

# The roads more or less traveled

A sequence analysis of family  
formation and parenthood for a cohort  
of Danish women born in the 1970s

BY TINNE STEFFENSEN

---

## ABSTRACT

Relatively low fertility and an increased age at first birth, along with the development of assisted reproduction technologies have increased attention to when and how many times Danish women give birth. While some argue that family formation has become increasingly plural and differentiated, others maintain that the nuclear family remains the ideal family for the majority of women. In this article, I investigate family formation trajectories for a random sample of 1,500 women born in 1973 and 1974. For this sample, I perform sequence analysis of longitudinal registry data on civil status, fertility, education and income through the ages 22 to 37. Focusing on timing, order and duration in the sequences studied, I identify seven distinct clusters (i.e. typologies) of family formations in Denmark. The majority (68 percent) of the women's trajectories represent varieties of the nuclear family. For all clusters, my results confirm the event of the first child as a constituting factor of the nuclear family, which often precedes marriage. However, the identified clusters also show great variation when it comes to age at birth of first child, socio-economic status and overall turbulence in their trajectories.

## KEYWORDS

*Fertility, family formation, sequence analysis, life course, longitudinal data, registry-based research.*

Tinne Steffensen holds a master in Sociology from University of Copenhagen. She currently works for the Centre for Public Innovation, and has previously worked on research in the field of disability and health. Her interest lies in longitudinal data, gender, health and demography.

The low total fertility rate in Denmark (1.7 in 2016) and most of Europe, together with the outlook to a growing dependency ratio between the working young and the older retired population, have contributed to an intensified public debate about fertility and family formation. As a result of the progress in fertility treatment and the fact that Denmark ranks among the highest in use of fertility treatment in Europe, doctors and the media have called infertility a “wide-spread disease” that will be costly for society at large (Nielsen 2015; Nielsen et al. 2016). In 2015, the public Danish Broadcasting Corporation (DR TV) aired a whole night ‘talk show special’ with the title “Bonk for Denmark” (In Danish: “Knald for Danmark”) urging young couples to have (more) children earlier. The same year Copenhagen Municipality, where the average age at first birth is two years older than the national average, and Rigshospitalet (largest hospital in Denmark) ran a much disputed campaign targeting young women and men in Copenhagen with billboards carrying the questions “Have you counted your eggs today?” and “Are they still swimming” reminding men and women that chances of getting pregnant decreases significantly from age 25 to age 35.

Often dominating in the public debate are the questions of whether men and especially women have ‘enough’ children, at the ‘right time’ and in the ‘right way’. Because of the relative nature of these factors and the complexity of family formation as a process, empirical investigations and thick descriptions are needed in order to balance the often-conflicting messages in popular media.

In this article, I argue that the application of sequence analysis, a pattern searching technique for life course trajectories, is fruitful for exploring family formation processes in a Danish context. Although se-

quence analysis is a known tool for demographic analysis of family formation, the method has yet to be applied on the life course trajectories of family formation in Denmark. The method is a ‘thick description’ method that was born out of a critique of some of the more causal quantitative methods and Andrew Abbott’s belief that “Before we can explain, we must describe” (Abbott 2001: 120). Thus, sequence analysis aims at finding varieties of patterns and processes in the data, while not making too many assumptions about the data.

I explore timing, order and duration in family formation through the ages 22-37 for a cohort of Danish women born in 1973/1974. In order to explore the particularly complex process of family formation, I ask: What constitutes women’s family formation trajectories and which observable factors are associated with the roads more or less traveled?

I will first introduce the theoretical background for looking at family formation sequences and shortly introduce the Danish welfare model context. After a short introduction to the method of sequence analysis, a presentation of the data follows before the main results are presented and discussed.

### SEQUENTIAL PROMISCUITY?

Historically demography has been on and off the public agenda, however with the rising aging population, low fertility in most of Europe, the US and Eastern Asia and the simultaneous and conflicting global overpopulation, demography and fertility have for definite reentered the public arena as both a worldwide and national concern.

In the 1980s Dick Van de Kaa and Ron Lesthaeghe coined the now central concept in modern demography and in studies of family formation in Europe - the Second

Demographic Transition. They used the term to describe the vital demographic changes in Europe: decreasing fertility, changes towards a more egalitarian and plural view on family, a rise in cohabitation and divorce as well as the widespread use of contraceptives (Lesthaeghe 2010). Although the term has received criticism for its lack of emphasis on the effects of globalization, two central consequences of the demographic changes described in the second demographic transition remains: The road to becoming a family and ways of being a family have changed, as has the concept of choice when it comes to becoming a parent.

Firstly, not only are the ways of living as a family diversifying or *pluralizing*, demographers also argue that family formation is increasingly *de-standardized*, defined as the process in which specific life courses, events and sequences are experienced by an increasingly smaller proportion of the population or occur at more scattered ages and with more scattered durations (Bruckner & Mayer 2005). As a result, the order of the past is no longer set and the dominance of a 'standard' or 'normal' life trajectory has become weakened (ibid.: 32).

However, the picture of a state of fragmentation and chaos is challenged by empirical studies that hold a more critical view upon the de-standardization thesis and argue that the level of de-standardization is not as high as often assumed. Elchardus (2006) and Thomson et al. (2013) find that in spite of vital changes in the diversity of life trajectories today, a standard life course remains a life-goal for the majority. Few studies are done specifically on a Scandinavian or Danish context and the discussion on the degree of de-standardization is challenged by the need for thorough empirical evidence and the relative nature of processes such as de-standardization and pluralisation. Elzinga and Liefbroer (2007) suggest a measurement of turbulence when looking at trajectories, a concept that seeks

to measure the 'chaos' by accounting for the number of distinct sequences found in the trajectory and the variation of time or period spent in the distinct states. In a study of 19 European countries they find that although younger generations' life trajectories are less standardized and more pluralized, the younger generations' family trajectories are *not* more turbulent – thus dominated by more and quicker changes over time.

Secondly, the second demographic transition includes a change in attitude towards having a child. Neither pregnancy nor infertility is accepted as a destiny, and both are to some extent a *choice* that can be made and unmade. However, as Knudsen and Wielandt (1995) warn: The greater the possibilities are for a negative control of fertility, the greater are the expectancies that it is also possible to practice a positive control of fertility. It is unknown exactly how many women and men who are involuntarily childless as a cause of infertility. Schmidt (2006) estimates that 16-26% of Danish couples will experience periods of infertility, defined in line with the World Health Organization as "a disease of the reproductive system defined by the failure to achieve a clinical pregnancy after 12 months or more months of regular unprotected sexual intercourse" (Zegers-Hocschild et al. 2009: 1522). To add to the question of choice and control over fertility, research based on a representative sample of pregnant women at a hospital in Denmark found that of the 2,611 pregnancies observed 29% could be categorized as not planned, but accepted, so called 'accepting non-planners' (Rasch et al. 2001), showing that although there are many options to control your own fertility, having a child might not always be planned in the first place.

However, family formation cannot be limited to a question of choice and control, it unfolds over time and is always embedded in the socio-cultural and economic context. Thus, before we look at how fami-

ly formation trajectories unfold, a short contextualisation of Danish Welfare state is needed. The Danish welfare model provides free and universal access to health care and family policies are extensive. All together parents in Denmark receive 52-weeks of paid parental leave and subsidized day care options, and the dual worker model is dominating with 68% of women in the fertile age (16-49 years old) taking part in the workforce in 2014 (almost identical rates to the 69% of the men) (StatBank Denmark 2017).

Together with the plurality of family forms and the subsidized fertility treatment for single women and lesbian couples (adopted in 2007), Denmark makes for an interesting case for looking at trajectories in family formation. In 2015, an estimated 8% of babies born in Denmark were born as a result of fertility treatment (Sundhedsdatastyrelsen 2015), which places Denmark among the highest-ranking fertility treatment providers in Europe.

