Same or Opposite? Parent Effects on Family Formation.

Anette Eva Fasang

Humboldt-University Berlin & Social Science Research Center Berlin (WZB)

Marcel Raab

University of Bamberg & Social Science Research Center Berlin (WZB)

Paper prepared for the 2012 EPC Meeting, Stockholm

Abstract

Research about parent effects on family behavior traditionally examines intergenerational transmission: whether children adopt the same family behavior as their parents. In this article, we draw attention to empirical regularities where children follow opposite family formation trajectories as their parents. Such regularities are stable parent effects on family behavior, even though there is no direct intergenerational transmission. We use data from the Longitudinal Study of Generations (LSOG, N=461 parent-child dyads) and propose an innovative application of multichannel sequence analysis and cluster analysis to analyze family formation trajectories between age 15 and age 40 holistically. Results show clusters of both 'same' and 'opposite' parent - child family formation. We conclude that, by focusing on average effects, regression-based methods for studying intergenerational transmission may loose sight of regularities where children shape their family formation in opposition to the template observed in their parents.

Note: figures 2 and 3 should be viewed in color

Research on parent effects on family behavior mostly examines intergenerational transmission of focal transitions, such as fertility (Barber 2000; Murphy 1999), divorce (Amato 1996; Wolfinger 2000) and, to a lesser extent, marriage (Feng et al. 1999; van Poppel et al. 2008). Generally, when children adopt the same behavior as their parents, this is taken as evidence for intergenerational transmission of family behavior. For instance, if children have their first birth at the same age as their parents, or if they have the same number of children as their parents, this is seen as an indicator of intergenerational transmission of fertility.

This focus on direct transmission, where children follow the same behavioral patterns as their parents, looses sight of other regularities in parent effects on their children's family formation. Such regularities are empirically stable patterns, where parents with a specific family behavior tend to have children with another family behavior, even if it is not the same for parents and children. In fact; they might be quite opposite, for example if parents who have many children tend to have children who have no or very few children themselves. Beyond direct intergenerational transmission, parents or significant role models can serve as a template that children either decide shape their own lives in opposition to. Such 'opposite' patterns would not indicate direct intergenerational transmission of family behavior, but they would suggest that some mechanism links a specific parental family behavior to another family behavior among their children.

We argue that, by focusing on average effects, traditionally applied regression-based methods and their extensions, such as event history and multilevel models, have limitations to study such 'opposite' parent effects on their children's family behavior. If the empirical distribution is indeed such, that children of specific parents cluster into ‘same’ or ‘opposite’ family formation patterns, average effects on isolated aspects of the family formation trajectory give a poor representation for both the ‘same’ and ‘opposite’ parent-child dyads. To study ‘opposite’ patterns of children’s family formation, it may be more conducive to first identify the ‘same’ and ‘opposite’ groups and then examine what determines children’s likelihood to follow a ‘same’ or ‘opposite’ pattern.

This study explores sequence analysis and cluster analysis to study ‘same’ and ‘opposite’ patterns of children’s family formation relative to their parents. We conceptualize family formation holistically as the process of union formation and child bearing between age 15 and age 40 of parent-child dyads. We use data from the Longitudinal Study of Generations (LSOG) (Bengtson et al. 2002) and multichannel sequence analysis (Pollock 2007; Gauthier et al. 2010) to show that, indeed, there are regularities in family formation of parent-child dyads beyond direct intergenerational transmission. Specifically, the children of parents that marry and start having children early cluster into one ‘same’ and two ‘opposite’ groups: either they follow their parents trajectory of relatively early marriage, early fertility and high parities (‘same’), as opposed to marrying late and remaining childless until age 40 (‘opposite’), or divorcing at a relatively early age (‘opposite’).

Moreover, our analytical approach proves useful for identifying intergenerational transmission against cohort trends in family formation, broadly captured in the notion of the ‘second demographic transition’ (SDT) (Elzinga and Liefbroer 2007, Lesthaghe, 2010). The core empirical regularities suggested in the second demographic transition are “sustained sub-replacement fertility, a multitude of living arrangements other than marriage, the disconnection between marriage and procreation and no stationary population.” (Lesthaghe 2010: 211). In essence, it is an argument about three trends: a (1) decline and (2) delay of fertility, and a general (3) de-standardization of family formation. In line with this assumption, we find two clusters where children follow the ‘same’ family formation pattern as their parents with the cohort-specific secular adjustment of starting everything a bit later and having one child less than their parents at age 40.

