Liao, T. F. & A. E. Fasang (2016)

Do State Policies Generate Different Life Courses? An Empirical Study of the Case of the Two Germanys via a Statistical Assessment

in G. Ritschard & M. Studer (eds), Proceedings of the International Conference on Sequence Analysis and Related Methods, Lausanne, June 8-10, 2016, pp 469-506.

Do State Policies Generate Different Life Courses?

An Empirical Study of the Case of the Two Germanys via a Statistical Assessment

Tim F. Liao, University of Illinois

Anette E. Fasang, Humboldt University Berlin & WZB

Abstract

1. Introduction

The life course is informative for social scientists to study because we can see the effects of individual characteristics such as education and occupation.[1] For sociologists, it is even more fascinating to analyze life courses because the state also impacts and structures them (Mayer 1989). The state may impact the life course in various ways: It may affect individual life courses by social structure such as educational and occupational systems (Shanahan 2000); it may directly influence life courses through institutional regulations and social policies (Mayer 2009; Neyer 2013).

As a quasi-experiment, the German reunification in 1990 offers an exceptional event for studying the imprints of the state on individual life courses (e.g. Goldstein and Kreyenfeld; Schnettler and Klüsener 2014). In this paper, we aim to formally assess how East and West German life courses differ by analyzing life course sequence data from a recent national survey (NEPS) with a retrospective history component, along the lines of life course research focused on contexts specified according to space and time (Mayer 2004). Specifically, we are interested in how Germans who were born and grew up in the former East and West Germanys lived different life courses in both the family and the employment dimensions even though they now live in the reunified Germany.

Life course and employment history observations often provide sequence data, a type of data more complex than the observation of a single aspect of the life course (e.g., Blanchard, Bühlmann, and Gauthier 2014; Gabadinho et al 2011). The timing and the duration of a life course event, such as the timing of first birth or the duration between completing education and first labor market entry, are examples of such single aspects. Statistical assessments of timing and duration variables are straightforward because mean timing and durations can be compared across samples, and their expected values can be used in further analysis.

With sequence data, however, there is no obvious single aspect to construct a variable for such statistical assessments. A life course sequence contains information on the timing and the duration of various life course events as well as on their sequence order (Brückner and Mayer 2005). Therefore, in this paper, we have twin objectives. In addition to the substantive goal above, we propose a medoid-based method for formally assessing the difference between sets of life course sequences, where the medoid is defined as the sequence in a given set of sequences that has the minimum sum of distances to all other sequences in the set. Using this method, we compute the Bayesian Information Criterion (BIC) and the Likelihood Ratio Test (LRT) statistic in the context of sequence comparisons. While the LRT is a common statistical test in many disciplines, the BIC has become a popular method for comparing and selecting models in sociology since Raftery's (1986) introduction into our discipline.

In the following, we first review the literature on the life course in the East and West Germanys, especially how their social policies and institutional regulations may differ. We then present the medoid-based method for analyzing sequence differences. The method is set up for the construction of the BIC and the LRT statistics for comparing sequence data. We then apply the method to the data from the 2009/2010 National Education Panel (NEPS), cohorts 1955-1965, to both illustrate the efficacy of the method and to assess the difference between East and West German life courses. We conclude by drawing some general conclusions about how the East and West German life courses differ initially and how they converge over cohorts and about the appropriateness of the proposed method.

2. Social Policies and Life Courses in East and West Germanys

Between 1949 and 1990 Germany was divided into two sub-societies with notable contextual differences in the welfare state, economic system and family policies between the communist German Democratic Republic (GDR) in the East and the democratic social market economy in the Federal Republic of Germany (FRG) in the West.

The GDR had a state socialist system, and a centrally planned economy. The regulative regime aimed at reconciling the political priorities of economic and population growth. To this end women were expected to work equal to men in a dual earner model and the family ideology was strongly *pro-natalist* (Engelhardt, Trappe & Dronkers, 2002). Family policies in the GDR conditioned access to state-controlled resources, such as housing, on marriage and parenthood to incentivize fertility. Normative pressure to have children in the early 20s was coupled with generous financial incentives for parenthood and practically universal public child care that widely enabled mothers' employment. As a result female labour market participation was around 90% (Huinink et al., 1995). Employment was practically universal but wages were rather low. In practice, marriage and parenthood often remained the quicker and more viable route to obtain housing and generous loans from the state than trying to accrue such privileges through labour market activity.

In contrast, in the period following World War II, the FRG had a democratic multiparty parliament, a market economy, and a corporatist conservative welfare state (Rosenfeld, Trappe, & Gornick, 2004). The arguably purest implementation of the male breadwinner model (Prince Cooke & Baxter, 2010) resulted in female labour market participation around only 50%, of which much was part time. Family policies where foremost *pro-traditional* and social policies comprehensively set strong incentives for a traditional male breadwinner–female carer household division of labour and the financial dependence of women on men (Brückner, 2004; Prince Cooke, 2011). These include joint taxation of married couples that discouraged employment of a second earner and the absence of public child care particularly for children under the age of three (Aisenbrey, 2009; Prince Cooke & Baxter, 2010).

In 1990, Germany was reunified by adapting the FRG model to the former GDR, which ceased to exist. Especially during the 1990s the reunification process was accompanied by severe economic recession in the East but also in the West that peaked in the mid-1990s. However, the West German institutional model was applied to a population with markedly different compositional features in the former East. For instance joint taxation for married spouses makes a large difference particularly when earnings between husbands and wives are very unequal. Joint taxation sets few incentives for marriage for couples with similar earnings, which was much more often the case in the former East also after reunification (Bastin, Kreyenfeld & Schnor, 2012). Further, the public child care infrastructure remained largely intact in the East maintaining a much more conducive environment for women's employment which remains less affected by motherhood and marriage (Matysiak and Steinmetz, 2008). East Germany is one of the only regions in Western societies that shows no motherhood penalty in wages, whereas this penalty remains among the highest in the West at 32 percent after in the early 2000s (Budig, Misra & Boeckmann, 2012). For an overview of institutional differences relevant to family formation around the reunification see Goldstein and Kreyenfeld (2011) and Fasang (2014).

Initial expectations of a quick convergence of the two Germanies thus did not materialize (Schneider, Naderi & Ruppenthal, 2012). Even in 2008, 18 years after the reunification, East Germany still had significantly lower rates of property ownership, lower average earnings, higher rates of female employment and a higher proportion of children in public care (Goldstein and Kreyenfeld, 2011). East Germany further continues to be one of the most secularized regions of Europe with 74 percent reporting no religious affiliation in 2008 compared to only 16 percent in the West (ibid: 457). While total fertility rates have converged (Goldstein and Kreyenfeld 2011), within Europe the two sub-societies remain on opposite ends of the continuum for a number of demographic indicators, including a proportion of nonmarital birth above 70 percent in the East compared with only around 12 percent in the West in 2007 (Klüsener, Perelli-Harris, Sanchez Gassen, 2013), as well as much higher rates of cohabitation in the East. Recently, Klüsener and Goldstein (2014) suggested that the long-standing divide in non-marital fertility in East and West Germany is attributable to regional variation that preceded the war in 1945 and was merely intensified during the division. Therefore the two subsocieties might plausibly never converge or at least not within a short time frame on some deeply rooted demographic and economic differences. Overall, there is more research on similarities and differences in family formation than on employment trajectories in East and West Germany. Studies on family formation are mostly located in demography and likely stimulated by the extreme patterns of difference in the two German sub-societies.