In Denmark, the first child often precedes marriage by a couple of years. In 2015, the average age at first child was 29 for women and 31 for men, while the average age at first marriage was 32 for women and 34.5 for men (StatBank Denmark 2017). Consequently, the birth of the first child has replaced the tradition of marriage as the constituting factor of family formation in Denmark. In 2011, 20,5% of 49-year old men and 13,6% of 49-year-old women did not have any children (Statistics Denmark 2011, 11) with a higher proportion of childlessness among higher educated women and lower educated men. In line with this, research suggests that childless people overall are a heterogeneous group who nevertheless still face taboos when taking the road less travelled (Hagestad & Call 2007).

I argue that in order to deal holistically with the heterogeneity of life course trajectories, it's necessary to look beyond cross-sectional data's focus on set points in time

and look closer at sequence patterns and order, timing and duration of events. This will help us better understand to what extent people go down the same path and who and when others depart from it, thus nuancing both the media's sometimes gloomy headlines and the question of choice and planning when it comes to family formation.

#### DESCRIBING AND EXPLORING SEQUENCE ANALYSIS

With the current growth in availability of data, and with the interest in and power of 'big data', pattern-searching techniques and visualization tools have become important tools to make sense of data. Additionally, sociological theory has in the last two decades seen a strong theoretical argument for *thick description* as something equally deep and fruitful as the often-preferred (more causal) *explanation*. As Bruno Latour puts it: "No scholar should find humiliating the task of sticking to description. This is, on the contrary, the highest and rarest achievement" (Latour 2005: 136) If we want to approach the questions of overall low fertility or timing and number of children, a thorough understanding of family formation patterns must first be approached.

Sequence analysis is a pattern searching method used to study ordered or sequenced events over time, especially useful for exploring life course trajectories (e.g. career, retirement, family formation). In spite of its advantageous use for studying family formation, the method has yet to be used on Danish data on family trajectories. By calculating dissimilarities between sequences, sequence analysis assigns individuals into groups or typologies of e.g. common pathways into parenthood, according to the proximity of their trajectories to others.

The method is based on some of the same methods used in DNA analysis, but developed and transferred to the field of

sociology by sociologist Andrew Abbott in the 1980s. Abbott critiqued sociology for abandoning curious and exploratory analysis in favor of too dominant belief in the technically advanced casual variable analysis; the eagerness to isolate the effects of a specific X on Y. Instead, Abbott argued for a more explorative approach, where trajectories are explored without a direct hypothesis or assumption about the data while context and process take the center stage: “Every social fact is situated, surrounded by other contextual facts and brought into being by a process relating it to other contexts” (Abbott 1997: 1152).

In spite of the holistic ‘trajectory’ dominating the *theory* of life course research, *applied* life course research remains dominated by the more specific ‘transition’ or single ‘event’. Thus, as Fasang and Aisebrey (2010) point out: Sequence analysis has the potential to “bring the trajectory, the actual ‘course’, back into research on the life course”. Sequence analysis stands in opposition to other time-dependent tools such as regression models and its extensions (event history analysis) that aim at modeling the specific likelihood of a transition or event occurring at a given time. This is not to say the different methods cannot be combined, however, I will argue that a thorough exploration of sequences is crucial before seeking more causal explanations and testing hypotheses (Lesnard 2010).

The main critique of sequence analysis remains to be sequence analysis’ descriptive nature and the methods’ early use of the Optimal Matching technique. Optimal Matching is an algorithm that calculates the cost of turning two strings or sequences of categories into the same, thus creating a dissimilarity measure for how alike two sequences are. In the end, the ‘lowest distance’ measure, the costs of transforming two sequences into the same, will help decide the numbers of groups that sequences should be sorted into. Traditionally, the

distance measures have to some extent been based on the researcher deciding on a numerical cost of either inserting or deleting a state (indel cost) or substituting a state with another (substitution cost). With advancement in computer programming however, the computation of distance matrices has become more accessible and the method has seen improved technical implementations during what Aisenbrey and Fasang (2010) call the ‘second wave’ of sequence analysis (post 2000). The ‘second wave’ mainly seeks to counter the critique of the abstract cost setting by developing new methods that are sensitive to the data and no longer an arbitrary choice by the researcher. I will make use of the Dynamic Hamming Distance measure, a part of the ‘second wave’, which will be introduced after a short introduction of the data and the central steps in Sequence Analysis.

## DATA

The data consists of a random sample of 1,500 Danish women taken from the COMPI DANAC - Danish National ART (Assisted Reproductive Technology) - Couple Cohort (Schmidt et al. 2013). COMPI DANAC includes registry-based information on partner status, fertility, education, income, employment and hospital visits for 42,915 couples who were registered with ART treatment in the Danish IVF registries in 1994-2009 as well as a representative sample of 215,000 couples in an age-matched comparison population.

In spite of the large population size of COMPI DANAC, I limit my data to a random sample of 1,500 women.<sup>1</sup> This is necessary for two reasons: a) to calculate the matrix distances for the sequences in a statistical program and b) for the visualization of sequences to be fruitful. The random sample of 1,500 women is representative and does not differ significantly from the actual population when it comes to central factors related to fertility as well as other

TABLE 1: AGE AND TIME OF SAMPLE

|             |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |     |      |
|-------------|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|-----|------|
| Cohort 1973 | 22 | 23 | 24 | 25 | 26 | 27 | 28 | 29 | 30 | 31 | 32 | 33 | 34 | 35 | 36 | 37 | Age |      |
| Cohort 1974 | 22 | 23 | 24 | 25 | 26 | 27 | 28 | 29 | 30 | 31 | 32 | 33 | 34 | 35 | 36 | 37 |     |      |
|             | 95 | 96 | 97 | 98 | 99 | 00 | 01 | 02 | 03 | 04 | 05 | 06 | 07 | 08 | 09 | 10 | 11  | Year |

important factors such as income, education and residence (tested with a chi-square test). The sample is thus highly representative of the actual population on all observable factors.

Information on fertility is from the Danish Fertility Database (FTDB) and the IVF registry. The FTDB covers fertility information from 1980 and on for the population of women and men who are permanently based in Denmark and who are in the fertile age of 13-49 (women) and 13-64 (men). FTDB is from the national population database and hence the “coverage is considered almost complete” (Blenstrup & Knudsen 2011: 81). The IVF registry covers information from 1994 to today as well, and it is compulsory to report to for both public and private clinics. The IVF registry covers all initiated IVF treatments for women with Danish registration (CPR number). Information from other national registries such as age, civil status, income, education, and city of residence are included from their respective registries (Statistics Denmark 2017).

Although the coverage and quality of Danish registry data must be emphasized, I note that the registries do not cover any subjective measures usually obtained by surveys or interviews. Hence, this article does not cover individual attitudes or decision-making when it comes to having a child. For a thorough qualitative biographical analysis of attitudes toward family formation for young adults in Denmark, I refer to Ottosen and Mouritzen (2013).

In order to take full advantage of the longitudinal data, I follow the cohort of

women born in 1973 and 1974 and their family status in the years 1995-2011 covering the age of 22 to 37 years. This covers most of the fertile age when babies in Denmark are born, in 1995-2011, 89,5% of all children were born by a mother between the age 22 and 37.<sup>2</sup> It is quite unique to be able to create and follow trajectories of this length with high quality registry information without the worries and problems of attrition and non-response.

A list of the included variables is shown in table 2 below.

PRESENTATION OF SEQUENCES

In order to examine the unfolding of family formation, I will shortly introduce the central structure and descriptive characteristics of the sequences. First, I will introduce the categorical variable (i.e. *the sequence alphabet*) that, together with age, forms the sequences. After this, I will introduce the Dynamic Hamming Distance measure that I use to calculate distance matrices between sequences before creating clusters and the typologies of trajectories are introduced.

