FP7 MULTILINKS Work package 6: **Report on methodology issues in the Multilinks project**

Arnstein Aassve

Department of Institutional Analysis and Public Management

and "Carlo F. Dondena" Research Centre for Research on Social Dynamics

Bocconi University

Via Roentgen 1

20136 Milano

Premise: The MULTILINKS research project touches on several dimensions of social science, including sociology, demography, social policy and economics. At the same time, it requires use of large-scale data analysis, both cross sectional and longitudinal. The key issue in the MULTILINKS project is to consider intergenerational linkages, measured in various ways, and also to relate them to certain outcomes of interest. This report, gives an overview of the technical aspects of the analyses implemented in the project. The report focuses on those papers that raise methodological issues from a statistical point of view. In other words, the papers included here are not exhaustive of the MULTILINKS project. Please consult the MULTILINKS website for a complete list of papers and presentations produced in the course of the project. This report builds on the shorter methodological report published earlier. The theme is the same, whereas in this report we highlight directly the papers that have applied the techniques outlined in that earlier report.

Data sources

The project members have used a range of data sources for their analysis. Examples include the Gender and Generation Survey (GGS), Survey of Health and Retirement in Europe (SHARE) and the European Social Survey (ESS). The nature and details of these surveys are well documented in the respective papers deriving from the project. The project has developed new data, in particular Wolfgang Keck and Chiara Saraceno have developed what is referred to as the MULTILINKS policy data base. In essence, this is a collection of legal and policy information for each European country, focusing on generational and gender interdependencies. For instance, they have been collecting information that concerns parental leave and intergenerational obligations. One important issue relevant from a statistical view, is that this information can be used as explanatory variables in regression models. However, the information is given at the country level, which would typically invite a Multilevel Set up of the regressions to be made. With the ESS (European Social Survey), this is certainly possible, given that there is a large number of countries, but more difficult with SHARE (Survey of Health and Retirement) and GGP (Generations and Gender Programme) for instance, where the number of countries available is smaller. We show, however, an example of how the information can be used in a simple way even when we do not implement explicitly a multilevel model.

Dealing with selection and endogeneity

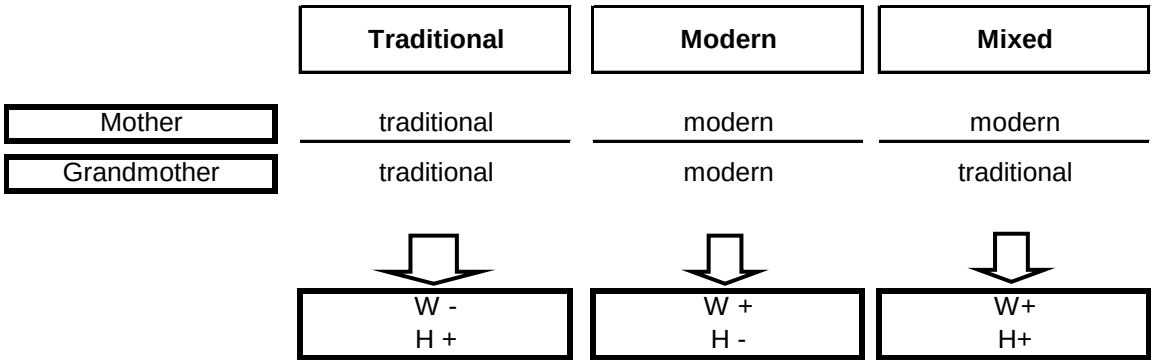
The issue of selection and endogeneity bias may arise from a variety of reasons. From a statistical point of view, the problem arises when in regression analysis we have omitted variables that may influence both the dependent and the explanatory variable, generating a correlation between the observed explanatory variables and the error term of the regression. Whereas this issue is not often dealt with when analysing inter-generational relationships, it may nevertheless be important for the statistical analysis. One possible omitted variable is that of attitudes and value orientation of the two generations involved in the relevant decision making. For instance, in terms of transfers between generations, preference, attitudes and values will matter for the outcome we observe.

In the MULTILINKS project this issue has been recognized, and one paper in particular has looked into this issue in detail. Aassve, Arpino and Goisis (2010) analysed the impact of grandparenting on mother's labour supply, but where they acknowledge that attitudes of both the middle generation and the grandparents may have an impact on the outcome – though such attitudes are unobserved in the data. Moreover, such attitudes will depend on the context in which families reside. We can think of different constellations of such attitudes. Grandparents might have rather traditional value orientations, thereby showing a high willingness to help out on childcare. Others might be of the opinion that this is mainly task for the state, in which case they might be less keen to provide childcare (Guzman, 1998). Whereas country characteristics matters for the average attitude, it is important to bear in mind that within countries – individual attitudes may differ substantially. Importantly, if provision of grandparenting depends on attitudes, then it is not unreasonable to assume that these very attitudes correlate with attitudes towards mothers working. The important point here is that mothers labour supply decision might be influenced by these unobserved attitudes of the grandparents. For instance, a grandparent with traditional attitudes will be more keen to provide childcare assuming the he or she believes that this is best undertaken by family members. At the same time, a traditional grandparent may have a negative attitude towards mothers of young children working. In the language of econometrics, grandparents' preferences for undertaking childcare are endogenous with respect to the mother's decisions to work.

The endogeneity problem is however more complex than most economic applications, simply because the outcome of interest is defined over the middle generation (i.e. the labour supply decision of the adult children), whereas the endogenous variable, childcare, is defined over the grandparents' actions (i.e. the parents of the middle generation). In the wake of these arguments, the key issue is that one needs to control for the potential bias that arises from not observing such preferences. On the backdrop of these arguments, it is useful to categorize "families" as follows:

- **Modern family** (modern woman with modern parents): characterised by highly motivated working women; weak family ties; working grandmothers (or hold a positive attitude towards women's employment); preferences for formal childcare.
- **Traditional family** (traditional woman with traditional parents): characterised by negative preferences for women working; strong family ties; positive preferences for family childcare and against formal childcare.
- **Mixed family** (modern woman with traditional parents): characterised by women highly motivated to work; grandmothers are available to help because they belong to a generation with low participation rates; mismatch across generations; strong family ties.

One can then classify family types and its relation to the observed outcomes (labour supply and grand childcare) in the following way.



Here "W" stands for work (i.e. mother's labour supply decision) and "H" stands for receiving help (i.e. childcare from grandparents). In modern types of families, ability and motivation can be a relevant omitted variable because it affects (positively) the probability to work and (negatively) the probability to receive help. This is due to the fact that highly motivated women are those living in more modern families, namely where women are not expected to stay at home caring for children. Within these families, it is more likely that both the mother and the grandmother want to work. These unobserved factors at the "family level" increase the probability of participation and decrease the probability of receiving help. If we do not control for these unobserved preferences we underestimate the grandparents' childcare effect (negative bias). Conversely, in more traditional families the probability that a woman participates to the labour market is lower and family ties are stronger. This implies higher probability to help each other. However, the probability to receive help may be lower because in these settings the mother is the primary child carer i.e. she does not need childcare support from grandparents. Moreover, for this type of families, unobserved preferences can create a downward bias if not

controlled for. A third scenario is characterized by the potential mismatch between generations. Modern women with traditional mothers (i.e. who have never worked or did so for a limited amount of time), for whom we expect to observe a higher probability of work (because of modern attitudes of young generations), and higher probability to receive help because of traditionalism (or strong family ties) in former generations. In this case, the bias created by unobserved preferences is positive, i.e. without taking them into account we would overestimate the effect of grandparents childcare.

