# Studying Family Formation Trajectories' Deinstitutionalization in Russia Using Sequence Analysis

Ekaterina S. Mitrofanova[1], Alyona V. Artamonova[1]

[1]National Research University Higher School of Economics

mitrofanovy@yandex.ru, alyona89152694371@yandex.ru

**Abstract.** This study focuses on changing family formation trajectories in the Russian Federation. In European countries, pathways to family ceased being stable several decades ago, while in Russia – as in any post-socialist country – such features of life course deinstitutionalization as postponement of marriage, rising cohabitation, and reordering of events were revealed only in the 1990s and explained from the perspective of the Second Demographic Transition (SDT). Our aim is to demonstrate how family formation trajectories of men and women from different generations were transforming with the incorporation of data mining. The three-wave panel data of the Russian part of the "Generations and Gender Survey" (2004, 2007, 2011; N=5321) and the retrospective data of the survey "Person, Family, Society" (2013; N=4477) are used for achieving this aim. Sequence Analysis shows that generations born after 1970 started to exhibit de-standardized family formation trajectories. As the proportion of Russians who raise children in cohabitation or while single rises, such models of behavior become more widely accepted and practiced in contemporary Russia. Women experience more events in the family trajectory, take steps toward family formation earlier, and stay alone with children more often than men. Matrimonial and reproductive behavior has become diverse, proving that Russia fully exhibits the SDT.

**Keywords:** family formation trajectories, matrimonial and reproductive behavior, Sequence Analysis

## 1 Introduction

People's family formation trajectories have considerably changed in recent decades. In many European countries, marital union with children has been the only acceptable method of family organization for a long time. Since the 1990s, a couple may be formed not only through marriage but also through cohabitation, people may postpone the birth of children or remain childfree, and a union may not be dissolved solely through divorce but also through separation; because of new freedom of thinking and behaving and people's orientation to individual self-development, this is one of the distinctive features of modern society [1].

The theorists of the Second Demographic Transition approach, headed by pioneers Lesthaeghe and Van de Kaa, explain the transformation in demographic behavior as the result of the broad and long-term changes in the norms and values that many countries witnessed between the mid-1960s and the end of the 1980s [2]. Mayer [3] claims that, since the 1960s, societies have embraced so-called "hedonistic individualism", which includes alternative lifestyles, emphasizing individual fulfillment and self-expression rather than sacrifices to the family, traditional values and altruistic orientations regarding children and the collective good. Instead of following the tradition of marriage, young people realize their own personal goals of self-expression and enjoyment [4].

All the SDT changes in paths to family formation started in Western European countries and followed the model of the European type of marriage prevailing west of Hajnal's line. Eastern European Russia displays demographic outcomes of the SDT in atypical fashion. Growing cohabitation rates alongside declining marital rates emerged in the Soviet Union in the middle of the 1980s, years before the fall of socialism [5]. Zakharov [6] revealed that Russians born after 1970s already started to demonstrate all features of SDT (e.g. the formation of partnerships outside marriage, the rise in non-marital childbearing, and the postponement of marriage). Mills showed that new pathways to family in Russia, contrary to SDT theory, are prevailing among less-educated people, reminiscent of a 'pattern of disadvantage' concept. It makes Russia look more like the United States than Europe with regards to life course deinstitutionalization.

Taking this complexity of matrimonial and reproductive behavior into consideration, we decided to trace the family formation trajectories' deinstitutionalization in Russia based on gender-generational differences using Sequence Analysis.

## 2 Hypotheses

The standardized trajectory of "Soviet" generation Russia starts from singlehood and includes universal marriage with at least one child. The proportions of those single with children and those secondly married were minimal. From the middle of the 1980s until the collapse of the Soviet Union, Russians turned to Western European countries' family lifestyles [7]. The average ages of marriage have been rising since the early 1990s. In 1993, the ages for men and women were 23.9 and 21.8 years, respectively. In 1999 and 2004 they consisted of ages 25.0 and 26.1 for men and ages 23.1 and 23.3 for women [1].

According to Mills and her co-authors, there is a high proportion of single parents in Russia (even higher than in some Western European countries),

which may be caused by a high divorce rate and particularly high adult male mortality, which is largely due to alcohol-related deaths [7].

Taking into consideration the information above, we decided to verify two groups of hypotheses.