DATA AND SAMPLE

We use data from the Longitudinal Study of Generations (LSOG) (Bengtson et al. 2002). The LSOG is a four-generation study that was administered in seven waves (1971, 1985, 1988, 1991, 1994, 1997, 2000). The sample consists of families who were drawn randomly from a subscriber list of 840,000 members of a California Health Maintenance Organization in Los Angeles. The sampling units were grandparents (generation 1) of three-generation families. We use data for generation 2, the ‘silent generation’ born in the 1920s and 1930s, and their children, generation 3 the ‘baby boomers’ born in late 1940s and 1950s.

The family formation sequences cover 26 years between age 15 and age 40. The categorical family formation states, the alphabet of sequence states, are a combination of marital status and parity:

SNC = single, no child;

SC=single, one or more children;

MNC=married, no child;

M1C=married, one child;
M2C=married, two children;
M3C=married, three children;
M4C=married, four children;
DNC=divorced, no child;
DC=divorced, one or more children.

One example for such a trajectory is:

Age	15	16	17	18	19	20	21	22	23	24	25
ID1	SNC	SNC	SNC	SNC	MNC	M1C	M1C	M2C	DC	DC	DC

This trajectory describes a family formation process of being single and childless until age 19, then marrying, having the first child at age 20, the second child at age 22 and divorcing at age 23 thus leading into the state 'divorced with one or more children' until the age of 25. Our sample consists of family formation sequences between age 15 and age 40 for 326 generation 2 parents and 291 generation 3 children that amount to 461 parent-child dyads.

METHOD

We use sequence analysis (Abbott, 1995) to analyze family formation trajectories of parent-child dyads. Sequence analysis was originally developed in biology for comparing strings of DNA and adopted in the social sciences by Andrew Abbott to study sequences of categorical states. Applications to categorical state sequences in the social sciences range from dance patterns (Abbott and Forrest, 1986), employment careers

(Stovel et al., 1996; Biemann et al., 2011), to retirement trajectories (Han and Moen, 1999, Fasang, forthcoming, see Aisenbrey and Fasang 2010). Optimal Matching was the first method for sequence comparison used in the social sciences (Abbott, 1995). Optimal Matching calculates the distance between two sequences as the ‘cost’ of turning one sequence into another (for an introduction see MacIndoe and Abbott 2004). This alignment is performed with three transformation operations: substitution of one state with another, and insertion or deletion of states along the sequence. Substitution, insertion and deletion are each assigned a cost by the researcher. Distance between a pair of sequences is calculated as the minimum possible cost of turning one sequence into the other.

In this study, we treat parent-child dyads as the unit of analysis, since we are interested in empirical regularities in dyad combinations. Instead of an individual sequence, the unit of analysis is the parent-child dyad. Are there groups of parent-child dyads where parents and children follow the same family formation trajectory and groups of dyads where children follow opposite family formation trajectories from their parents? We propose to conceptualize the parents’ family formation sequence as one dimension and the child’s family formation as another dimension of parent-child dyads to analyze them holistically and take the dyad seriously as the unit of analysis. We therefore compare every parent-child dyad with every other parent-child dyad.

Several approaches have been suggested in the literature to deal with multidimensional sequences (Stovel et al., 1996; Blair-Loy, 1999; Han and Moen, 1999; Pollock, 2007; Gauthier et al., 2010). Relatively early on, Stovel et al. (1996) proposed to combine states from different dimensions into one multidimensional state space. For example if employment trajectories, family trajectories and residential trajectories are treated as three separate dimensions, possible combined states are ‘employed, divorced, house owner’ or ‘unemployed, married, homeless’ (see also Blair-Loy, 1999). However, this quickly leads to a very high numbers of possible state combinations that become practically unmanageable. Further, with combined state spaces it is impossible to distinguish how much each dimension contributes to the overall distance between sequences. Another proposal was to calculate pair-wise distance separately for each dimension and then post hoc add them to one combined pair wise distance matrix (Han and Moen, 1999).