By adopting this multi-generational perspective, we can infer the strength of such preferences and confront them across countries. The overall direction of the bias will depend on the share of families of the three types outlined before. If a country is characterised by a prevalence of consistent preferences across generations (modern or traditional families) the bias is negative. On the contrary, if there is a considerable share of "mixed" families then the bias is expected to be positive. This situation is more likely to happen in countries that experienced strong economic and cultural changes in the recent decades in the sense that such changes have created a mismatch in preferences across generation.

Aassve et al. take their data from the Gender and Generations Survey (GGS), which is a set of comparative surveys that deals with topics related to children and childbearing, partners, parents, work and everyday life. The GGS survey includes several question concerning relations between generations. These include frequency of contact, monetary and emotional support and, importantly for us, transfers in the form of time spent on childcare. Taking a sample of women in working age (aged 15-55), they consider work decision among those who have who have at least one child in the age range 0-14[1]. The dependent variable is a binary indicator taking the value one if the woman declares to be part of the labour force at the time of the interview, zero otherwise. Labour force participation is then modelled as an underlying unobserved propensity to participate W^* as a function of a set of covariates X^w and the childcare help received by grandparents H, is a dummy variable taking value 1 if woman receives childcare help from her parents and 0 otherwise:

$$W^* = X^W \beta^W + \delta H + \varepsilon^W$$

with $\varepsilon^w \sim N(0.1)$. The key technical problem here is that the error term may be correlated with the key explanatory variable H, which renders δ biased. The problem can be solved with the use of Instrumental Variable (IV) methods. Instruments, Z, are variables associated with the endogenous covariate (H) and are supposed to influence the outcome (W) only through the effect

[1] We exclude lone mothers because they are a quite different sub-group. The proportion of lone mothers is in any case very low in most of our national samples.

on the grandparents help. That is, they should not have a direct effect on the outcome. The IV method is implemented as a recursive model in which grandparents' childcare is assumed to influence the probability that a woman works as follows:

$$\begin{cases} W^* = X^W \beta^W + \delta H + \varepsilon^W \\ H^* = X^H \beta^H \qquad\quad + \varepsilon^H \end{cases} \tag{2}$$

The error terms of the two equations are allowed to be freely correlated in order to account for the possibility that some unobserved factors influence both decisions to work and receive grandparents help[2]. If the correlation between the two errors is significantly different from zero this means that there is endogeneity and the bivariate probit model has to be used in order to get unbiased estimates and the sign of the correlation gives the direction of the bias in the naive probit model. The instruments are defined as Z_1 (a dummy variable taking value 1 if the woman's mother is alive at the time of the interview. 0 otherwise) and Z_2 (number of siblings the woman has). The variables Z_1 and Z_2 are included in the set of covariates X^H but not in X^W. Both variables can be a priori considered good instruments since they are expected to be associated with the childcare received (relevance) but not with the labour supply decision (after help and other covariates are controlled for) (validity). In particular, Z_1 is expected to be positively (and strongly) correlated with "Help". This argument is strongly supported by many sociological studies, whereby it is indicated that not only grandmothers (as opposed to grandfathers) are the primary source of supply of child-care but also that care for grandchildren is more likely to affect the career prospects of the daughter as opposed to those of the sons (Tobio 2001). Moreover, after controlling for age effects, the mother being alive should not produce direct effects on the decision to participate in the labour market. As for Z_2, one would expect that having more siblings reduces the likelihood to receive childcare help by your parents because a higher number of siblings imply a potentially higher number of nephews whom the grandparents could take care of. Again, we do not expect direct effects on participation in the labour market[3].

As Aassve et al. demonstrate in their analysis, controlling for endogeneity makes a quite large difference on the estimates of grandparenting on women's labour force participation. An extract of their results are reported in the following table:

[2] More precisely, in the bivariate probit model the error terms in the two equations follow a bivariate normal distribution. Each error has zero mean and a variance equal to 1.

[3] To evaluate the validity of the instruments we use the functional form as identifying restrictions: first, we test the null hypothesis of insignificance of the instruments in the participation equation (test for validity) and then we test the null hypothesis of insignificance of the instruments in the help received (test for relevance).

Probit versus bivariate probit (controlling for endogeneity), by country.

	Bulgaria		France		Georgia		Germany		Hungary		Netherlands		Russia
Probit	0.066		0.088	***	0.444	***	0.351	***	0.147	*	0.164	*	0.167
Biprobit	0.996	***	0.657	*	-0.155		1.401	***	0.877	***	0.133		-0.246

Source: own computations on GGS data. Estimates of grandparents childcare effects from Table 2 and Table 3. *** p-value < 0.01; ** p-value < 0.05; * p-value < 0.10.

This Table compares the estimates of the effect of childcare on mothers' employment decision for the naive estimator and for the case when controlling for endogeneity. In several countries the effect of grandparents' childcare increases once endogeneity is properly controlled for. This is the case for Bulgaria, France, Germany and Hungary, and the change in magnitude is substantial. There is in other words a negative bias in the probit estimate for these countries as also shown by the negative sign of the estimated correlation among the error terms. Importantly, for these countries the effect of informal childcare in the bivariate probit model is statistical significant (however, for France only at the 10% level). Why do the estimates differ across countries in this way? The negative bias in the first set of countries is due to the prevalence of attitudes toward grandparents' childcare and against mothers' participation (traditional families) or to the prevalence of attitudes toward formal childcare and mothers participation (modern families). In societies with prevalence of modern attitudes, grandparents might have negative (unobserved) preferences for providing childcare, and these may arise if there is a norm towards public childcare, which in turn might arise because of good supply of services. In other words, if grandparents observe in the society where they reside that childcare is well provided by the state, they may feel less inclined to provide childcare. In modern contexts is also likely that mothers have positive attitudes toward participation in the labour market.

The analysis by Aassve et al (2010) indicates that endogeneity may matter a great deal when considering inter-generational relationships, and it needs to be taken into account. Techniques based on Instrumental Variables (IV) are widespread in econometrics, but appear equally relevant in other disciplines. The benefit of the approach is that it can reveal unbiased estimates even when there are omitted variables, which typically cannot easily be found in surveys.

Modelling life course trajectories.

The idea behind modelling and analysing life course *trajectories* is to improve our understanding of those individuals who undertake an "orderly" or "traditional" transition to adulthood (i.e. completion of education – obtaining stable employment – marriage – childbearing). Alternatively, if data are available, one can compare life course trajectories for different (family) generations. The analysis of life course trajectories as conceptual units represents a considerable challenge, even when focussing on a specific section of the life course such as the young adult years (Giele and Elder 1998). This happens because one needs to simultaneously take into account the timing, sequencing, and quantum of life course events.

A promising approach is to use sequence analysis (Abbott and Tsay 2000; Elzinga 2003). The benefit of this approach is that it enables us to study a complex set of life course trajectories as they actually take place, providing ideal-types of trajectories that can be interpreted and analysed in a meaningful way both in terms of theoretical perspectives and policy implications (Aassve et al 2007; Billari and Piccarreta 2006). Given a sequence representation of individuals' life trajectories, one can apply standard clustering algorithms to identify groups with similar life course trajectories (Abbott and Tsay 2000), and importantly, hold it up against the characteristics of the older generation and their life course behaviour. The comparative aspect brings further insight to this issue. It might for instance be the case that some transition patterns appear disadvantageous in some countries, but not in others. Using information on country specific policies and regulations, we can make inference on why such differences arise.