**Group 1. Gender:**

- Women take steps to family earlier than men;
- Women stay alone with children more often than men;
- Women experience more family formation events than men;

**Group 2. Generations:**

- De-standardization of family formation trajectories was demonstrated first by representatives of the first "Modern" generation (1970-79 birth cohort);
- "Modern" generations experience more varied matrimonial and reproductive events than the representatives of "Soviet" generations.
- To test these hypotheses, we decided to apply Sequence Analysis, which requires longitudinal or retrospective data.

## 3 Data

We used the panel data of the Russian part of the Generations and Gender Survey (GGS-panel: 2004, 2007 and 2011) and retrospective data of the "Person, Family, Society" survey (PFS: 2013). We choose these surveys because their designs apply the Life Course approach, which tends to understand different types of demographic events as a chain of interconnected processes. The questions about life course events were asked in a very accurate and detailed way. Most of the dates contain not only years but also months of starts and ends of events. We should mention that the questions about children were asked so as to show our interest in the biological children of respondents.

To work correctly with sequences, it was necessary to constrain the ages of events. 15 years as the lower age point was chosen because it is the beginning of possible reproductive behavior. Obviously, there were respondents who enter into their first union or have their first child before reaching this age but such atypical cases are outside the scope of our study. In the samples of used datasets there are respondents who, at the time of the survey, were 25 years old (GGS-2011, third wave) and even 18 years old (PFS-2013). Marriages in Russia were early and universal for a long time, and almost all representatives of the Soviet generations started their unions by the age of 25. We supposed that younger generations demonstrate a delay in the start of their first unions in comparison with the Soviet generations. That is why, if we want to trace the change in the age of the first union formation, we should analyze a wide range of ages. However, the representatives of the older generations have lived longer lives than the youth, and some unique cases of the first unions at ages over 40 years can shift the average age. Moreover, it is not correct to compare the full

matrimonial biographies of people who reached the age of final celibacy and people who only started their union histories. Taking into account all these arguments, we decided to impose a limit on the age of matrimonial and reproductive events occurring. After considering several options, we limited the age of entry into first union at 35 years, no matter whether not all respondents finished the transition to family life by the age of 35.

The final GGS and PFS datasets contain 5321 and 4477 cases, respectively.

In order to analyze the generational aspect of matrimonial behavior, we divided our samples into two key groups: the "Soviet" generations (1930-39, 1940-49, 1950-59, 1960-66 in GGS and 1960-69 in PFS), who socialized before the collapse of the Soviet Union, and the "Modern" generation (1970-79, 1980-86 in GGS and 1970-79, 1980-89, 1990-95 in PFS), who socialized after it [8]. The proportions of men and women in different generations of GGS and PFS can be found in the Table 1.

**Table 1.** Proportions of men and women in Russian generations

| Generation | Gender | GGS | | PFS | |
|---|---|---|---|---|---|
| | | Absolute numbers | Percentages | Absolute numbers | Percentages |
| 1930-1939 | Men | 192 | 25% | - | - |
| | Women | 585 | 75% | - | - |
| 1940-1949 | Men | 214 | 28% | - | - |
| | Women | 552 | 72% | - | - |
| 1950-1959 | Men | 387 | 30% | - | - |
| | Women | 923 | 70% | - | - |
| 1960-1969 | Men | 423 | 36% | - | - |
| | Women | 761 | 64% | - | - |
| 1970-1979 | Men | 325 | 36% | 798 | 48% |
| | Women | 585 | 64% | 855 | 52% |
| 1980-1986(89) | Men | 158 | 42% | 939 | 49% |
| | Women | 216 | 58% | 988 | 51% |
| 1990-1995 | Men | - | - | 473 | 53% |
| | Women | - | - | 424 | 47% |

## 4 Methodology

In recent years, there has been a strongly growing interest in the study of life course trajectories to describe life trajectories, to classify individuals according to them by using the Sequence Analysis (SA) method [9, 10, 11, 12]. SA is

based on data mining approaches, namely on the measures of dissimilarity or distance between individual trajectories. It is entirely non-parametric.

The majority of papers devoted to SA highlights both certain socio-demographic phenomena and the methodological development of the method. There are some papers about the deinstitutionalization of the life course [13], starting events are postponed [14, 15, 16], women are more proactive in social life and they are postponing maternity [17], and the number of social roles are growing for both sexes [18]. Matrimonial trajectories are becoming more diverse and less predictable [19, 20, 21].