Pollock (2007) suggests multiple sequence analysis that creates combined states of multiple dimensions but specifies different substitution costs for each dimension. Multiple sequence analysis was developed with multiple channels for the same person in mind, for example parallel employment trajectories, family trajectories, and residential trajectories (Pollock, 2007). We adopt the multiple sequence analysis approach to study parent-child dyads. Each parent-child dyad is associated with one intergenerational family formation sequence. The parents’ family formation trajectory is one dimension of this dyadic sequence and the child’s family formation trajectory is another.

Pollock (2007) proposes to specify a combined state space. For example, if the parent is single and has no children (SNC) at age 20 and the same is true for the child, this is combined to the state [SNC SNC]. The first three letters correspond to the parent’s state and the second three letters correspond to the child’s family formation state at this particular age. If the parent already was married and had 1 child at this age but the child was still single and had no children the combined state would be [M1C SNC]. For example with combined states one can describe two parent-child dyads, *A* and *B*, as follows:

Age	16	17	18	19	20
Dyad *A*	[MNC MNC]	[M1C M1C]	[M2C M2C]	[M3C M3C]	[M4C M4C]
Dyad *B*	[MNC SNC]	[M1C SNC]	[M2C SNC]	[M3C SNC]	[M4C SNC]

A shows a parent-child dyad of perfect transmission, or in other words, perfectly 'same' family formation behavior of parent and child. Both the parent and the child experience the same states of 'single, no child', 'married, no child', 'married, one child', 'married, two children', and 'married, three children' at the same ages. *B* shows a parent-child dyad of 'opposite' family formation. The parent in dyad *B* goes through the same states as dyad *A*, from single and being childless to being married and having four children, whereas the child in dyad *B* remains single and childless during the entire time.

In multiple sequence analysis, transformation costs are specified separately for each dimension (Pollock, 2007). In our application this corresponds to separate cost specifications for the parent and child dimensions of the dyadic sequences. We use Optimal Matching with substitution costs derived from transition rates between two states (Rohwer and Trappe, 1995; Gabadinho et al., 2011). We do not use indel operations because the sequences all are of equal length and timing is substantively important in family formation sequences: it is not the same whether one has 4 children at age 22 or at age 36 (see MacIndoe and Abbott 2004 on the specification of indel and subcosts to emphasize sequence similarity in terms of order and timing).

The transition rate between two states is the probability to transition from one state to another. Table 1 shows transition rates between family formation states for each generation. For example, the probability to transition from 'married, 1 child' (M1C) to 'married, 2 children' (M2C) is .34 in the parent generation. It is considerably lower at .18 in the child generation.

TABLE 1: Transition rates between family formation states for the parent and child generation

Based on these transition rates, substitution costs *SC* between state *i* and state *j* are calculated as:

$$SC_{ij} = 2 - p_{ij} - p_{ji}$$

where p_{ij} denotes the transition rate from state *i* to state *j*, and p_{ji} denotes the transition rate from state *j* to state *i*. Substitution costs based on transition rates are bound by 1 and 2. 1 is the lowest possible substitution cost when the probability of transition between two states is 100 percent. 2 is the highest possible substitution cost when the probability to transition from one state to another is zero. Table 2 shows the substitution costs based on transition rates between states separately for the parent and the child generation. The diagonal is set to zero since substituting a state with itself should of course not generate costs/distance. Continuing the example above, table 2 shows that substituting the states 'married, 1 child' (M1C) and 'married, 2 children' (M2C) comes at a cost of 1.66 (2 - .34 - .00) in the parent generation and at a cost of 1.82 (2 - .18 - .00) in the child generation. This is the case because transitioning from 'married, 1 child' to 'married, 2 children' was more frequent in the parent generation than in the child generation.