Precisely because we see notable convergence in some indicators of family formation and employment trajectories, but persistent and sizeable differences in others, it is particularly useful to examine holistic life course trajectories. In addition to a detailed view on single outcomes this enables us to broaden our analytical scope and assess to what extent overall differences emerge form diverging patterns in single indicators. In this paper we therefore analyze whether and how men's and women's employment and family life courses were different and remain different in divided and reunified Germany conceptualizing them as longitudinal "process outcomes" (Abbott 2005).

3. A Medoid-Based BIC Method for Sequence Assessment

Comparing sequences across samples is essentially a problem of assessing differences across groups. Such assessments can be conducted by computing distance measures both within and between sets of sequences belonging in specific groups. On the one hand, distances within specific groups can be computed summarizing the degree of homogeneity or standardization of the life courses in this set of sequences. On the other hand, distance between sequences belonging to different groups can be computed (Fasang 2014). The discrepancy analysis proposed by Studer et al (2011) is a good method for analyzing differences between sets of sequences. However, as reported later, a discrepancy analysis of the East and West German life course data can be very insensitive to any variation in the degree of differences between sets of sequences, as also shown through our simulation study discussed later, thus failing to distinguish between cohorts for either men or women family formation and employment history data.

In this section, we consider the BIC and the LRT as an alternative method that is also based on distance measures. There are many viable candidates as distance measures for making statistical comparisons, including optimal matching, dynamic hamming matching or the so called subsequence metrics (e.g. Aisenbrey and Fasang 2010 for an overview of several distance measures). One can compute the classical Levenshtein or another statistical distance between all possible pairs of sequences of the groups under comparison. Alternatively, one can compute a distance of each sequence to the medoid sequence of the sequence set, a popular method used for complexity reduction in recent research on sequence visualization and comparison (Aassve et al 2007; Fasang and Liao 2013; Gabadinho et al 2011; Piccarreta 2012). One important advantage of the medoid based approach is that there is only one distance value for each sequence to the medoid instead of $N(N$-1)/2 pariwise comparisons thus yielding higher efficacy. For reason that we detail below, we use the medoid for computing distances within and between groups.

Let s_i denote the sum of squared distance of each sequence group i to the sequence group medoid of group i:

$$s_i = \sum_{j=1}^{n_i} {q_{ij}}^2 \tag{1}$$

where q_{ij} is the distance between each sequence j in the ith sequence group and the ith sequence group medoid, which can be understood as the best representative observed sequence in the ith group (i.e., giving the shortest distances to all group members overall). Note that a set of sequences may have several – i.e. tied – medoid sequences, where more than one sequence share a minimum distance to all other sequences in the group. The statistic in (1) can be computed for all G number of sequence groups.

Because a group medoid is the observed sequence that has the shortest distance to all the sequences in the same group, it is theoretically possible that the medoid for a sequence from another group may have a shorter distance than its own medoid, a rare situation we have empirically verified. To avoid this empirical problem because of a single unexpected better sequence, we use a selection of best representative sequences from a group, or collective medoids, in (1) for computing BIC and LRT statistics. For example, we may select a small percentage or number of best representative sequences, such as *m* number of or percent sequences from a sequence group. While any small percentage or number of representative sequences may work, we have found in our simulation study that using 5% can sufficiently or smaller avoid misrepresentation (for each of the groups with the *m* for the combined group as the sum of the *m*s). The ultimate choice of *m* depends on the heterogeneity and unevenness of the sequence groups to be compared.

3.1 BIC

Assuming the distances in (1) are normally distributed – an assumption we test below – we follow Burnham and Anderson (1998, 2004) in expressing least squares as likelihoods for constructing BICs as well as the closely related Akaike Information Criterion).[2] Thus, –2 log-likehood for (1) can be expressed as

$$-2\ \text{log-likelihood}_i = n_i \log\left(\frac{s_i}{n_i}\right) \qquad (2)$$

where n_i denotes the number of sequences in sequence group i. The BIC for the ith sequence group can then be expressed as

$$\text{BIC}_i = n_i \log\left(\frac{s_i}{n_i}\right) + K \log(n_i) \qquad (3)$$

For the ith sequence group, K=1 because the only degree of freedom involved is that associated with the ith medoid. For comparing the hth and the ith sequence groups, or more generally, G number of groups, we propose a nested model approach, which is consistent with the typical application of BIC when models are nested. We define BIC_A as the BIC based on the sum of squared distances using the overall medoid and BIC_G as the BIC based on the sum of squared distances using group-specific medoids:

$$\mathrm{BIC}_A = \frac{n_A}{w}\log\left(\frac{s_A}{n_A}\right) + k\log\left(\frac{n_A}{w}\right) \quad (4)$$

$$\mathrm{BIC}_G = \frac{n_A}{w}\log\left(\sum_{i=1}^{G}\frac{s_i}{n_i}\right) + k\log\left(\frac{n_A}{w}\right) \quad (5)$$

where n_A is the total number of distance computations between the sequences in all groups combined and the reprentative sequences, and w is a weight that equals to $(e \times \log(s) \times n_A)^{0.5}$ and is applied once in each of the additive items where n_A appear. This weight is applied to the entire BIC because the distance computation of (1) is unnecessarily increased by the multiplicative factor of the number of episodes, e, and the number of states, s, and the number of sequences n_A (the number of distance computations between sequences and their representatives) though the contribution by s can be reasonably conceived as a modest nonlinear function.[3] The parameter K gives the number of degrees of freedom, which equals to 1 in (4) because only the overall medoid (or an overall set of representative sequences) is used and equals to G in (5) because G medoids (or G sets of representative sequences) are used in the computation.[4] The BIC difference between (4) and (5), or (4)–(5), forms the criterion for assessing differences between sets of (life course) sequences.

To assess the statistical properties of the BIC difference based on (4) and (5), we conducted a simulation study, as explained and reported in Appendix A. The simulation results show that the statistic of BIC differences using (4) and (5) can capture actual differences between two sets of life course sequences very well though it still is, to a degree, sensitive to sample sizes. The original BIC is sensitive to sample size increases; the weight in (4) and (5) reduces some of that sensitivity. We also applied discrepancy analysis methods to the simulation data. Regardless of the degree of difference between the two simulated sets of sequences, the *F*-test in the discrepancy analysis is always significant at least at the 0.001 level. Thus, we do not report the results of the simulation study regarding the behavior of the discrepancy analysis.

3.2 LRT

While the BIC should provide a proper statistical assessment of differences between groups or samples of sequences, it does not give a significance test. Additionally, we can construct a Likelihood Ratio Test (LRT) using (2). Let ll_i stand for the -2 log-likelihood of the *i*th group. Following Liao's (2002, 2004) discussion of the LRT for generalized linear or logit models, the LRT for testing the null hypothesis that all groups are equal in their sequences is obtained as the difference between the *ll* from the restricted model and the *ll* from the unrestricted model. In the current application of comparing *G* number of sequence groups, the restricted model is the situation when we assume all groups are equal, with a single medoid (i.e., the global medoid), and the unrestricted model is the case where all groups are considered unique, with their individual medoids. Following Liao (2002: (6.13) & (6.14)), we further define ll_{R} as the -2 log-likelihood from the restricted model and ll_{U} as the -2 log-likelihood from the unrestricted model where

$$ll_{\mathrm{U}} = \sum_{i=1}^{G} ll_i \qquad (6)$$

Therefore, the LRT for testing sequence group differences is given as

$$\mathrm{LRT} = ll_{\mathrm{R}} - ll_{\mathrm{U}} = ll_{\mathrm{R}} - \sum_{i=1}^{G} ll_i \sim \chi^2 \qquad (7)$$

with *G*-1 degrees of freedom that follows the chi-squared distribution. Applying (7) to sequence data by using (4) and (5), we obtain

$$\text{LRT} = \frac{n_A}{w}\log\left(\frac{s_A}{n_A}\right) - \frac{n_A}{w}\log\left(\sum_{i=1}^{G}\frac{s_i}{n_i}\right) \sim \chi^2 \qquad (8)$$

Note that the LRT of (8) is also based on the nested model principle, just like the BIC difference given by (4) and (5). Thus, it can be considered a significance test equivalent to the nested BIC assessment. Because of the similarity between the LRT in (8) and the BIC difference based on (4) and (5), we did not conduct another simulation study to verify its statistical behavior. The results in Appendix A should apply here (as the actual data analysis confirms later).