The development of methods goes in two directions: development with the sources of mathematical statistics and Data Mining [22, 23, 24]. The researchers not only discover typical sequences for different classes, but also cluster them [25, 26], evaluate their resemblance [27], creat classifiers [28], define the transaction costs [16], and build the decision trees [14].

The representation of life course trajectories in SA is similar to the code of DNA molecules [9]. It focuses on a time window with chosen ages of start and finish, inside of which studied events (e.g. entry to first and second cohabitations (P1 and P2), marriages (M1 and M2), and birth of first and second child (C1 and C2)) can occur. As was explained above, in our research, the first point of the time window is 15 years (when the majority of Russians do not have any matrimonial (i.e. single – S) or reproductive (i.e. childless – C0) events) and the last point is 35 years. We deal with so-called 'non-recurrent sequences', where an event may not repeat at all.

As individual life courses can be represented as a sequence of events, we are able to code every event with a letter and build the "word" that describes the state of an individual at every point of a chosen time window. Table 2 shows all possible states of partnership and fertility trajectory.

**Table 2.** Alphabet of partnership and fertility states

| Code | State | Code | State |
|---|---|---|---|
| SC0 | Single, no children | M1C0 | First marriage, no children |
| SC1 | Single, 1 child | M1C1 | First marriage, 1 child |
| SC2 | Single, 2 children | M1C2 | First marriage, 2 children |
| P1C0 | First cohabitation, no children | M2C0 | Second marriage, no children |
| P1C1 | First cohabitation, 1 child | M2C1 | Second marriage, 1 child |
| P1C2 | First cohabitation, 2 children | M2C2 | Second marriage, 2 children |
| P2C0 | Second cohabitation, no children | | |
| P2C1 | Second cohabitation, 1 child | | |
| P2C2 | Second cohabitation, 2 children | | |

In our study, we used TraMineR (R-package) to mine and visualize sequences of matrimonial and reproductive events [29]. The first tool we used

was chronograms. A chronogram is the representation of all the sequences of a group at each age. It is a summary of individual trajectories. We used the graphs representing the entropy – the measure of disorder of sequences – at each time period. We calculated the mean time spent in statuses which, that is, how long every member of each group, on average, was in each status. And finally, we calculated the number of family formation events, which mean how many events each member of each group experienced in his or her life.

## 5 Empirical Results

We first show the results of the first group of tested hypotheses and then move to the second group.

**Gender**. In order to prove that women take steps to family earlier than men, we compared distributions of partnerships and fertility statuses, all sequences, and entropy by gender.

On the horizontal axis of the plots, there are the ages of the respondents between 15 and 35 years. The youngest respondents have not yet reached the upper age limit: this is why we had to work with censored data (indicated in gray). On the vertical axes of the first and third plots, the proportions of individuals belonging to each state at a given age are shown. On the vertical axis of the second plot there are respondents, so we can observe individual family formation trajectories of men and women.

The plots reveal that either in GGS or in PFS, men start to experience family formation events at the age of 17, while women do it earlier. In fact, 80% of women have at least one event at 23, while among men this proportion is reached at 26.