TABLE 2: Substitution costs based on transition rates between sequence states for the parent and child generation

Note that the substitution costs derived from transition rates take into account the generation-specific prevalence of different family formation states. Within each generation, substitution of two states will be cheaper if transitions between these states occur frequently. For instance in the parent channel *p* substitution of the states 'married, 3 children' and 'married, 4 children' will be cheaper and thus generate less distance than in the child channel *c*, because transitions between these two states occur more frequently in the parent generation (see table 1 and 2). We thus take into account the generation-specific prevalence of different family formation states in the cost specification. The principle of deriving substitution costs is the same for the parent and child generation, but the absolute values will differ depending on the generation-specific empirical distributions of family formation states.

The dyadic sequences are aligned as follows: the parent state of one dyad is aligned with the parent state of another dyad at the parent-specific substitution costs, and the child state is aligned with the child state using the child-specific substitution costs. For example, substituting the combined state [M1C M1C] with [M1C M2C] for two parent-child dyads would come at a cost of 1.82. In both dyads the parents are married with one child at this age, therefore there are no costs associated with aligning the parent component of the combined states. However in one dyad the child also is married and has one child (M1C) at this age, whereas the child in the other dyad is married and has two children (M2C). To align them we therefore substitute the states M1C with M2C at the child-specific substitution costs of these two states of 1.82 (see table 2). Substituting [M1C M1C] with [M2C M1C] comes at a cost of 1.66 and substituting [M1C M1C] with [M2C M2C] comes at a cost of 3.54. Intergenerational family formation sequences of parent-child dyads are thus considered most similar, if both parents follow the same trajectory, and both children follow the same trajectory. They are more distant, when either the parents are similar, or the children are similar, and most distant when both parents and children have very different family formation trajectories. The crucial point for finding patterns of 'opposite' parent child dyads is that similarity within the dyad, between the parent and the child, does not contribute to the calculation of distance between parent-child dyads.

To date, sequence analysis has mainly been used to compare individual sequences to one another in pair-wise comparisons. Recent approaches explore constructing sequence distances against an ideal type (Wiggins et al., 2007). We only know of two applications that explore sequence analysis for studying dyads. Brückner et al. (2006) compare adolescents' ideal romantic relationship sequences to their actual romantic relationship experience. Liefbroer and Elzinga (2008) construct distances between family formation of parent-child dyads. They compare parents to children within dyads, rather than treating the dyads as the units of analysis to be compared to each other, as in our approach. Overall, empirical studies exploiting the potential of sequence analysis for dyadic research are sparse. To our knowledge this is the first application of multichannel sequence analysis to dyadic sequences to study intergenerational continuities and discontinuities in life course processes.

RESULTS

All analyses were conducted using the TraMineR package for sequence analysis in R (Gabadinho et al., 2011). After calculating pair wise sequence distances by comparing every parent-child dyad with every other parent-child dyad as described above, the resulting distance matrix is entered into a ward cluster analysis. Figure 1 shows the dendogram to determine the number of clusters [other cluster cut-off criteria will be added]. The dendogram starts out with one line for each case at the bottom and shows how much additional distance, displayed on the y-axis, one has to allow for joining two dyads into a joint cluster. The cluster dendogram suggests a five cluster solution at the cut-off indicated by the dashed horizontal line. Each vertical line above the dashed horizontal line corresponds to one cluster. One would have to allow for a large jump in additional distance to further combine the clusters to four and three groups. Therefore, the dendogram suggests the five cluster solution as the best trade-off between accuracy and general pattern search.

FIGURE 1: Cluster Dendogram

Figure 2 shows the resulting five clusters as sequence index plots (Scherer, 2001). Sequence index plots display the order in sequences, usually time or age, on the x-axis.

The N of cases is displayed on the y-axis. Each horizontal line in the graph corresponds to one family formation trajectory. Different family formation states are indicated by different colors. Figure 2 shows the parent channel of the parent-child dyads on the left hand side, and child channel on the right hand side for each cluster. For both channels, the sequences are sorted according to the timing of first birth of the parent. The child line next to the parent line corresponds to the same dyad. Figure 3 shows modal state plots (Gabadinho et al., 2011) for each cluster that display the modal state, i.e. the most common state at each age. The height of the bars corresponds to the proportion of the cluster that experiences the respective family formation state at this age.