The sequence alphabet is made up of six different *states* (categories) of family status, (from now on: *family status*), and is based on registry information on cohabitation and fertility (see variable list in table 2). I am primarily interested in when and if women have their first child and whether or not they have a partner with whom they are cohabiting. I do not distinguish between having one or more children or between being single or divorced, since I see little conceptual difference between the

TABLE 2: VARIABLE LIST

| Variable name                    | Description  | Values     |
|----------------------------------|--|------------|
| Age                              | Age in years   | 22;37      |
| Family status                    | <ul style="list-style-type: none"> <li>· Single, no child: Does not reside with a partner (defined by Statistics Denmark as a man that is less than 14 years younger or older); does not have a registered child in fertility database (FTDB).</li> <li>· Single, with child: Does not reside with a partner (defined by Statistics Denmark as a man that is less than 14 years younger or older); has a registered child in FTDB</li> <li>· Cohabitation, no child: Resides with a partner (defined by Statistics Denmark as a man that is less than 14 years younger or older); does not have a registered child in FTDB</li> <li>· Cohabitation, with child: Resides with a partner (defined by Statistics Denmark as a man that is less than 14 years younger or older); has a registered child in FTDB</li> <li>· Married, no child: Is registered in a marriage; does not have a registered child in FTDB</li> <li>· Married, with child: Is registered in a marriage; has a registered child in FTDB</li> </ul> | 1:6        |
| Income                           | Personal yearly income in total Danish kroner excluding rent value of residence and before deduction of interest rate expenses   | DKK        |
| Higher university degree         | Holds a BA, MA or Ph.D. degree   | 0;1        |
| Loose attachment to labor market | Not employed for 3 years or more in 2002-2009  | 0;1        |
| Divorce                          | Had a divorce at some point between the age 22-37  | 0;1        |
| Main partner                     | Partner with longest period of cohabitation in the age 22-37   | Identified |
| IVF treatment                    | Is registered in the IVF registry at some time in the period 1995-2009   | 0;1        |
| Turbulence                       | Degree of turbulence in sequences. Calculated in R according to Elzinga's (2006) quantification of number distinct states, number of transitions and variance of duration spend in each state. Turbulence is calculated both on an individual level and as the mean for each cluster.  | 0-10,4     |
| Frequent hospital visits         | Is registered in the Danish National Patient Registry (LPR) in 6 or more years during the period 1995-2009   | 0;1        |

two states. Additionally, the creation of too many categories in the sequence alphabet will result in a myriad of possible family states and decrease the clarity of patterns and clusters without adding to the overall analysis.

Furthermore, I make no distinction between whether the woman lives with the father of her child or with another partner since I am merely interested in whether the household that provide the primary socialization for the child is a two-parent household or not. Unfortunately, a strong shortcoming of the data registries is that same-sex couples are impossible to identify before their partnership is formalized by marriage (or civil union). As a consequence, any women taking part in un-married same-sex relationships will be categorized as single in the analysis. As a result, it is not possible to say much about same-sex family formation from the results presented here. In total, there are only eight of the 1,500 women who at one point are in a same-sex marriage. In spite of the quality of registry data, there is no way to observe if people in the single-category have a partner with whom they do not live.

The sequence alphabet of family status in this analysis is: S = Single, SWC = Single, with child, C = Cohabitation, CWC = Cohabitation, with child, M = Married, MWC = Married, with child.

Figure 1 shows an example of how a typical trajectory *can* look like: Woman 1 follows a non-turbulent pattern of being single for four years, entering into a cohabitating relationship at 26, and getting married the same year as she gives birth to her first child at 29. Using the sequence alphabet her trajectory is as follows: S4, C4, MWC8.

## MATCHING AND CLUSTERING

To pair sequences and assign them into clusters I apply the Dynamic Hamming Distance (DHD) measure that calculates

time sensitive substitution costs based on the data (Lesnard 2010). The strength of the DHD measure is that the substitution costs, the cost of turning two different states in two different sequences into the same, depend on the dynamic change of events over time found in the data. Put in another way, the substitution costs depend on the point in time at which two sequences differ in states and the data-based proportional transition frequencies. The advantage for the case of analyzing family formation is apparent: The timing of transition has social meaning, e.g. being a married woman with a child and transitioning to a single mother at the age of 22 will most likely be different than making the transition at 37 where more women go through the same experience and the child is also significantly older.

In this way, DHD allows to take the social reality of the data into account while at the same time making no clear assumptions about the data and belongs to the ‘second wave’ of sequence analysis that counters the critique of optimal matching’s abstract cost setting being too detached from the sociological understanding of events.

## CLUSTERING

On the basis of Ward’s hierarchical clustering, I have created *seven* clusters of family formation patterns for the women in cohort 1973/1974. Ward’s hierarchical clustering is a graphical presentation of the calculated distances between groups in data (see figure 2 on page 56). The tree-shaped diagram below (called dendrogram) shows the arrangement of potential clusters. The two extremes are to have only two clusters (top) or one cluster for each observation (bottom). I chose seven clusters since the length of the vertical lines (on the dissimilarity axis in figure 2) decreases drastically after this point and hence there is less of a need for separate clusters.<sup>3</sup> The bold hori-



zontal line represents the ‘cutting line’ of further numbers of clusters.

## RESULTS

Figure 3 on page 57 shows the proportion of women in the different family status states for each year, for the cohort 1973/1974 through the age 22-37. Up until the age of 29, the most likely status is to be single, after which it becomes most likely to be married and have at least one child.

The state distribution plot tells us something about the overall picture, but does not show individual sequences, when most transitions take place or how turbulent their trajectories are - e.g. even though it is most common to be cohabiting and have one child *before* getting married, cohabitation with one child is never the most common state. The individual sequences are shown in the sequence index plot (figure 4, page 57).

The data supports the notion of the first child as the establishing factor of the family replacing the institution of marriage (illustrated with the small proportion of women in in the “married” category - blue line in figure 3). The two most common specific order of sequences are: Single -> Cohabitation -> Cohabitation with child -> Married with child (6%) *or* Cohabitation -> Cohabitation with child -> Married with child (6%). However, in total only 12% follow this most common path without any abruptions.

In figure 4, all 1,500 women are represented with each their horizontal line showing their trajectory from the age of 22-37. The Y-axis shows each id-number from 1-1,500 representing each woman in the data. Many colors over time for each line (horizontally) signify that the woman undergo many transitions. The visualization is somewhat messy, yet it is clear to see the general trajectory going from ‘single’ / ‘cohabitation’ / ‘cohabitation, with child’ towards ‘married, with child’ and that most

women undergo several transitions. However, the unsorted index plot fails to adequately depict “minority states”, e.g. the substantial part of women who are single mothers periodically in their mid-30s (represented by the light purple color, figure 3), and the index plot needs an overall sorting into clusters for a clearer picture to emerge.

The individuals are grouped into clusters using the DHD measure. Seven different trajectory patterns that describe Danish women’s family formation for the cohort 1973/1974 are found. I assign each cluster a name based on some of the central characteristics of trajectories (e.g. timing of first child, timing of marriage, whether married, relationship history, etc.). I use the term ‘nuclear family’ as a description for clusters where a majority of the women either for a majority of the time or towards the end of the time period traced live in a household made up of a mother, her child(ren) and the father or live in a household made up of a mother, her child(ren) and the mother’s partner.

I name the seven clusters: 1) Modern nuclear family, 2) Early Nuclear family, 3) Early nuclear family, delayed marriage, 4) Late nuclear family, non-married, 5) Single, no child, 6) Single mothers and 7) Cohabiting or married, no child. The state distribution plots for the clusters are shown in figure 5. The size of the clusters is shown in their titles, the percentage of the overall sample of 1,500 women that are represented in the clusters.