In the MULTILINKS project, there have been a number of applications of these methods. Aassve and Robette (2011) use the Gender and Generation Surveys (GGS), the British Household Panel Survey (BHPS) and the National Longitudinal Survey of Youth (NSLY) to provide sequence analysis of family trajectories for sample of 18 to 35 year olds. This means that they are able to provide sequences, and hence clustering of family trajectories for the following countries: Bulgaria, France, Georgia, Germany, Hungary, Romania, Russia, USA and the UK. Whereas there is no direct link to inter-generational support in this approach, the surveys do include a set of background information – including that of parents. This information is used in a final regression framework to assess how the different trajectories translate into different levels of wellbeing. In other words, using happiness and relative income as measures of wellbeing, the clusters of family trajectories are used as explanatory variables together with available family background information. The family histories are defined by marital status, for which the states

are “single”, “unmarried union”, “married” and “divorced (or separated)”and by parental status, where the states are “no children”, “one child”, “two children” and “three children or more”.

Aassve and Robette use *Optimal Matching Analysis* (OMA) techniques as a means to identify distinct patterns of family histories, which are then classified into typologies. OMA was first developed in computer science (Hamming, 1950; Levenshtein, 1966) and in molecular biology (more specifically the study of DNA strings), and was introduced in social sciences by the sociologist Andrew Abbott during the 80’s (Abbott & Forrest, 1986). Its principle consists in measuring the dissimilarity between pairs of sequences. Every trajectory is coded as a sequence of states and then the dissimilarity is measured for each pair of sequences. This leads to a distance matrix used as an input for further analyses, such as dimension reduction techniques (multidimensional scaling...) or clustering techniques. “Matching” pairs of sequences is done by means of three kinds of elementary operations: insertion, deletion and substitution. Each operation is assigned a cost, and the dissimilarity (or distance) between two sequences is equal to the minimal cost needed to transform one sequence into the other. The choice of the cost scheme is a crucial step in OMA. This has been a source of criticism (Wu 2000; Elzinga 2003) since in early applications, the cost scheme was often chosen arbitrarily. The scheme applied can be derived from theoretical considerations concerning the states composing the sequences, but for many applications, the relationship between theoretical predictions and the chosen cost scheme is not necessarily obvious. An alternative is that the cost scheme is derived from the data itself, which is the approach taken here. This means that substitution costs are derived from the transition likelihoods between the states. The more frequent transitions between two states occur, the smaller the substitution cost between them will be. Moreover, following Macindoe and Abbott (2004), the insertion and deletion cost (for short “indel”) is set to slightly more than half the maximum substitution cost. This prevents a too frequent use of *indel* operations and given higher importance to the moment when transitions occur.

Performing OMA and clustering on these large samples is highly computer intensive. As a means to reduce the computational burden Aassve and Robette first built a typology of family histories from a subsample of individuals and then assigned individuals (i.e. from the whole sample) to the cluster whose *medoid*[4] is the closest (Elzinga & Liefbroer, 2007; Wiggins et al, 2007). In other words, OMA is performed on a subsample of 2700 individual sequences (i.e. 300 randomly chosen sequences within each country), from which a 10-cluster typology is built[5]. Corresponding medoids are identified to which the complete sample of individuals is assigned.

[4] The *medoid* of a cluster is the individual whose average distance to all the individuals of the cluster is minimal.
[5] Practically, each country has been investigated separately so as not to miss some country-specific patterns of family histories. So 10 clusters corresponds to the level of partition that enlightens every major patterns, whether they are country-specific or more equally spread.

The assignment is done by comparing the cluster sequence by the medoids. The individual is assigned to the group whose medoid is closest. As a result, every one of the sample of 12933 women and 11059 men are assigned to one (out of 10) pattern of family history.

Women and men's trajectories are summarized in Tables 1 and 2. It is immediately clear that qualitatively they are quite similar. What differs is of course the age in which transitions occur, since family formation usually occurs about two or three years later for men. The medoid is represented by states whose meaning is as follows: S=single never been in partnership, C=in cohabitation, M=married, D=separated or divorced. The number following the letter corresponds to the number of children (3 meaning three or more) and the number following the semi-colon corresponds to the number of years spent in the state. So for instance (S0;5),(M0;13) corresponds to an individual who stayed single with no child for 5 years (from 18 to 22) and then married at 23. In a similar way for the second cluster for women, we have the medoid sequence (S0;2),(M0;2),(M1;14) which means that medoid stayed single without children for two years, then married and remained married without children for another two years, before having the first child. At the time of the interview, the medoid had been married with one child for fourteen years.

The first six clusters for women have medoids that start with marriage. Again, this is not to say that none of these respondents did not cohabit prior to their marriage, but it does not appear as dominant state in any of these clusters. The key difference between the medoids of the first six cluster are in terms the age when they married and the number of children they had by age 35. Three of these groups consist of individuals who have two children at the age of 35 but yet differ according to the age of family formation. A first pattern is made of women who marry and have children quite early, a second one of women who form a family approximately at the same age as those who end with one child or three or more children, and the last one of women who enter into family roles later than the others. For women as for men, there is one cluster labelled cohabitation. Here the medoid represents a non-marital union, and consist of individuals who have children without getting formally married. Taken all countries together this is a relatively small group. As we will see, there is heterogeneity in this group across countries.

The last clusters are composed of singles and differ depending on whether these singles are parents or not and, if they are, according to the number of children they have at 35. "Single with children" seem quite heterogeneous as they may have two, three or more children at 35.

Table 1. Women's 10-cluster typology of family histories

Pattern label	Medoid*	N	%	d**
Married with no child	(S0;5),(M0;13)***	1185	9,2	9,7
Married with one child	(S0;2),(M0;2),(M1;14)	1332	10,3	6,5
Married with two children	(S0;3),(M0;1),(M1;4),(M2;10)	3154	24,4	7,5
Married with three children or more	(S0;2),(M0;1),(M1;2),(M2;4),(M3;9)	1392	10,8	9,2
Late family formation	(S0;8),(M0;2),(M1;4),(M2;4)	1780	13,8	9,0
Cohabitant	(S0;7),(C0;6),(C1;3),(C2;2)	1127	8,7	16,0
Separated or divorced	(S0;2),(M0;1),(M1;6),(D1;9)	565	4,4	12,2
Single with no child	(S0;18)	1581	12,2	2,8
Single parent	(S0;7),(S1;6),(S2;5)	817	6,3	13,5
Total		*12933*	*100,0*	

*: The medoid is the individual whose average distance to all the individuals of the cluster is minimal.

**: Average distance to medoid. This gives a measure of the cluster heterogeneity.

***: S=single never been in partnership, C=in cohabitation, M=married, D=separated or divorced; the number following the letter corresponds to the number of children (3 meaning three or more) and the number following the semi-colon corresponds to the number of years spent in the state. So for instance (S0;5),(M0;13) corresponds to an individual who stayed single with no child during 5 years (from 18 to 22) and then married at 23.

Table 2. Men's 10-cluster typology of family histories

Pattern label	Medoid*	N	%	d**
Married with no child	(S0;6),(M0;12)	1087	9,8	8,4
Married with one child	(S0;4),(M0;2),(M1;12)	791	7,2	5,5
Married with two children	(S0;6),(M0;1),(M1;3),(M2;8)	2449	22,1	7,3
Married with three children or more	(S0;5),(M0;1),(M1;2),(M2;3),(M3;7)	875	7,9	9,2
Late family formation	(S0;8),(M0;2),(M1;8)	1129	10,2	6,4
Cohabitant	(S0;9),(C0;4),(C1;3),(C2;2)	1175	10,6	14,1
Separated or divorced	(S0;8),(M0;2),(M1;2),(D1;6)	343	3,1	12,5
Single with no child	(S0;18)	2720	24,6	2,5
Single parent	(S0;8),(S1;6),(S2;4)	490	4,4	12
Total		*11059*	*100,0*	

*: The medoid is the individual whose average distance to all the individuals of the cluster is minimal.

