ISSN 2042-4116

ESRC Centre for Population Change • Working Paper 71 • October 2015

centre for population change

Multilevel models for cross-national comparisons: The association between individual and national-level demographic characteristics in fertility and partnerships

Mark Lyons-Amos

E·S·R·C ECONOMIC & SOCIAL RESEARCH COUNCIL

Improving our understanding of the key drivers and implications of population change

ABSTRACT

Multilevel models are increasingly used in social sciences and demography to both account for clustering within higher level aggregations and evaluate the interaction between individual and contextual information. While this is justifiable in some studies, the extension of multilevel models to national level analysis- and particularly cross-national comparative analysis- is problematic and can hamper the understanding of the interplay between individual and country level characteristics. This paper proposes an alternative approach, which allocated countries to classes based on economic, labour market and policy characteristics. Classes influence the profiles of three key demographic behaviours at a sub-national level: marriage, cohabitation and first birth timing. Individual data are drawn from a subset of the Harmonized Histories dataset, and national level information from the GGP contextual database. In this example, three country classes are extracted reflecting two Western patterns and an Eastern pattern, divided approximately along the Hajnal line. While Western countries tend to exhibit higher levels of family allowances albeit accounting for a lower share of spending which is associated with lower marriage and later fertility, Eastern countries generally show a higher share of spending but at lower absolute levels with lower cohabitation rates and early fertility.

KEYWORDS

Multilevel models; Latent class analysis; Cross-national; Marriage; Cohabitation; Fertility.

EDITORIAL NOTE

Mark Lyons-Amos is a Fellow in Population Health at the London School of Economics and Political Science. Mark previously worked on Fertility themed projects within the Centre for Population Change.

Correspondence to: Mark Lyons-Amos, m.lyons-amos@lse.ac.uk

ESRC Centre for Population Change

The ESRC Centre for Population Change (CPC) is a joint initiative between the Universities of Southampton, St Andrews, Edinburgh, Stirling, Strathclyde, in partnership with the Office for National Statistics (ONS) and the National Records of Scotland (NRS). The Centre is funded by the Economic and Social Research Council (ESRC) grant numbers RES-625-28-0001 and ES/K007394/1.

This working paper series publishes independent research, not always funded through the Centre. The views and opinions expressed by authors do not necessarily reflect those of the CPC, ESRC, ONS or NRS.

The Working Paper Series is edited by Teresa McGowan.

Website | Email | Twitter | Facebook | Mendeley

MULTILEVEL MODELS FOR CROSS-NATIONAL COMPARISONS: THE ASSOCIATION BETWEEN INDIVIDUAL AND NATIONAL-LEVEL DEMOGRAPHIC CHARACTERISTICS IN FERTILITY AND PARTNERSHIP

TABLE OF CONTENTS

1. INTRODUCTION 1

2. METHODOLOGICAL DISCUSSION.. 3

2.1. FIXED EFFECTS.. 6

2.2. LATENT CLASS ANALYSIS.. 7

3. NUMBER OF CLASSES 9

4. DATA AND MODEL 10

5. RESULTS 14

5.1. CLASS CHARACTERISTICS.. 14

5.2. INDIVIDUAL CHARACTERISTICS.. 17

6. CONCLUSIONS 21

REFERENCES.. 24

1. INTRODUCTION

Multilevel modelling is increasingly popular within both general social science and demography. The importance of the interaction between macro and micro level is recognised within demographic literature (e.g. Neels et al. 2013, Billingsley and Ferrarini 2014, Perelli-Harris and Sanchez-Gassen 2012). Multilevel models seem a natural method to be applied for these types of research questions prima facie, due to their ability to attribute partition variation in demographic outcomes between individual and contextual influences, and explicitly model the relationship between individual and macro level characteristics.

Within the longitudinal context, these models will frequently be applied to examine the effect of national level policies on lifecourse events. In demography, this frequently takes the form of examining the effect of either policies on fertility and partnership behaviour or the effect of some other exogenous factor (e.g. Neels et al. 2013, Billingsley and Ferrarini 2014). The major advantage of this approach is that it allows the integration of policy indicators as contextual variables with more general clustering parameters to capture unobserved or unspecified country level characteristics (e.g. cultural variation). Country level variation is also of interest in many other research areas, such as politics and sociology (Hox et al. 2012)

Unfortunately, standard multilevel (random effects) models and the aim of examining micro-macro interactions are not coherent when the higher level clusters are countries. Specifically, the model will complicate the interpretation of country specific variation, the small number of countries can hamper model estimation, and the fundamental assumptions of the multilevel model are not compatible with national level data. Random effects multilevel modelling is therefore not an appropriate solution when examining the effect of national level characteristics on individual demographic behaviour. Unfortunately, fixed effects models provide little by way of solution, due to their limited ability to provide inference and statistical inefficiency. Fixed effects models provide some advantages over random effects models in this respect, but have other issues such as the inability to be generalised beyond their country level sample, potentially inefficient estimation for large numbers of countries and, critically where country level policies are of research interest, the inability to include covariate information at the cluster level.

Some authors have attempted to overcome the limitations of a purely data driven approach by using a priori specifications of country typologies. Esping-Anderson (1990) classified countries as belong to different social welfare regimes within Western Europe; countries belonging to either Liberal, Corporatist-Statist or Social Democratic welfare regime. Blossfeld and Drobonic (2001) extended this to incorporate former Socialist countries. Further attempts at linking welfare regimes to demographic behaviour have been made by Hofacker et al. (2006) characterising countries as being Conservative (Germany and the Netherlands), Southern (Italy and Spain), Liberal (UK and US), Social Democratic (Sweden, Norway, Denmark) and Post-Socialist (Czech Republic, Estonia, Hungary), as well as other demographic examples (Korpi et al. 2013, Korpi 2000, Kalwij 2010) There are considerable advantages to the typology driven approach. The typologies derived will show few of the disadvantages of a purely empirically based approach, since typologies can be derived from a small sample of countries and by their nature are interpretable. Additionally, the grouping derived will be conceptually valid and consistent with existing theoretical understandings; this potentially may not occur in a purely empirical approach. That said, the major drawback of the typology based approach is that typologies have to be specified a priori by the researcher, and the ability to validate these grouping can often be neglected when linking typologies to the variable of interest.

In this paper I propose an alternative means for the analysis of individual and national interactions, through the use of two-level latent class growth models. These models provide the ability to generate clusters at the national level, but based on observed characteristics, rather than the distributional characterisation of random effects models. The observed characteristics used to define the class are interpretable, providing a substantive explanation. This has the additional advantage of being a means by which theoretically derived country level typologies (Esping-Anderson 1990, Hofacker et al. 2006) can be validated empirically. I apply this model to data from the Harmonized Histories dataset and Gender and Generations Programme (GGP) contextual database, which captures individual level demographic and country level welfare data in the European context. Individual level demographic behaviour is measured through the three processes of the timing of first marriage, the timing of first cohabitation and the timing of first birth. I classify countries based on relevant socio-economic (family allowance, social support) and legal (recognition of cohabitation within the legal framework) characteristics, and allow the timing of demographic behaviour to vary by class. This provides results which explicitly model micro-macro level interactions, without the loss of information associated with traditional multilevel models, and countries clustered within clusters which afford a substantive interpretation.