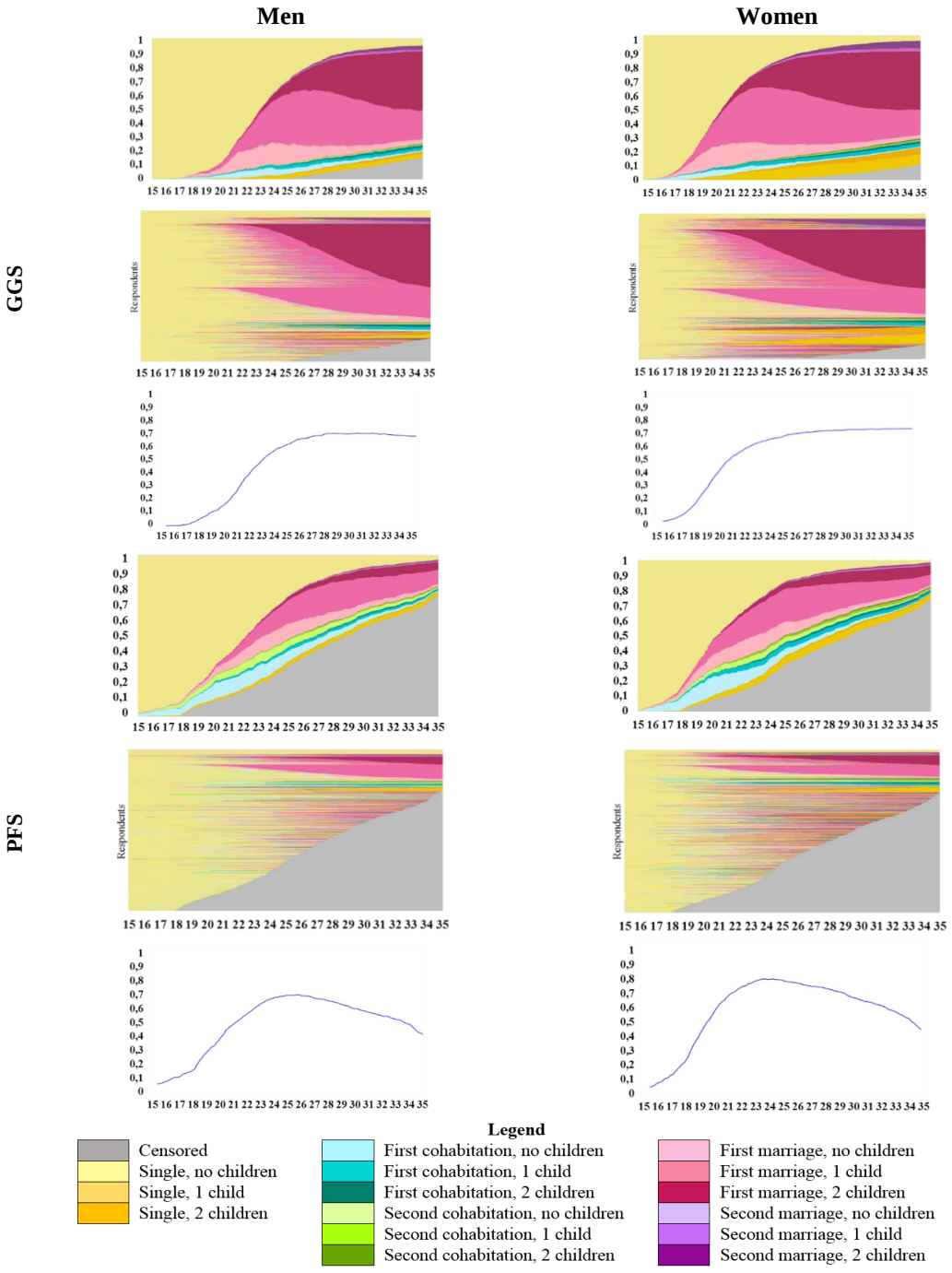



**Fig. 1.** Family formation trajectories of Russians

One more evidence for our hypothesis is the mean time spent in singlehood and without children (Figure 2). Men spend about 100 months after 15 years in this status, while women spend only about 80 months.