**: Average distance to medoid. This gives a measure of the cluster heterogeneity.

***: S=single never been in partnership, C=in cohabitation, M=married, D=separated or divorced; the number following the letter corresponds to the number of children (3 meaning three or more) and the number following the semi-colon corresponds to the number of years spent in the state. So for instance (S0;6),(M0;12) corresponds to an individual who stayed single with no child during 6 years (from 18 to 23) and then married at 24.

The summary of the Aassve and Robette analysis is as follows[6]:

- The traditional family pattern of marriage and childbearing is still highly persistent
- There is huge variation across countries in family behaviour
- Cohabitation appears to replace marriage in many countries, but not all. In countries where cohabitation is rare – it is also associated with lower wellbeing.
- Single parenthood is always associated with lower income and lower subjective wellbeing

These results can be seen from Tables 3 and 4 below. The first column describes the predominant family path of the group. The first four groups (i.e. the four first lines) consist of "traditional" family paths starting with marriage and in most cases ending up with childbearing. The only difference is that they differ in terms of the number of children. For estimating difference in income and subjective wellbeing, the third group is taken as the reference category – all estimates of income or subjective wellbeing are relative to this group. A negative number means a relative worsening – a positive value reflects a relative improvement, and the blue shade indicates that the difference is statistically significant. The percentage sign indicates how frequent the group appears in the sample.

After the fourth line, we have another six family types. They are: 5) **Late family formation**, consisting of individuals who are characterized by starting family formation at a later age – independent which family formation this might be (marriage, cohabitation, or childbearing), 6) **Cohabitants**, individuals whose path from 18 to 35 is dominated by being in a cohabiting union, 7) **Separated or divorced**, individuals whose path is dominated by having split with their partner, 8) "**Single with no children**", and 9) "**Single parent**", consisting of individuals whose path is dominated by being single parent to one or more children.

Table 3 refers to women's family paths and the outcome is defined over relative income. The general pattern is that having more than two children always puts women lower on the income distribution. The effects are worst in Russia, Romania, Bulgaria and Hungary – countries known to have low fertility rates. A surprising result is that married women with more than two children in France are also significantly worse off than if only having two children. Those who start family formation late are not so different than the reference category, apart from the UK and

[6] See also: Policy brief # 5: A Aassve, N. Robette (2011). Family paths and their implications for economic wellbeingand happiness. http://www.multilinks-project.eu/uploads/papers/0000/0039/Fifth_Policy_Brief_Multilinks.pdf

Table 3: Women's relative household income by country and family-history pattern

Household income (quantiles)	**BUL**		**FRA**		**GEO**		**GER**		**HUN**		**ROM**		**RUS**		**GB**		**USA**	
	estim.	%	estim.	%	estim.	%	estim.	%	estim.	%	estim.	%	estim.	%	estim.	%	estim.	%
Married with no child	-0,36	3,2	1,59	3,1	0,44	2,8	1,75	6,4	0,09	4,6	0,63	6,3	0,35	3,6	1,50	8,0	1,42	24,5
Married with one child	0,59	18,8	0,67	5,0	-0,16	6,6	0,98	9,6	0,48	9,9	0,69	25,5	0,32	18,1	0,91	3,7	0,62	2,9
Married with two children	*4,91*	*43,0*	*4,42*	*15,9*	*5,47*	*31,3*	*4,84*	*16,3*	*3,42*	*38,0*	*4,56*	*32,1*	*6,03*	*33,8*	*4,22*	*11,8*	*4,59*	*10,2*
Married with three children or more	-1,19	7,0	-1,10	12,1	-0,86	20,2	-0,89	7,5	-1,39	13,1	-1,44	9,9	-1,37	7,3	-0,73	12,3	-0,88	10,5
Late family formation	-0,05	9,5	0,05	18,1	-0,13	10,8	0,28	22,6	-0,03	12,3	0,78	12,4	-0,07	11,1	0,72	17,3	0,89	12,7
Cohabitant	-1,42	4,0	0,60	28,3	0,21	8,1	0,48	10,8	-0,36	5,5	-1,17	3,0	0,82	6,2	0,58	16,5	-0,83	4,6
Separated or divorced	-0,54	3,6	-0,55	2,9	-0,57	2,6	-0,87	2,5	-1,38	4,5	-0,17	3,3	-0,26	9,1	-0,57	4,0	-0,66	5,4
Single with no child	0,01	8,2	0,83	11,4	-0,43	15,8	0,20	14,1	-0,93	6,0	0,74	6,4	0,03	6,3	2,11	14,7	1,08	18,9
Single parent	-1,21	2,6	-1,11	3,2	-0,07	1,8	-1,05	10,2	-0,58	6,0	0,16	1,1	-0,42	4,6	-0,14	11,7	-1,87	10,4

USA where they are better off. Postponing family formation does not indicate any worsening in income compared to the traditional pattern.

Cohabitation is a highly interesting family group in cross-national perspective. Cohabitants are better off in France, Germany, Russia and Great Britain. They are worse off in Bulgaria, Romania and the USA. These differences reflect the different roles of cohabitation. In the USA they are a small group (4.6%) characterized by relative disadvantage. In fact, in all countries where cohabitants are worse off, they constitute small parts of the overall family pattern. In France, they are the largest group (28.3%) and they are better off than the "traditional" family group defined as married with two children.

Those separated or divorced are in all countries, bar Russia, small groups. They are worse off in Germany, USA and particularly in Hungary. As for those defined as single with no children, country differences arise. These are women without partner and without children and are clearly better off in Great Britain and the USA, but also in Romania and France. They represent a disadvantaged group in Georgia and Hungary. The last group concerns single mothers, and they are worse off in almost all countries. Worst off, relatively speaking, are single mothers in the USA, and interestingly, there is no significant effect for single mothers in Great Britain. However, in all other countries, single motherhood is associated with disadvantage.

Comparing family path types across countries according to their economic wellbeing is of course important. However, income is based on net equivalised household income, meaning that an adjustment is done for the number of adults and children living in the household. In line with standard measures of income and poverty, we used the modified OECD equivalence scale. This means that we have used an average weighting scheme for all households and also used the same weights for all countries. Whereas this is the standard approach, it might have the effect of making large households worse off almost by construction. An alternative is to use a subjective measure of wellbeing. In all samples apart from the NLSY individuals are asked about their level of happiness using a scale ranging from 0 (extremely unhappy) to 10 (extremely happy). Similarly to the previous analysis we estimate the average level of happiness for each of the family groups.

The picture is now different. The "traditional" family paths are often associated with higher happiness than those reflecting "new" behaviour. For instance, those paths dominated by separation or divorce, or single motherhood are typically associated with lower levels of happiness. There are exceptions to the rule however. For instance in Great Britain and France there is almost no significant difference in terms of happiness for the nine different family paths. The most likely reason for this is that in these two countries, the "new" family paths are also more common. For instance, in France, more than 28% of the respondents are classified into the

cohabitation group – and they do not differ statistically compared to the other family paths. In contrast, in countries such as Bulgaria or Romania, the cohabitation is group is very small, but they are also considerably worse off in happiness compared to the other groups.