2. METHODOLOGICAL DISCUSSION

Multilevel models are a form of regression, which in its most basic form functions by capturing deviations in clusters from the overall regression equation by partitioning error terms. The most basic model is described by equation 1.

$$y_{ij} = \boldsymbol{\beta}'\boldsymbol{x_{ij}} + u_j + \varepsilon_{ij}$$

$$\boldsymbol{u_j} \sim \boldsymbol{N}(\boldsymbol{0}, \boldsymbol{\sigma_j^2})$$
$$\boldsymbol{\varepsilon_{ij}} \sim \boldsymbol{N}(\boldsymbol{0}, \boldsymbol{\sigma_{ij}^2})$$

Eq. 1

In equation 1, the overall population line is captured by $\boldsymbol{\beta}'\boldsymbol{x_{ij}}$, which is composed of a vector of coefficients $\boldsymbol{\beta}$ with a corresponding vector of covariate information $\boldsymbol{x_{ij}}$ for each individual in the dataset (including a vector of 1 to capture the intercept term).

In a longitudinal context, y will typically be a series of observations at the individual level, either in the form of repeated measurement for a growth curve model, or a string of 0s followed by a 1 for s survival model (followed by censoring). Under certain circumstances, individuals can have their own random effects, which allow disturbance from their cluster specific regression line (e.g. growth curve, survival model for repeated events) via an additional random effect.

Deviations from the overall population line due to country level variation are captured by the term u_j, which is a draw from a normal distribution. Deviation is assumed to be due to a sampling process with an approximately normal sampling distribution- the stochastic nature of this model leads to models like that in equation 1 being termed random effects models. Residuals (cluster level deviations) will typically be shrunk or precision weighted, to take account of the fact that *j* level units with few individuals (*i* level observations) will be unreliably estimated (Efron and Morris 1973). Individual (within country) level variation is captured by $\boldsymbol{\varepsilon_{ij}}$.

This model is attractive to social scientists wishing to investigate the interaction between individual and higher order clusters. The model will present corrected standard errors (and in certain circumstances corrected β estimates) to correct for the dependence between individuals within a *j* unit. Moreover, the model allows for the estimation of the correct effect of individual and contextual covariates, which may not occur with ordinary least squares (OLS) models (ecological fallacy). Secondly, the model presented explicitly partitions variation from the overall pattern into individual and higher level variation, and provides interpretable value such that the researcher can examine the proportion of variation attributable to higher order clustering. This interpretation is often related to the research question of interest and is an advantage over other techniques which seek merely to correct for correlation structures such as generalised estimating equations, or other so called 'correlation as a nuisance' corrections (Snijders and Bosker 1999). Finally, the model can be extended to allow random effects for β coefficients (random slopes model) and interactions between *j* and *i* level information (cross-level interactions). This is perhaps of most interest to scholars when investigating the association between individual and higher level characteristics.

That said, standard multilevel models have some limitations which means that they may not be methodologically germane to research questions which specify certain types of cluster, for example countries, at the *j* level.

Technically, the fact that many analyses will typically be limited in the number of countries analysed (for example Neels et al. 2013 analyse only 14 countries, Billingsley and Ferrarini 2014 use 21). This can be problematic when trying to obtain estimates for σ_j^2, due to a lack of precision and since many iterative methods (such as IGLS) will assume normality, which is difficult to verify with such a small sample. Small samples sizes can typically result in underpowered analysis, Hox et al. (2012) finding that a sample size of at least 20 higher order units is required for the accurate interpretation of regression coefficients, and a sample size of at least 50 higher order units is required for sufficiently powered analysis of variance parameters. The use of Bayesian estimation techniques can produce more reliable estimates for a far lower number of higher level units (Hox et al. 2012). However, the small number of j units can mean that reliable estimates require exceptionally long model runs (when using MCMC, Browne 2009) or the use of informative priors, although there is protection afforded against certain estimation problems e.g. the production of negative variance which can result from ML estimation. An additional problem is that the assessment of convergence for Bayesian models is more demanding than for ML estimation, requiring intensive consideration of MCMC runs.

Interpretation can also be difficult when trying to establish country (or cluster) specific effects. The deviations of individual countries from the overall population line are now captured by the parameter u_j. This means that interpreting the deviation for a particular country requires interpretation of posterior or empirical Bayes residuals, and is not intuitive. Where the model is more complicated than equation 1 (for example through the addition of random slopes), country specific estimates can become increasingly obtuse.

Finally, it is questionable whether this model is conceptually valid. The fundamental assumption of the random effects model is that $\boldsymbol{u_j}$ approximates variation that is characteristics of a sampling process, where the countries in the observed dataset were drawn at random from a larger population. For many researchers, the selection of countries within a dataset is non-random and may often be purposive. At best, it is unclear what the population of countries to which inference is being made actually is, since enumeration of higher order units will tend to be complete or approaching the finite population from which they are drawn (Stegmueller 2013). Further, it is assumed that the j level deviations from the population line can be well approximated by a draw from an *i.i.d.* Normal distribution. Evidence in demographic literature of the existence of country typologies (Korpi et al. 2013, Korpi 2000, Kalwij 2010) or groupings of higher order units within multilevel models (Billingsley and Ferrarini 2014) would indicate that this assumption is shaky when countries are specified as the level 2 unit. The ability to classify countries according to characteristics could indicate that a) countries are not taken from an independent draw, due to the demonstrated similarities and b) the fact that discrete groups are formed would indicate that a Normal (or indeed any continuous) distribution is not appropriate due to the 'clumpyness' of these groupings.

2.1. FIXED EFFECTS

A common means of overcoming some of the disadvantages of random effects models is the use of fixed effects models. Fixed effects models take the form of equation 2.

$$y_{ij} = \boldsymbol{\beta}'\boldsymbol{x_{ij}} + \sum_{\boldsymbol{j=1}}^{J} \boldsymbol{\gamma_j}\,\boldsymbol{D_j} + \varepsilon_{ij}$$

$$\varepsilon_j \sim N(0, \sigma_{ij}^2)$$

Eq. 2

In this equation, membership of a particular cluster is indicated by the vector of binary variables $\boldsymbol{D_j}$ which takes the value 1 when an individual is a member of cluster *j*, and 0 otherwise. The magnitude of deviation from the overall population line is described by the vector $\boldsymbol{\gamma_j}$. To identify the model either one D_j or one γ_j is constrained to zero.

This has some advantages over, and can overcome some of the conceptual limitations, of random effects models. Since the deviations are now not described by a continuous distribution, the model in equation 2 makes no distributional assumptions and hence is protected against misspecification. This also allows deviation from the overall population line to exhibit the clumpy characteristics expected under circumstances where regime typologies exist (e.g. Elzinga and Liefbrouer 2007). Moreover, the fixed effects model is commonly applied under circumstances where causal inference are required to be made about some $\boldsymbol{\beta}$. Since the model will remove via differencing both observed and unobserved j level variation, fixed effects will go a long way toward removing potential correlations between both *i* and *j* level error terms and any $\mathbf{x}$.

