">636</td></tr>
<tr><td>16</td><td colspan="8">525</td><td colspan="6">576</td><td colspan="4">486</td><td colspan="6">414</td></tr>
<tr><td>17</td><td colspan="11">311</td><td colspan="3">389</td><td>253</td><td>624</td><td>1062</td><td>1050</td><td>527</td><td colspan="5">276</td></tr>
<tr><td>18</td><td colspan="4">390</td><td colspan="2">465</td><td colspan="2">461</td><td colspan="2">390</td><td colspan="2">520</td><td colspan="2">635</td><td colspan="2">441</td><td colspan="8">301</td></tr>
</table>

* Connected cells show for which hours of the day the same transition probability has been assumed. The numbers in the cells show upon how many observations the estimates of the transition probabilities are based.
Source: Own illustration, based on German Time Use Survey 2001/2002.

With respect to the possible dependency of the matrices M(h,s) on the vector s, we first use the sample of 426 single people described in the introduction. This means that we assume initially that the matrices M(h) are identical for all singles belonging to this sample.[7] The number of observations which determines each estimate $\hat{m}_{ij}(h)$ with respect to Equation (11) is also given in Table 2.

Given all these steps, we are able to estimate the matrices M(h), and then evaluate their attractors u_1 and second eigenvalues λ_2. The results are described in the next section.

7 In a further step, we test whether the M(h) assume different values if the sample of 426 single people is split up into sub-samples that are characterised by elements of the vector s, such as age, sex, or income. The results are mentioned at the end of Section 4.

4 Empirical results

Due to space limitations, it is not possible to show all estimated Markov matrices M(h) in Figure 2, but rather their attractors. λ_2=1 does not hold for any of the estimated matrices. Accordingly, the attractors are unique. They give a quite credible and accurate picture of a typical working day. The highest probability is obtained during the late evening and night-time by sleeping. During the morning, travelling time, paid work, and education predominate, whereas in the late afternoon and early evening, many people seem to be watching TV. Figure 3 shows Cramer's V for the cross tables that one needs in order to subsequently derive the Markov matrices according to Equation (10). The measures of contingency have a large magnitude and are different from zero to a significant degree, providing evidence that the estimated matrices involve nontrivial information.

Subsequently, the entropies of the attractors were evaluated.[8] They are shown in Figure 4 together with the norms of the second eigenvalues. The entropy of an attractor provides a measure of the diversity of time use. The higher the entropy, the less monotonic (consistent) the selection of activities. On the other hand, the norm of the second eigenvalue is a measure of the stability of time use. The lower $\|\lambda_2\|$, the faster the vectors $x^{<t>}$ will approach the attractors u_1(h). Therefore, the smaller the norm of the second eigenvalue, the greater degree of stability.

Given these considerations, we can see from Figure 4 that the higher the entropy, the lower the value of $\|\lambda_2\|$. Hence, the more diverse the time use, the greater its stability. This relationship holds with respect to varying the time of day. In Figure 5, we can see that it holds more generally. Figure 5 was generated as follows. For each member of our sample of 426 singles, we can evaluate the entropy i of individual time use during the two observed working days.[9]

We divide our sample into two sub-samples:

ILOW: 217 persons with entropy of individual time use i<2.

IHIGH: 209 persons with entropy of individual time use i≥2.

For each sub-sample, Markov matrices were estimated, eigenvalues and eigenvectors and the entropies of the attractors were computed. Our concern is again with respect to the relation-

