Time Use Data and Contact Patterns

Francesco Billari, *Università Bocconi, Milan*
Piero Manfredi, *Università di Pisa*
Emilio Zagheni, *University of California, Berkeley*
Joel Mossong, *Laboratoire National de Santé, Luxembourg*

February 27, 2006

Abstract

0.1 Background

Social contact patterns are relevant factors in the explanation of the spread of several infectious diseases, such as influenza, measles and tuberculosis. Analyses of mixing patterns provide decision-makers with useful information concerning the specific groups in a population that should be targeted by vaccination measures. In addition, they can help to determine the efficacy of prevention programmes. In spite of their relevance, especially in terms of policy implications, the dynamics of contact patterns in contemporary societies are still unclear and further investigation is needed.

The problem of relating the patterns of disease spread to the mechanisms of the spread of the disease agent was, in the first phase of epidemiological modelling, confronted by assuming that the host population was homogeneously mixing. Nonetheless, it was recognised that populations are made up of subgroups and that heterogeneity must affect the process of disease transmission.

The early attempts to introduce heterogeneity led to the proportionate mixing approach (Nold, 1980): within this framework, contacts between individuals are considered random, but in a heterogeneous population. Age, geographical separation and population density are some of the main sources

of heterogeneity and they are taken into account when modelling disease transmission.

The mathematical tools that have been used in the modelling of heterogeneity consist of three types of matrices: contact matrices, mixing matrices and transmission parameter matrices. Contact matrices give the number of contacts that persons in group i have with persons in group j, per unit of time. Mixing matrices give the fraction of the contacts that persons in group i have with persons in j, per unit of time. Transmission parameter matrices include the effects of the probability of transmission per single contact between a susceptible person in group i and an infectious individual in group j.

The proportionate mixing approach represents a first step towards the improvement of the homogeneous mixing scheme, but it does not take into account assortativeness (the fact that people tend to have more contacts with individuals similar to them, for instance in terms of age). In order to overcome the limitations of proportionate mixing and to model heterogeneity in a more realistic way, several approaches have been developed. On the one hand, Anderson and May (1985) suggested an approach based on the 'WAIFW'(Who Acquires Infection From Whom) matrix. On the other hand, several authors (e.g. Nold (1980), Jacquez et al. (1988) and Hethcote (1989)) have contributed to the development of the 'preferred' and 'structured' mixing models. Finally, a further line of development concentrated on directly filling in the elements of a contact matrix (e.g. Edmunds et al. (1997)).

Anderson and May (1985) focused on indirectly assessing the patterns of contacts within and between age groups. They defined a n x n WAIFW matrix, the elements of which represent effective contacts (those contacts that are likely to result in the transfer of infection) between individuals in age group i and age group j. Estimates of this matrix were derived from estimates of the force of infection for each of the n age groups. This approach has been widely used, but it shows several limitations: for example, the number of distinct mixing rates is constrained to be n, instead of n x n. As a consequence, the researcher is required to choose between certain constrained matrix structures that are unlikely to accurately represent the true contact patterns.

The preferred mixing approach was developed to improve the proportional mixing scheme. It is an attempt to introduce assortativeness: the contact matrix is built as a convex combination of a proportionate mixing matrix and a diagonal (fully assortative) mixing matrix. The basic idea behind this scheme is that people are allowed to reserve an arbitrary fraction of their group's contacts for within-group contacts while the remaining contacts are subject to proportional mixing. This type of matrix was used, for instance, by Jacquez et al. (1988) in their studies about HIV/AIDS transmission.

As for the direct approach, an example is given by Edmunds et al. (1997) who estimated age-specific contact rates on the basis of data collected from respondents who were asked to write down in a diary information about the people they had conversations with throughout the day.

0.2 Objectives

Our main goal is to investigate contact patterns by means of time use data. We assume that the time of exposure to people, by age, is a good measure for the contact patterns that could lead to the spread of airborne infections. In that sense our aim is to build a contact matrix on the basis of time use data and to get insights on mixing patterns, such as the level of assortativeness by age.