FIGURE 2: Sequence index plots (should be viewed in color)

FIGURE 3: Modal state plots

Figure 2 and 3 show that cluster 1, cluster 2 and cluster 4 reflect patterns of 'same' family formation of parents and children, once the cohort trend of the second demographic transition is taken into account.[1] In these three clusters, the children have on average one child less than their parents at age 40 and all focal transitions within the family formation process occur later. In contrast, cluster 3 and 5 show patterns where children have 'opposite' family formation trajectories than their parents. For cluster 3, the modal state in the parent channel is being married and having four children by age 32 (Figure 3). Quite the opposite, the most common family formation state for children of cluster 3 is being single and childless until age 39.

---

[1] For regression based analysis of transmission of fertility patterns Anderton et al. (1987) suggested to take into account cohort specific fertility pattern by using relative rather than absolute fertility as key indicator.

Note that the parents in cluster 2 and cluster 3 show a similar pattern of being married and having 4 children in their thirties. However, their off-spring shows very different family formation behavior: for cluster 2 there is transmission of relatively early marriage and having 2-3 children - intergenerational transmission adjusted for the cohort trend of the second demographic transition. In contrast for cluster 3, we find the opposite pattern of unmarried and childless children. Children who had many siblings themselves seem to cluster into two extremes in terms of their own family formation: either ‘same’ or ‘opposite’ to their parents. Cluster 4 and 5 show a similar pattern where the parents in these two clusters are similar in marrying early and having three children through most of their late twenties and thirties. The children in cluster 4 follow a ‘same’ cohort-adjusted pattern of being married and having two children throughout their thirties. In contrast, the children of cluster 5 follow an ‘opposite’ pattern of divorcing relatively early on, in their late twenties and early thirties.

DISCUSSION AND NEXT STEPS

The objective of this paper is to draw attention to regularities in ‘opposite’ patterns in parent-child family formation. The first results presented in this paper shows that the holistic multichannel sequence analysis approach is fruitful for establishing the existence of such patterns in the LSOG data. If children indeed polarize into either ‘same’ or ‘opposite’ patterns to their parents, traditional regression-based methods for studying these effects come with limitations. Since regression coefficients summarize averages across groups, they would be misleading in case of such a dichotomous structure into extremes. They would underestimate transmission for one group and overestimate transmission for another. As a result, the average coefficient would be far from correctly capturing the relationship between parents' family behavior on their children's for any real parent-child dyad. In this case, it might be more conducive to analyze what determines membership in the 'same' or 'opposite' group to advance our understanding of parent effects on family formation. Our findings suggest that the holistic pattern search approach of sequence analysis carries some promise for examining such regularities.

Next steps:

The obvious next step is to examine what determines, whether children follow a 'same' or 'opposite' pattern as their parents. We will examine education and collectivists and individualists values in the parent-child clusters next. To substantiate the patterns of 'same' and 'opposite' parent-child dyads we will further analyze the sequence distances by looking at: (1) distances within each parent cluster and within each child cluster, (2) distances between parents across clusters, (3) distances between children across clusters, and (4) distance between parents and children within each cluster.

REFERENCES

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

Abbott, A., and Tsay, A. (2000) Sequence Analysis and Optimal Matching Methods in Sociology. *Sociological Methods and Research* **29**(1), 3-33.

Abbott, A. (1995) Sequence Analysis: New Methods for Old Ideas. *Annual Review of Sociology* **21**, 93-113.

Amato, P. R. (1996). Explaining the Intergenerational Transmission of Divorce. *Journal of Marriage and the Family*, *58*(3), 628–640.

Aisenbrey, S., and Fasang A.E. (2010). New Life for Old Ideas: The "Second Wave" of Sequence Analysis. *Sociological Methods & Research* **38**, 420-462.

Bengtson, V. L., Biblarz, T. J., & Roberts, R. E. L. (2002). *How families still matter. A longitudinal study of youth in two generations*. Cambridge: Cambridge University Press; Cambridge Univ. Press.

Elzinga, C., & Liefbroer, A. (2007). De-standardization of Family-Life Trajectories of Young Adults: A Cross-National Comparison Using Sequence Analysis. *European Journal of Population*, *23*(3), 225–250.

