

Zagel, Hannah; Van Winkle, Zachary

Article — Accepted Manuscript (Postprint)

Women's Family and Employment Life Courses Across Twentieth-Century Europe: The Role of Policies and Norms

Social Politics: International Studies in Gender, State & Society

Provided in Cooperation with:

WZB Berlin Social Science Center

Suggested Citation: Zagel, Hannah; Van Winkle, Zachary (2022) : Women's Family and Employment Life Courses Across Twentieth-Century Europe: The Role of Policies and Norms, Social Politics: International Studies in Gender, State & Society, ISSN 1468-2893, Oxford University Press, Oxford, Vol. 29, Iss. 2, pp. 446-476,
<https://doi.org/10.1093/sp/jxz056>

This Version is available at:

<https://hdl.handle.net/10419/261285>

Standard-Nutzungsbedingungen:

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

Terms of use:

Documents in EconStor may be saved and copied for your personal and scholarly purposes.

You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.

If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.

Women's family and employment life courses across twentieth century Europe: the role of policies and norms

Online Supplementary Material

1. Sequence & Cluster Analysis

Sequence analysis can be traced back to Vladimir Levenstein in 1965 as a means to enumerate the similarity of character strings and was initially used in the natural sciences to identify similarities and regularities in DNA strings (Needleman and Wunsch 1970). Sequences are defined as empirically observed traces of ordered and categorical states. In the case of DNA, there are four states, the nucleobases cytosine (C), guanine (G), adenine (A), and thymine (T), which are ordered along polynucleotide strands. The similarity between two DNA strands can be expressed in terms of an Optimal Matching (OM) distance, which is the minimum number of edits required to change one DNA strand into another. A pair of DNA sequences with a large OM distance value are less similar a pair with a low distance value.

Three steps are necessary when using sequence analysis to empirically establish patterns: 1) generate sequences by defining the categorical states relevant to the process being described as well as the time unit and sequence length, 2) calculate pair-wise sequence dissimilarities, and 3) subject the pair-wise sequence distance matrix to a cluster algorithm. Our family and employment life course trajectories are operationalized as sequences composed of 35 consecutive annual states from age 15 to 50. Therefore, our sequences capture not only early life course events, e.g. leaving the parental home and marriage or entering the labor market after leaving school, but also later life events, e.g. divorce and remarriage or entering early-retirement. Our family life sequences combine annual residential, union, and fertility information to give a holistic representation of individuals' family formation process over the course of their lives. At any given stage, i.e. age, an individual can either be in the parent home (P), living outside the parental home and not cohabiting (S), cohabiting (Co), married (M), or divorced (D); each with or without a child, e.g. (SC). After excluding individuals with missing information, we are able to construct family sequences for 26,474 women

Retrospective information from SHARELIFE and sequence analysis is also used to generate our labor market outcome: employment trajectory patterns. Specifically, we reconstruct women's employment histories between age 15-50 years. At any given age, women can either be in education (E), retired (R), unemployed, (UE), inactive (OLF), employed part-time (PT), or employed full-time (FT). Full-time and part-time employment is defined subjectively by the respondents and not based on an objective number of working hours, which is not available in SHARELIFE. We filled in missing employment states between the years 1939 and 1955 with a WW II gap state and included general gap states for persons with missing state information for a maximum of six years. This allows us to generate employment sequences for 29,363 women.

We then proceed to calculate distances between every pair of family sequences and every pair of employment sequences, which enumerate the similarity between two sequences. A large number of distance metrics have been developed in sequence analysis that highlight various

differences between sequences (see Studer and Ritschard 2016 for a review). We opted for calculating sequence distances based on the Longest Common Subsequence (LCS) algorithm (see Elzinga, Rahmann, and Wang 2008). LCS distances stress the prevalence and ordering of sequence states over the timing of transitions between sequences. A contrasting measure would be Dynamic Hamming Distance (DHD), which highlights timing differences rather than differences in the ordering or prevalence of states (see Lesnard 2014, 2010, 2008). However, we are especially interested in the prevalence of family states, such as divorce or single parenthood, and employment states, such as not participating in the labor force and part-time employment, that may be less common. DHD would be less inattentive to differences generated by the occurrence of those states. LCS distances are calculated in three steps: First, subsequences for each subsequence are extracted. Subsequences contain all the elements, i.e. single states and an order of states without durations, within each sequence that can be obtained by deleting any element within that sequence. Second, the length of the longest common subsequence that every sequence pair shares is quantified as a count metric. That count reflects the similarity of each sequence pair. In a third and final step, this count is transformed to reflect dissimilarity.