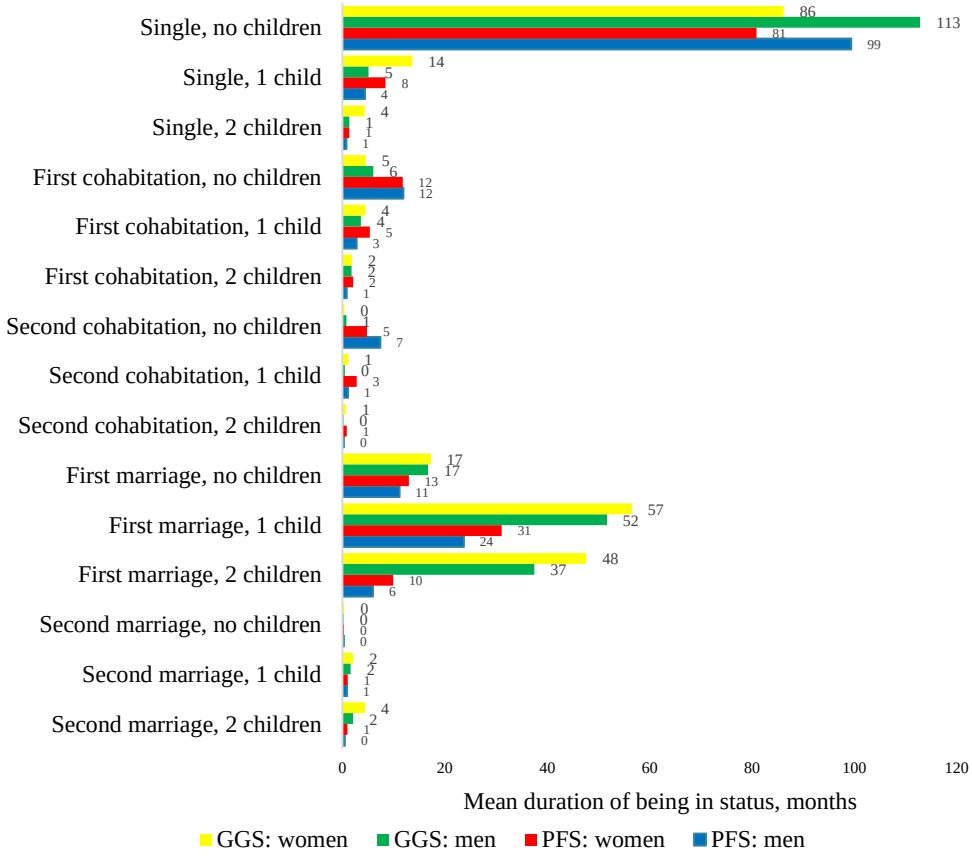
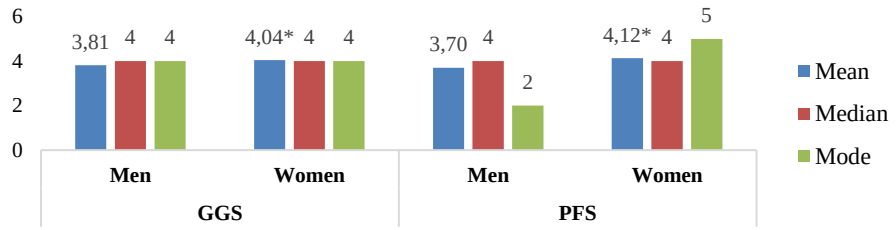



**Fig. 2.** Mean time spent in status

In order to prove that women stay alone with children more often than men do, we should look at Figures 1 and 2. The distribution of partnerships and fertility statuses plot demonstrates that the proportion of single women with children at the age of 35 (25%) is more than twice the proportion of such men (10%). Mean time spent in these two statuses is higher for women (about 14 months) than for men (about 5 months) as well.

In order to prove that women experience more family formation events than men, we compare mean, median, and mode number of family formation events for men and women. The mean demonstrates that women have significantly more events than men but, according to two other figures, the numbers are the same.




*The difference is statistically significant (p<0,001)

**Fig. 3.** Number of family formation events by gender

**Generations**. In order to prove that the de-standardization of family formation trajectories was demonstrated first by representatives of first "Modern" generation (1970-79 birth cohort) we compared the entropy of different generations (Figure 4) and the distribution of partnerships and fertility statuses by gender and generation (Appendix 1).