4. Data

We use the newly released German National Education Panel data (NEPS), starting from cohort 6 (NEPS) (Leopold, Skopek, & Raab, 2011). This part of NEPS contains retrospective life course information for 11,649 individuals born between 1944 and 1986 who were surveyed in 2009/2010. The survey instruments contain detailed questions about education, work and work interruptions, as well as family formation, including the formation and dissolution of marital and cohabiting unions. We use data for family and employment trajectories from ages 15 to 40 of East and West German men and women born 1944-1970 measured in monthly intervals. We thereby start with the oldest available cohort born in 1944 and end with the cohort born in 1970, which is the last birth cohort that we can observe until age 40. Examining sequences until age 40 is important to assume that individuals have reached occupational maturity (Aisenbrey and Brückner 2008) and have largely completed the active family formation phase. For these cohorts we have 6,578 sequences of complete family and employment information from ages 15 to 40, 5,432 for West Germany and 1,146 for East Germany. For the cohort comparison, we group birth cohorts in 4 year-intervals: 1944-1949, 1950-1953, 1954-1957, 1958-1961, 1962-1965, 1966-1970. The oldest and youngest birth cohort groups are slightly larger and comprise 5 birth years.

This grouping provides sufficient case numbers per cohort group for East and West Germany and is sufficiently fine grained to detect change over time in the birth cohorts' life courses.

Figure 1 shows a Lexis diagram that places our study cohorts in historical context. The cohorts born 1944-1949 experienced their entire life courses until age 40 in divided Germanys. The cohorts born 1950-1953 and 1954-1957 had already largely completed family formation and reached occupational maturity when they experienced the reunification in the 30ies. For these cohorts childhood and young adulthood happened during very different state systems that had been clearly established in the 1960s and 1970s, whereas the 1940s and 1950s were arguably foremost characterized by rather universal post-war situation of scarcity and then an emerging new economy. Finally the cohorts born in the 1960s already experienced their active family formation phase and labor market entry after the reunification in 1990.

Figure 1: Lexis diagram on study cohorts

The family sequences are specified with nine states of "single, no child" (SNC), "single, with child" (SC), "cohabiting, no child" (CNC), "cohabiting, with child" (CC), "married, no child" (MNC), "married with child" (MC). Being single is defined as not being in a cohabiting relationship and thus includes persons who were never married as well as divorcees. Only biological children are included. The percentage of reported adopted and foster children is very low (below 1 percent each) and legislation and practice of adoption and foster parenting differed substantially in East and West Germany.

The employment sequences are operationalized based on different stages out of the labor force and Erickson-Goldthorpe-Portocarero (EGP) classes of employment. EGP classes are hierarchical but categorical in nature and thus ideally suited to examine in a sequence analysis framework while maintaining the ability to distinguish upward and downward mobility. We specify 13 employment states including “out of the labor force/gap”(OLF), “unemployment” (UE), “military” (M) “education” (EDU)”, and “parental leave” (PL). The EGP classes are included as “Higher grade professionals [I]”, “Lower grade professionals [II]”, “Routine non manual employees, higher grade [IIIa]”, “Routine non manual employees, lower grade [IIIb], “Small Proprietors, farmers [IVa,b,c], “Lower grade technicians; supervisors of manual workers [V]”, “Skilled manual workers [VI]”, “Manual worker in primary production [VIIa,b]”. Further, to increase rigor in our East-West comparison, we only included respondents in the East West samples, who were born in the respective region and where still living there at the time of the interview in 2009/2010. We excluded foreign-born persons, persons who migrated between East and West and people living in Berlin. This was necessary because the survey did not distinguish between the former East and West regions of Berlin at the time of the interview in 2009/2010 and thus they could not unambiguously be assigned to either the East or West regions.The data are weighted using a calibrated design weight provided in the NEPS. This weight includes a sampling design weight and a calibration factor (multiplier) to adjust the sample to the means of the German Microcensus 2009 (Aßmann & Zinn, 2011).

Note that in addition to this substantive application it could be useful to test the proposed methods on simulated sequence data. However, to date it is not clear how to simulate appropriate life course sequences. Simply randomly assigning orders of states in random duration with transitions at random timing are not appropriate benchmarks, because patterning in life course sequences is generally high and will always be more structured that such completely randomly simulated sequences. If random assignment of the core properties – order, duration timing – is not an option, this requires specifying other properties. This is difficult to do in a sensible way for the combination of order duration and timing, especially considering that sequence analysis is embedded in an algorithmic data modeling culture that abstains from parametric assumptions.

4.2 Visual description of the example data: relative frequency sequence plots

The sample sizes of the life course sequences measured over 25 years are too large to plot in conventional sequence index plots, due to overplotting that distorts the visual impression of the graphs. Therefore we use relative frequency sequence plots to graphically display the sequences (Fasang & Liao 2014). For the relative frequency sequence plots, first the sequences of each subgroup are sorted according to a score of the first factor derived with multidimensional scaling (see also Piccareta and Lior 2010). The multidimensional scaling is based on a sequence distance matrix that was generated with Optimal Matching with constant substitution costs of 2 and indel costs of 1 (see MacIndoe and Abbott 2004). Then each sample of the four comparison groups is divided into 100 equally sized frequency groups. For each of these frequency groups, the medoid sequence, also calculated using optimal matching with constant substitution costs of 2 and indel costs of 1, is chosen as the representative of these medoids (see Fasang an Liao 2014 for details).

Figure 2 shows the relative frequency plots for men and women in East and West Germany for the family life courses. Figure 3 displays the employment trajectories for men and women in East and West Germany. The distance to medoid plots on the right indicate the average distance from the chosen medoid that is displayed in the RF sequence plot.

Figure 2: RF Sequence Plots of family trajectories for East and West German men and women, k=100 sorted by score of first factor derived with multidimensional scaling

The two upper panels in Figure 2 for East Germany show a polarization between individuals who are married with children at an early age, visualized by the blue colors. Notably, this early marriage with child pattern for East Germany is highly standardized, as indicated by the very low distances in the distance to medoid-plots for this region of the plot. Marriage and parenthood are highly coupled with only short periods of childless marriage for both men and women. This pattern accounts for the majority of East and West Germany men in our sample (about 2/3). At the bottom of the two upper panels, we see a very different pattern of unmarried parenthood in cohabitation for men and women (dark purple), and single motherhood for women in East Germany (green). The family life courses represented in the lower region of the RF Sequence plots are far more heterogeneous with larger distances from the medoid.