0.3 Data

Our analysis mainly relies on data from the "American Time Use Survey (ATUS) - 2003". However, we plan on extending our analysis to European countries, especially for comparative purposes: in which case the main source of data is the " Multinational Time Use Study (MTUS)", that represents a harmonised set of time use surveys composed of identically recoded variables.

The ATUS is a continuous survey on time use in the United States, whose goal is to measure how people divide their time among life's activities. ATUS covers all residents living in households in the United State that are at least 15 years old[1]. The ATUS sample is composed of approximately 40500 households annually.

For our purposes, the ATUS is particularly interesting because it does not only provide data about the amount of time each respondent spends in his daily activities. For most activities, the survey specifies where the activity took place (e.g. respondent's home, respondent's workplace, restaurant,

[1]The survey does not include active military personnel and people residing in institutions such as nursing homes and prisons.

school, mall, car, bus, etc.) and with whom each activity occurred (e.g. alone, spouse, own household child, roomates, friends, coworkers, etc.).

0.4 Methods

We consider two different approaches to analyze contact patterns by means of time use data and we try to synthesize them, in order to get insights on the structure of contact matrices.

Several time use data sets provide us with some information about the presence of other people during each activity done by the respondent. This is useful information, since it allows us to compute, for each activity, the average time spent by the respondents alone or in the presence of someone else. We can use this information to obtain the proportion of time each age group spends in the activity considered, compared to other age groups. By assuming random mixing between the age classes during the activity, it is possible to estimate a contact matrix by age group in which the ij cell is the product of the average proportion of time spent on the activity considered by the age group i and the average time spent doing the activity by the age group j. By summing up all the activities, it is then possible to obtain a contact matrix by age group for the day as a whole. Figure 1 shows a matrix computed by applying this method, for the activity 'public transportation', in the United States.

The approach we just mentioned may be very useful in describing contact patterns for those activities for which we have data about the presence of other people when the activity occurred. On the other hand, the approach relies on the assumption of the proportional mixing scheme, that might not hold in all cases.

An alternative and complementary method exploits diary-structured data: we divide the 24-hour day into many time slots, each of which consists of a few minutes. We then make the hypothesis that the number of contacts between people in age group i and age group j, for each time interval and for each activity or location considered, is proportional to the product of the amount of people in age group i and age group j that either simultaneously participate in the activity or simultaneously share the same location. By summing up all the slots first and then all the activities considered, it is possible to aggregate the 'intensities of contacts' by age. Figure 2 shows a matrix computed by applying this method, for the activity 'public transportation': the fact that people of the same age tend to use public transportation during the same time slots increases the probability that they have contacts among themselves, compared to what is predicted by a proportional mixing model.

This method is based on the assumption of random mixing within each time slot, but the process of aggregation of time slots allows a departure from random mixing, due to the fact that people heterogeneously structure their day according to their age: students, for instance, schedule their day's activities according to their scholastic engagements. As a consequence, they share the same locations for several hours a day: they use the public transportation system mainly during the same time intervals and they share classrooms. With this approach one is able to grasp heterogeneities in daily schedules by giving more weight to activities simultaneously carried out.

The two methods can be combined: The time use contact matrices that we obtain by assuming proportional mixing may be weighed by matrices of 'intensities of contacts' that we can get from diary-structured data, in order to take into account heterogeneity and assortativeness. Figure 3 shows a matrix computed by weighing the time use contact matrix and the 'intensities of contacts', for the activity 'public transportation'. It is interesting to note that the values obtained on the main diagonal of the matrix are all bigger than the respective values of the time use contact matrix based on the proportional mixing assumption: as a matter of fact, by weighing the time use contact matrix using the 'intensities of contacts', we introduced an element of assortativeness into the the time use contact matrix.