Table 4: Women's happiness by country and family-history pattern

Happiness	BUL		FRA		GEO		ROM		RUS		GB	
	estim.	%	estim.	%	estim.	%	estim.	%	estim.	%	estim.	%
Married with no child	-0,19	3,2	0,07	3,1	-0,76	2,8	0,14	6,3	0,07	3,6	0,29	8,0
Married with one child	-0,17	18,8	-0,13	5,0	-0,70	6,6	0,22	25,5	-0,06	18,1	0,26	3,7
Married with two children	*0,21*	*43,0*	*0,04*	*15,9*	*-0,32*	*31,3*	*0,41*	*32,1*	*0,46*	*33,8*	*5,19*	*11,8*
Married with three children or more	-0,21	7,0	-0,40	12,1	0,05	20,2	0,23	9,9	0,09	7,3	-0,17	12,3
Late family formation	-0,17	9,5	-0,11	18,1	-0,14	10,8	0,06	12,4	-0,48	11,1	0,30	17,3
Cohabitant	-0,82	4,0	-0,38	28,3	-0,69	8,1	-1,22	3,0	-0,44	6,2	0,11	16,5
Separated or divorced	-0,72	3,6	-0,75	2,9	-2,56	2,6	-1,32	3,3	-0,38	9,1	-0,06	4,0
Single with no child	-0,85	8,2	-0,18	11,4	-0,75	15,8	-1,18	6,4	-0,40	6,3	0,24	14,7
Single parent	-2,36	2,6	-1,03	3,2	-1,89	1,8	-3,28	1,1	-0,62	4,6	-0,21	11,7

Multi-level modelling

The comparative perspective is important for our understanding of why intergenerational relationships differ across different socio-economic and socio-political landscapes. Typically comparative analysis is undertaken by comparing regression coefficients across countries (Aassve et al 2007). In other cases, countries are pooled in one sample and country differences are estimated through a set of binary variables (Uunk 2004), thereby creating country level fixed effects. Observed differences are then often explained in light of welfare regime theory (Esping-Andersen 1990; Trifiletti 1999). Whereas this is informative, it is nevertheless a crude approach to comparative analysis.

There are two main objections to this tradition of comparative analysis. First, controlling for country differences through binary variables, possibly supplemented by some interactions, produces estimates that have little direct meaning. That is, apart from indicating that countries differ in the intercept of a regression line, we cannot say much about what lies behind these differences. Second, behaviour is heterogeneous within countries (Testa and Grilli 2007). For instance, demographic behaviour and hence their linkages between generations, might be very different in South of Europe than what is the case in North Europe. Moreover, observed

behaviour in certain regions in one country might be very similar to what is observed in some regions in other countries, whereas at the same time overall country differences remain significant. In essence, this means that behaviour may not only be clustered by countries, but also by regions within countries. Such features of the surveys call for appropriate multilevel modelling (Skrondal and Rabe-Hesketh 2004). In general, these models facilitate a much more rigorous analysis of the underlying causes for the outcomes of interest. A multilevel can be expressed in the following way:

$$Y_{irc} = \alpha_0 + \beta X_{irc} + \mu S_{rc} + \gamma D_c + u_{0c} + \eta_{0rc} + \varepsilon_{irc}$$

Here Y_{irc} represents the outcome of interest measured for individual *i* in region *r* of country *c*, X_{irc} is a vector of individual characteristics, S_{rc} represents covariates measured at the regional level and D_c are variables measured at the country level. Importantly, the error term is decomposed into three different components, the first at the individual-level ε_{irc}, regional-level η_{0rc} and country-level u_{0c}. The country specific error term represent unobserved country characteristics that determine the differences in the outcome across countries. Similarly, η_{0rc} represents unobserved difference across regions within countries. One important issue here is that the decomposition will capture the fact that observations within countries and regions may not be independent of each other. Thus, in case of hierarchical data structure, failing to decompose the error term in the proper way, will produce biased estimates of the standard errors, generally leading to an overestimate of the statistical significance (Goldstein, 2003).

The multilevel structure also enables us to distinguish the variance of the response variable at country level, i.e. the variance across countries, the variance across regions and variability within countries and regions. Through the intra-class correlation coefficient ρ one can compute which is the proportion of total variance accounted for between-country variation. The coefficient is given by:

$$\rho = \frac{Var(u_{0c})}{Var(u_{0c}) + Var(\eta_{0rc}) + Var(\varepsilon_{irc})}$$

where $Var(u_{0c})$ is the variance across countries, $Var(\eta_{0rc})$ across regions in country c and $Var(\varepsilon_{irc})$ among individuals in region r and country c. The typical approach in applied analysis, is to start by estimating a model where no explanatory variables are included (also

called a null model). From this one get estimates of variance of the error terms. Thus, one can assess to what extent the total variance in the outcome is factored into country, region and individual levels. Importantly, one can now include country and region level variables to understand how country and region characteristics determine the outcome of interest. Moreover, given the decomposition of the error term we can be sure to have correct estimates of the standard errors, thus providing correct significance levels for the estimated coefficients. As variables are included at country, region and individual levels, one can assess to what extent the observed variables can explain variations of the error term – all along comparing them when estimated in the null model (when no explanatory variables were included).

There are however, some practical limitations to the use of multilevel models. For instance, with the GGS, information from eleven countries is currently available. In general, this is a too small sample to estimate country effects safely. The situation is much better for the ESS where one might have access to 25 countries (depending on which round one is using). The SHARE surveys currently include 12 wave-one countries, which again is a small number if considering country effects in a multilevel setting. However, sample size is generally large enough at the region level. Thus, in so far one has information measured at the regional level this can be easily embedded in the analysis. However, so far our empirical analysis seem to suggest that for many phenomena of interest, variation appear to take place predominantly at the country and individual levels.

The paper by Van Bavel et al (2010), entitled "Regional family cultures and childcare by grandparents in Europe", is a particularly nice example of multi-level modelling, where the focus is on inter-generational support. In contrast to the majority of previous work, this paper, rather than classifying countries in a limited number of categories, uses sub-national regions of Western as well as Eastern Europe. The paper considers a number of dimensions underlying family culture, including practices as well as attitudes. They do this by creating a range of indicators about gendered family norms and intergenerational exchange of support within the family on the level of the NUTS1 regions of Europe. The constructed regional indicators of family culture are then used in a multilevel analysis to explain outcomes and behaviours on the individual level. In doing so, they are able to consider the extent to which European mothers rely on grandparents as their main childcare providers, and how this is influenced by aspects of the family culture of the region these mothers live in. Childcare provided by grandparents for their grandchildren is one particular, and important, form of intergenerational family support. The increased participation of women in the labour market implies that the demand for childcare is increasing. As for informal care, evidence suggests that grandparents are by far the preferred informal carers when parents cannot look after the children themselves. A considerable share of European

grandparents does indeed help their children by providing childcare for the grandchildren (Smith Koslowski 2009; Hank and Buber 2009; Wheelock and Jones 2002; Lewis, Campbell, and Huerta 2008; Settles et al. 2010).