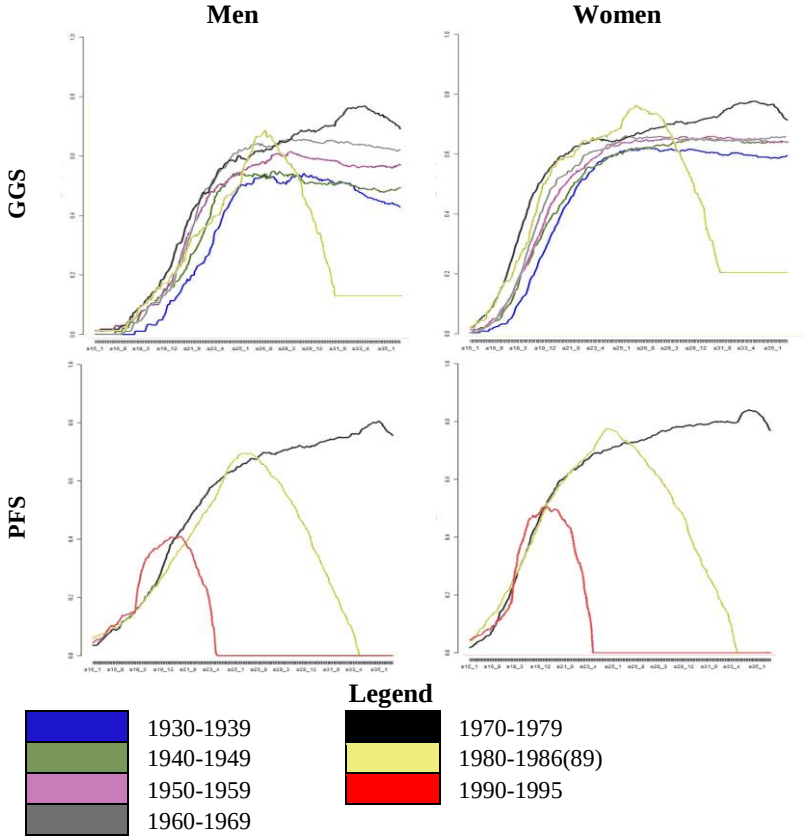



**Fig. 4.** Entropy by generations

It is apparent from the Figure in Appendix 1 that the proportion of married people with at least one child decreased while the proportions of cohabited (blue pallet) and single people with children (yellow pallet) have increased dramatically. The visible changes started with the generations born after 1970.

In order to prove that "Modern" generations experience more varied matrimonial and reproductive events than the representatives of "Soviet" generations, we counted mean, median, and mode number of family formation events for different generations (Table 3).

**Table 3.** Number of family formation events by generation and gender

| | **Men** | | | **Women** | | |
|---|---|---|---|---|---|---|
| | **Mean** | **Median** | **Mode** | **Mean** | **Median** | **Mode** |
| **GGS** | | | | | | |
| 1930-1939 | 3.70 | 4 | 4 | 3.68 | 4 | 4 |
| 1940-1949 | 3.82 | 4 | 4 | 3.84 | 4 | 4 |
| 1950-1959 | 3.81 | 4 | 4 | 4.07 | 4 | 4 |
| 1960-1969 | 3.89 | 4 | 4 | 4.07 | 4 | 4 |
| 1970-1979 | 4.10 | 4 | 4 | 4.54 | 4 | 4 |
| 1980-1986 | 3.15 | 3 | 2 | 3.95 | 4 | 4 |
| **PFS** | | | | | | |
| 1970-1979 | 4.12 | 4 | 4 | 4.42 | 4 | 4 |
| 1980-1989 | 3.94 | 4 | 2 | 4.44 | 5 | 5 |
| 1990-1995 | 2.51 | 2 | 2 | 2.79 | 2 | 2 |

The figures demonstrate that the number of events for men and women in generations do not differ.