The two lower panels of figure 2 show family life courses for West German men and women. We see an orderly pattern of short cohabitation, relatively long childless marriage followed by parenthood, which suggests clear “normative clocks” for structuring family formation. Men and women who marry at a later age on average show longer periods of cohabitation before. Overall, the coupling of marriage and parenthood is weaker in terms of timing compared to East Germany. In contrast to East Germany, the prevalence of parenthood out of wedlock is much lower in the West reflecting the strong male breadwinner norms for our study cohort. In West Germany family life courses characterized by long periods of cohabitation are the most heterogeneous, indicated by the distance to medoid plot.

Figure 3: RF Sequence Plots of employment trajectories for East and West German men and women, k=100 sorted by score of first factor derived with multidimensional scaling

Figure 3 shows the RF plots for the corresponding employment sequences. The EGP classes are depicted using heat colors, such that light yellow represents the lowest class and red represents the highest class. For all four comparison groups employment trajectories were most orderly and homogenous in the highest prestige occupations visualized in red. It is a known feature of the EGP class scheme that the lowest classes primarily comprise male occupations, which is also visible in our graphs. In West Germany we see more uninterrupted episodes of employment or non-employment (women) compared with more frequent movements between classes and in and out of the labor force in East Germany. This likely reflects the strong insider outsider segmentation and high employment protection in the conservative corporatist West Germany welfare model. We now turn to assessing quantifying and assessing the difference between these for groups and assess how differences between East and West German men and women developed across birth cohorts. In particular, we are interested in observing a converging, diverging or stable difference between family and employment life courses in East and West Germany based on the proposed BIC and LRT statistics.

4.3 Assessing the normality assumption and making adjustment

Equation (2) given earlier relies on the normality assumption for the distances between individual sequences and the medoid. That is, to reasonably consider sequence distances as errors in a linear model, normal distribution of such distances is assumed. To assess the assumption visually, a common plot, known as the quantile-quantile (QQ) plots of the distance data can be employed. A QQ plot presents the standardized observed data against the standard normal distribution. There are also statistical tests for testing distributions, such as the Anderson-Darling test, the Kolmogorov-Smirnov test, and the Shapiro-Wilk test.[5]

Through a series of preliminary analysis, we found a mild degree of deviation from normality in the German data. To deal with the general problem of deviation from normality regardless of its severity, we propose an adjustment to the BICs from (3), (4), and (5) by using the Shapiro-Wilk test statistic. Of four common statistical tests of normality, the Shapiro-Wilk test is the most powerful (Razali and Wah 2011). The BIC can be adjusted by applying a function of the Shapiro-Wilk statistic to obtain aBIC, or adjusted BIC:

$$\mathrm{aBIC} = \frac{\mathrm{BIC}}{1-\log_{10}(SW_y)} \qquad (8)$$

where SW_y is the Shapiro-Wilk test statistic of the distribution of variable y. We conducted a simulation study over a range of sample sizes and degree of deviation from normality, and found (8) an effective way of compensating BIC in the presence of deviation from normality.[6]

5. Findings from the Comparion of East and West German Life Course Sequences

We report in Tables 1 and 2 the results from an application of discrepancy analysis on the East and West German life course sequence data. As shown clearly in all the statistics presented in the two tables, discrepancy analysis is very sensitive to any differences between the two data sources, producng rather small pseudo-R^2s and highly significant pseudo-F tests for either the employment history or the family history life course comparisons between the East and the West.

Table 1: Pseudo-R^2, pseudo-F, and Levene statistics analyzing the German employment history data

Table 2: Pseudo-R^2, pseudo-F, and Levene statistics analyzing the German family history data

Because of this lack of differentiation in comparing the two sequence groups, we conducted BIC and LIT analysis of the East and West German life course data, and present the findings in Tables 4 and 5. As summarized in Table 3, a BIC differences between 0 and 2 indicates negligible differences between two sets of sequences, 2 to 6 suggests moderate differences whereas BIC difference above 6 can be considered strong and above 10 very strong.

Table 3: BIC comparison guide

Table 4: Bayesian Information Criterion differences and Likelihood Ratio statistics of the German family history data, the East vs the West (nested model approach, BIC_a, BF_a, and LR_a are the adjusted BIC, BF, and LR using Shapiro-Wilke statistic)

Table 5: Bayesian Information Criterion differences and Likelihood Ratio statistics of the German employment history data, the East vs the West (nested model approach, BIC_a, BF_a, and LR_a are the adjusted BIC, BF, and LR using Shapiro-Wilke statistic)

Overall we find significant differences between east and West German family (Table 4) and employment life courses (Table 5). The difference indicated by BICa is larger for family life courses (27.3) than for employment life courses (17.7). Moreover on both life course dimension, East and West German women are more different from one another than East and West German men as is visible in the larger BICa for women.

In addition to these overall differences, we find notable cohort variation that support divergence after the life courses of the 1944-1949 cohort followed by convergence for the post-reunification cohorts for family and employment life courses. BIC differences are largest for the cohorts born 1950-1953 and remain sizeable for the cohorts born 1954-1958. Beginning with the cohorts born 1960 differences between east and West German life course diminish and are no longer significant for the youngest birth cohorts. We visualize the adjusted BIC differences by cohort in figures4 (based on values displayed in tables 4 and 5). Also, men's life courses appear to differ less between the East and the West than women's life courses. This pattern is true for both the family formation and the employment history comparisons.

Figures 4: Visualization of BICa

6. Discussion

One may believe that the 1944-1949 birth cohort should exhibit the greatest difference in BIC because these women and men lived almost their entire lives between age 15 and 40 in the divided countries. This understanding, however, is incorrect. According to prior life course research, a period of planful orientation during mid-adolescence (around ages 14 and 15) is important for one's later life course because this planful orientation helps with realistic decisions about adult roles and relationships (Shanahan 2000).

The building of the Berlin Wall started in August, 1961 though improved wire fences were added during the three years of 1962 to 1965 and improved concrete walls were added after that. Thus, even the youngest of the 1944-1949 cohort did not spend their mid-adolescence in an entirely closed-off and separated Germanys. Perhaps that is why the differences for the oldest cohort are rather modest in size, even showing statistically significant (at the 0.05 level) LRTs for men's family history and employment history comparisons.

The middle cohorts of either men or women show the greatest differences between the East and the West for both their family formation and employment history life course comparisons. This is not surprising because these women and men spent most of their post-mid-adolescent years in the separate Germanys, with their different social welfare systems and availability of childcare facilities. The extent to which men's and women's life courses differ is rather small for these cohorts in either the family formation or the employment history dimension.

The youngest male or female cohort, however, had a different life course experience from their older compatriots. These women and men spent their entire early, mid- and late adolescent years in the separate Germanys. However, when they became adults, the two Germanys had been unified; similar employment and welfare opportunities began to become available to them.

Indeed, these women and men lived much of their adult years between 18 and 40 in the unified Germany, with their life courses, in terms of either family formation or employment history, shaped by the new unified social structure.

To return to our research question of whether state policies generate differences in life courses, we offer a rather affirmative answer because the middle cohorts demonstrate consistent differences in family formation and in employment history for both men and women of the East and the West origins. The smaller differences in the oldest as well the youngest cohorts further confirm this answer. Whereas similar macro-situations during mid-adolescent years helped reduce men's life course differences among those of the oldest cohort, the identical state policies promoted a divergence of the earlier differing life courses between those from the former East Germany and those from the former West Germany, especially among men.