That said, there are drawbacks to this approach. Firstly, $\boldsymbol{D_j}$ refers only to the sample under analysis by necessity (unobserved countries do not have an estimated γ_j). Therefore, it is difficult to make inference beyond the observed data. Secondly, where there are relatively few observations within each j unit then the estimated values of $\boldsymbol{\gamma_j}$ will tend to be unreliably estimated, with the result that estimates may be extremely large or extremely small compared to the parameter (Goldstein 1997). This contrasts with the shrunken residual estimated in the random effects model. Indeed, the random effects model in general will tend to be more efficient, since the deviations from the population line are captured by one parameter ($\boldsymbol{u_j}$), while the fixed effect model requires the estimation of *J*-1 dummy variables (for a model with *J* clusters). Finally, the use of fixed effects models can preclude the use of *j* level covariates since these are confounded with $\boldsymbol{D_j}$. This is a distinct disadvantage when seeking to understand the effect of country level information, such as national level policy. Moreover, it precludes the use of more complicated models, such as those including random slope parameters or cross-interactions.

2.2. LATENT CLASS ANALYSIS

To overcome the limitations highlighted, I propose the use of Two-level Latent Class models, which I will demonstrate go some way to overcoming the limitations of both random and fixed effects models for longitudinal data analysis. Two-level Latent Class models can be thought of as structural equation models, and are represented as such in Figure 1.

Within the figure, we observe the response variable y at a series of timepoints t=1,t=2,…t=T observed at the level of the individual. The level of the response variable at each time point is determined by intercept (i) and slope (s). The shape of the slope trajectory can take a variety of shapes (quadratic, cubic) to fully describe the data. The values of the intercept and slope are allowed to vary according to membership of a higher level class, C, which is determined at the cluster (country level). In this instance, class membership is determined by a set of relevant cluster level variables here denoted by the vector **x**.

This specification is somewhat limited, in that it assumes that country level variation is an adequate descriptor of variation within trajectories. Extensions can be made to include classes at the individual level, which describe between individual variation in lifecourse trajectory (e.g. Dariotis et al. 2011), but in order to ensure comparability to the random and fixed effects model already described these are omitted in the current analysis.

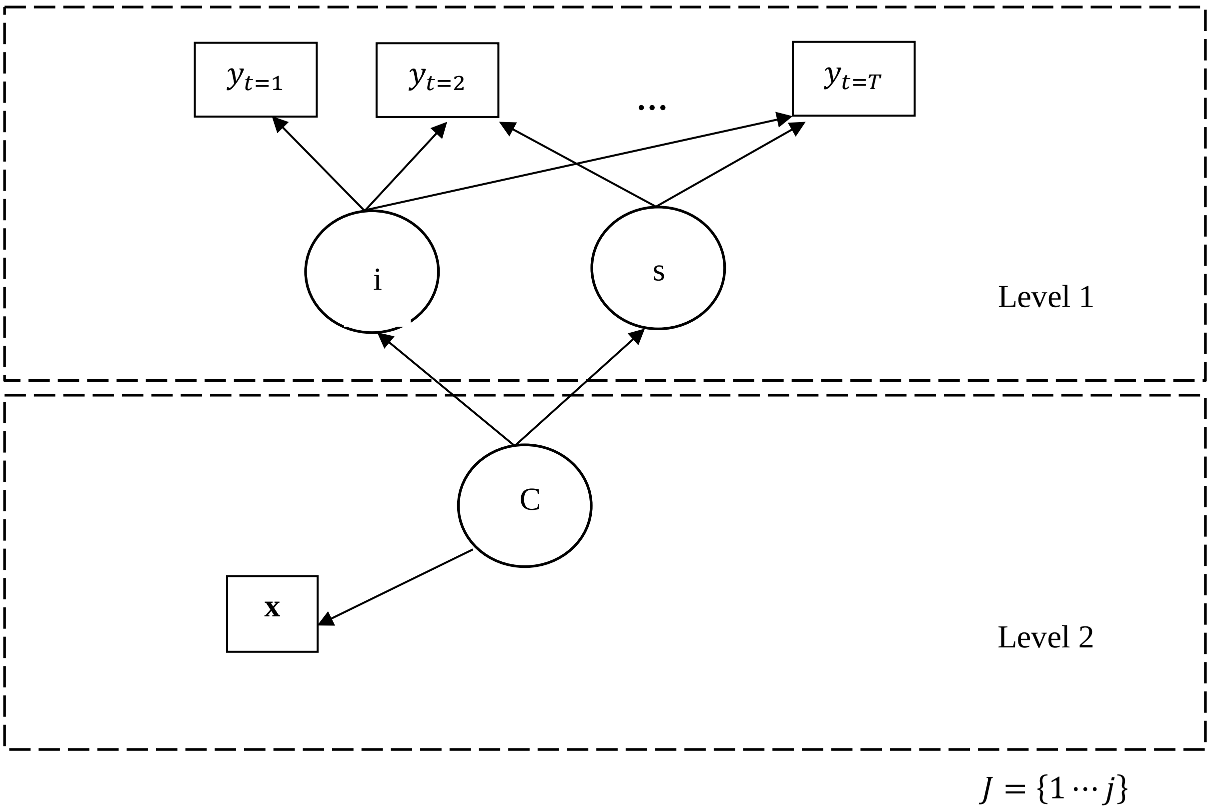

Figure 1: Structural equation representation of two-level latent class model employed

There are a number of advantages to this specification of the model over both the random and fixed effects models described. Firstly, the fact that the cluster level effect is described by classes means that the requirement of specifying a Normal (or indeed any other distribution) is no longer required, and hence clumpyness within the level 2 distribution can be accurately captured. Moreover, the fact that the level 2 random effect is now generated from manifest variables (**x**) the class can be easily interpreted and ascribed qualitative meaning in terms of the level of the contextual information. The model has the advantage over the fixed effect model that since C is a 'random' effect, it is possible to make inferences beyond the data within the sample at cluster level (assuming that clusters belong to one of the estimated classes). There is also greater opportunity to include contextual information, since x is unconfounded with C and include interactions between cluster and individual level information, by interacting **x** with *i* or *s*.

3. NUMBER OF CLASSES

A disadvantage of the latent class approach is that the researcher needs some means by which to decide on the number of classes. One approach to specify classes a *priori* according to typologies identified from theoretical literature (e.g. Esping Andersen 1990, Blossfeld and Drobnic 2001, Hofacker et al. 2006, Korpi et al. 2013, Korpi 2000, Kalwij 2010). Whilst there are theoretical clusters which are germane to this analysis, such as the welfare typologies identified by Esping-Andersen (1990), Hofacker et al. 2006 or Korpi et al (2013), I prefer data a driven approach which will provide an overall best fitting solution, and does not preclude qualitative interpretation of classes consistent with those in existing literature a *posterori* as well as empirical validation of theoretical clusters.

The major requirement therefore is to identify the number of classes which provide parsimonious model fit. A variety of statistics are available based on overall goodness of fit (Akaike/AIC, Bayesian Information Criterion/BIC, Sample-size adjusted BIC) and adjusted likelihood ratio tests (LMR-LRT; Lo Mendel Rubin 2001). Nylund et al (2007) provide a simulation evaluation of the performance of these tests of goodness of fit, and find that in general the LMR-LRT will tend to be conservative. In this instance, I therefore make use of the AIC, BIC and SS-BIC to decide on the number of classes.