0.5 Findings

The methods we suggested allow us to describe contact patterns by age by means of time use data. The approach, based on the assumption that the time of exposure to people, by age, is a measure of contact patterns that is relevant in modelling the spread of infectious diseases, provides insights on the shape of contact matrices. In particular, we showed the effect of assortativeness on the contact patterns by introducing an element of heterogeneity into a proportional mixing scheme.

0.6 Tables and figures

	15-19	20-24	25-29	30-34	35-39	40-44	45-49	50-54	55-59	60-64	65-69	70-74	75-79		totals
15-19	**1,72**	0,71	0,82	0,75	0,70	0,91	0,66	0,87	0,71	0,42	0,35	0,47	1,26		10,34
20-24	0,71	**0,29**	0,34	0,31	0,29	0,37	0,27	0,36	0,29	0,17	0,14	0,19	0,52		4,26
25-29	0,82	0,34	**0,40**	0,36	0,34	0,44	0,32	0,42	0,34	0,20	0,17	0,23	0,61		4,97
30-34	0,75	0,31	0,36	**0,32**	0,30	0,40	0,29	0,38	0,31	0,18	0,15	0,20	0,55		4,50
35-39	0,70	0,29	0,34	0,30	**0,29**	0,37	0,27	0,36	0,29	0,17	0,14	0,19	0,52		4,23
40-44	0,91	0,37	0,44	0,40	0,37	**0,48**	0,35	0,46	0,38	0,22	0,18	0,25	0,67		5,48
45-49	0,66	0,27	0,32	0,29	0,27	0,35	**0,26**	0,34	0,27	0,16	0,14	0,18	0,49		4,01
50-54	0,87	0,36	0,42	0,38	0,36	0,46	0,34	**0,44**	0,36	0,21	0,18	0,24	0,64		5,25
55-59	0,71	0,29	0,34	0,31	0,29	0,38	0,27	0,36	**0,29**	0,17	0,14	0,19	0,52		4,27
60-64	0,42	0,17	0,20	0,18	0,17	0,22	0,16	0,21	0,17	**0,10**	0,08	0,11	0,31		2,52
65-69	0,35	0,14	0,17	0,15	0,14	0,18	0,14	0,18	0,14	0,08	**0,07**	0,10	0,26		2,10
70-74	0,47	0,19	0,23	0,20	0,19	0,25	0,18	0,24	0,19	0,11	0,10	**0,13**	0,35		2,83
75-79	1,26	0,52	0,61	0,55	0,52	0,67	0,49	0,64	0,52	0,31	0,26	0,35	**0,93**		7,61
totals	10,34	4,26	4,97	4,50	4,23	5,48	4,01	5,25	4,27	2,52	2,10	2,83	7,61		

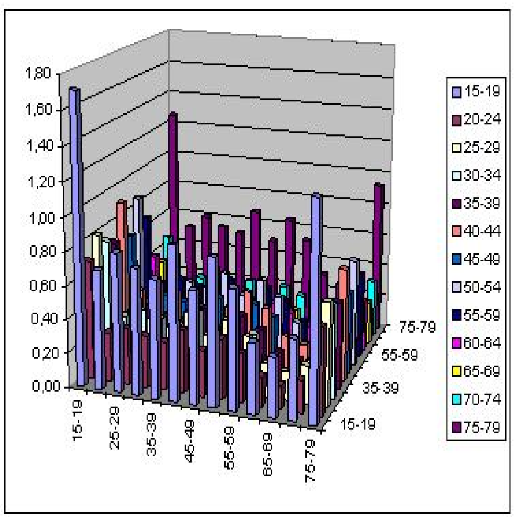

Figure 1: Time use contact matrix by age, and its graphical representation, for the activity 'public transportation (bus, subway/train, boat/ferry, taxi/limousine, airplane)' in the U.S. The entries represent the average daily amount of time that people spend together on public transportation, under the assumption of proportional mixing. Data source: ATUS 2003.