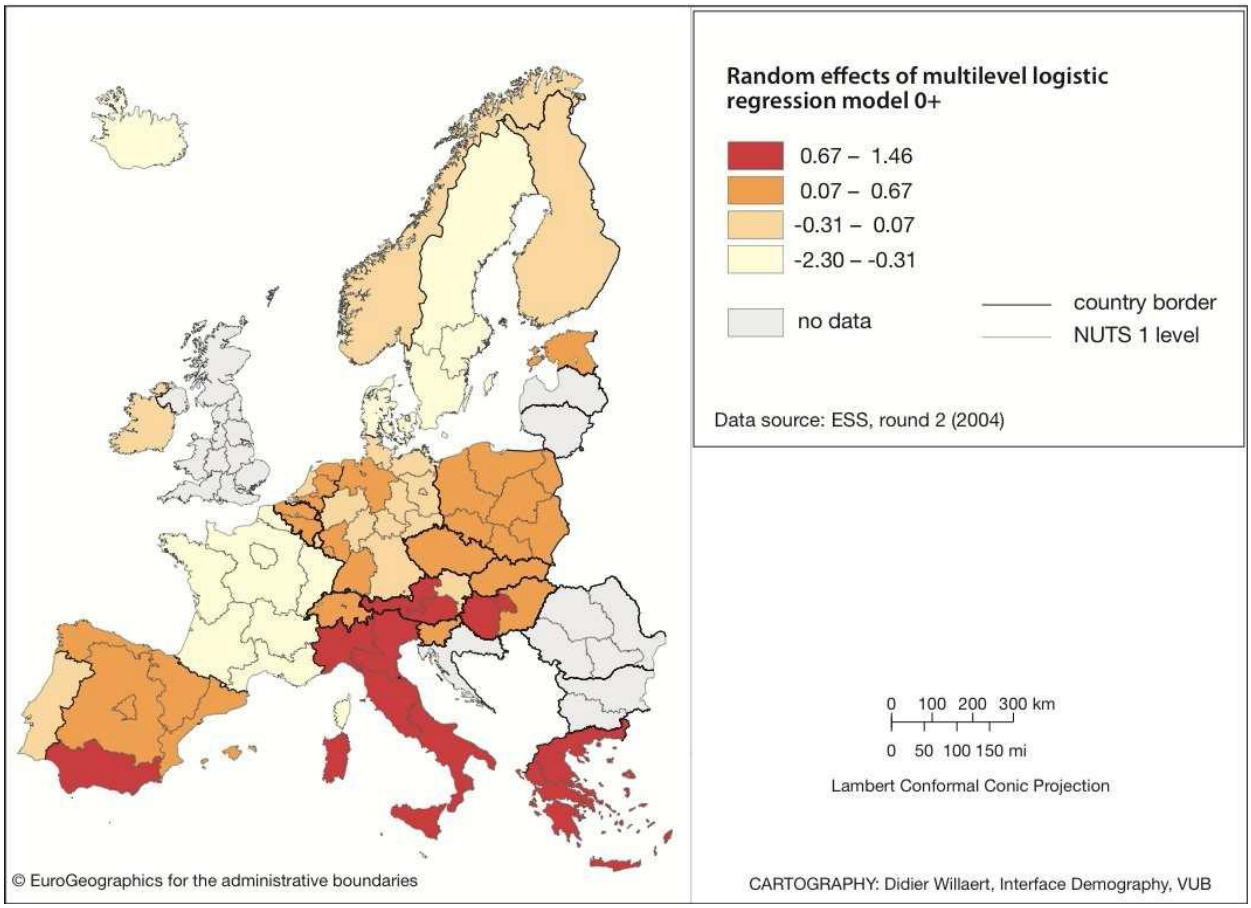
The analysis is based on the second round of the European Social Survey (ESS2), with fieldwork carried out in 2004 and 2005. ESS2 questioned respondents about attitudes towards gendered family roles, the use of childcare and intergenerational exchange of financial support and care. In addition to the ESS2, they use GDP data from Eurostat and statistics about formal childcare provision from the MULTILINKS database on intergenerational policy indicators (Keck, Hessel, and Saraceno 2009).

The European Social Survey contains information on regions within the countries surveyed, but as they point, it is not straightforward to combine this information with EU regional statistics because the ESS2 regional classification differs from the standard regional classification used by Eurostat (the EU statistical office). The ESS2 regional classification is therefore harmonized with Eurostat's NUTS system. NUTS is the Nomenclature of Territorial Units for Statistics, a coherent and standardized system for referencing sub-national regions within European countries, created and regulated by the European Union. NUTS is a hierarchical system, with three levels of NUTS defined. ESS2 includes 26 countries: Austria, Belgium, Czech Republic, Denmark, Estonia, Finland, France, Germany, Greece, Hungary, Iceland, Ireland, Italy, Luxembourg, Netherlands, Norway, Poland, Portugal, Slovakia, Slovenia, Spain, Sweden, Switzerland, Turkey, UK and Ukraine. Data from Italy are not included in the main ESS2 dataset, because the rotating modules were administered as split ballot there (ESS Central Coordinating Team 2010). This also means that only half of the respondents were asked the questions in the rotating module about family, work and wellbeing, but data for Italy were added to the dataset and included into the following analyses. Data from Turkey, Ukraine and the United Kingdom, however, were dropped from the analysis. In the end, they use data from 23 countries and 74 NUTS 1 regions. To adjust for the differences in sample design between countries, cases have been weighted by the design weight provided by ESS.

Summing up, they use multilevel logistic regression analysis to explain the extent to which grandparents are said to play a role as the main source of childcare, with explanatory variables on three levels: the individual, the NUTS 1 regional level, and the country level. Independent variables on the NUTS 1 level consist of indicators about gendered family norms and intergenerational exchange of support within the family: an indicator pointing to more of less conservative family norms in the region, the percentage of older parents coresiding with children, the percentage of older parents providing care or support with housework to adult

children outside the household, and the percentage of older parents receiving care or support from children outside the household.

Figure 3: Estimates of NUTS 1-level random effects in baseline model



The map above presents the variation in random effects between NUTS 1 regions. NUTS 1 random effects were summated with the random effect of their respective countries. It is clear that, net of the individual covariates, the probability that mothers make use of their children's grandparents for childcare differs not only between but also within countries. Nevertheless, the country appears to be the main level of variation for our dependent variable: all regions in Italy and Greece fall within the highest end of the distribution of childcare by grandparents; all regions of France are found in the low end of the distribution.

Another highly relevant paper from the MULTILINKS project exploiting Multi-level analysis, is the one by Jan van Bavel and Tom de Winter (2011). The paper is entitled "Becoming a grand parent and Early retirement in Europe", and is an application of an event history analysis with a Multilevel structure. The key idea in the paper is that the literature on ageing and retirement tends to ignore the fact that retirement decisions may depend on becoming a grandparent. Whereas this is largely ignored in the existing literature – it does make a great

deal of sense – particularly in those countries where childcare of young children is undertaken either by the mother or by grandparenting. Again, they use the ESS (third round) and estimate a discrete time event history model of the time when respondents retire as a function of becoming a grandparent. Part of the covariates is measures of the policy environment where such behaviour takes place. As before, such information is derived from the MULTILINKS database on intergenerational policy indicators.

The research questions they pose are: 1) does becoming a grandparent speed up retirement, i.e. does the rate of retirement before reaching the standard age increase upon becoming a grandparent? 2) is the effect of grandparenthood on early retirement, if there is any, stronger for women than for men? 3), is the effect of grandparenthood, if there is any, on early retirement stronger in countries where the family is supposed to be the provider of childcare and where formal childcare facilities are highly developed?

They follow all people from age 50 onwards until their current age at the time of the survey. They then start by presenting the results of a purely descriptive life table approach. Next, they turn to multivariate regression analysis. More specifically, they fit multilevel hazard models, using a logistic specification of the rate of entry into retirement, with a random effect added on the country level. The discrete time hazard regression model can be specified as follows:

$$\ln\left(\frac{h_{ic}(t,t+1)}{1-h_{ic}(t,t+1)}\right) = \alpha_t + \beta_1 G_{ict} + \beta_2 X_{ic}$$

where h_{ic} is the discrete time hazard rate of retirement at age t, α_t is the intercept at age t (with $t=0$ corresponding to age 50 in all cases). B_1 is the country effect on the rate of retirement, assumed to be a random variable drawn from a normal distribution and with mean and variance to be estimated from the data. β_2 is the main parameter of interest, i.e. the effect of being a grandparent on the rate of retirement, with G_{ict} a time-varying dummy variable indicating for person i in country c whether he or she has already become a grandparent at age t. Finally, β_2 represents the effect of the time-constant control variables.