We then apply the family and employment sequence distance matrixes to a clustering algorithm. Cluster analysis is a tool to simplify complex data and empirically identify groups with sequences that are maximally similar to one another within their group and maximally dissimilar to sequences from other groups (Studer 2013). We use the partitioning around medoids (PAM) clustering algorithm that separates an initial set of sequences in subgroups characterized by the highest possible within-group homogeneity and between-group heterogeneity (see Gabadinho et al. 2011). Medoids are representative sequences that have the smallest dissimilarity to the other sequences of the cluster they belong to. We extract a five cluster solution for our family sequences and a three cluster solution for our employment sequences after a close consideration of both the average width silhouette (AWS) values and substance (see Figures A1 and A2). ASW is an indicator for the coherence of a cluster solution. High ASW values signal that distances between the sequences of different clusters are large, while distances between the sequences within each cluster are small.

Figure A1: Cluster Solution Quality Criteria for Family Sequences

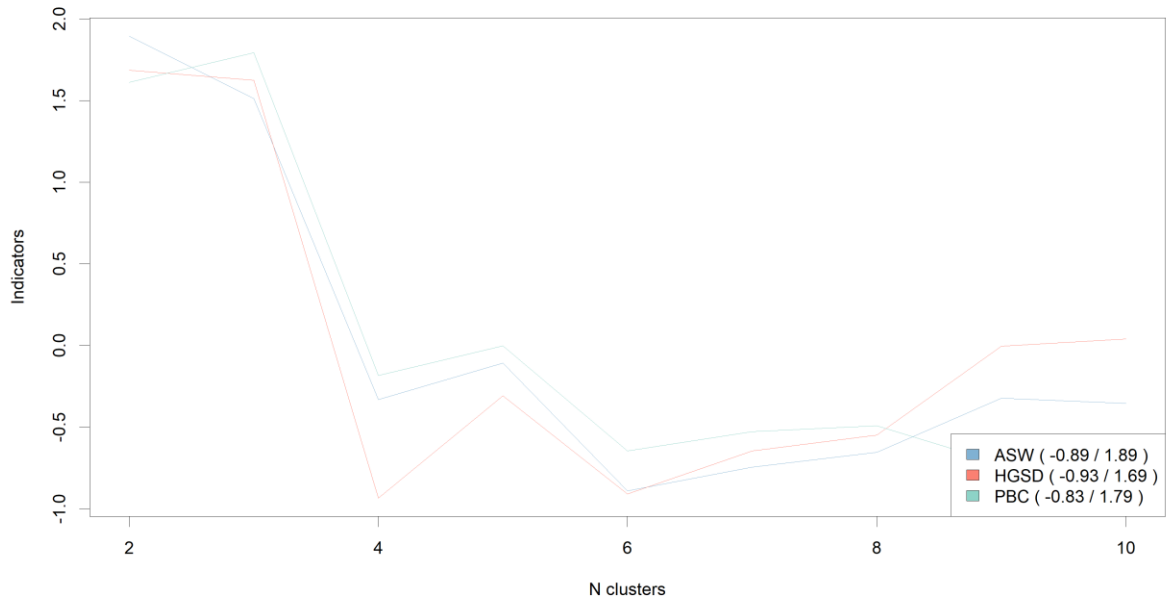
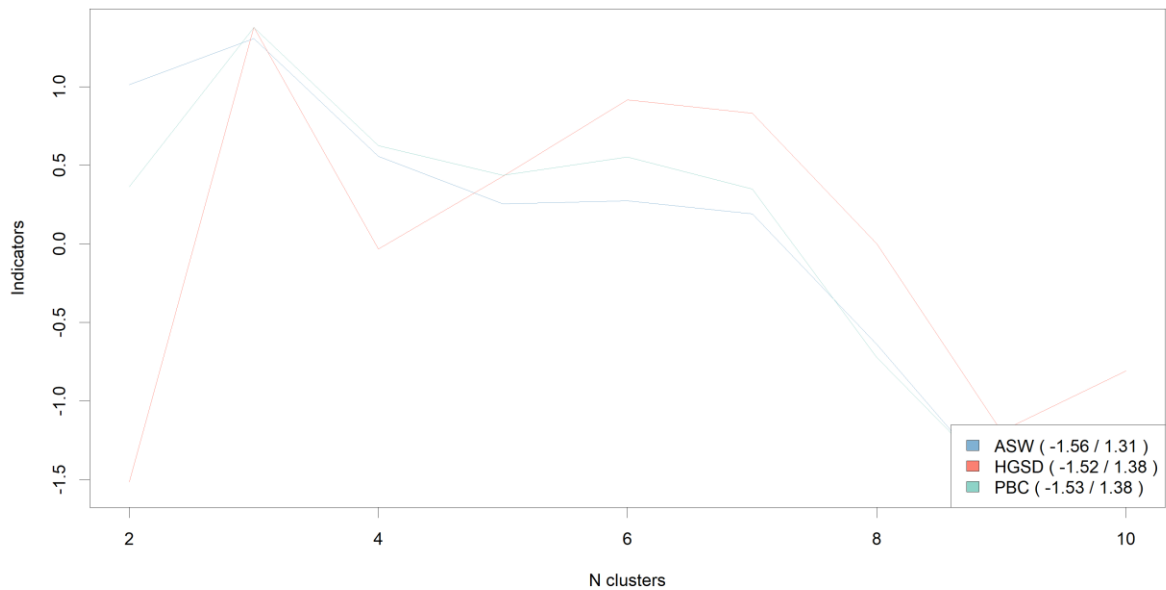