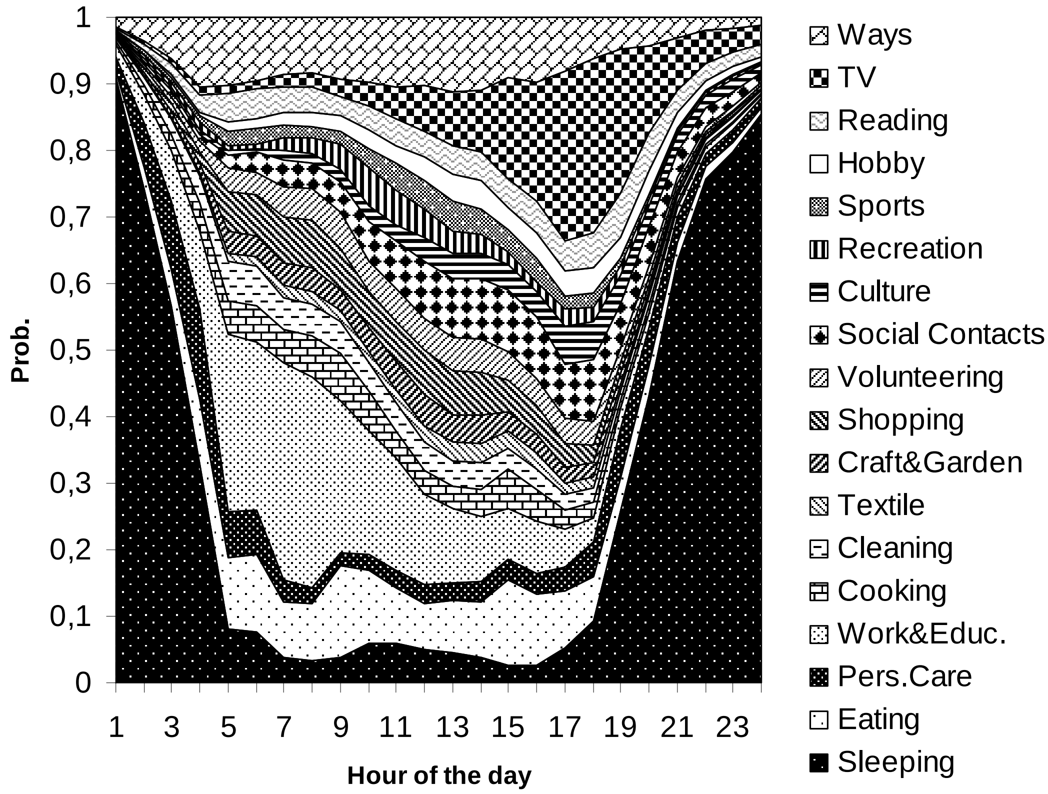
---

8 Let $u_1(h) = \left({}_1u, \ldots, {}_{18}u\right)$ be the first eigenvector of the matrix M(h). We then evaluate $i(h) = \sum_{j=1}^{18} - {}_j u \cdot \ln({}_j u)$ and denote it by "entropy of the attractor u_1 (h)".

9 For a given person, the observed relative frequency of Activity Group j during the two observed working days shall be denoted by p_j. We set $i = \sum_{j=1}^{18} - p_j \cdot \ln(p_j)$ and denote i by "entropy of individual time use" of this given person.

ship of the norm of the second eigenvalue and the entropy of the attractor. The result is shown in Figure 5.

Figure 2
Attractors of the Markov matrices M(h), for all singles (n=426)



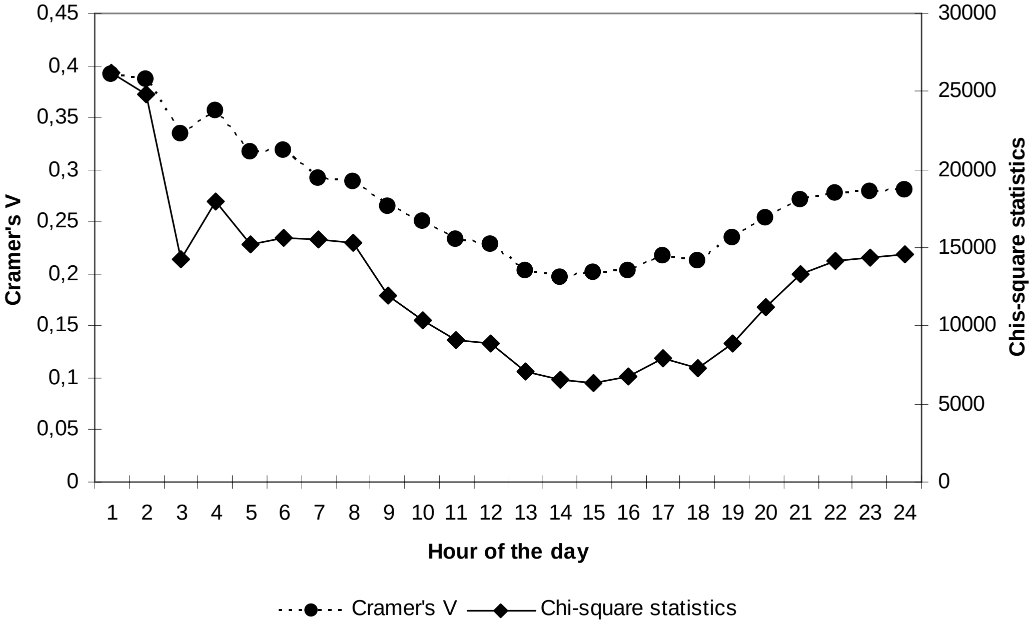
The probability of each activity group is shown as length relative to the ordinate for each hour of the day.
Source: Own illustration, based on German Time Use Survey 2001/2002.