The major drawback about the ESS for the current research is that it is a cross-sectional survey with no retrospective information about health and wealth. Respondents are asked about their current health and wealth, but not about the situation prior to retirement. As a result, the effects of health related factors and of the financial affordability on the retirement rate cannot be estimated, even if these are well-known determinants of exit from the paid labour market. Yet, to the extent that the correlation between health and wealth on the one hand and the rate of

grandparenthood on the other hand is limited, the omission of health and income from the regression analysis will not bias the estimates for the effect of becoming a grandparent.

Figure 4: Life table survivor function for entry into retirement, all 22 countries pooled, by gender and grandparenthood status.

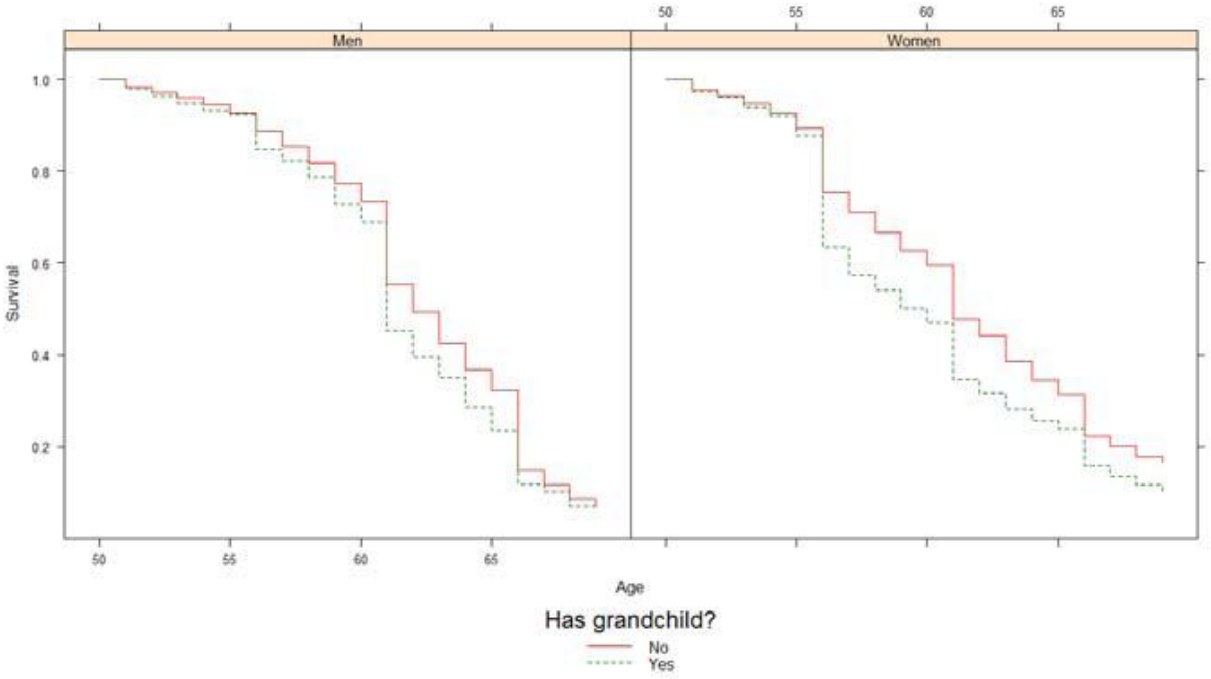


Figure 4 gives the life table survival curves separately for men and women. This plot suggests that the effect of grandparenthood on early retirement was small for men but markedly larger for women, as expected given the greater role played by grandmothers in childcare. In the pooled sample, just over 50% of the grandmothers who were previously active on the labour market are retired before reaching age 60, as compared to 37% of the women without any grandchildren a difference of 13%. Among men, the difference between grandfathers and men without any grandchildren is much smaller: before reaching age 60, 27% of the grandfathers have retired already, while this is 23% of the men without grandchildren - a difference of 4% only.

There is a positive effect of becoming a grandparent on the speed of exit from the labour market before reaching the standard retirement age. Once people have at least one grandchild, the likelihood that they retire early increases significantly. The odds of retirement are estimated to increase with 14% upon becoming a grandparent (i.e. the odds ratio associated with the effect of grandparenthood can be calculated exp(0.128) = 1.14). This effect of grandparenthood applies as long as people have not yet reached the standard age of retirement applicable to their gender. If grandparents have not retired before reaching that age, their odds of subsequent retirement are significantly lower than for peer without grandchildren. Indeed, there is a statistically significant

interaction effect between the status of grandparenthood and having reached the standard age: for elderly people without grandchildren, the effect of reaching the standard age on the retirement rate is positive, but this is not the case for grandparents. In sum, the overall picture across Europe is that, for grandparents, the highest retirement rates are observed before the standard age, while for people without grandchildren, the highest retirement ages are observed upon reaching the standard age. Finally, the considering the childcare coverage rate, they show that people aged 50 and more tend to retire later in countries with higher childcare coverage rates. A 10% increase in formal childcare coverage is associated with a decrease in the retirement rate of about 23%. Yet, the parameter for the interaction with the status indicator for grandparenthood is very small and not significant, indicating that the effect of childcare coverage is not typical for those who have already become a grandparent.

Several other papers from the MULTILINKS project use multi-level techniques. The paper by Schenk et al (2010) considers the role of welfare states on the likelihood of money transfers between generations. They use SHARE, and since they consider both the characteristics of the older generation and their children (i.e. sibling pairs), that use a multilevel structure that is defined over grandparents and their children – the latter coming from the same parents. The idea here is that one controls for the fact that there might be family specific (unobserved) heterogeneity that can influence the results. Another application by Schenk et al (2011) considers older adults' network and public care receipt, which asks specifically whether partners and adult children tend to substitute for unskilled public care. The multilevel structure lies here in that partners and the adult children belong to the same family – thus the approach controls for family specific effects.

Multi-level techniques are also applied in a paper by Moor, Komter and De Graaf (2011): 'Family, welfare state generosity and the vulnerability of elderly: a cross-national study'. The authors not only aim to indentify factors that affect subjective well-being of older people, but also try to indentify the conditions under which these factors have an impact. The focus of this paper is on the impact of living arrangements and family structure on subjective well-being. Because ageing takes place in specific sociological contexts, they explicitly take a cross-national perspective. They aim to contribute to the discussion about the moderating influence of welfare state services on the relationship between family structure and life satisfaction. The question is raised whether formal services can compensate for the absence of family resources and do or do not crowd out family care. For measuring welfare state generosity directed to the elderly, the authors made use of the Social Policy Indicators database, which is part of the Multilinks project. Instead of using a rather general measurement of welfare generosity, such as the percentage of GDP spent on social services, or categorizing counties by type of welfare regimes, a more substantive

measurement of welfare state generosity is used: the *net income replacement* granted by the most general pension scheme in the respective country to a standard pensioner, and *the share of older adults (65+) living in institutional care homes*. Additional use was made of the fourth wave of the European Values Study (2008). In order to answer the research question, the authors performed a linear regression analysis in which 7.397 older people are nested in 24 European countries. The results demonstrate that the relationship between family resources and life satisfaction is not moderated by welfare state arrangements. Family resources are equally important for older people who live in countries with more and with less generous welfare state arrangements aimed at elderly. In other words, welfare state services do not crowd out the importance of family resources for older people's subjective well-being.