4. DATA AND MODEL

Country level data are drawn from two sources. I take economic and social data from the GGP contextual database. Most recent values (2009) are taken in all cases. Note that this clearly makes it difficult to establish the causal direction of the variation between class and demographic behaviour (the country level variable are measured after the demographic behaviour), and I proffer no such claims in this paper. It should also be noted that there is some missing information for some of the economic variables in the GGP contextual dataset. An advantage of the Two-level Latent Class approach is that estimation can be performed even when there is some missing data at either level. Full Information Maximum Likelihood, implemented in programmes such as Mplus (Muthen and Muthen 2006), can effectively account for missing data under MAR circumstances, and can perform better than multiple imputation for some multilevel models (Larsen 2011). Despite some degree of missingness in the economic variables in the GGP database therefore, I am effectively able to incorporate these countries into the modelling process.

Country level information is obtained from the GGP contextual database, using the most recent year as the source of information on the grounds that this provides the most complete set of data for the selected countries. This is a limitation of this analysis. Clearly, the use of policy data from a point in time either contemporary or following the lifecourse processes under study introduces an ambiguity into the causal direction of the variables under consideration: policies may influence lifecourse behaviour; both the lifecourse processes and the policies pertaining to them may be endogenous with respect to broader cultural trends within the country of study, or policies may be derived to reflect social pressure within the country (Perelli-Harris and Sanchez-Gassen 2012). However, similar to Neels et al., (2013) it is not the aim of this paper to make causal statements about the particular economic variables included; rather the indicators are taken as manifest representations of the generosity of the welfare system in a country. It may be possible for researchers applying this method to construct more robust indicators and thus make causal statements, but this is beyond the scope of the current analysis.

I include three indicators designed to capture the generosity of the welfare state with regard to childrearing consistent with (Kalwij 2010). The absolute value of child allowances provided by the state in 2005 US$ (PPP adjusted) is included as an indicator of welfare provision. Consistent, with Neels et al. (2013) this is not intended as a robust economic indicator and hence is included for one time point only. I also include three indicators designed to capture the 'effort' that countries are providing in terms of child support: the percent of GDP devoted to family allowances and the percent of GDP devoted to state funded childcare indicate the direct prioritisation of family behaviours within the welfare system, while the percent of GDP devoted to social support in general captures the degree of support for the welfare system within the country context.

As noted by Billinglsey and Ferrarini (2014) and Kuehner (2007), merely including expenditure data is unlikely to adequately capture the design of welfare policies, and are particularly inadequate in the context of fertility variation which is strongly related to gender equality. I also include two variables designed to indicate the integration of mothers into the workplace. This has been shown to influence the fertility rate in certain settings, due rising female employment and increased difficultly in fulfilling multiple roles (Mason and Oppenheimer 1998, McDonald 2000, Mills 2010.) The female labour force participation rate is included as a measure of the extent to which women are integrated within the workplace. The age at starting school is also included as a measure of how easy it is to combine childbearing with childbearing, since it is typically easier to combine childbearing with labour force activity where the child is in a formal care setting for at least part of the day.

I also draw on work by Perelli-Harris and Sanchez-Gassen (2012) for indicators of the legal status of cohabitation. Specifically, I use the proportion of policies in which cohabitation was mentioned and the proportion of policies in which marriage and cohabitation were afforded equal status. Finally, I include a binary indicator of whether the country in question has a legally recognised cohabiting state distinct from marriage (such as the PActe Civil de Solidarité or PACS).

Individual level data for this analysis are drawn from the Harmonized Histories (Perelli-Harris, Kreyenfeld, and Kubisch 2009, and see www.nonmarital.org). The Harmonized Histories is a dataset containing consistent retrospective demographic histories from 16 countries across Europe. The data for Austria, Belgium, Bulgaria, Estonia, France, Hungary, Italy, Norway, Romania, and Russia come from the Generations and Gender Surveys (GGS), which interviewed nationally representative samples of the resident population in each country. Because the GGS is not available for all countries (or the retrospective histories were not adequate for our purposes), we also relied on other data sources. The Dutch data come from the 2003 Fertility and Family Survey (FFS). The data for the UK are from the British Household Panel Survey (BHPS). The Spanish data come from the Survey of Fertility and Values conducted in 2006, and the Polish data are from the Employment, Family, and Education survey conducted in 2006. The U.S. data are from the National Survey of Family Growth, conducted between 2006 and 2008.

I extract a subset of 9 countries from the full Harmonized Histories dataset to ensure that each country in the dataset has at least some contextual level information available. As noted, contextual information need not necessarily be complete, but at least some information is required to allow allocation to a class. Therefore, I select respondents from Austria, Bulgaria, Estonia, France, the Netherlands, Norway, Romania, Russia, Spain and the United Kingdom.

Three individual level processes are used in this analysis, capturing the processes of marriage formation, the formation of non-marital cohabitating unions and becoming a mother (first birth). The processes are modelled accurate to the nearest year. Although data in the Harmonized Histories are available accurate to the nearest month, it was too computationally intensive to produce models based on such a fine gradation, and so years are used as a reasonably accurate approximation. I model these processes between the ages of 16 and 45, delimiting the effective exposure to the three processes modelled (women cannot marry before the age of 16 in the selected countries, and are generally speaking post-menopausal by the age of 45).

Marriage and cohabitation processes are both modelled as growth curves where the response variables $\boldsymbol{y_{tj}^{M}}$ and $\boldsymbol{y_{tj}^{C}}$ take the value 0 for years *t* when the respondent is not in a marital or cohabitating relationship respectively, and 1 where they are. Note that these are not cumulative rates: women can exit both marriage and cohabiting relationships reflected in the falling probability of being in either of these states. This allows trends such as divorce/union dissolution or the transition of cohabitation into marriage to be reflected in the modelling process. In contrast, entry into motherhood is modelled as a cumulative growth curve (consistent with Dariotis et al. 2011). The response variable $\boldsymbol{y_{tj}^{F}}$ therefore takes the value 0 for all years where the respondent is nulliparous, and 1 in all years including and following where the respondent had first birth. Unlike marriage and cohabitation, it is not necessary to include the possibility of falling probabilities since this growth curve is modelling a one way transition.

All growth curves are modelled jointly and depend on the higher level class of which the individual is a member. The process is reflected in Equation 5. The probability of being married, in a cohabiting relationship and having ever had a first birth is a function of a third order polynomial of the age of the respondent. Note that this will depend on the class membership of the respondent's country: the profile of marriage, cohabitation and fertility will differ between countries which belong to different classes.

$$\begin{cases} logit\left(y_{tj}^{M}\right) = \beta^{M} age_{tj} + \beta^{M} age_{t}^{2} + \beta^{M} age_{t}^{3} \\ logit\left(y_{tj}^{C}\right) = \beta^{C} age_{t} \; + \beta^{C} age_{t}^{2} + \beta^{C} age_{t}^{3} \\ logit\left(y_{tj}^{F}\right) = \beta^{F} age_{t} \; + \beta^{F} age_{t}^{2,F} + \beta^{F} age_{t}^{3} \end{cases}$$

$$J = \{1 \cdots j\}$$

Eq. 5

5. RESULTS

5.1. CLASS CHARACTERISTICS

Table 1 presents fit statistics for models with differing numbers of latent classes. Broadly speaking, all three fit statistics show a consistent pattern. There are dramatic falls in the values of the AIC, BIC and Sample size adjusted BIC when increasing the number of classes from 1 to 2, and 2 to 3. This indicates and improvement in model fit. The addition of a fourth class increased the value of all fit statistics marginally, indicating that best model fit was afforded by a three class model.