The attractor of the sub-sample with lower *individual* entropy has a lower entropy than the attractor of the sub-sample with higher *individual* entropies.

The norms of the second eigenvalues for the sub-sample with lower individual entropy are greater than the norms of the second eigenvalues for the sub-sample with higher individual entropies. These results are not equally convincing for all hours of the day. Therefore, 90% confidence intervals were computed for $||\lambda_2||$.[10] These are shown in Figure 6. At least between 5 o'clock in the morning and 4 o'clock in the afternoon, we can say that $||\lambda_2||$ is significantly greater for the sub-sample with low entropy.

[10] The method is described in Oberhofer and Kmenta (1973).

Figure 3
Cramer's V and χ^2-statistics of the Markov matrices by hour of day



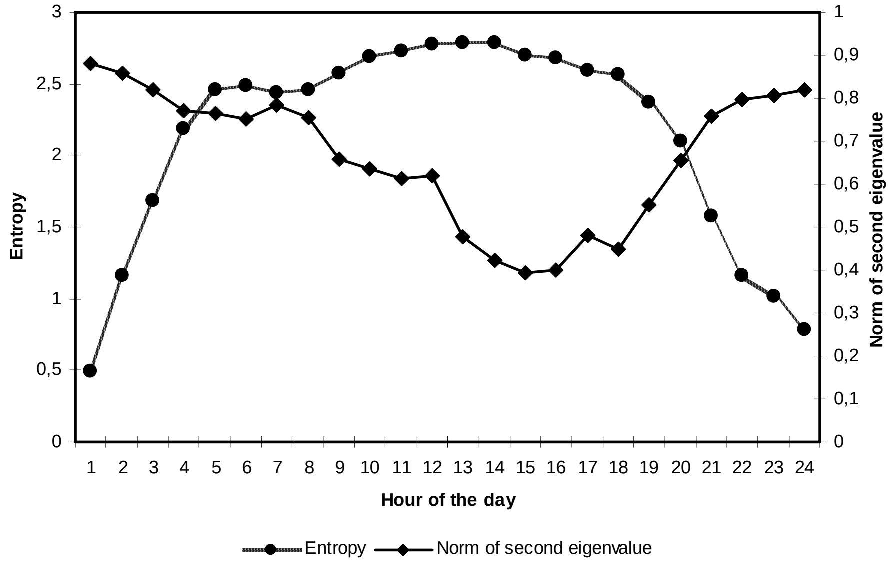
Cramer's V and χ^2 are evaluated for the cross tables, which are the basis for calculating Markov matrices according to Equation (10).
Source: Own illustration, based on German Time Use Survey 2001/2002.

Given these restrictions, one can summarize that the greater the entropy of time use, the greater the degree of stability. This result holds when the time of day is varied and if the sample is divided into two sub-samples, i.e. when the vector of personal characteristics s, is varied with respect to entropy in individual time use.

It remains necessary to report that other variations of s, such as with respect to age, income, or sex, did not produce any significant differences in the stability of the attractors of the subgroups defined in these terms.

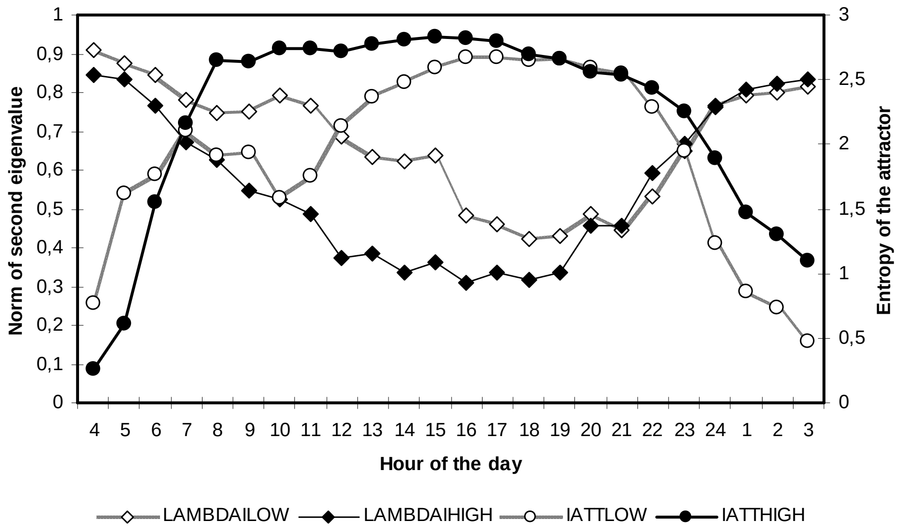
Finally, it is important to point out that our finding that high entropy is related to high stability, is not a priori. Indeed, it is easy to give some counter-examples of matrices that connect an attractor with high entropy with a high norm of the second eigenvalue. Thus, our observation that high entropy is related to high stability, is a "real" empirical finding, which could not have been deduced a priori from the model. In the following section, some possible consequences of the observed regularity, are discussed.