7. Conclusion

We have shown in this paper that statistically comparing samples of life course sequences is a difficult task: Comparing sets of single value measurement such as income would require simply testing mean or log-mean differences of the two samples. Life course sequences, on the other hand, contain a set of complex measurements including the timing of life course events, the duration of such events as well as their ordering. In this proper, we have proposed a method based on distances of sequences from a set of representative medoid sequences (instead of a single medoid) for generating BICs and LRTs for properly assessing differences between sets of sequences statistically.

Because distances can be a function of sequence length, the number of qualitative states, and the number of comparisons being made, we devised a weight for compensating the inflation effect of this function. Our simulation study assessed the relative adequacy of the weighted BIC and LRT computations. Such weighted BICs and LRTs were used in the assessment of the differences in family and employment history sequences between East and West German men and women.

The analysis of the German life course sequence data provided some clear support for macro-level, institutional effects on micro-level, individual life course differences. A pattern of diverging trends between the two Germanys in individual life courses took shape with the erecting of the Berlin Wall and converging trends in these individual life course sequences occurred with the tearing down of the Berlin Wall and the reunification, based on our analysis of the six birth cohorts of East and West German data. Such diverging and converging trends are more observable among men than among women. This gender difference suggests that men's life courses are more moldable than women's by societal and institutional changes exemplified by the postwar German history of two Germanys and their reunification.

References

Aassve, Arnstein, Francesco C. Billari and Raffaella Piccarreta. 2007. Strings of Adulthood: A Sequence Analysis of Young British Women's Work-Family Trajectories. *European Journal of Population* 23: 369-388.

Aisenbrey, Silke, and Anette E. Fasang. 2010. "New Life for Old Ideas: The 'Second Wave' of Sequence Analysis. Bringing the 'Course' Back into the Life Course." *SociologicalMethods & Research* 38:420-62.

Aßmann, C., & Zinn, S. 2011. NEPS Data Manual. Starting Cohort 6: Adult Education and Lifelong Learning [Supplement C] Weighting. *National Eucation Panel Study Data Center (NEPS) Documentation, University of Bamberg*.

Blanchard, Philippe, Felix Bühlmann and Jacques-Antoine Gauthier (eds). 2014. *Advances in Sequence Analysis: Methods, Theories and Applications*. Spinger.

Brückner, Hannah and Karl Ulrich Mayer. 2005. De-standardization of the Life Course: What it might mean? And if it means anything, whether it actually took place? *Advances in Life Course Research* 9: 27-53.

Burnham, Kenneth P. and David R. Anderson. 1998. *Model Selection and Inference: A Practical Information-Theoretical Approach*. New York: Springer-Verlag.

______. 2004. "Multimodel Inference: Understanding AIC and BIC in Model Selection." *Sociological Methods & Research* 33: 261-304.

Fasang, Anette E. 2014. New Perspectives on Family Formation: What Can we Learn from Sequence Analysis? In: Blanchard, Philippe, Felix Bühlmann and Jacques-Antoine Gauthier (eds). 2014. *Advances in Sequence Analysis: Methods, Theories and Applications*. Spinger.

Fasang, Anette E. and Tim F. Liao. 2014. “Visualizing Sequences in the Social Sciences: Relative Frequency Sequence Plots.” *Sociological Methods & Research* 43(4): 643-676.

Gabadinho, Alexis, Gilber Ritschardm Nicolas Müller and Matthias Studer. 2011. Analyzing and Visualizing State Sequences in R with TraMineR. *Journal of Statistical Software* 40(4): 1-37.

Goldstein, Joshua and Michaela Kreyenfeld. 2011. Has East Germany Overtaken West Germany? Recent Trends in Order-specific Fertility. *Population and Development Review* 37(3): 453-472.

Kass, Robert E., and Adrian E. Raftery. 1995. “Bayes Factor.” *Journal of the American Statistical Association* 90: 773-795.

Leopold, T., Skopek, J., & Raab, M. (2011). NEPS Data Manual. Starting Cohort 6: Adult Education and Lifelong Learning. *National Eucation Panel Study Data Center (NEPS) Documentation, University of Bamberg*.

Liao, Tim F. 2002. *Statistical Group Comparison*. New York: Wiley.

Liao, Tim F. 2004. “Comparing Social Groups: Wald Statistics for Testing Equality among Multiple Logit Models.” *International Journal of Comparative Sociology* 45: 3-16.

Oh, Man-Suk, and Adrian E. Raftery. 2001. “Basyesian Multidimensional Scaling and Choice of Dimension.” *Journal of the American Statistical Association* 96: 1031-1044.

Piccarreta, Raffaella. 2012. “Graphical and Smoothing Techniques for Sequence Analysis.” *Sociological Methods & Research* 41: 362-380.

Marshall, Victor W., and Margaret M. Mueller. 2003. “Theoretical Roots of the Life-Course Perspective.” Pp. 3-32 in *Social Dynamics of the Life Course: Transitions, Institutions,* *and Interrelations*, edited by Water R. Heinz and Victor W. Marshall. New York: Aldine de Gruyter.

Mayer, Karl Ulrich, and Urs Schoepflin. 1989. "The State and the Life Course." *Annual Review of Sociology* 15: 187-209.

Mayer, Karl Ulrich. 2004. "Why Lives? How History, Societies, and Institutions Define and Shape Life Courses." *Research in Human Development* 1(3): 161-187.

______. 2009. "New Directions in Life Course Research." *Annual Review of Sociology* 35: 413-433.

Neyer, Gerga. 2013. "Welfare State, Family Policies, and Fertility in Europe." Pp. 29-53 in *The Demography of Europe*, edited by Gerda Neyer, Gunnar Andersson, Hill Kulu, Laura Bernardi, and Christopher Büller. London: Springer.

Raftery, Adrian E. 1986. "Choosing Models for Cross-Classifications." *American Sociological Review* 51(1): 145-146.

Razali, Nornadiah Mohd, and Yap Bee Wah. 2011. "Power Comparisons of Shapiro-Wilk, Kolmogorov-Smirnov, Lilliefors and Anderson-Darling Tests." *Journal of Statistical Modeling and Analytics* 2(1): 21-33.

Schnettler, S., & Klüsener, S. (2014). Economic stress or random variation? Revisiting German reunification as a natural experiment to investigate the effect of economic contraction on sex ratios at birth. *Environmental Health*, 13(1), 117.

Shanahan, Michael J. 2000. "Pathways to Adulthood in Changing Societies: Variability and Mechanisms in Life Course Perspective." *Annual Review of Sociology* 26: 667-692.

Studer, Matthias, Gilbert Ritschard, Alexis Gabadinho, and Nicolas S. Müller. 2011. “Descrepancy analysis of State Sequences.” *Sociological Methods & Research* 40(3): 471-510.

Yazici, Berna, and Senay Yolacan. 2007. “A Comparison of Various Tests of Normality.” *Journal of Statistical Computation and Simulation* 77(2): 175-183.

Appendix A

We conducted a simulation study comparing two random samples of family history sequences drawn with replacement from the early births and late first births life history sequence samples from the 1966 women birth cohorts sorted by the timing of first births at varying sample sizes of different mixing percentages. These two birth-timing groups were chosen to ensure their life courses are different enough. The purpose of the simulation study is to investigate the effect of two variations—that of different mixing proportions from the two samples of origin and that of sample size on BIC differences (because BIC is a function of sample size). Without doubt, we are primarily interested in finding out whether BIC differences can reflect true sample differences in their life course sequences.