	15-19	20-24	25-29	30-34	35-39	40-44	45-49	50-54	55-59	60-64	65-69	70-74	75-79
15-19	154,01	25,66	38,47	41,25	38,82	48,47	25,04	36,82	19,38	12,08	9,67	6,03	18,04
20-24	25,66	14,64	12,65	13,35	13,26	15,33	10,62	11,80	7,88	3,91	2,58	3,20	5,81
25-29	38,47	12,65	20,66	16,21	16,13	17,92	10,13	15,62	8,68	5,12	3,71	3,08	8,53
30-34	41,25	13,35	16,21	18,65	16,52	19,62	11,52	15,22	8,88	5,37	3,70	3,31	7,59
35-39	38,82	13,26	16,13	16,52	18,53	19,94	11,29	15,02	8,81	4,81	3,28	3,67	8,09
40-44	48,47	15,33	17,92	19,62	19,94	27,48	15,98	18,54	12,24	6,28	3,74	4,25	9,38
45-49	25,04	10,62	10,13	11,52	11,29	15,98	11,66	10,90	7,53	3,42	1,90	2,91	5,42
50-54	36,82	11,80	15,62	15,22	15,02	18,54	10,90	15,39	8,65	4,78	3,06	3,02	7,78
55-59	19,38	7,88	8,68	8,88	8,81	12,24	7,53	8,65	7,31	3,00	1,35	1,87	4,48
60-64	12,08	3,91	5,12	5,37	4,81	6,28	3,42	4,78	3,00	2,17	1,10	0,85	2,30
65-69	9,67	2,58	3,71	3,70	3,28	3,74	1,90	3,06	1,35	1,10	1,34	0,69	1,88
70-74	6,03	3,20	3,08	3,31	3,67	4,25	2,91	3,02	1,87	0,85	0,69	1,35	1,97
75-79	18,04	5,81	8,53	7,59	8,09	9,38	5,42	7,78	4,48	2,30	1,88	1,97	5,53

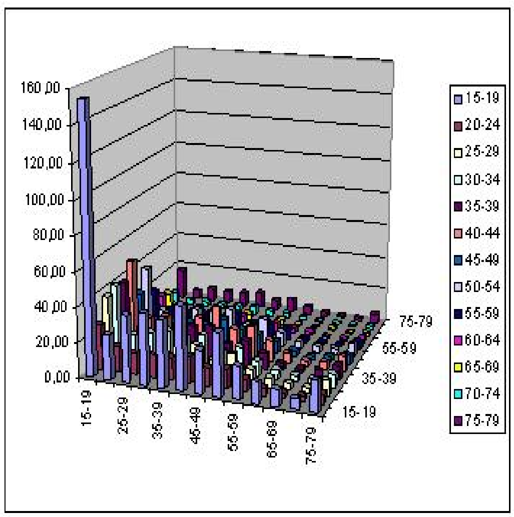

Figure 2: Matrix of the intensities of contacts by age, and its graphical representation, for the activity 'public transportation (bus, subway/train, boat/ferry, airplane)' in the U.S. The matrix captures the heterogeneity in the way people structure their activities throughout the day, according to their age. Data source: ATUS 2003.