	AIC	**BIC**	**Sample Size adjusted BIC**
1	1581496.061	1581751.337	1581659.174
2	1552354.465	1552803.398	1552641.319
3	1534286.142	1534928.733	1534696.738
4	1534330.274	1535166.522	1534864.611

Table 1: Fit statistics for latent class models.

The allocation of countries to classes is presented in Figure ~~3~~2. Broadly speaking, there is an East-West divide, with the Hajnal (Hajnal 1965) demarking the geographic clustering of Eastern European countries (Bulgaria, Estonia, Romania and Russia) which are all members of class 1, and clustering of Western countries, which are members of classes 2 and 3. Of the western European countries, only the Netherlands and Spain are members of class 2, while Austria, France, Norway and the United Kingdom are members of class 3.

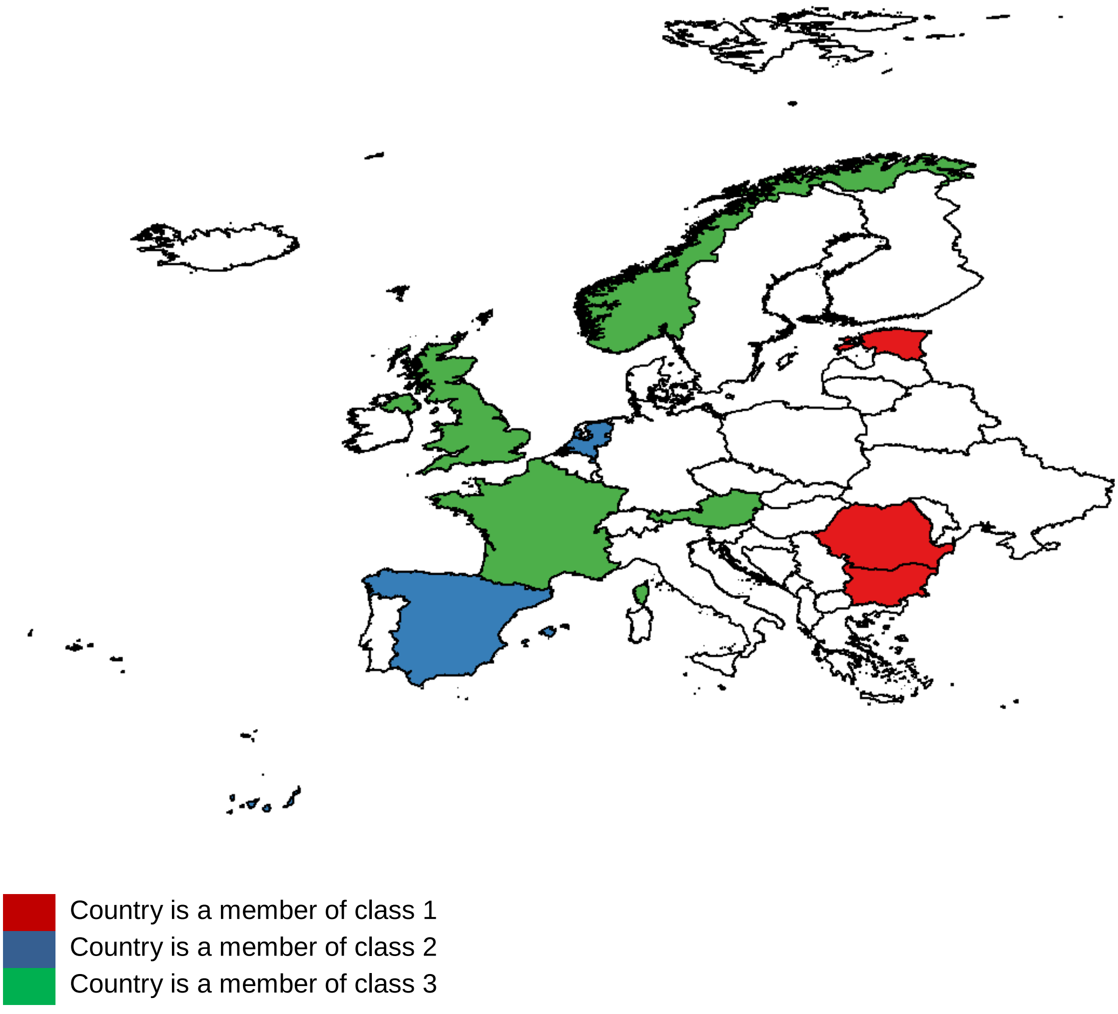

Figure 2: Allocation of countries to latent classes
Note: Russia not shown in full for visual clarity.

Table 2 presents the estimated characteristics for the latent classes. Class 1 (Eastern European class) is characterised by a low level of family support, with the lowest absolute value of family remittances, and relatively low level of support for family as a proportion of GDP in terms of both family allowance , social expenditure and public expenditure on childcare. This cluster can roughly be seen to incorporate the post-Socialist typology of Hofacker et al. (2006).

In terms of the ease of childcare domain, the female labour force participation is rather lower in this class than the other two, while the school entry age is somewhat comparable. There is little support for cohabitation in the legal frameworks of these countries, with cohabitation mentioned in only 26.7% of legislation, and equivalent to marriage in only 30% of legislations (lower than the other two classes). There is no legal recognition of cohabitations as a partnership form within this class.

Class 2 presents a somewhat intermediate picture and incorporates countries included in Conservative (Netherlands) and Southern (Spain) welfare regimes (Hofacker et al. 2006). The absolute value of family allowances are higher than in class 1 but lower than in class 3 (by a considerable margin). This is reflected in the percent of GDP devoted to family allowances, which is the lowest of all classes. In contrast, the value of social expenditure is high, and the degree of public expenditure of childcare is the highest of all classes. This pattern therefore reflects a family support regime which is focussed on in-kind benefits; the value of family allowances are moderate, but women can expect to receive a relatively high degree of support through subsidised childcare, for example.

Female labour force participation is high in this class at over 73%: this is considerably higher than in the Eastern European class (Class 1) and comparable to class 3. School entry age is rather higher than other classes. The legal status of cohabitation is rather mixed in this class. Cohabitation is mentioned in 16% of legislations, and legally equivalent to marriage in 38% of these- the highest in all classes. However, the existence of a legally distinct non-marital unions is somewhat mixed, with a 50% probability.

Class 3 is characterised by the highest level of family support incorporating both Liberal (UK) and Socio-Democratic welfare regimes (Hofacker et al. 2006). The absolute value of remittances are considerably higher than in either class 1 or class 2, and are similarly higher in terms of percent of GDP. General social expenditure is high, although the support for childcare as a percent of GDP is somewhat lower than in class 2 and consistent with Eastern European levels. Female labour force participation is broadly speaking high at in excess of 70%, albeit it slightly lower than in class 2. School entry for countries in this class is the lowest seen in all classes.

In general this class demonstrates a high degree of support for cohabitation. Cohabitation is mentioned in nearly one third of relevant laws, and legally equivalent in nearly 35% of these. The existence of a legally recognised cohabiting relationship is highly prevalent for countries in this class, with a near universal existence of a formal non-marital union within the law.