The different variations of a nuclear family (clusters 1-4) account for all in all 68% of the women. For a large majority, the women in these clusters end up in cohabitation or marriage with at least one child at the age of 37. However, while the women in the *modern nuclear families* start forming families at around 30, the women in the *early nuclear families* start several years earlier. There are also clear differences in socio-economic and geographical back-

FIGURE 1: EXAMPLE OF A SEQUENCE

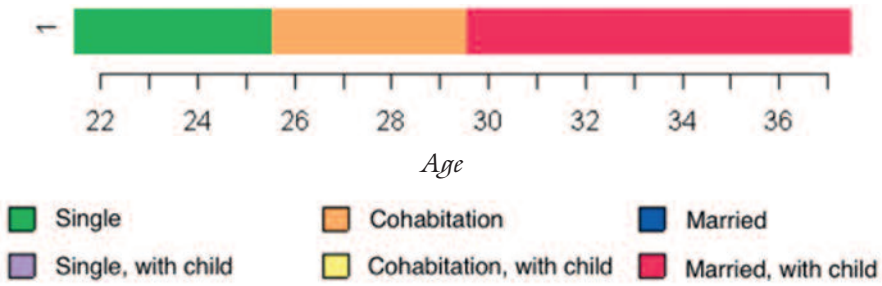
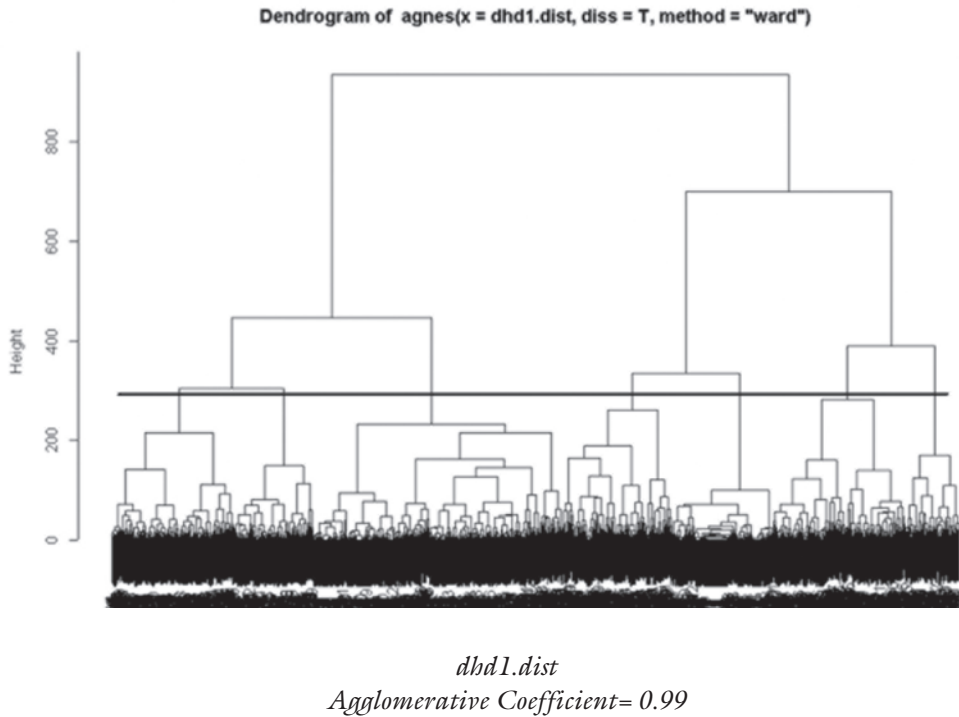


FIGURE 2: WARD DENDROGRAM FOR HIERARCHICAL CLUSTERING



grounds and the women’s life trajectories are shown in table 3. The differences and similarities will be further explored in the presentation of clusters.

The last three clusters make up a total of 32% of the observed sequences. The

women’s trajectory patterns are very different, however, they do have one thing in common: Their trajectories do not fall into varieties of the *nuclear family* category as the women have no or sparse periods of living in a cohabiting or married relationship

FIGURE 3: STATE DISTRIBUTION PLOT FOR FAMILY STATUS 22-37 YEARS OLD

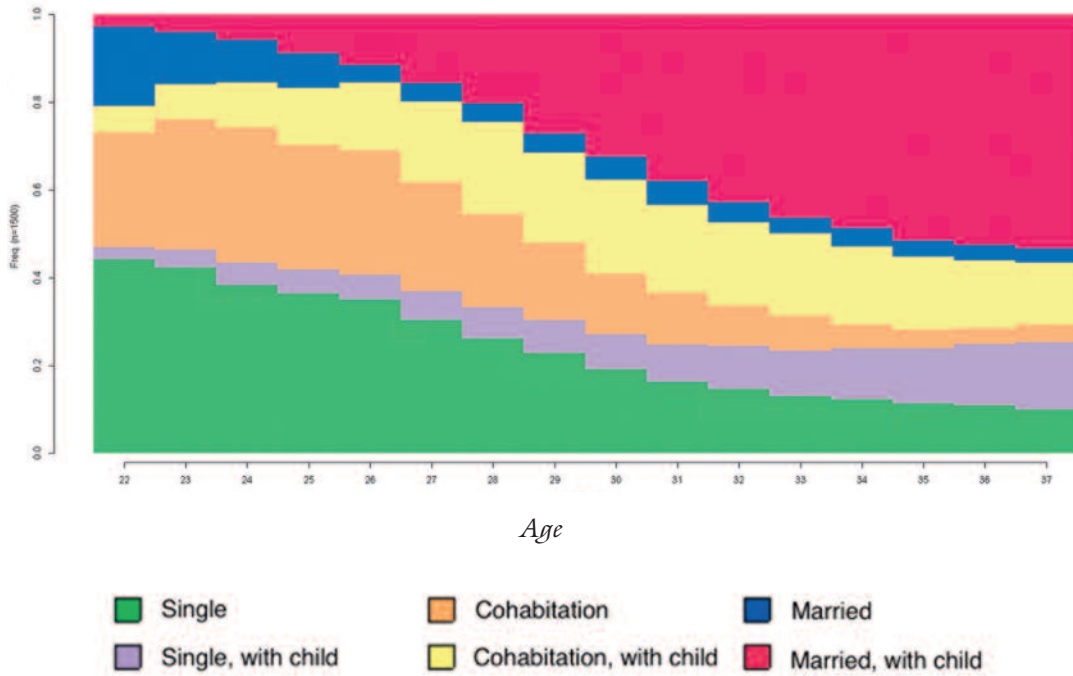
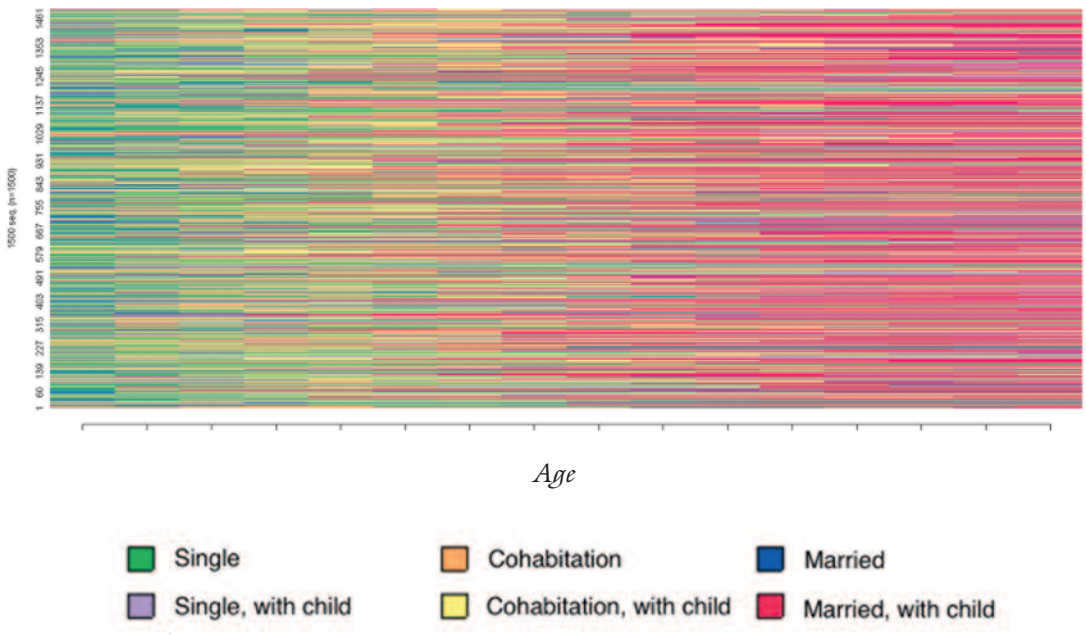


FIGURE 4: SEQUENCE INDEX PLOT FOR FAMILY STATUS 22-37-YEARS OLD



as a parent. Like the first four groups there are also clear socio-economic and geographical differences both within the three clusters and between the *nuclear family* clusters.