Fasang, A.E. Retirement Trajectories and Income Inequality. *Social Forces*, forthcoming.

Feng, D., Giarrusso, R., Bengtson, V. L., & Frye, N. (1999). Intergenerational Transmission of Marital Quality and Marital Instability. *Journal of Marriage and the Family*, *61*(2), 451–463.

Gabadinho, A., Ritschard, G., Studer, M., and Müller, N. S. (2009) *Mining sequence data in R with the TraMineR package: A user's guide.* Technical report, Department of Econometrics and Laboratory of Demography, University of Geneva, Geneva, 2009b. URL http://mephisto.unige.ch/traminer/.

Gabadinho, A., Ritschard, G., Studer, M., and Müller, N. S. (2011a) Analyzing and Visualizing State Sequences in R with TraMineR. *Journal of Statistical Software* **40**(4), 1-37.

Gauthier, J.-A., Widmer, E. D., Bucher, P., and Notredame, C. (2010) Multichannel Sequence Analysis Applied to Social Science Data. *Sociological Methodology* **40**, 1-38.

Liefbroer, A. C. & Elzinga, C. H. (2008). *Intergenerational Transmission of Behavioral Patterns: Similarity of Parents' and Children's Family-Life Trajectories.*

Levine, J. H. (2000) But What Have You Done for Us Lately? Commentary on Abbott and Tsay. *Sociological Methods & Research* **29**(1), 34-40.

MacIndoe, H. and Abbott, A. (2004) Sequence Analysis and Optimal Matching Techniques for Social Science Data. In *Handbook of Data Analysis*, Bryman, A. and Hardy, M. (eds.) London: Sage Publications.

Pollock, G. (2007) Holistic trajectories: a study of combined employment, housing and family careers by using multiple sequence analysis. *Journal of the Royal Statistical Society: Series A* **170**, 167-138.

Scherer, S. (2001) Early Career Patterns: A Comparison between Great Britain and West Germany. *European Sociological Review* **17**, 119-114.

Stovel, K., Savage, M. and Bearman, P. (1996) Ascription into Achievements: Models of Career Systems at Lloyds Bank 1890-1970. *American Journal of Sociology,* **102**(2), 358-399.

Wiggins, R. Erzberger, C. Hyde, M., Higgs, P., and Blane, D. (2007). Optimal Matching Analysis Using Ideal Types to Describe the Life Course: An Illustration of How Histories of Work, Partnerships and Housing Relate to Quality of Life in Early Old Age. *International Journal of Social Research Methodology* 10(4), 259-278.

Wolfinger, N. H. (2000). Beyond the Intergenerational Transmission of Divorce: Do People Replicate the Patterns of Marital Instability They Grew Up With? *Journal of Family Issues*, *21*(8), 1061–1086.

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

von Poppel, F., Monden, C., & Mandemakers, K. (2008). Marriage Timing over the Generations. *Human Nature*, *19*(1), 7–22.

TABLES

TABLE 1: Transition rates between family formation states for parent and child generation

PARENT GENERATION									
	D1C	DNC	M1C	M2C	M3C	M4C	MNC	S1C	SNC
D1C	0.81	0.00	0.05	0.02	0.08	0.04	0.00	0.00	0.00
DNC	0.00	0.44	0.06	0.00	0.00	0.00	0.50	0.00	0.00
M1C	0.01	0.00	0.65	0.34	0.00	0.00	0.00	0.00	0.00
M2C	0.01	0.00	0.00	0.84	0.15	0.00	0.00	0.00	0.00
M3C	0.01	0.00	0.00	0.00	0.89	0.10	0.00	0.00	0.00
M4C	0.01	0.00	0.00	0.00	0.00	0.99	0.00	0.00	0.00
MNC	0.00	0.01	0.43	0.00	0.00	0.00	0.55	0.00	0.00
S1C	0.02	0.00	0.02	0.01	0.00	0.00	0.00	0.95	0.00
SNC	0.00	0.00	0.01	0.00	0.00	0.00	0.13	0.00	0.85
CHILD GENERATION									
D1C	0.91	0.00	0.04	0.04	0.01	0.00	0.00	0.00	0.00
DNC	0.03	0.87	0.01	0.00	0.00	0.00	0.09	0.00	0.00
M1C	0.04	0.00	0.77	0.18	0.00	0.00	0.00	0.00	0.00
M2C	0.03	0.00	0.00	0.93	0.05	0.00	0.00	0.00	0.00
M3C	0.03	0.00	0.00	0.00	0.91	0.06	0.00	0.00	0.00
M4C	0.02	0.00	0.00	0.00	0.00	0.98	0.00	0.00	0.00
MNC	0.00	0.03	0.19	0.00	0.00	0.00	0.78	0.00	0.00
S1C	0.00	0.00	0.08	0.02	0.00	0.00	0.00	0.90	0.00
SNC	0.00	0.00	0.00	0.00	0.00	0.00	0.09	0.00	0.90