Figure 4
Entropy and stability of the attractors for all singles (n=426)



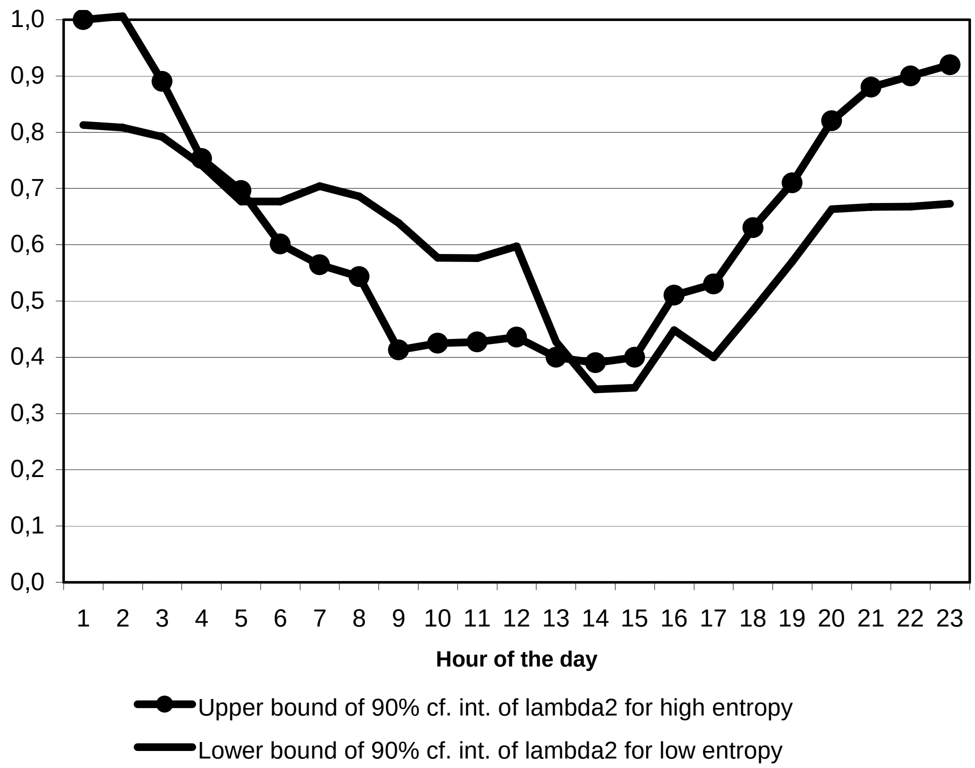
Source: Own illustration, based on German Time Use Survey 2001/2002.

Figure 5
Norm of second eigenvalue and entropy for sub-samples with low and high individual entropy



Source: Own illustration, based on German Time Use Survey 2001/2002.

Figure 6
90% asymmetric confidence intervals for λ_{2high} and λ_{2low}



Source: Own illustration, based on German Time Use Survey 2001/2002.

5 Conclusions

That flexibility ensures stability will not surprise anyone who knows the basics of systems theory. Nevertheless, this investigation provides a specific example of this theorem in the field of household science.

The results seem interesting and useful enough to look for generalizations. Firstly, it is worth investigating whether the observed regularity can also be found in other data sets. Extensions at the theoretical level can then be made. Models for households consisting of several individuals should be formalized and the iteration process include greater time lags and eventually non-linear relationships.

However, this paper certainly provides a useful introduction and starting point for such investigations. The fundamental preliminary insights are:

1. With the present data and methods, the central object of analysis should be vectors depicting probabilities of time use and connected by transition matrices.

2. These objects have further properties, such as entropy, stability and determination.
3. There are meaningful empirical relationships between these properties.
4. These properties are only weakly associated with the usual socio-economic and demographic characteristics. Thus, new dimensions of socio-economic research are appropriate objects of investigation.

These seem to be rather abstract results of basic research. Assuming that further empirical evidence will be obtained to bolster and extend these findings, practical consequences will surely emerge in due course.