The clustered index plot (figure 6 page 60) shows individual sequences for all 1,500 women in their assigned clusters. As in figure 4, each horizontal line represents the trajectory of one woman (identification number on the y-axis) through the age 22-37 (x-axis). The clusters show a relative high difference in homogeneity and common trajectories (e.g. cluster 5 shows high homogeneity and stable sequences, while cluster 6 show a low homogeneity and a higher degree of turbulence).

The ten most frequent sequences for each cluster in figure 7 further exemplifies the overall differences within clusters. While the ten most frequent sequences for the *single mothers* in cluster 6 only describes 5,6% of the women's overall trajectories, the ten most frequent sequences describe an overall 39% of the women who are *single, without child*. The coverage level is shown on the top of the y-axis.

From the results in figure 5-7 and the central characteristics shown in table 3, I will now describe the seven typologies of family formation in more detail.

#### CLUSTER 1: MODERN NUCLEAR FAMILY (27%)

These 'modern' women make up the biggest group. Most live a single-life going in and out of a couple of relationships until their late twenties, after which they enter a new relationship and give birth to their first child at around the age of 30. A large majority marries shortly after becoming parents: 42% are married in the year of their first child, which is more than double of the population in general (18%). The women in this group are in general highly educated. At the age of 37, more than half of the women have a higher university degree.

In this cluster, the women have the second highest rate of IVF treatment (12%), about double the population average. Together with the cluster *single mothers*, the *modern nuclear family* women have the most turbulent trajectories, however, the index plot (figure 6) shows that for the *modern women*, turbulence is mostly concentrated around the age 22-30. Having a higher education increases turbulence significantly and so does living in Copenhagen or Aarhus even when education is taken into account. The big city effect can be a form of 'neighborhood effect' of being exposed to a greater plurality of family forms, greater number of potential partners and whatever else the bigger cities have to offer.

#### CLUSTER 2: EARLY NUCLEAR FAMILY (19%)

The second biggest cluster is made up of women who follow more traditional family trajectories. The women only have a few years as singles in their early twenties and become mothers pretty early, on average at the age of 25. At the age of 30, almost *all* the women in this cluster are married and have at least one child, and a notable 94% are together with the father of the child at 37. In general, these women have stable, non-turbulent trajectories and the group has a high degree of homogeneity in trajectories. Only one in four go through a divorce at some point, mostly following a quite early marriage in their early twenties. Regionally the women differ from the average and especially the *modern nuclear family*: Nine out of ten of the women live outside Copenhagen or Aarhus, well below the population average. This group represents the most traditional family trajectory, but there is no carry over of the housewife role, which is to be expected, since the women are born right in the middle of the largest entry of women into education and the workforce.<sup>4</sup>

TABLE 3: CENTRAL CHARACTERISTICS OF THE SEVEN FAMILY FORMATION CLUSTERS

|   | Cluster 1: Modern Nuclear Family | Cluster 2: Early Nuclear Family | Cluster 3: Early Nuclear Family | Cluster 4: Modern Nuclear Family, delayed marriage | Cluster 5: Single, no child | Cluster 6: Single mothers no child | Cluster 7: Cohabiting or married, | Total population |
|---|----------------------------------|---------------------------------|---------------------------------|--|-----------------------------|------------------------------------|-----------------------------------|------------------|
| Cluster %   | 27 %                             | 19 %                            | 12 %                            | 10 %   | 14 %                        | 12 %                               | 7 %                               | -                |
| Child by age 37   | 100 %                            | 100 %                           | 100 %                           | 100 %  | 18 %                        | 100 %                              | 17 %                              | 83 %             |
| Age at first birth (avg.)   | 30<br>(2,2)                      | 25<br>(1,8)                     | 25<br>(2,2)                     | 31<br>(2,4)  | 35<br>(1,1)                 | 26<br>(2,8)                        | 36<br>(1,1)                       | 28<br>(3,6)      |
| Partner's age at year of first birth  | 32<br>(3,8)                      | 29<br>(3,6)                     | 29<br>(4)                       | 34<br>(4)  | 36<br>(5,6)                 | 29<br>(4)                          | 37<br>(5)                         | 31<br>(4,4)      |
| Single in year of first birth   | 8 %                              | 6 %                             | 20 %                            | 18 %   | 46 %                        | 33 %                               | 16 %                              | 15 %             |
| Has received IVF-treatment  | 12 %                             | 3 %                             | 0 %                             | 6 %  | 3 %                         | X*                                 | 24 %                              | 7 %              |
| Is or has been married by age 37  | 100 %                            | 100 %                           | 68 %                            | 43 %   | 43 %                        | 62 %                               | 78 %                              | 77 %             |
| Is or has been divorces by age 37   | 40 %                             | 25 %                            | 23 %                            | 26 %   | 38 %                        | 59 %                               | 45 %                              | 36 %             |
| Number of partners (avg.)   | 1,4<br>(0,6)                     | 1,2<br>(0,5)                    | 1,6<br>(0,8)                    | 1,6<br>(0,9)                                       | 1,6<br>(0,6)                | 2,2<br>(1)                         | 1,6<br>(0,9)                      | 1,5<br>(0,7)     |
| Main partner has child from previous relationship                               | 7 %                              | 6 %                             | 11 %                            | 17 %   | 37 %                        | 15 %                               | 23 %                              | 14 %             |
| Together with father of child at age 37   | 73 %                             | 94 %                            | 74 %                            | 85 %   | 93 %                        | 92 %                               | 25 %                              | 81 %             |
| Loose attachment to labor market  | 10 %                             | 12 %                            | 23 %                            | 14 %   | 12 %                        | 37 %                               | 11 %                              | 16 %             |
| Frequent hospital visits<br>(Observed for 6+ years in LPR)                      | 15 %                             | 17 %                            | 21 %                            | 8 %  | 9 %                         | 31 %                               | 14 %                              | 17 %             |
| Yearly income in 1,000 DKK at age 35  | 322<br>(137)                     | 308<br>(117)                    | 301<br>(108)                    | 311<br>(108)                                       | 314<br>(224)                | 288<br>(87)                        | 314<br>(210)                      | 310<br>(145)     |
| Higher university degree at age 37  | 51 %                             | 27 %                            | 24 %                            | 45 %   | 41 %                        | 23 %                               | 31 %                              | 37 %             |
| Higher university degree in last observed year of relationship                  | 49 %                             | 26 %                            | 20 %                            | 47 %   | 25 %                        | 15 %                               | 32 %                              | 31 %             |
| Higher university degree for main partner in last observed year of relationship | 34 %                             | 19 %                            | 21 %                            | 33 %   | 30 %                        | 4 %                                | 18 %                              | 26 %             |
| Grew up in Copenhagen or Aarhus   | 6 %                              | 10 %                            | 11 %                            | 8 %  | 7 %                         | 17 %                               | 6 %                               | 9 %              |
| Lives in Copenhagen or Aarhus   | 33 %                             | 11 %                            | 16 %                            | 34 %   | 39 %                        | 26 %                               | 25 %                              | 26 %             |
| Turbulence measure  | 6,5<br>(1,2)                     | 4,3<br>(1,7)                    | 6<br>(1,9)                      | 6,1<br>(1,3)                                       | 4,3<br>(2,5)                | 6,3<br>(1,6)                       | 5,8<br>(1,7)                      | 5,6<br>(2)       |

\* Too few observations

### CLUSTER 3: EARLY NUCLEAR FAMILY – DELAYED MARRIAGE (12%)

These women show similar trajectories to the *early nuclear family*, but have more

transitions and a little higher turbulence. Like the *early nuclear family*, the average age at first birth is 25 years, but these women do not get married before well into their thirties. 20% are living alone at the