Figure A2: Cluster Solution Quality Criteria for Employment Sequences



Note on Relative Frequency Sequence Plots:

Relative frequency sequence plots¹ are generated by (1) sorting the sequences, (2) dividing the sorted sample into subgroups, (3) choosing medoid sequences from the subgroups to represent them, (4) plotting the medoid sequences, and (5) plotting dissimilarities from the medoid sequences as boxplots with R² and F statistics that evaluate the goodness of fit of the chosen medoid sequences. We sort the sequences using multidimensional scaling and divide the sample into 100 medoid sequences. The dissimilarities are calculated using LCS distance.

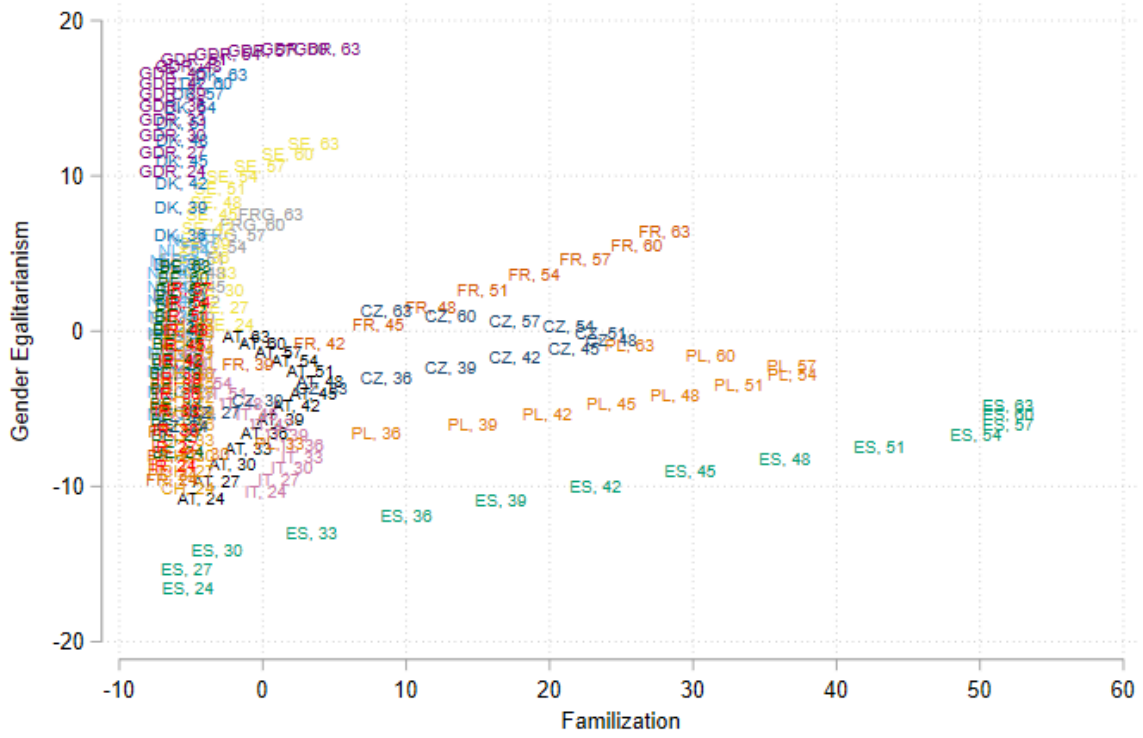
¹ The plots were created with the seqplot.rf function developed by Matthias Studer, Anette Fasang, and Tim Liao implemented in the TraMineRextras package using R, version 3.2.0.