Each pair of samples sum up to a combined sample size of 100, 200, 500, 1,000, and 2,000, with the mixing percentages from the west and the east samples of 5% and 95%, 10% and 90%, 20% and 80%, 30% and 70%, 40% and 60%, and 50% and 50%, respectively. Each of the data situations of the mixing percentages at the various sample sizes is simulated 1,000 times.

Figure A.1 presents the simulation results. Clearly, BIC differences, to a degree, are still a function of sample size, with larger sample sizes generating greater BIC differences (going down the rows of the violin plots), despite of the weight *w* applied in (4) and (5). It is reassuring to know that these BIC differences truly reflect the actual differences in the percentages of the samples drawn from each of the sources. That is, a 50%-50% should give rise to the greatest BIC difference while a 5%-95% should generate an insignificant amount of difference. The rather small spread shown by the violin plots indicate statistical efficiency of BIC difference caculations.

---Figure A.1 about here---

Table A.1 shows the median values of the BIC difference simulations when comparing these two groups of samples. These median values indicate when a BIC difference is above, 2, 4, 6, or 10, the various guideline threshold values, given the different mixtures of the two birth-timing cohorts and the sample size.

---Table A.1 about here---

The simulation study shows that the computation of BIC difference discussed in the methods section provides an efficient method for assessing sequence differences. Relatively speaking, small sample sizes are less efficient than larger sample size BIC calculations.

Figures and Tables

Figure 1: Lexis Diagram of Life Course from Ages 15 to 40 of the Study Cohorts (1944-1970) Placed in Historical Time (1959-2009), red line vertical marks German reunification

Figure 2: RF Sequence Plots of family trajectories for East and West German men and women, k=100 sorted by age of first childbirth

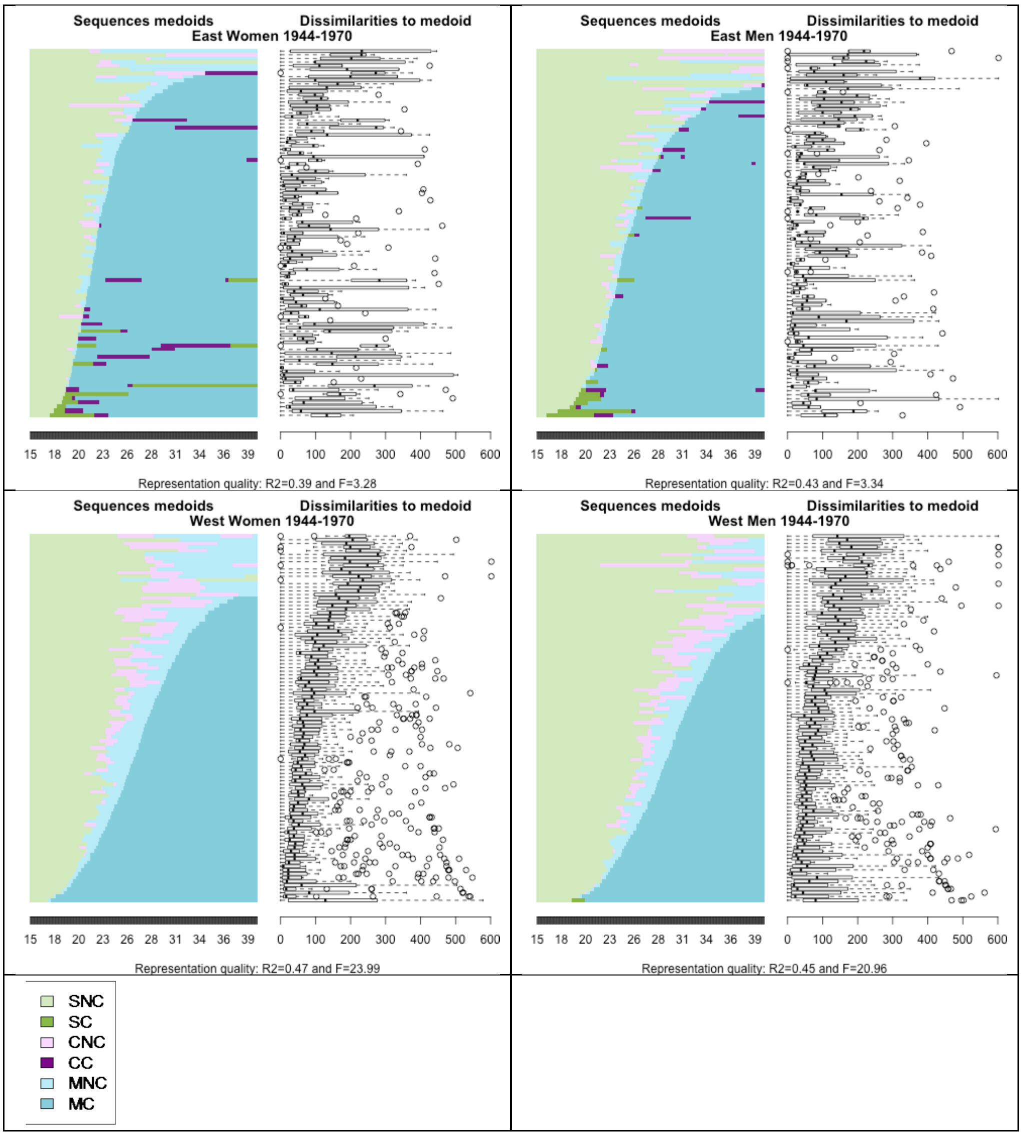


Figure 3: RF Sequence Plots for employment trajectories for East and West German men and women, k=100 sorted by age of first job