FIGURE 5: STATE DISTRIBUTION PLOT FOR CLUSTERS

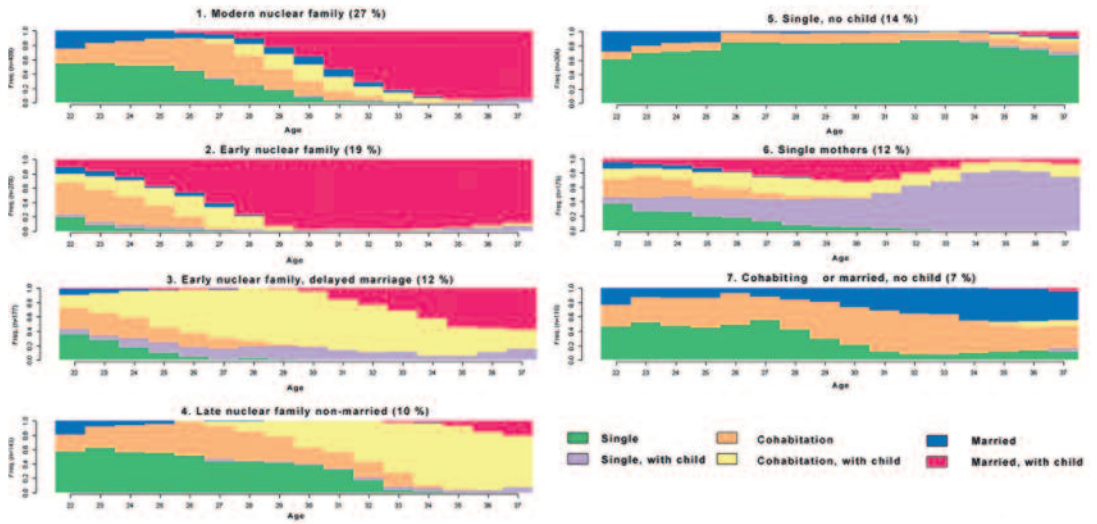
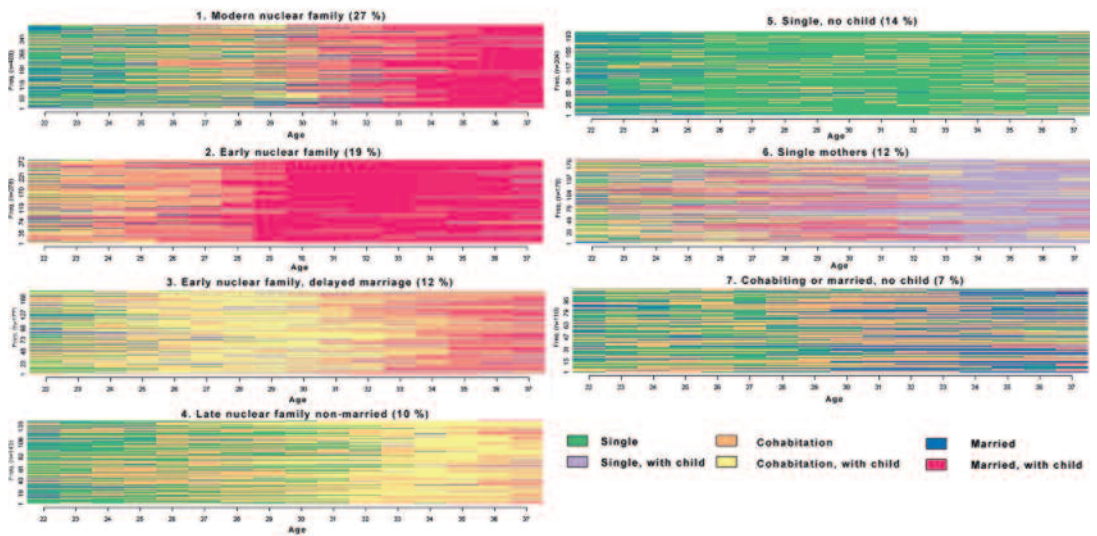


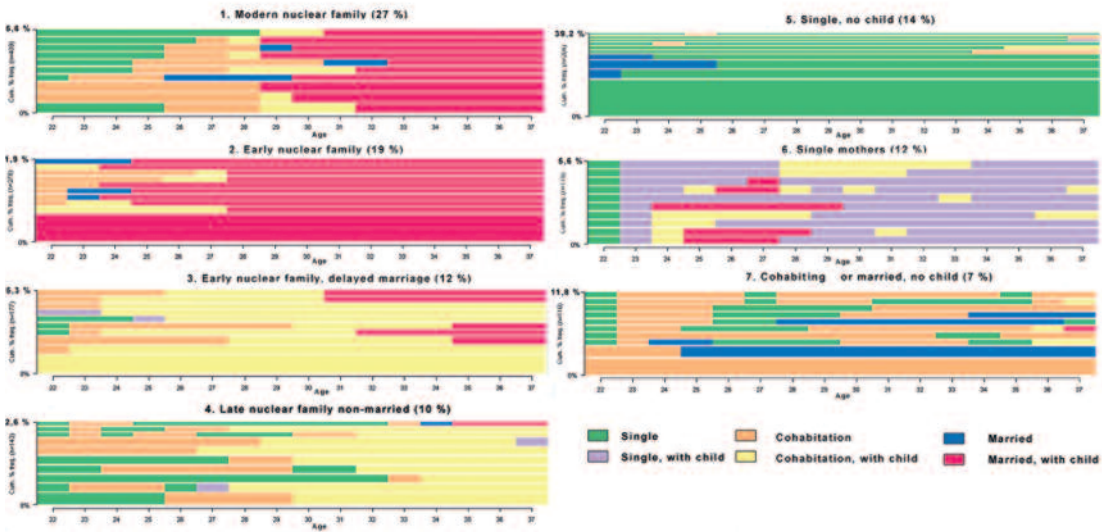
FIGURE 6: INDEX PLOT FOR FAMILY STATUS 22-37 YEARS



year of first birth and only 1% of the women are married at the time of first birth. However, at last year of observation, 74% of the women are cohabiting with the

father of their child, which is similar to the *modern nuclear family*. This cluster is the only cluster in which there is not a single woman who has been

FIGURE 7: TEN MOST FREQUENT SEQUENCES AND THEIR COVERAGE



in fertility treatment. Together with the relative low age at first birth, the high proportion of single mothers at the year of first birth and research of pregnancy types (Rasch et al. 2001: 1032), it is likely that a significant proportion of these women had non-planned but accepted pregnancies.

CLUSTER 4:  
MODERN NUCLEAR FAMILY,  
NON-MARRIED (10%)

Like the *modern nuclear family women*, the *non-married* women are highly educated, one out of three are living in Copenhagen or Aarhus. They are single and in and out of one or to two relationships until they start having children around the age of 29, with an average age of first birth at 31. Also, similar to the *modern nuclear family women*, the women in average more often than their partners have a higher university degree: 30% more women than men have a higher university degree at the last observed year they were together.

Although most of the couples stay to-

gether until the age of 37, only 20% are married. This does not mean that they from start have no belief in the institution of marriage: 43% have been married at some point in the period and 25% go through a divorce. 17% of the women live to some extent in a reconstituted family as their main partner has a child from a prior relationship, which is more than double as many as the women in the other *modern nuclear family* cluster.

CLUSTER 5:  
SINGLE, NO CHILD (14%)

The *single, no child women* are in average single for a total of 13 years out of the 16 years that they are observed. The group is the most homogeneous in their sequences: The ten most common sequences, which are mainly living as single throughout most of the period, cover almost 40%. However, 43% of the women have been married at some point in the period, which is almost equal to the percentage that experience going through a divorce (38%).

For the one out of five of the women who *do* become mothers, the average age of first birth is 35. While relatively few become mothers themselves, 37% of the women are in their longest relationship with a man who has at least one child from a prior relationship, which is more than double the population average.