2. Familization

As an additional analysis, we estimated our models with a familization index rather than a defamilization indicator. This indicator is calculated in the same manner as the defamilization index, but includes annual information on the availability of unconditional family allowances (Gauthier and Monna 2004) as well as the standardized duration of unpaid parental leave, and parental leave reserved for mothers (DICE 2015). We find that familizing policies vary across country-cohorts to a lesser degree than defamilizing policies (see Figure A3). Similar to our analyses with defamilization, our findings show that familization and gender egalitarianism tend to influence the association between family and employment life course patterns independently of one another (see Figure A4). However, familization is associated with a higher probability for women in the early family formation, singlehood, and divorce clusters to exit the labour market in low egalitarianism contexts. Moreover, familization is associated with a lower probability for women in the singlehood and divorce clusters to remain in stable full-time employment in low egalitarianism contexts.

Figure A3: Country-Cohorts by Familization and Gender Egalitarianism

Panel A



Panel B

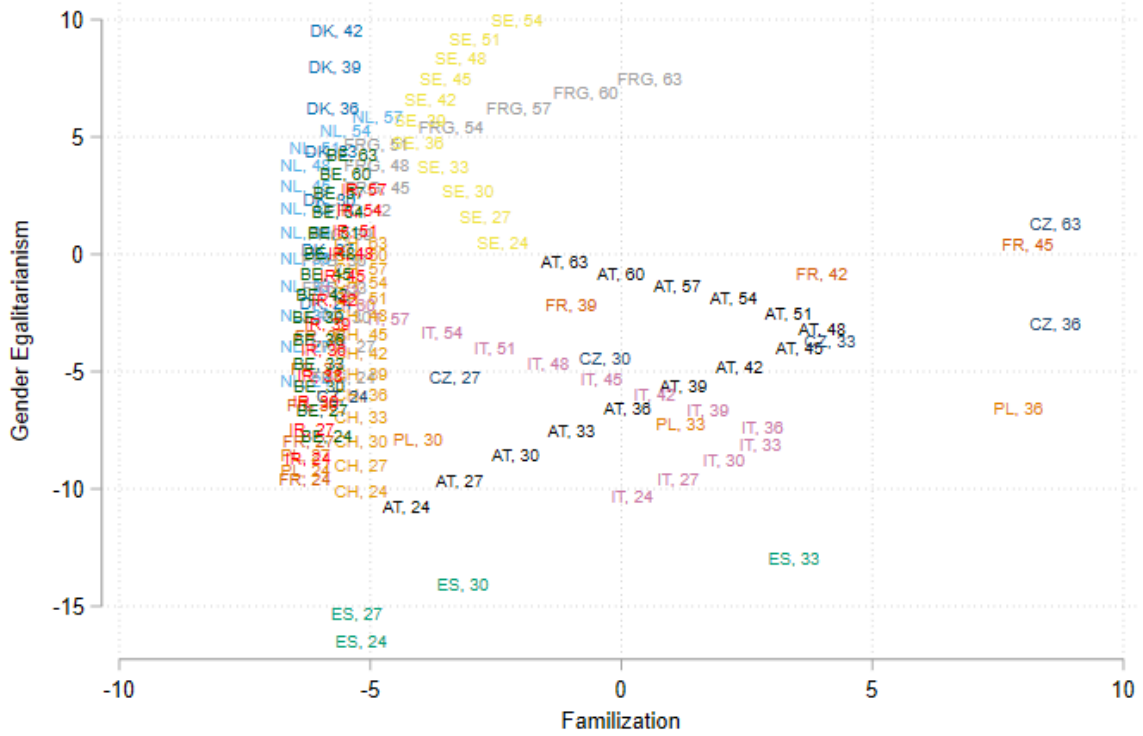
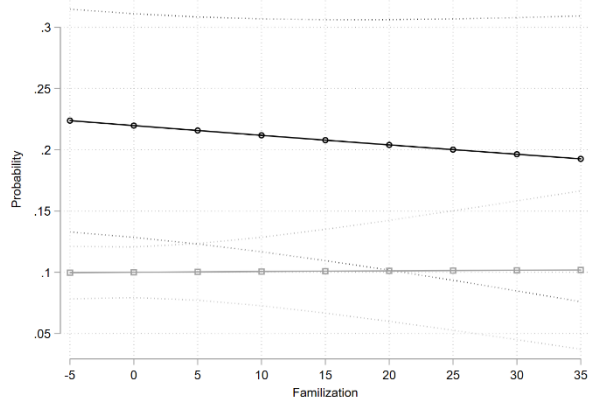