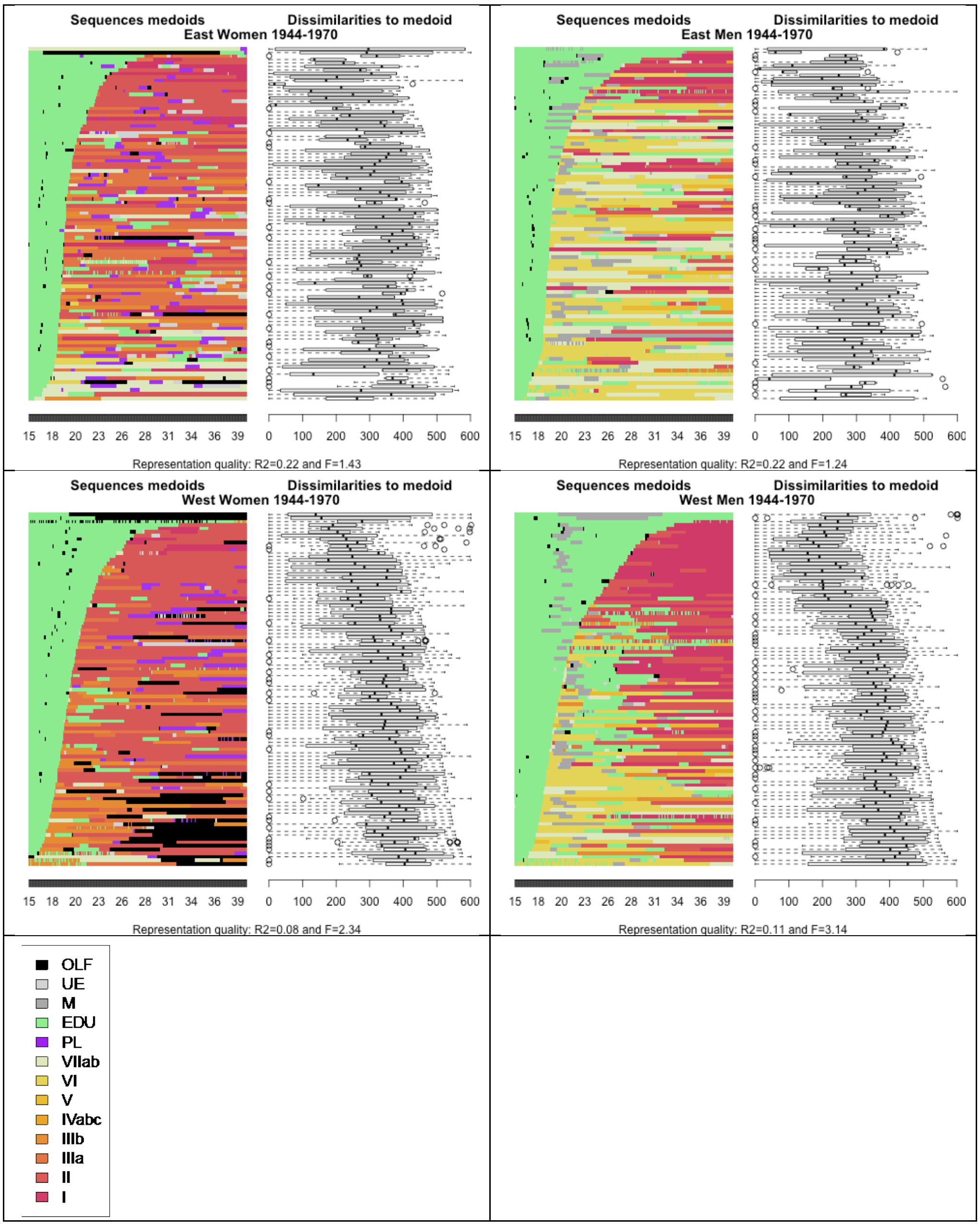


Figure 4: Line Graph of BIC difference

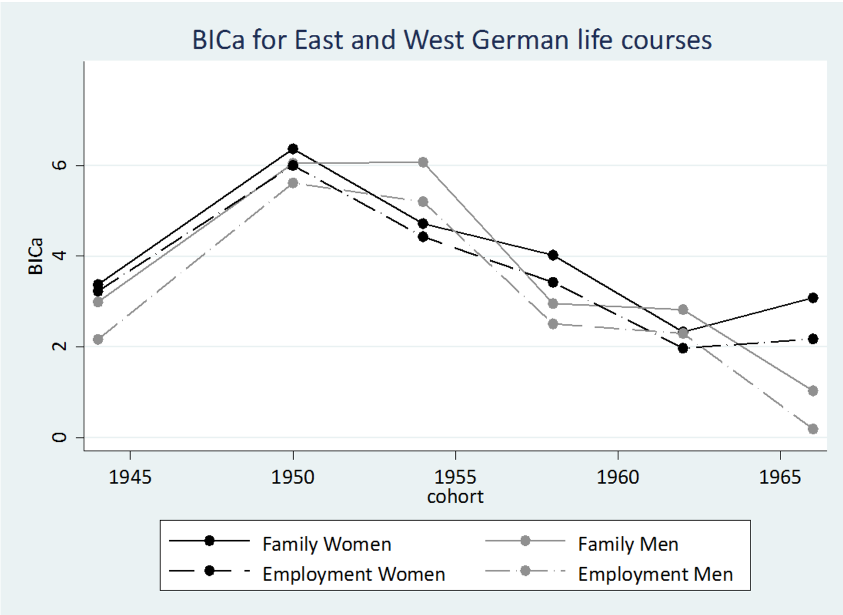


Table 1: Pseudo-R^2, pseudo-F, and Levene statistics analyzing the German employment history data

East–West	Pseudo R^2	Pseudo F	Sig. level	Levene	Sig. level
overall	0.005	30.187	0.001	0.812	0.380
men	0.008	23.918	0.001	7.407	0.008
women	0.007	23.906	0.001	0.048	0.806
cohort 44, m	0.004	2.005	0.018	3.559	0.054
cohort 50, m	0.004	2.005	0.018	3.559	0.054
cohort 54, m	0.013	6.377	0.001	14.285	0.001
cohort 58, m	0.012	7.642	0.001	5.983	0.014
cohort 62, m	0.007	4.529	0.001	0.451	0.487
cohort 66, m	0.008	4.287	0.001	0.375	0.535
cohort 44, w	0.008	4.287	0.001	0.375	0.535
cohort 50, w	0.008	4.287	0.001	0.375	0.535
cohort 54, w	0.008	4.287	0.001	0.375	0.535
cohort 58, w	0.008	4.287	0.001	0.375	0.535
cohort 62, w	0.008	4.287	0.001	0.375	0.535
cohort 66, w	0.008	4.287	0.001	0.375	0.535

Table 2: Pseudo-R^2, pseudo-F, and Levene statistics analyzing the German family history data

East–West	**Pseudo R^2**	**Pseudo F**	**Sig. level**	**Levene**	**Sig. level**
overall	0.019	128.267	0.001	19.277	0.001
men	0.018	58.57	0.001	1.730	0.170
women	0.021	73.253	0.001	21.423	0.001
cohort 44, m	0.021	73.253	0.001	21.423	0.001
cohort 50, m	0.021	9.234	0.001	5.506	0.015
cohort 54, m	0.029	14.74	0.001	4.859	0.028
cohort 58, m	0.022	14.109	0.001	0.071	0.795
cohort 62, m	0.028	17.311	0.001	0.342	0.576
cohort 66, m	0.010	5.094	0.001	6.275	0.009
cohort 44, w	0.020	9.035	0.001	10.52	0.001
cohort 50, w	0.016	6.525	0.001	6.311	0.021
cohort 54, w	0.018	10.000	0.001	7.397	0.007
cohort 58, w	0.028	20.777	0.001	1.218	0.272
cohort 62, w	0.026	18.922	0.001	0.491	0.472
cohort 66, w	0.026	15.647	0.001	0.030	0.866

Table 3: BIC comparison guide

Evidence	**BIC difference**	**Bayes factor**
Not worth a bare mention	0 to 2	1 to 3
Positive	2 to 6	3 to 20
Strong	6 to 10	20 to 150
Very strong	>10	>150

Table 4: Bayesian Information Criterion differences and Likelihood Ratio statistics of the German family history data, the East vs the West (nested model approach, BIC_a, BF_a, and LR_a are the adjusted BIC, BF, and LR using Shapiro-Wilke statistic)

East–West	BIC	BIC_a	BF	BF_a	LR	Sig. level	LR_a	Sig. level
overall	28.014	27.292	1211309	843871.3	31.486	0.000	30.678	0.000
men	15.061	14.920	1863.655	1737.08	17.808	0.000	17.641	0.000
women	24.284	23.376	187613.8	119107.5	27.094	0.000	26.08	0.000
cohort 44, m	2.212	2.206	3.022	3.014	3.173	0.075	3.165	0.075
cohort 50, m	6.994	6.957	33.018	32.415	7.594	0.006	7.554	0.006
cohort 54, m	5.971	5.925	19.793	19.344	6.708	0.010	6.656	0.010
cohort 58, m	2.272	2.248	3.115	3.077	3.425	0.064	3.389	0.066
cohort 62, m	3.293	3.235	5.189	5.042	4.42	0.036	4.343	0.037
cohort 66, m	0.197	0.191	1.103	1.100	1.26	0.262	1.226	0.268
cohort 44, w	4.592	4.521	9.933	9.59	5.351	0.021	5.269	0.022
cohort 50, w	7.326	6.796	38.969	29.898	7.842	0.005	7.275	0.007
cohort 54, w	5.592	5.470	16.380	15.411	6.563	0.010	6.420	0.011
cohort 58, w	4.955	4.696	11.912	10.463	6.228	0.013	5.902	0.015
cohort 62, w	2.120	2.072	2.887	2.817	3.44	0.064	3.360	0.067
cohort 66, w	3.563	3.448	5.939	5.606	4.673	0.031	4.521	0.033