TABLE 2: Substitution costs based on transition rates for parent and child generation

PARENT GENERATION									
	D1C	DNC	M1C	M2C	M3C	M4C	MNC	S1C	SNC
D1C	0.00	2.00	1.93	1.97	1.91	1.96	2.00	1.98	2.00
DNC	2.00	0.00	1.94	2.00	2.00	2.00	1.49	2.00	2.00
M1C	1.93	1.94	0.00	1.66	2.00	2.00	1.57	1.98	1.99
M2C	1.97	2.00	1.66	0.00	1.85	2.00	2.00	1.99	2.00
M3C	1.91	2.00	2.00	1.85	0.00	1.90	2.00	2.00	2.00
M4C	1.96	2.00	2.00	2.00	1.90	0.00	2.00	2.00	2.00
MNC	2.00	1.49	1.57	2.00	2.00	2.00	0.00	2.00	1.87
S1C	1.98	2.00	1.98	1.99	2.00	2.00	2.00	0.00	2.00
SNC	2.00	2.00	1.99	2.00	2.00	2.00	1.87	2.00	0.00
CHILD GENERATION									
D1C	0.00	1.97	1.92	1.94	1.96	1.97	2.00	2.00	2.00
DNC	1.97	0.00	1.99	2.00	2.00	2.00	1.88	2.00	2.00
M1C	1.92	1.99	0.00	1.82	2.00	2.00	1.81	1.92	2.00
M2C	1.94	2.00	1.82	0.00	1.95	2.00	2.00	1.98	2.00
M3C	1.96	2.00	2.00	1.95	0.00	1.94	2.00	2.00	2.00
M4C	1.97	2.00	2.00	2.00	1.94	0.00	2.00	2.00	2.00
MNC	2.00	1.88	1.81	2.00	2.00	2.00	0.00	2.00	1.91
S1C	2.00	2.00	1.92	1.98	2.00	2.00	2.00	0.00	2.00
SNC	2.00	2.00	2.00	2.00	2.00	2.00	1.91	2.00	0.00

FIGURES

FIGURE 1: Cluster Dendogram

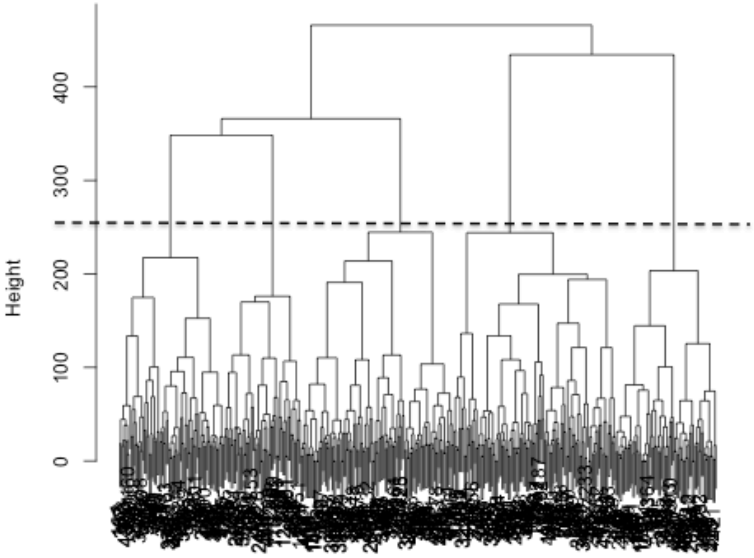

FIGURE 2: Sequence index plots of clusters, sorted by parent's age of first birth (view in color)