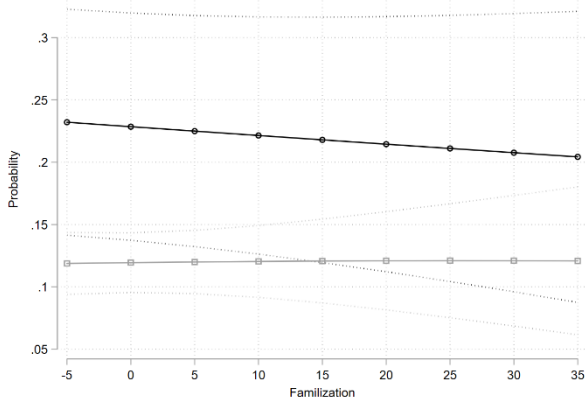
Figure A4: Predicted Probability of Employment Cluster Membership by Family Cluster, Defamilization, and Gender egalitarianism

A) Re-Entry into Part-Time

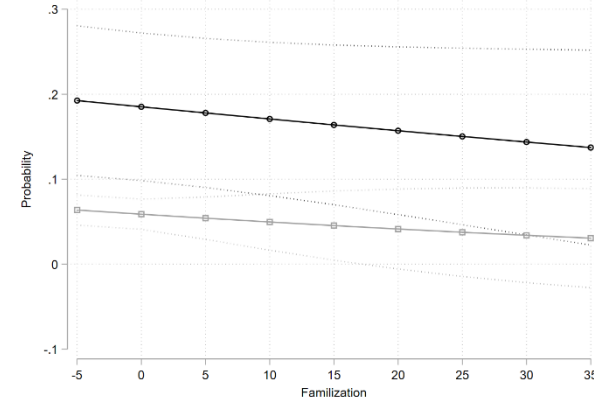
Delayed Family Formation



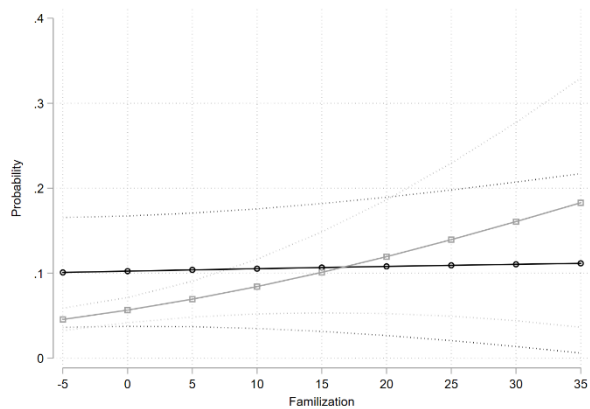
Early Family Formation



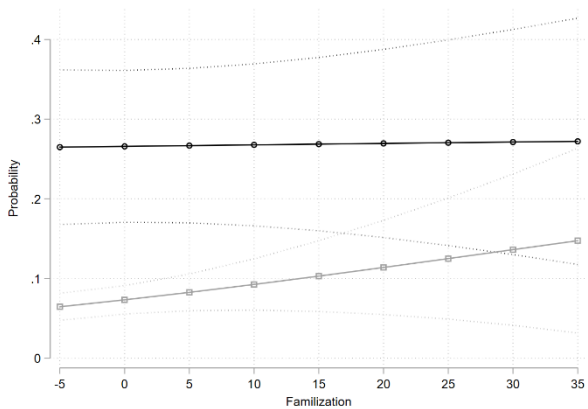
Childless Marriage



Singlehood



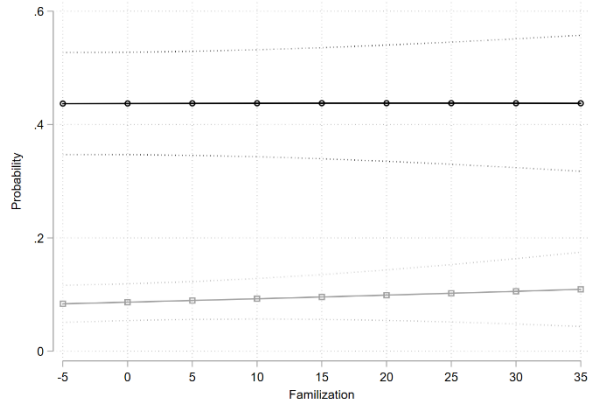
Divorce with Children



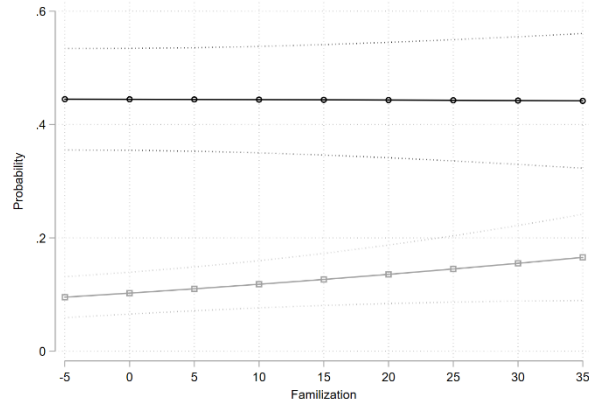
○ Low Egalitarianism □ High Egalitarianism