39% of these women live in Copenhagen or Aarhus and only 3% have been in IVF treatment in the period, which is low compared to the other *childless* cluster. There is no way to know if these women have undergone other types of fertility treatment, however IVF treatment is the *next* step after insemination. Thus, compared to the last cluster of women without child, these women are to some extent likely to be ‘voluntarily’ childless.

#### CLUSTER 6: SINGLE MOTHERS (12%)

The *single mothers* in average become mothers at the age of 26 with most of the women living alone with their child for the majority of the years 22-37 years. The relative low age at first birth, lower rate of education and high rate that are single at the year of birth also point toward that a substantial part of the pregnancies in this cluster could be made up of non-planned, but accepted pregnancies (Rasch et al. 2001: 1032).

There is a high degree of turbulence as a result of the many transitions and short relationships. Eight out of ten of the most turbulent individual sequences are from this group. When it comes to income, education, workforce connection and health these women score lowest. One out of three has a loose connection to the workforce and likewise one out of three has been in contact with or had treatment at a hospital for an average of six years in total, the highest percentage of all groups.

Comparing the low work force participation with the proportion of women living

for longer as single-mothers there may be some women in this group who are living in non-cohabiting relationships. In Denmark, a range of benefits are dependent on whether you are deemed ‘single’ meaning whether you ‘share a household’ with a partner (Ankestyrelsen 2015). This does not mean that the women are not single mothers; unlike those in the *nuclear families* it is expected that a significant proportion of the women do not have a partner who contributes financially or homemaking-wise to their household.

The size of this cluster (12%) is almost equal to the percentage of children who, according to Statistics Denmark, were living alone with their mother in 2011 (13%) (StatBank Denmark 2017). However, there is a big difference between being single at the year of birth or in *one* given year and living alone with your child for the majority of its childhood. One of the strengths of longitudinal data and sequence analysis is to be able to differentiate between these groups on the basis of their full trajectories.

#### CLUSTER 7: COHABITING OR MARRIED, NO CHILD (7%)

These women have low turbulence of sequences, but show similar patterns to the *modern nuclear family* and *late nuclear family, non-married* in that transitions are centered around the age of 22-27 years; a good part of them are single while others are going in and out of one or two longer relationships before the majority enter long and stable relationships in their end-twenties (see figure 6 & 7).

One out of four of the women have undergone IVF treatment at some point. As in the other *childless* group, few have children (17%) and at an average late age (36). However, since some of the women might continue or start IVF treatment after the age of 37, the percentage of women in IVF treatment or having a child might increase.



The cluster has a high rate of marriage compared to the other clusters pre-first birth, which suggests that marriage persists as a constituting factor of the family, when there is no child.

## DISCUSSION

In this article, I studied the family formation pattern of a representative sample of 1,500 women born in 1973/1974 through the age of 22 to the age of 37 in order to get a data-based picture of the many pathways to family formation that Danish women undertake. The results of the sequence analysis revealed a myriad of pathways into parenthood. Nevertheless, the majority of the women, 68%, follow trajectories that can be described as a nuclear family form with two parents and at least one child, a state occurring either in the majority of the years 22-37 of age *or* appearing as a later outcome. The analysis confirms the child as the over-all constituting factor of family formation, but for the majority of women this event is followed by a marriage some years later.

But the road more travelled is not uniform: The four nuclear family groups differ substantially in mean timing of first birth (from 25 years to 31 years) as well as in the degree of turbulence, the measurement for the ‘chaos’ or ‘complexity’ of the trajectories (based on the number of distinct states, number of transitions and variance of duration spent in each state).

The two *modern nuclear family* clusters together account for 37% of the women and show that modern nuclear family also includes break-ups, postponements of first child and reconstituted families: 17% of *non-married modern nuclear family* live in reconstituted families by having a partner who has a child from a prior relationship, while 40% of the women in the *modern nuclear family* cluster experience going through a divorce and many remarrying between the ages 22-37.

For a more in depth analysis of pluralization of family formation, the sequence alphabet could be broadened to include reconstituted families, e.g. whether the women live with a partner who is not the biological father of her child or has children from a previous relationship. This was attempted in a separate analysis, however it resulted in too messy a picture, showing that, although sequence analysis reveals substantial outliers in the data better than alternative tools, e.g. survival analysis (Aisenbrey & Fasang 2010: 422), there *is* a limit to how many categories the sequence alphabet can consist of (while still producing meaningful plots) and thus how much complexity or pluralization sequence analysis can handle.

The results also call for a contextualization of ‘turbulence’ as a concept; the polar opposites of *modern nuclear family* and the *single mothers*, when it comes to education, income and cluster size, share a high degree of turbulence. Although turbulence is indeed an important tool for looking at “chaos” in trajectories, the measure of turbulence is in itself technical and does not account for the social circumstances and resources that might give turbulence different meaning to different people. Thus, in spite of their significantly more turbulent trajectories, the turbulence measure in itself has little meaning without context. Where turbulence for the women following *modern nuclear family* trajectories is centered around their twenties when about half of them are getting their higher university degree, the turbulence of the *single mothers* cluster, *single mothers* occurs later in their trajectories with most of them bringing up their child(ren) alone for a majority of the years.

The results also call for more research into existence of neighborhood effects of living in bigger cities in Denmark including the social, cultural and economic effects. In the two *modern nuclear family* clusters as well as the two *childless* clusters, there are

4-5 times more women living in the two major metropolitan areas of Denmark (Copenhagen and Aarhus) than women who grew up there, in strong opposition to the two *early nuclear family* clusters where the percentage who grew up there and live there are almost equal (and lower). Furthermore, on top of a correlation between high turbulence and obtaining a higher university degree, I find a correlation between *living* in Copenhagen or Aarhus (*not* growing up there) and high turbulence.

With some municipalities and regions struggling to retain the youth, especially in the childbearing age, more and more creative experimentation with welfare services and general branding is needed to withhold and attract the missing demographic balance. A better understanding of neighborhood effects and their character will be relevant in order to get a better understanding of how regions and municipalities can target and shape the future of their demographics and challenge the pull-effect of 'big cities'.

It is impossible to make a clear distinction between voluntarily childlessness and childlessness stemming from infertility, and this was not within the scope of this article. However, we can observe a significant difference in the proportion undergoing IVF treatment in the two *childless* clusters (4% for *single, no child* and 25% for *cohabiting or married, no child*). The high rate in the last cluster might result from the proportion of women who experiencing break ups at crucial times (late twenties), but who otherwise show similar trajectories to the typical trajectory of the *modern nuclear family*. It would be interesting to take a closer look at the association between the timing of break ups at crucial points, postponement of childbearing and solo-motherhood.

However, in order to adequately address the question of choice and pluralism in fertility and family formation, a supplement of survey information and/or qualitative in-

terviews would be beneficial. As with all data analysis there are limits to both the data and statistical models. Though information on partners were included in this article, relationship and family history of the partners is not traced. Having a child is a decision for all parties in a partnership, thus for a deep understanding of the relationship between men and women's family formation trajectories a comparative sequence analysis would have to be performed. The women following trajectories of the *single mothers* and *single, without child* can include solo-mothers or rainbow-families who have had inseminations, have found other ways of getting pregnant or whose same sex-partners, if not married, are not included in Statistics Denmark's categorization of civil status. When it comes to IVF treatment other things are not traceable in the data: In 1997-2007 (when the women were aged 23-34) doctors were forbidden by law from performing fertility treatment on single and same-sex couples, who had to seek treatment elsewhere (Retsinformation 1997). Very little is known about the extent to which single women and lesbians have been treated in spite of the ban in this period. Additionally, due to the fact that Statistics Denmark does not register cohabiting same-sex couples that are not married as a couple, there is little way to add to the knowledge of family formation for same-sex couples. This shows that when categories do not account for *other* ways of living, knowledge on other family forms will always-already be erased. However, sequence analysis is not blind to outliers or 'others'. In opposition to many other quantitative methods that focus mainly on the aggregated probabilities and effects, sequence analysis represents *all* individual trajectories revealing outliers, while at the same time showing that even typical trajectories, such as the nuclear family, are made up of very different pathways.