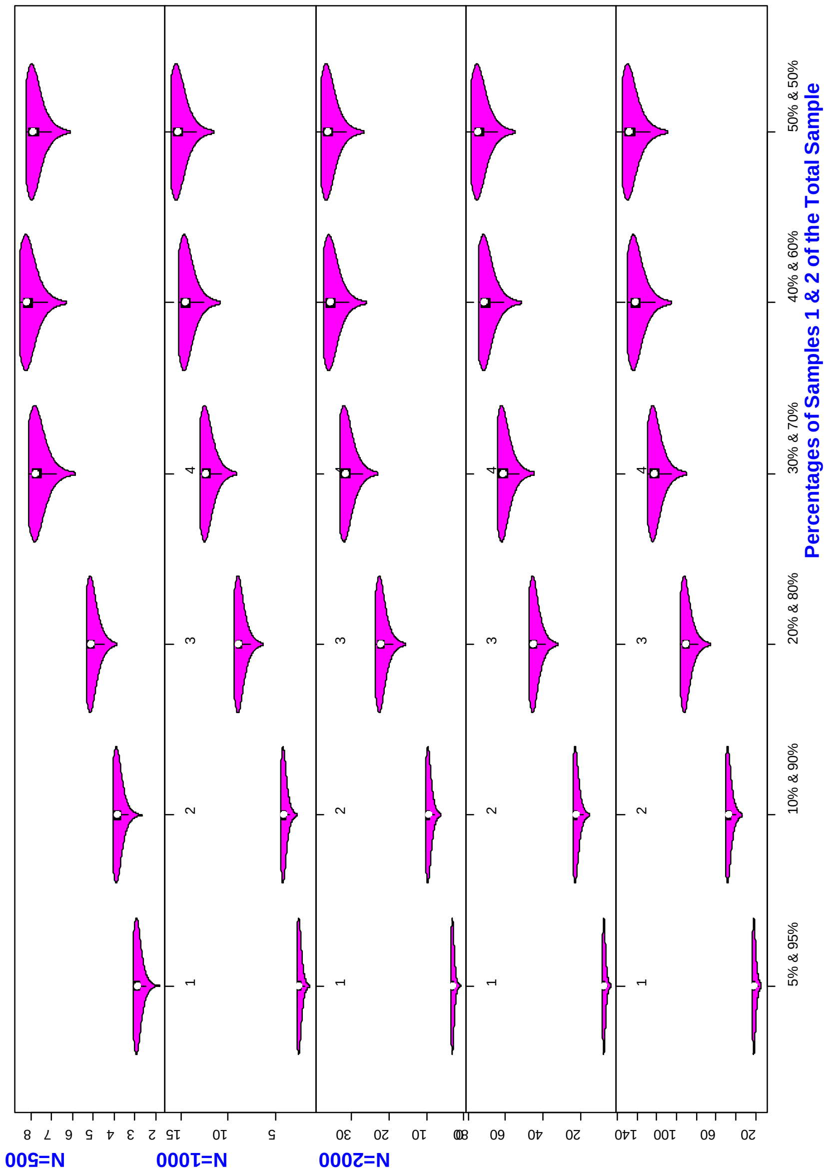
Note: The computation used a multiple representative medoid approach, in this case, 5% representative medoids.

Table 5: Bayesian Information Criterion differences and Likelihood Ratio statistics of the German employment history data, the East vs the West (nested model approach, BIC_a, BF_a, and LR_a are the adjusted BIC, BF, and LR using Shapiro-Wilke statistic)

East–West	BIC	BIC_a	BF	BF_a	LR	Sig. level	LR_a	Sig. level
overall	15.032	14.746	1837.476	1592.316	18.325	0.000	17.670	0.000
men	11.5	11.391	314.24	297.455	14.068	0.000	13.934	0.000
women	12.972	12.84	655.926	613.998	15.602	0.000	15.443	0.000
cohort 44, m	2.696	2.617	3.850	3.700	3.494	0.062	3.391	0.066
cohort 50, m	6.604	6.468	27.170	25.38	7.024	0.008	6.879	0.009
cohort 54, m	6.378	6.24	24.269	22.647	6.936	0.008	6.786	0.009
cohort 58, m	3.586	3.499	6.006	5.753	4.559	0.033	4.449	0.035
cohort 62, m	2.883	2.803	4.227	4.062	3.831	0.05	3.725	0.054
cohort 66, m	0.577	0.567	1.335	1.328	1.461	0.227	1.435	0.231
cohort 44, w	3.453	3.384	5.621	5.431	4.033	0.045	3.952	0.047
cohort 50, w	6.443	6.24	25.063	22.645	6.780	0.009	6.567	0.010
cohort 54, w	4.895	4.854	11.557	11.325	5.686	0.017	5.639	0.018
cohort 58, w	4.268	4.215	8.447	8.227	5.361	0.021	5.295	0.021
cohort 62, w	2.527	2.509	3.538	3.506	3.667	0.056	3.641	0.056
cohort 66, w	3.135	3.055	4.795	4.607	4.065	0.044	3.962	0.047

Note: The computation used a multiple representative medoid approach, in this case, 5% representative medoids.

Figure A.1: Simulation Adjusted BIC Results of Comparing Varying Percentages of Samples 1 and 2 of a Total Sample Size of 100 to 2000 Using Random Samples of the 1966 Early and Late Birth Women's Cohort Sequences with Replacement



36

Table A.1: Medians of the Simulated BICs Reported in Figure A.1

N	**5% & 95%**	**10% & 90%**	**20% & 80%**	**30% & 70&**	**40% & 60%**	**50% & 50%**
100	2.908040	3.874113	5.125869	7.782945	8.210573	7.929981
200	2.521137	4.169957	8.932371	12.454760	14.622118	15.463125
500	3.079883	9.587429	22.545704	31.932319	35.956990	36.619731
1,000	7.538608	22.734413	45.291397	61.664360	71.330921	74.745039
2,000	21.636605	48.013100	91.704994	123.490443	142.945227	148.732730

Endnotes

[1] Indeed, the status attainment model that focuses on education and occupation is one of the origins of life course research (Marshall and Mueller 2003).
[2] In a similar tradition, Oh and Raftery (2001) adapted BIC for assessing dimension choice in multidimensional scaling.
[3] While increasing the number of states theoretically increases then chance of variability exponentially, which can be further multiplied by the number of sequences in the sample, empirically many observations tend to change states similarly, thus decreasing the total variability. This is reflected by the natural logarithm function and the square root function of the overall weight.
[4] The treatment of a set of representative sequences as a single medoid actually makes statistical sense. Say we use 5% representative sequences for both the combined and the separate samples. As a result the actual number of representative sequences used in the overall, combined sample and the number of representative sequences used in the separate (say the East and the West) samples are equal but in actuality the latter situation provides a better fit to the data. Thus, here the collective representation set replaces the single medoid.
[5] For a general comparison of various normality tests, see Yazici and Yolacan (2007).
[6] In the interest of space, we do not present the simulation results, which can be obtained upon request.