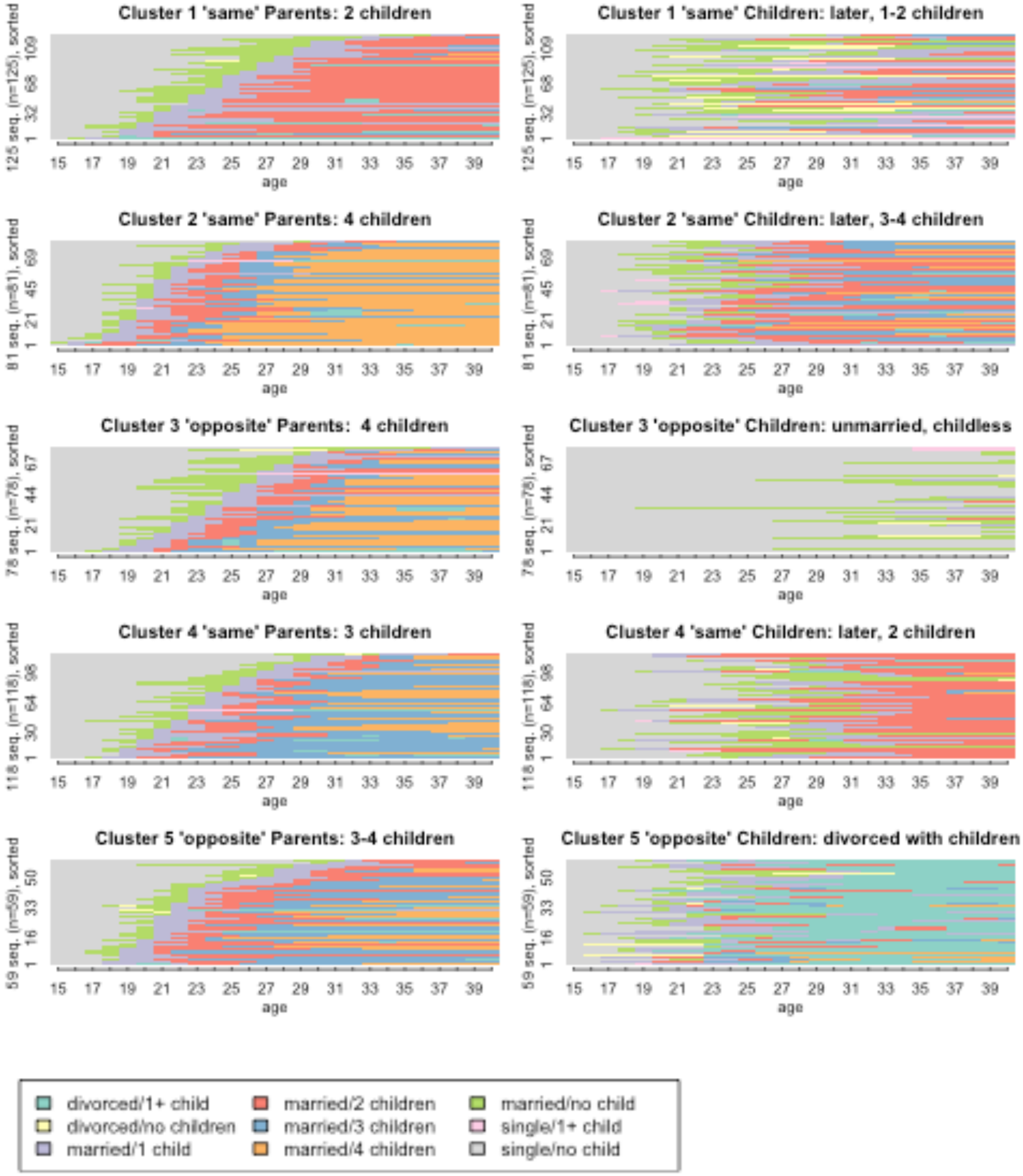

FIGURE 3: Modal state plots of clusters