## 6 Conclusions

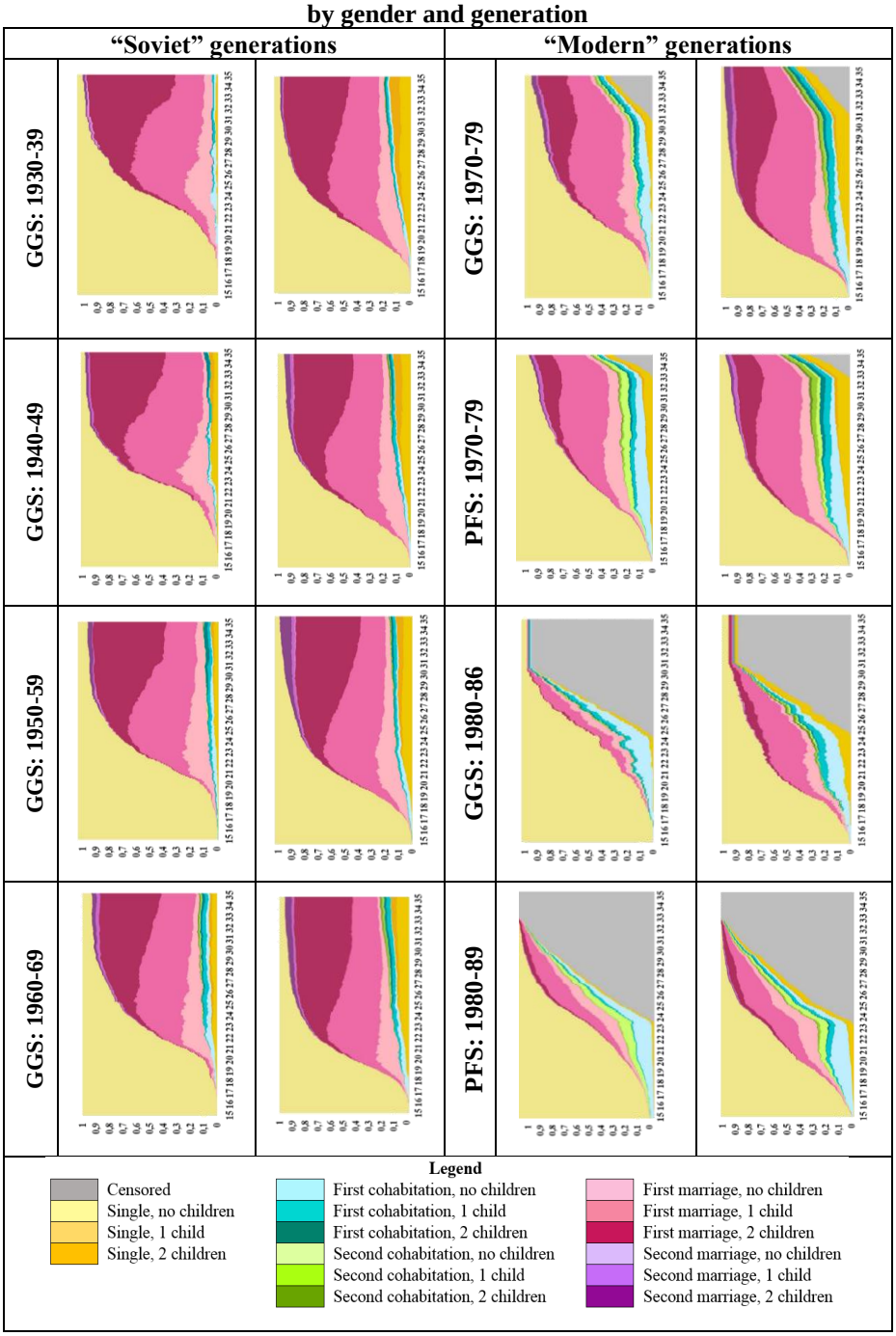
In this paper we revealed several points about family formation trajectories of Russians:

- women start to entry into first matrimonial events earlier than men;
- women stay alone with children more often than men do;
- women and men experience equal number of family formation events;
- generations born after 1970 started to exhibit de-standardized family formation trajectories;
- the number of events for men and women in different generations remains stable.

Matrimonial and reproductive behavior is becoming diverse, proving that Russia fully displays Second Demographic Transition.

**Acknowledgements**. The article was prepared within the framework of the Academic Fund Program at the National Research University Higher School of Economics (HSE) in 2016 (grant № 16-05-0011 "Development and testing of demographic sequence analysis and mining techniques") and supported within the framework of a subsidy granted to the HSE by the Government of the Russian Federation for the implementation of the Global Competitiveness Program. The authors also want to thank Thomas H. Espy for inestimable help in the preparation of this paper.