B) Labor Market Exit

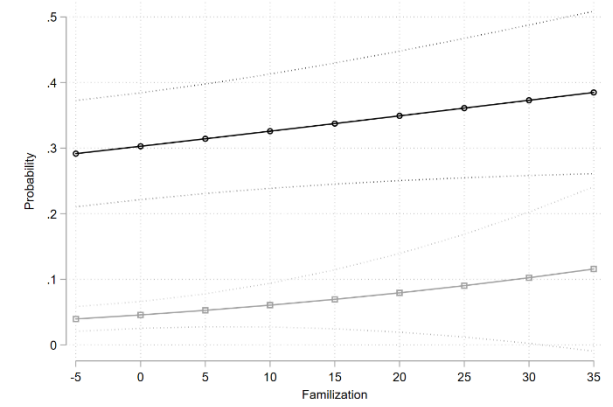
Delayed Family Formation



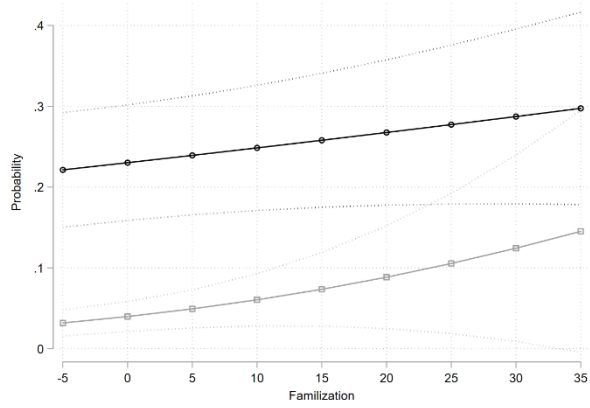
Early Family Formation



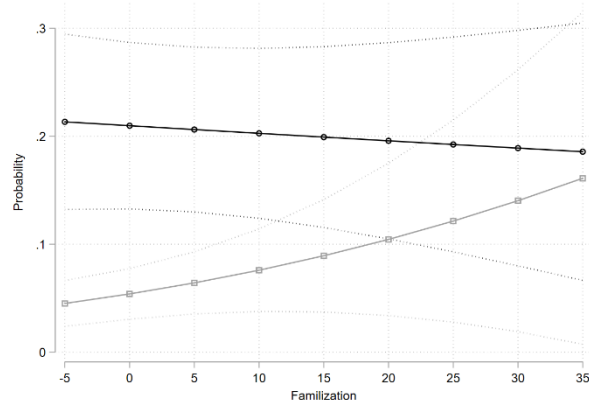
Childless Marriage



Singlehood



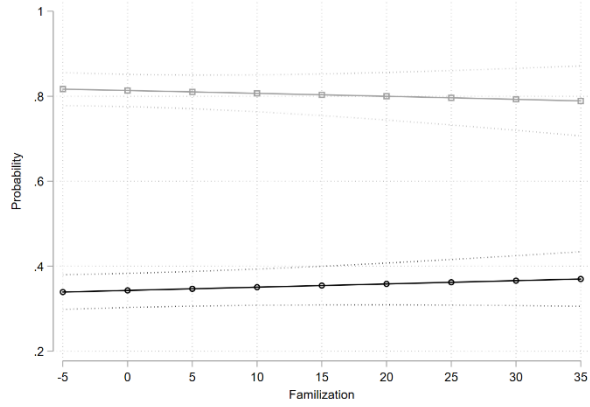
Divorce with Children



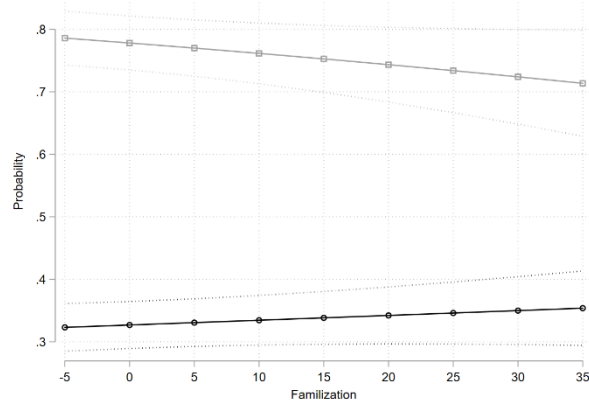