Latent Class Analysis

Latent Class Analysis is an approach which relates a set of observed discrete multivariate variables to a set of latent variables. Importantly, one assumes probabilistic rather than deterministic relationships between the latent construct. A basic assumption of LCA is conditional independence, which means that associations between manifest indicators exist only insofar as they measure the same latent construct. LCA has the advantage that the classes of the latent construct are discrete and need not be ordered along a continuum (Clogg 1995). As part of the MULTILINKS project, Pearl Dykstra and Tineke Fokkema applied such techniques to the analysis of relationships between parents and their adult children, thereby constructing a West European typology of late-life families. Their analysis motivates from earlier work, which argues that family solidarity patterns are divided between an individualistic North (Reher 1998). Through LCA, Dykstra and Fokkema, use multiple dimensions of intergenerational solidarity drawn from the Survey of Health, Ageing and Retirement in Europe, and develop a typology of late-life families which is robust across northern, central and southern regions. The four types are: (a) descending familialism: living nearby, frequent contact, endorsement of family obligation norms, and primarily help in kind from parents to children, (b) ascending familialism: living nearby, frequent contact, endorsement of family obligation norms, and primarily help in kind from children to parents, (c) supportive-at-distance: not living nearby, frequent contact, refutation of family obligation norms, and primarily financial transfers from parents to adult children, (d) autonomous : not living nearby, little contact, refutation of family obligation norms, and few support exchanges. Importantly, the four types are common in each European country, though the distributions differ, thereby questioning the idea that a particular country can be characterised by a single dominant type of late-life family.

Latent Transition Analysis

An extension of Latent Class Analysis is Latent Transition Analysis. This approach enables researchers to determine the changes in latent class membership over time. LCA is performed on datasets following respondents over two or more time points. LTA estimates transition probabilities, indicating the probability that, given membership of a certain latent class at time T-1, respondents are classified as a member of a certain latent class at time T. Niels Schenk and Pearl Dykstra use LTA on data from the Netherlands Kinship Family Study. They start off by replicating a study by Van Gaalen and Dykstra (2006) which uses LCA to develop a typology of child-parent relationships. LTA is then used to describe how these relationships shift over a three-year period. Contrary to their expectations, child-parent relationships hardly shifted from one type to another. Only 5% of relationships shifted type and most did so towards a relationship that was more harmonious than three years before. Shifts were explained by transitions in the lives of both children and parents. Findings indicated that parental widowhood prompted shifts towards a harmonious relationship whereas parental repartnering led to shifts towards discordant relationships.

References

Aassve, A and N Robette (2011) "Family paths and their implications for economic well-being and happiness", MULTILINKS Policy brief #5

Aassve, A, B Arpino and A Goisis (2011) "Grandparenting and mothers' labour supply: A comparative analysis using the Gender and Generations Survey", Dondena Working Paper no 36.

Aassve A, FC Billari and R. Piccarreta (2007) "Strings of Adulthood: A Sequence Analysis of Young British Women's Work-Family trajectories", *European Journal of Population*

Aassve A, G Betti, S Mazzuco and L Mencarini (2007) "Marital Disruption and Economic Wellbeing: A comparative analysis", *Journal of the Royal Statistical Society Series A* 170(3): 781-799

Abbott, A. and J Forrest (1986) "Optimal Matching Methods for Historical Sequences", *Journal of Interdisciplinary History*, 16(3), p. 471-494.

Abbott, A and A Tsay (2000) Sequence Analysis and Optimal Matching Methods in Sociology, *Sociological Methods & Research*, 29, 3-33.

Billari, FC and R Piccarreta (2005) Analysing Demographic Life Courses through Sequence Analysis, *Mathematical Population Studies*, 12, 81-106.

Clogg, CC (1995) Latent class models. In Arminger, G., Clogg, C. C. and Sobel, M. E. (eds), Handbook of Statistical Modeling for the Social and Behavioral Sciences. Plenum, New York, 311–59.

Dykstra, P and T Fokkema (2011) "Relationships between parents and their adult children: a West European typology of late-life families", *Ageing and Society*, 31:pp545 - 569

Elzinga CH (2003) Sequence Similarity – A Non-Aligning Technique, *Sociological Methods & Research*, 31, 3-29.

Esping-Andersen G (1990) *The Three Worlds of Welfare Capitalism*. Princeton, NJ: Princeton University Press.

Gender and Generations Program, UNECE, at: http://www.unece.org/pau/ggp/Welcome.html

Goldstein, H (2003) Multilevel Statistical Models, Third edition, London: Edward Arnold.

Giele JZ and GH Elder Jr. (Eds.) (1998) *Methods of Life Course Research. Qualitative and Quantitative Approaches*, Thousand Oaks: Sage.

Guzman, L. (1999). The use of grandparents as childcare providers, *NSFH Working Paper*. University of Wisconsin-Madison: Center for Demography and Ecology.

Jappens M and J van Bavel (2011) "Regional family cultures and childcare by grandparents in Europe", Interface Demography, Vrije Universiteit Brussel, mimeo.

Hamming R.W. (1950), "Error-detecting and error-correcting codes", *Bell System Technical Journal*, 29(2), p. 147-160.

Keck W., Hessel, P. and Saraceno, C. 2009. Database on Intergenerational Policy Indicators Methodological Report (available at http://www.multilinks project.eu/uploads/papers/0000/0016/technical-report-version-1-2.pdf).

Levensthein, VI. (1966[1965]) "Binary codes capable of correcting deletions, insertions, and reversals", *Soviet Physics Doklady* (10) pp 707-710.

Macindoe Heather, Abbott Andrew, 2004, « Sequence analysis and optimal matching techniques for social science data », in Hardy Melissa, Bryman Alan *Handbook of Data Analysis*, London, Sage, p. 387-406.

Moor, N., Komter, A.E., & De Graaf, P. (2011). Family, welfare state generosity and the vulnerability of the elderly: A cross-national study. Working paper No ?

Reher, DS (1998) "Family ties in Western Europe: persistent contrasts". *Population and Development Review*, 24, 2, 203–34.

Schenk, N, PA Dykstra and I Maas (2010) "The role of European welfare states in intergenerational money transfers: a micro-level perspective", *Ageing and Society* 30: pp 1315 - 1342

Schenk, N, PA Dykstra and I Maas (2011) "Older adults' networks and public care receipt: Do partners and adult children substitute for unskilled public care?" Erasmus University Rotterdam, Department of Sociology, mimeo.

Schenk N and PA Dykstra (2011) "Continuity and change in intergenerational family relationships: An examination of typology shifts over a three-year period" Erasmus University Rotterdam, Department of Sociology, mimeo

Skrondal A and S Rabe-Hesketh (2004) *Generalized Latent Variable Modelling: Multilevel, Longitudinal and Structural Equation Models*, Boca Raton, FL: Chapman & Hall/CRC.

Testa MR and L Grilli (2006) "The influence of childbearing regional context on ideal family size in Europe", Population, 61(1-2): pp109 - 138

Tobio, C. 2001. *Elever ses petits-enfants*. In: C Attias-Donfut & M.Segalen (Eds.). *Le siècle des grands-parents. Une generation phare. ici at ailleurs.* Paris: Editions Autrement.

Trifiletti R (1999) Southern European Welfare Regimes and the Worsening Position of Women, *Journal of European Social Policy*, 9(1), 49-64.

Van Bavel, J and T de Winter (2011) "Becoming a Grandparent and Early Retirement in Europe" TransEurope Social Sciences research network, Working paper No 28

Van Gaalen, RI and PA Dykstra (2006) "Solidarity and conflict between adult children and parents: A latent class analysis." *Journal of Marriage and Family*, 68, 947-960.

Wu L, (2000) "Some comments on Sequence analysis and optimal matching methods in sociology: review and prospects", *Sociological Methods and Research*, 29(1), p.41-64.

Uunk W (2004) "The Economic Consequences of Divorce for Women in the European Union: The Impact of Welfare State Arrangements". *European Journal of Population*, 20, 251 285.