With increasing availability of longitudinal data, sequence analysis provides demog-

raphy and family sociology with an important tool for exploring processes, such as inter-generational transmission of patterns, differences in regional attitudes to family or different cohort or period effects on family formation. This article provides a needed base for looking at timing, order and duration in sequences of family formation in a Danish context and shows that in spite of differences both between and within groups, the majority of Danish women in the cohort 1973/1974 show clear overall patterns and variants of the nuclear family, however different the path to this state of family might be.

## NOTER

1. Due to the structure of the data and registries and in order to mirror the actual population of women in the fertile age, the random sample of 1,500 women is made up of 93.5% from the control population (the representative sample of Danish women who have not received IVF treatment) and 6.5% from the IVF population (the full population of Danish women who have received IVF treatment). The 6.5% is calculated on the basis of the percentage of women in the Danish population who have had IVF treatment in 1994-2009 (period of registration).
2. Calculated on the basis of numbers from Statistics Denmark's database that contains all births by mother's age (10-64 years). Located on 29/01/2017 at Statistics Denmark's website: <http://www.statistikbanken.dk/>.
3. Choosing the amount of clusters is like the cost-setting an abstract part of sequence analysis that is also a dynamic process in which you produce more or less clusters to see how the different options add to the overall picture. Furthermore, the Ward's clustering has a tendency towards making clusters of equal size, which is useful if clusters are needed for other statistical treatment (regressions etc.).
4. In 1960, 43.5% of women aged 15-69 years were part of the work force, in 1970, it had grown to 54.1% and in 1980, it was 68.1% (Statistics Denmark 2008: 10).

## LITERATURE:

- Abbott, Andrew (1997): Of Time and Space: The Contemporary Relevance of the Chicago School, in: *Social Forces* 1997/75(4): 1149-82.
- Abbott, Andrew (2001): *Time matters: On Theory and Method*. Chicago: University of Chicago Press.
- Aisenbrey, Silke & Fasang, Anette (2010): New Life for Old Ideas: The 'Second Wave' of Sequence Analysis Bringing the 'Course' Back Into the Life Course, in: *Sociological Methods & Research* 2010/38(3): 420-462.
- Ankestyrelsen (2015): *Hvornår er man enig?* Available at: <https://ast.dk/social/ydelser-til-enlige/hvornar-er-man-enlig>, accessed 15/01 2017.
- Blenstrup, Lene & Knudsen, Lisbeth (2011): Danish Registers on Aspects of Reproduction, in: *Scandinavian Journal of Public Health* 2011/39: 79-82.
- Brückner, Hannah & Mayer, Karl Ulrich (2005): De-Standardization of the Life Course: What It Might Mean? and if it Means Anything, Whether it Actually Took Place?, in: *Advances in Life Course Research* 2005/9: 27-53.
- Elchardus, Mark & Smits, Wendy (2006): The Persistence of the Standardized Life Cycle, in: *Time & Society* 2006/15: 303-326.
- Elzinga, Cees H. (2006): *Turbulence in Categorical Time Series*. Available at: [www.researchgate.net/publication/228722703\\_Turbulence\\_in\\_categorical\\_time\\_series](http://www.researchgate.net/publication/228722703_Turbulence_in_categorical_time_series), accessed 19/09 2015.
- Elzinga, Cees H. & Aart C. Liefbroer (2007): De-Standardization of Family-Life Trajectories of Young Adults: A Cross-National Comparison Using Sequence Analysis, in: *European Journal of Population / Revue Européenne de Démographie* 2007/23(3-4): 225-250.
- Hagestad, Gunhild & Call, Vaughn (2007): Pathways to Childlessness: A Life Course Perspective, in: *Journal of Family Issues* 2007/28: 1338-1361.
- Knudsen, Lisbeth & Wielandt, Hanne (1995): *På vej mod abort*. Copenhagen, Frydenlund Grafisk.
- Latour, Bruno (2005): *Reassembling the Social*. Oxford, Oxford University Press.
- Lesthaeghe, Ron J. (2010): *The Unfolding Story of the Second Demographic Transition*. Research Report 10-696. Population Studies Center, University of Michigan.
- Lesnard, Laurent (2010): Setting Cost in Optimal Matching to Uncover Contemporaneous Socio-Temporal Patterns, in: *Sociological Methods & Research* 2010/38: 389-419.

- Nielsen, Mathias (2015): Overlæge: Infertilitet er en overset folkesygdom. *Berlingske Tidende*. 04/10/2015. Available at: <https://www.b.dk/nationalt/overlaege-infertilitet-en-totalt-overset-folkesygdom>, accessed 15/01 2017.
  - Nielsen, Henriette S.; Schmidt, Lone; Andersen, Anders N.; Petersen, Kathrine B.; Gyrd-Hansen, Dorte; Jensen, Tina K.; Juul, Anders & Knudsen, Lisbeth B. (2016): *Forebyggelse af nedsat frugtbarhed*. Vidensråd for Forebyggelse, Copenhagen.
  - Ottosen, Mai Heide & Mouritzen, Sofie Skovdal (2013): Patterns of partnership and parenthood: experiences, approaches and readiness towards commitment and creating a family, in: Anne Lise Ellingsæter; An-Magritt Jensen & Merete Lie (eds.): *The Social Meaning of Children*. New York, Routledge.
  - Rasch, Vibeke; Knudsen, Lisbeth & Wielandt, Hanne (2001): Pregnancy planning and acceptance among Danish pregnant women, in: *Acta Obstetricia et Gynecologica Scandinavica* 2001/80: 1030-1035.
  - Retsinformation (1997): *Cirkulæreskrivelse vedrørende ny lov om kunstig befrugtning og nye krav for adgang til behandling*. Available at: <https://www.retsinformation.dk/Forms/R0710.aspx?id=10233>, accessed 20/01 2017.
  - Schmidt, Lone (2006): *Infertility and assisted reproduction in Denmark, Epidemiology and psychosocial consequences*. Copenhagen, Lægeforeningens Forlag.
  - Schmidt, Lone; Hageman, I.; Hougaard, C. Ø.; Sejbaek, C. S.; Assens, M.; Ebdrup, N.H. & Pinborg, A. (2013): Psychiatric disorders among women and men in assisted reproductive technology (ART) treatment. The Danish National ART-Couple (DANAC) cohort: protocol for a longitudinal national register-based cohort study, in: *BMJ Open* 2013/3(3): e002519.
  - StatBank Denmark (2017): *StatBank Denmark*. Available at: <https://www.statbank.dk/statbank5a/default.asp?w=1280>, accessed 15/01 2017.
  - Statistics Denmark (2017). Register og variabeloversigt. Available at: [http://www.dst.dk/da/TilSalg/Forskningsservice/Data/Register\\_Variabeloversigter](http://www.dst.dk/da/TilSalg/Forskningsservice/Data/Register_Variabeloversigter), accessed 15/01 2017.
  - Statistics Denmark (2008): *60 år i tal – Danmark siden 2. Verdenskrig*. Statistics Denmark, Copenhagen.
  - Sundhedsdatastyrelsen (2015): *Assisteret reproduktion. Analyse og tal*. Sundhedsdatastyrelsen, Copenhagen.
  - Thompson, Elizabeth; Winkler-Dvorak, Maria & Kennedy, Sheela (2013): The Standard Life Course: An Assessment of Variability in Life Course Pathways, in: Ann Evans & Janeen Baxter (eds.): *Negotiating the Life Course*. New York, Springer.
  - Zegers-Hochschild, F. et al. (2009): The International Committee for Monitoring Assisted Reproductive Technology (ICMART) and the World Health Organization (WHO) Revised Glossary on ART Terminology, in: *Human Reproduction* 2009/24 (11): 1520-1524.
- The analyses presented in this paper are made in R using the TraMineR package. A special thanks to Lone Schmidt, Assoc. Prof., DMSci, PhD, Department of Public Health, University of Copenhagen, for providing access to the data that made this article possible.