○ Low Egalitarianism □ High Egalitarianism

C) Stable Full-Time Employment

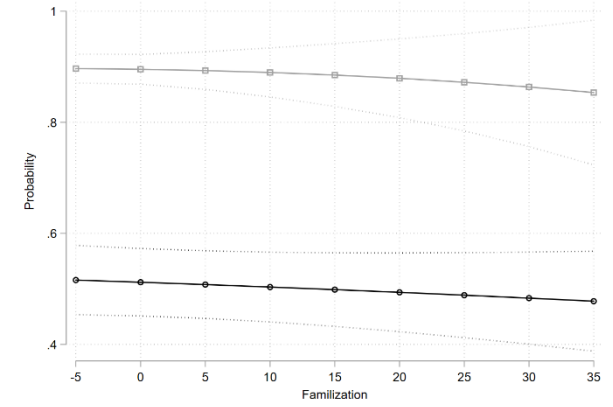
Delayed Family Formation



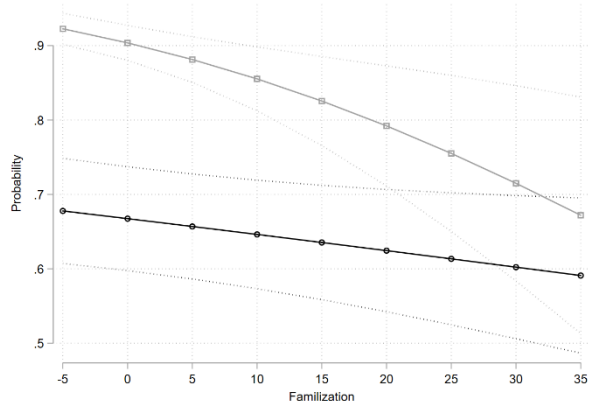
Early Family Formation



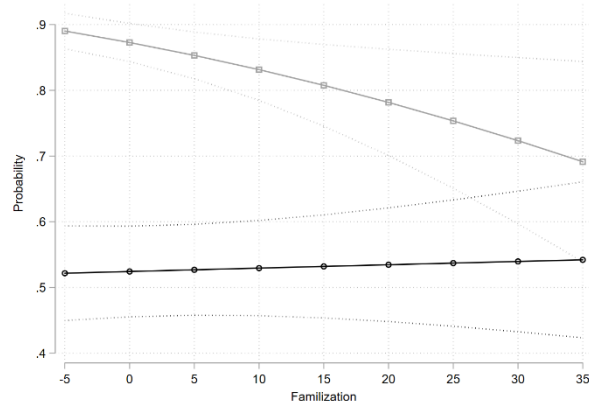
Childless Marriage



Singlehood



Divorce with Children



○ Low Egalitarianism □ High Egalitarianism