Indicator (Number of country members)	*Class 1: Eastern Europe (4)*	*Class 2: Western Europe lesser support (2)*	*Class 3: Western Europe higher support (4)*
Family support			
Value of family allowance (PPP adjusted 2005 $)	82.21	92.26	133.0
Family allowance (% of GDP)	1.38	0.11	1.78
Social expenditure (% of GDP)	13.16	26.8	26.13
Public expenditure on childcare (% of GDP)	0.38	0.64	0.38
Ease of childcare			
Female labour force participation (%)	64.40	73.40	70.46
School entry age	3.250	4.000	3.000
Legal status of cohabitation			
Cohabitation mentioned (%)	26.7	26.0	29.7
Legal equivalence (%)	30.7	37.5	34.5
Legally recognised (prob)	0.00	0.50	0.99

Table 2: Class characteristics at national level from 3 class result

5.2. INDIVIDUAL CHARACTERISTICS

Growth curves for three demographic processes were extracted, with the growth curve varying by class. The growth curves for the probability of marriage are presented in Figure 3. Broadly speaking, the overall pattern of marriage is consistent across all classes, increasing from relatively low levels at early ages, and peaking around the early 30s. Thereafter, there is some evidence in the decline of the probability of being married, due in the most part to the dissolution of marital unions. There are some differences in the overall levels of marriage between classes, as well as marriage timing. Class 1 (Eastern European pattern) shows the overall highest propensity for women to have a marital union, with the highest probability of marriage at all ages. Class 2 (Western Europe- limited support) presents a more mixed progression. The probability of marriage is low at younger ages- 10% points lower than in class 1 at age 16. However, the rate of increase across ages is rapid and increases to the extent that the probability of marriage is nearly at the level seen in class 1 by age 36. Class 3 (Western European- extensive support) shows the overall lowest propensity for marriage, which is consistent across the lifecourse. The overall probability is low at early ages, consistent with the levels seen in class 2. However, in class 3 the increase in the probability of marriage is not as dramatic, and the peak in the probability of marriage is slightly below 70%, in comparison to classes 1 and 2 which are both in excess of 85%. Overall then, class 3 demonstrates the lowest prevalence of marital behaviours, while classes 1 and 2 show a greater recourse to the institution of marriage. That said, class 2 demonstrates some postponement of entry into marriage, while class 1 is characterised by high levels of marriage at early ages.

The estimated growth curves for cohabitation are presented in Figure 3. The probabilities of cohabitation are somewhat lower than they were for marriage (Figure 3) reflecting the preponderance of formalised unions across all classes. The probability of living in a cohabiting relationship follows a similar trajectory for both class 1 (Eastern European) and class 2 (Western Europe- limited support). There is a gentle increase in the probability of cohabitation, reaching a peak around age 30 at approximately 5%. Thereafter, there is a fall in the probability of being in a cohabiting union, to just over 1% in class 2, while the probability remains above 3% in class 1. This reflects that in Eastern Europe and Western regimes with limited social support the probability of being in a non-marital union is still low.

In contrast the probability of cohabitation is considerably higher in class 3 (Western European- extensive support), increasing to peak at age 30 at around 11% (over double the other classes). Thereafter the probability of cohabitation decreases rapidly, falling to around 3% by the age of 45. This is indicative that in the Western

European regimes with high levels of social support, there is a far greater prevalence of non-marital partnership behaviour.

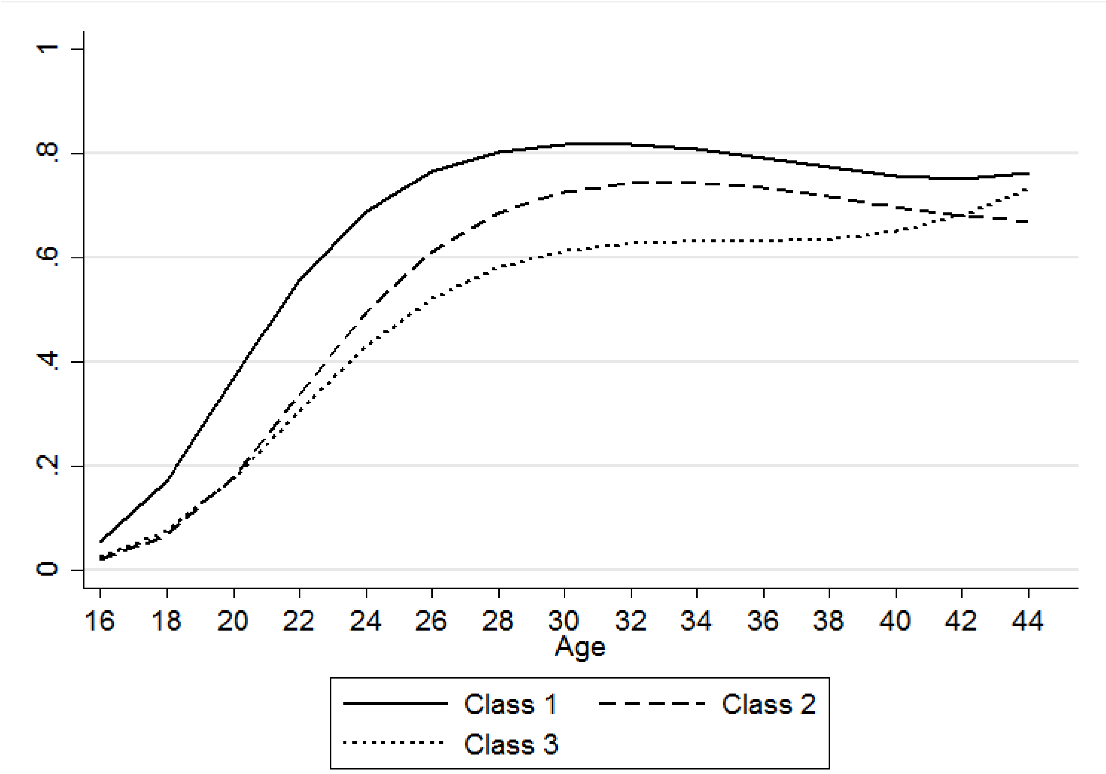

Figure 3: Predicted growth curves for probability of marriage for Western extensive and limited and Eastern classes