We are used to judging our behaviour and the behaviour of others. In the broadest context, drug abuse, addiction, criminal or selfish behaviour, running risks of many kinds should be reduced and pollution avoided. Corpulent people should eat less, anorexic people should eat more, and for both, more exercise and sport is recommended. It is possible to add a long list of merit or demerit goods and activities. (see Musgrave and Musgrave 1984, 78). We will constrain ourselves to activities. The frequency of merit activities ought to be increased and that of demerit ones, decreased. This is the field of didactics, which uses educational advertising, training and other methods for this purpose. The author is well aware that Markov-matrices as presented are a very simple behavioural model, but let us accept them for the moment at least as the starting point for more elaborate ones and think along these lines. Changing behaviour would mean changing the attractor of time use. And this again would mean changing the transmitting matrix M. Whether this is achieved by information, arguments or training may remain open for the present considerations. The mapping from matrices to attractors is not isomorphic, so for a given attractor, there are, in general, many matrices M as inverse images. Among these matrices, one would prefer the more stable ones, if the attractor represents desirable behaviour. It is a result of this paper, however, that empirically stability is associated with higher entropy.

Accordingly, the positive relationship between diversity in time use and its stability could have implications for teaching Home Economics and such related fields as Public Health or Financial Literacy. It might not be sufficient merely to implement behaviour that is generally thought to be adequate and appropriate. In order to stabilize appropriate behaviour, it is necessary to acquire competency in reacting flexibly to disturbances, or, viewed from another perspective, we should tolerate variant behaviour, because it includes the benefit of giving stability to the sequencing of our everyday activities.

References

Falk, G. (1990), *Physik - Zahl und Realität - Die begrifflichen und mathematischen Grundlagen einer universellen quantitativen Naturbeschreibung*, Birkhäuser, Basel et al.

Gille, M. and J. Marbach (2004), Arbeitsteilung von Paaren und ihre Belastung mit Zeitstress, in: Statistisches Bundesamt (ed.), *Alltag in Deutschland - Analysen zur Zeitverwendung*, Schriftenreihe Forum der Bundesstatistik, Band 43, Wiesbaden, 86-113.

Grossmann, S. (2007), Empirische Ermittlung von Haushaltsführungsstilen mit Daten der Zeitbudgeterhebung 1991/1992, IÖB Discussion Paper No. 3/2007, Institut für ökonomische Bildung der WWU Münster, www.wiwi.uni-muenster.de/ioeb/forschen/IOEB_DP_03_2007.html.

Hufnagel, R. (2000), Die Emergenz von Ordnung im Haushalt, in: *Hauswirtschaft und Wissenschaft*, No. 2/2000, 110-122.

Hufnagel, R. (2004), Empirische Ermittlung von Haushaltsführungsstilen mit Daten der Zeitbudgeterhebung 2001/2002, in: Statistisches Bundesamt (ed.), *Alltag in Deutschland - Analysen zur Zeitverwendung*, Schriftenreihe Forum der Bundesstatistik, Band 43, Wiesbaden, 274-303.

Kahle, I. (2004), The catchword. Time Use Survey 2001/2002, in: *Methods, Approaches, Developments – Materials of the Federal Statistical Office*, No. 2/2004, 3-4.

Merz, J. and D. Burgert (2004), Arbeitszeitarrangements - Neue Ergebnisse aus der Zeitbudgeterhebung 2001/02 im Zeitvergleich zu 1991/1992, in: Statistisches Bundesamt (ed.), *Alltag in Deutschland - Analysen zur Zeitverwendung*, Schriftenreihe Forum der Bundesstatistik, Band 43, Wiesbaden, 204-336.

Musgrave, R. and P. Musgrave (1984), *Public finance in theory and practice*, McGraw-Hill, New York.

Oberhofer, W. and J. Kmenta (1973), Estimation of standard errors of the characteristic roots of a dynamic econometric model, in: *Econometrica*, No. 41, 171-177.

Statistisches Bundesamt (ed.) (2004), *Alltag in Deutschland - Analysen zur Zeitverwendung*, Schriftenreihe Forum der Bundesstatistik, Band 43, Stuttgart.

Wilson, C. (1999), Sequence Alignment Analysis of Daily Activitites, in: Merz, J. and M. Ehling (eds.), *Time use - Research data and policy*, Nomos, Baden-Baden, 505-516.

Wooldridge, J. (2006), *Introductory econometrics - A modern approach*, 3rd ed., Thomson South-Western, Mason, Ohio.