	15-19	20-24	25-29	30-34	35-39	40-44	45-49	50-54	55-59	60-64	65-69	70-74	75-79		totals
15-19	**2,54**	0,56	0,78	0,74	0,67	0,88	0,56	0,83	0,59	0,40	0,38	0,32	1,10		10,34
20-24	0,56	**0,42**	0,33	0,32	0,30	0,36	0,31	0,35	0,32	0,17	0,13	0,23	0,46		4,26
25-29	0,78	0,33	**0,51**	0,36	0,34	0,40	0,28	0,43	0,32	0,21	0,18	0,20	0,63		4,97
30-34	0,74	0,32	0,36	**0,37**	0,31	0,39	0,28	0,38	0,30	0,19	0,16	0,19	0,50		4,50
35-39	0,67	0,30	0,34	0,31	**0,33**	0,38	0,26	0,35	0,28	0,16	0,14	0,20	0,51		4,23
40-44	0,88	0,36	0,40	0,39	0,38	**0,55**	0,39	0,46	0,41	0,23	0,16	0,25	0,63		5,48
45-49	0,56	0,31	0,28	0,28	0,26	0,39	**0,36**	0,34	0,31	0,15	0,10	0,21	0,45		4,01
50-54	0,83	0,35	0,43	0,38	0,35	0,46	0,34	**0,48**	0,36	0,22	0,17	0,22	0,65		5,25
55-59	0,59	0,32	0,32	0,30	0,28	0,41	0,31	0,36	**0,41**	0,18	0,10	0,19	0,50		4,27
60-64	0,40	0,17	0,21	0,19	0,16	0,23	0,15	0,22	0,18	**0,14**	0,09	0,09	0,28		2,52
65-69	0,38	0,13	0,18	0,16	0,14	0,16	0,10	0,17	0,10	0,09	**0,13**	0,09	0,27		2,10
70-74	0,32	0,23	0,20	0,19	0,20	0,25	0,21	0,22	0,19	0,09	0,09	**0,24**	0,39		2,83
75-79	1,10	0,46	0,63	0,50	0,51	0,63	0,45	0,65	0,50	0,28	0,27	0,39	**1,24**		7,61
totals	10,34	4,26	4,97	4,50	4,23	5,48	4,01	5,25	4,27	2,52	2,10	2,83	7,61		

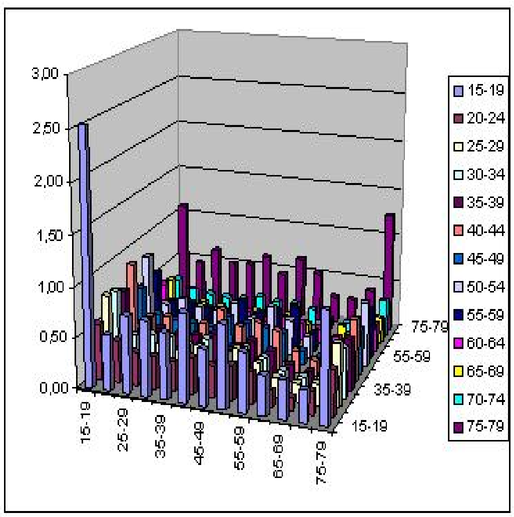

Figure 3: Time use contact matrix by age, and its graphical representation, for the activity 'public transportation (bus, subway/train, boat/ferry, taxi/limousine, airplane)' in the U.S. The matrix is obtained as a ponderation of the contact matrix based on the proportional mixing assumption and the the matrix of intensities of contacts based on how people structure their activities throughout the day. Data source: ATUS 2003.

References

[1] Anderson, R.M. and May, R.M. (1985), Age-related changes in the rate of disease transmission: implications for the design of vaccination programmes, *J. Hyg., Camb.*, 94, 365-436.

[2] Edmunds, W.J., O'Callaghan, C.J. and Nokes, D.J. (1997), Who mixes with whom? A method to determine the contact patterns of adults that may lead to the spread of airborne infections, *Proc. R. Soc. London B Biol. Sci.* 264, 949-957.

[3] Hethcote, H.W. (1989), Three basic epidemiological models. In *Applied Mathematical Ecology*, L. Gross, T.G. Hallam and S.A. Levin (eds.), Springer-Verlag, 119-144.

[4] Jacquez, J.A., Simon, C.P., Koopman, J., Sattenspiel, L. and Perry, T. (1988), Modelling and the analysis of HIV transmission: the effect of contact patterns, *Mathematical biosciences* 92, 119-99.

[5] Nold, A. (1980), Heterogeneity in disease transmission modeling, *Mathematical Biosciences*, 52, 227-240.