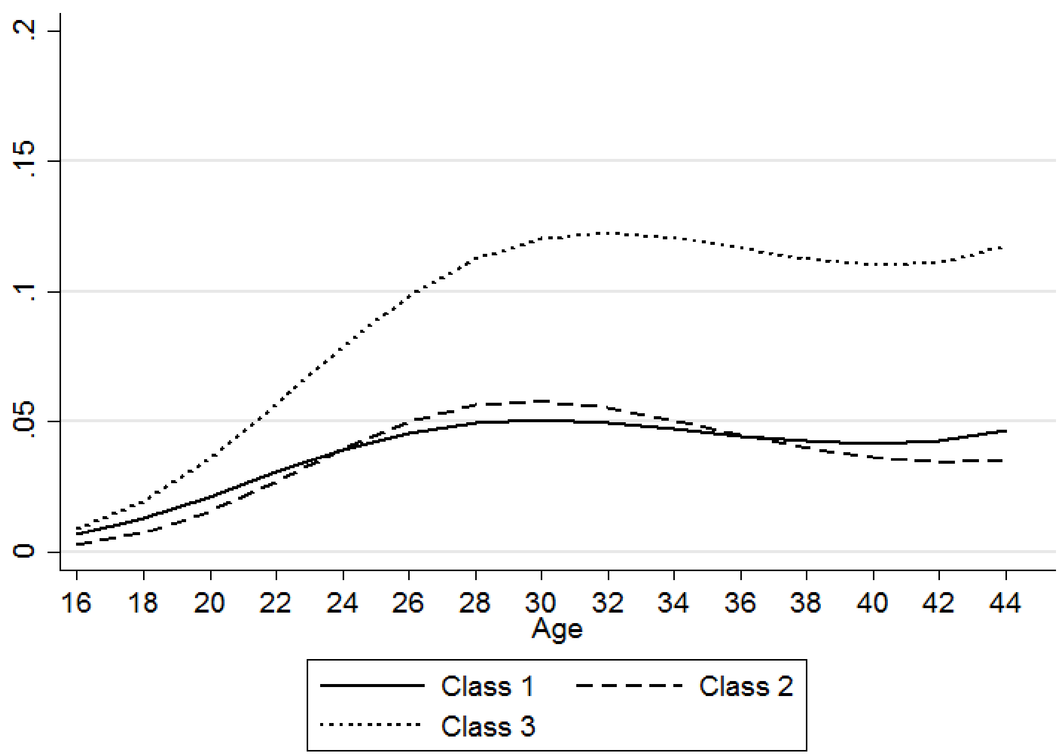

Figure 4: Predicted growth curves for probability of cohabitation for Western extensive and limited classes.

Figure 5 presents predicted growth curves for the cumulative probability of birth. Class 1 reflects a pattern of a rapid transition into motherhood, with the incidence of first birth rising rapidly at the start of the lifecourse. For instance, 50% of women have experienced their first birth before the age of 22 and 80% by the age of 26. However, this pattern reflects the persistence of historic trends in socialist countries toward early birth (Sobotka 2003), since there is little increase in the proportion of women who have a first birth after the age of 30, reflected in the flattening of the curve toward the end of the reproductive lifecourse.

In contrast, classes 2 and 3 (Western Europe- limited support, Western Europe- extensive support) show a rather later transition to motherhood. In class 2 the median age at motherhood is age 25, and in class 3 slightly earlier at age 24. It is worth noting that the increase in cumulative fertility for these classes persists to rather later ages than in class 1 resulting in a fall in the gap between the proportion of women who have had first birth, which amounted to nearly 20% points at age 22, but had fallen to 10% points by age 28 and 5% points by age 36. This is indicative of the increasing postponement of fertility in Western settings, and a persistence of entry into motherhood even toward the end of the reproductive lifecourse.

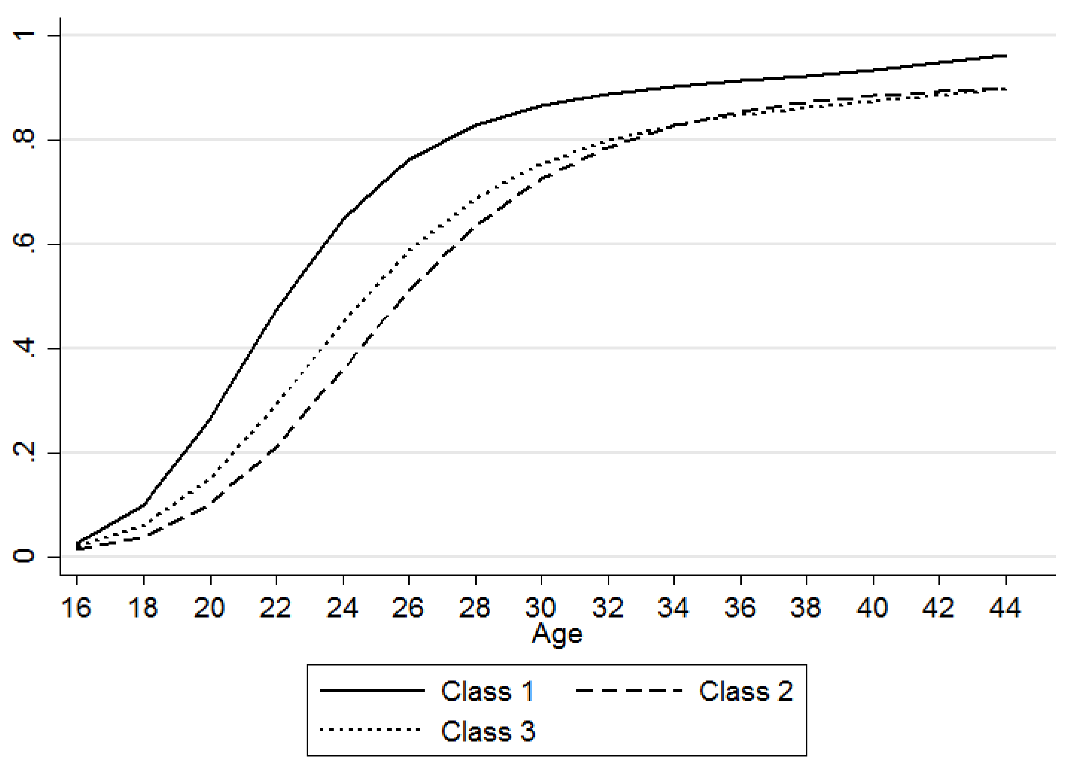

Figure 5: Predicted growth curves for probability of birth for Western extensive and limited and Eastern classes.

6. CONCLUSIONS

This paper considered different strategies for incorporating higher level information into longitudinal modelling. The specific focus in this context was the use of countries as higher level units: country context is of wide interest to the demographic and social science community. However the usual multilevel modelling techniques (random and fixed effects) will often be severely limited where countries are a higher unit, either due to the characteristics of the method themselves or due to the sampling structure at the country level. The implementation of latent classes at the country level is an attempt to overcome the limitation of both of these methods. The additional advantage of the latent class based approach is that in forming clusters based on empirical data, it provides the opportunity to validate *a priori* theoretical clusters or typologies (e.g. Esping-Andersen 1990, Hofacker et al. 2006)

The advantages of this approach are demonstrated by evaluating the relationship between fertility and partnership behaviour and country clusters as determined by a set of policy relevant indicators. This analysis extracted three distinct country level clusters in the European settings. Firstly, Eastern European countries were distinct exhibiting relatively low welfare support in an absolute and relative term, but with high levels of female labour force participation and a fair degree of child support. There was no protection for cohabitation as a childbearing union. Unions tended to be marital in these settings, with an early incidence of fertility behaviour. The uniqueness of the Eastern fertility pattern and extraction as a discrete cluster distinct from Western fertility regimes has been remarked upon in demographic literature and links to theoretical understanding of fertility patterns (Hofacker et al. and Drobnic 2001).

Western Europe was characterised by two classes: lesser and greater support. The lesser support class (The Netherlands and Spain) comprised countries in conservative and Southern welfare state typologies; both characterised by welfare systems designed to either support the family institution or rely on familial support (Hofacker et al. 2006). This class had relatively low financial support for childbearing and welfare in general, but relatively high levels of female labour market engagement. There was more limited legal protection for cohabiting unions. Again, in terms of partnership behaviour, unions are predominately marital, with little evidence of cohabitation, but with fertility behaviour somewhat delayed. The final class captures generous European welfare states, which afford considerable welfare provision as well as a high degree of support for childcare and legal protection for cohabitation. In these settings, cohabitation is far more common and fertility behaviour occurs relatively late in the lifecourse. Countries in this class belong to either Liberal or Social Democratic welfare regimes (Hofacker et al. 2006).