**Appendix 1. Distribution of partnerships and fertility statuses by gender and generation**



## References


1. Avdeev, A., Monnier, A.: Marriage in Russia: A Complex Phenomenon Poorly Understood. Popul. Engl. Sel. 12, pp. 7–49 (2000)
2. Kaa, D.J. van de, Lesthaeghe, R.: Two Demographic Transitions? Population: Growth and Decline, pp. 9–24 (1986)
3. Mayer, K.U.: Whose Lives? How History, Societies, and Institutions Define and Shape Life Courses. Research in Human Development 1 (3), pp. 161–187 (2004)
4. Gerber, T.: Changing family formation behavior in post-socialist countries: Similarities, divergences, and explanations in comparative perspective. Draft for presentation at the "1989: Twenty Years After Conference" (2009)
5. Gerber, T.P., Berman, D.: Entry to Marriage and Cohabitation in Russia, 1985–2000: Trends, Correlates, and Implications for the Second Demographic Transition. Eur. J. Popul., pp. 26: 3–31 (2010)
6. Zakharov, S.: Russian Federation: From the first to second demographic transition. Demographic Research, 19, pp. 907–972 (2008)
7. Mills, M., Lesnard, L., Potarca, G.: Family Formation Trajectories in Romania, the Russian Federation and France: Towards the Second Demographic Transition?. European Journal of Population, 29, pp. 69-101 (2013)
8. Levada, Y.: Generations of XX Century: Opportunities of Studies. In Y. Levada, T. Shanin. M. Fathers and Children: Analysis of Contemporary Russian Generations. Moscow, pp. 39-60 (2005)
9. Billari, F.C.: Sequence Analysis in Demographic Research. Special Issue on Longitudinal Methodology. Canadian Studies in Population Vol. 28(2), pp. 439-458 (2001)
10. Ignatov, D., Mitrofanova, E., Muratova, A., Gizdatullin, D.: Pattern Mining and Machine Learning for Demographic Sequences. KESW 2015, pp. 225-239 (2015)
11. Blockeel, H., Fürnkranz, J., Prskawetz, A., Billari, F.C.: Detecting Temporal Change in Event Sequences: An Application to Demographic Data. PKDD, pp. 29-41 (2001)
12. Low-Kam C., Raïssi C., Kaytoue M., Pei J.: Mining Statistically Significant Sequential Patterns. ICDM, pp. 488-497 (2013)
13. Worts, D. et al.: Individualization, Opportunity and Jeopardy in American Women's Work and Family Lives: A Multi-State Sequence Analysis. Advances in Life Course Research, 18, pp. 296–318 (2013)
14. Billari, F. C., Fürnkranz, J. and Prskawetz, A.: Timing, Sequencing, and Quantum of Life Course Events: A Machine Learning Approach. European Journal of Population, 22, pp. 37–65 (2006)
15. Mouw, T.: Sequences of Early Adult Transitions: How Variable Are They, and Does It Matter. Frontier of Adulthood: Theory, Research, and Public Policy, pp.256–91 (2005)
16. Billari, F.C., Rosina, A.: Italian "Latest-Late" Transition to Adulthood: An Exploration of Its Consequences on Fertility. Genus (2004)
17. Aassve, A., Billari, F.C. and Piccarreta, R.: Strings of Adulthood: A Sequence Analysis of Young British Women's Work-Family Trajectories. European Journal of Population, 23, pp. 369–88 (2007)
18. Jackson, P.B. and Berkowitz, A.: The Structure of the Life Course: Gender and Racioethnic Variation in the Occurrence and Sequencing of Role Transitions. Advances in Life Course Research, 9, pp. 55–90 (2005)
19. Elzinga, C.H. and Liefbroer, A.C.: De-Standardization of Family-Life Trajectories of Young Adults: A Cross-National Comparison Using Sequence Analysis // European Journal of Population/Revue Européenne de Démographie, 23, pp. 225–50 (2007)
20. Oris, M. and Ritschard, G.: Sequence Analysis and Transition to Adulthood: An Exploration of the Access to Reproduction in Nineteenth-Century East Belgium. Advances in Sequence

Analysis: Theory, Method, Applications. Springer International Publishing, pp. 151–167 (2014)
21. Piccarreta, R., Billari, F.C.: Clustering work and family trajectories by using a divisive algorithm. J. R. Stat. Soc. Ser. A Stat. Soc. T. 170. № 4, pp. 1061–1078 (2007)
22. Dong, G., Pei, J.: Sequence Data Mining. New York: Springer (2007)
23. Zaki, M.J. and Meira, W. Jr.: Data Mining and Analysis: Fundamental Concepts and Algorithms. Cambridge University Press, New York, NY, USA (2014)
24. Han, J., Kamber M.: Data Mining: Concepts and Techniques. Morgan Kaufmann (2006)
25. Elzinga, C.H., Studer, M.: Spell sequences, state proximities and distance metrics. Sociol. Methods Res. T. 44. № 1 (2015)
26. Elzinga, C.H.: Distance, Similarity and Sequence Comparison. Advances in Sequence Analysis: Theory, Method, Applications. Springer, pp. 51–74 (2014)
27. Elzinga, C.H.: Sequence Similarity: A Nonaligning Technique. Sociol. Methods Res. T. 32. № 1, pp. 3–29 (2003)
28. Barban, N., Billari, F.C.: Classifying life course trajectories: A comparison of latent class and sequence analysis. J. R. Stat. Soc. T. 61. № 5, pp. 765–784 (2012)
29. Gabadinho, A., Ritschard, G., Muller, N. S., Studer, M.: Analyzing and Visualizing State Sequences in R with TraMineR. Journal of Statistical Software, 40(4), pp. 1-37 (2011) URL: http://www.jstatsoft.org/v40/i04/