These findings demonstrate the major advantages of the latent class approach. By relaxing the assumption of a Normal/continuous distribution among higher order units, I am able to identify clusters of similar higher level units, which is informative in itself due to the neat division along the Hajnal line. Another major advantage is that the higher level groups are formed using contextual information, which can provide an additional point of interpretation. It should be noted that this is also a limitation of the study: the requirement for contextual information is now greater than for either random or fixed effects multilevel models. Additionally, a more cautious approach should be applied for a researcher wishing to make causal statements: the fact that much country level information is relatively recent means that some of these indicators are likely to be endogenous with fertility and partnerships behaviour (Perelli-Harris and Sanchez-Gassen 2012). The final advantage is that this approach allows straightforward comparison of individual level behaviour between clusters: the lateness of fertility behaviour in Western European higher support settings compared to Eastern Europe is intuitive and straightforward and links with existing understanding of different fertility norms in Western and post-Socialist countries (Sobotka 2003, Blossfeld and Drobnic 2001).

REFERENCES

Billingsley, S. and Ferrarini, T. (2014). "Family Policy and Fertility Intentions in 21 European Countries." *Journal of Marriage and Family*, 76, 428-445.

Browne, W.J. (2009). "MCMC Estimation in MLwiN. Version 2.13." Bristol, Centre for Multilevel Modelling, University of Bristol.

Blossfeld H.P and Drobnic, S. (2001). "Careers of Couples in Contemporary Societies: From Male Breadwinner to Dual Earner Families." Oxford, Oxford University Press.

Dariotis, J.K., Pleck, J.H., Astone, N.M. and Sonenstein, F.L. (2011). "Pathways of Early Fatherhood, Marriage, and Employment: A Latent Class Growth Analysis." *Demography,* 48, 593-623.

Efron, B. and Morris, C. (1973). "Stein's estimation rule and its competitors – an empirical Bayes approach." *Journal of the American Statistical Association,* 68 (1), 117-130.

Elzinga, C.H. and Liefbrouer, A.C. (2007). "De-standardization of Family-Life Trajectories of Young Adults: A Cross-National Comparison Using Sequence Analysis." E*uropean Journal of Population*, 23, 225-250.

Esping-Andersen, G. (1990). "The three worlds of welfare capitalism." Princeton, New Jersey, Princeton University Press.

Goldstein, H. (1997). "Methods in school effectiveness research." *School Effectiveness and School Improvement*, 8 (4), 369-395.

Hajnal, John (1965): European marriage pattern in historical perspective in D.V. Glass and D.E.C. Eversley, (eds.) Population in History, Arnold, London.

Hofacker, D., Buchholz, S. and Blossfeld, H.P. (2006). "Late careers in a globalizing world: a comparison of changes in twelve modern societies." Chapter 15 in Blossfeld, H.P., Buchholz, S. and Hofacker, S. (eds.) Globalization, Uncertainty and Late Careers in Society. Routledge, London.

Hox, J., van de Schoot, R. and Matthijsse, S. (2012). How few countries will do? Comparative survey analysis from a Bayesian perspective. *Survey Research Methods,* 6 (2), 87-93.

Kalwij, A. (2010). "The impact of family policy expenditure on fertility in western Europe." *Demography,* 47, 503–519.

Korpi, W. (2000). "Faces of inequality: Gender, class, and patterns of inequalities in different types of welfare states." *Social Politics,* 7, 127–191.

Korpi, W., Ferrarini, T. and Englund, S. (2013). "Women's opportunities under different family policy constellations: Gender, class and inequality tradeoffs in western countries re-examined." *Social Politics,* 20, 1–40.

Kuehner, S. (2007). "Country-level comparisons of welfare state change measures: Another facet of the dependent variable problem within the comparative analysis of the welfare state?" *Journal of European Social Policy,* 17, 5–18.

Larsen, R. (2011). "Missing data imputation versus Full Information Maximum Likelihood with Second-Level Dependencies." *Structural Equation Modeling: A Multidisciplinary Journal*, 18 (4), 649-662.

Lo, Y., Mendell, N.R. and Rubin, D.B. (2001). "Testing the number of components in a normal mixture." *Biometrika*, 88, 767-778.

Mason, K. and Oppenheim, K. (1998). "Gender and demographic change: What do we know?" In Jones, G.W., Douglas, R.M., Caldwell, J.C. and D'Souza, R.M. (eds.) *The continuing demographic transition,* 158–182, Oxford, UK, Clarendon Press.

McDonald, P. (2000). Gender equity, social institutions and the future of fertility. *Journal of Population Research,* 17 (1), 1-16.

Mills, M. (2010). "Gender roles, gender (in)equality and fertility: An empirical test of five gender equity indices." *Canadian Studies in Population,* 37, 445–474.

Muthén, L.K. and Muthén, B. (2006). "Mplus user's guide (Version 4)." Los Angeles, CA, Muthén and Muthén.

Neels, K., Zita, T. and Wood, J. (2013). "Economic recession and first births in Europe: recession-induced postponement and recuperation of fertility in 14 European countries between 1970 and 2000." *International Journal of Public Health,* 58 (1), 43-55

Nylund, K.L., Asparouhav, T. and Muthen, B. (2007). "Deciding on the Number of Classes in Latent Class Analysis and Growth Mixture Modeling: A Monte Carlo Simulation Study." *Structural Equation Modeling: A multidisciplinary Journal*, 14 (4), 535-569.

Perelli-Harris, B. and Sanchez-Gassen, N. (2012). "How similar are Cohabitation and Marriage? Legal Approaches to Cohabitation across Western Europe." *Population and Development Review,* 38 (3), 435-467.

Perelli-Harris, B., Kreyenfeld, M. and Kubisch, K. (2009). "Technical Manual for the Harmonized Histories Database." *MPIDR Working paper 2010-011*, Rostock.

Snijders, T.A. and Bosker, R.J. (1999). "Multilevel analysis: An introduction to basic and advanced multilevel modeling." Thousand, Oaks, CA, Sage.

Sobotka, T. (2003). "Re-emerging diversity: Rapid fertility changes in Central and Eastern Europe after the collapse of the communist regimes" *Population*, 58 (4-5), 451-486.

Stegmueller, D. (2013). "How Many Countries for Multilevel Modeling? A Comparison of Frequentist and Bayesian Approaches." *American Journal of Political Science,* 57, 748–761.

ESRC Centre for Population Change
Building 58, Room 2001
Faculty of Social and Human Sciences
University of Southampton
SO17 1BJ

T: +44 (0)2380 592579
E: cpc@soton.ac.uk
www.cpc.ac.uk

To subscribe to the CPC newsletter and keep up-to-date with research activity, news and events, please register online: www.cpc.ac.uk/newsletter

For our latest research updates you can also follow CPC on Twitter, Facebook and Mendeley:

www.facebook.com/CPCpopulation

www.twitter.com/CPCpopulation

MENDELEY www.mendeley.com/groups/3241781/centre-for-population-change

The ESRC Centre for Population Change (CPC) is a joint initiative between the University of Southampton and a consortium of Scottish universities including St Andrews, Edinburgh, Stirling and Strathclyde, in partnership with the Office for National Statistics and National Records of Scotland.

University of Southampton | University of St Andrews | The University of Edinburgh | University of Strathclyde Glasgow | University of Stirling | National Records of Scotland | Office for National Statistics