Mechanisms of family formation

Han, Sapphire Yu; Liefbroer, Aart C.; Elzinga, Cees H.

Published in:
Advances in Life Course Research

DOI:
10.1016/j.alcr.2019.03.001

Advances in Life Course Research 43 (2020) 100265

Contents lists available at ScienceDirect

Advances in Life Course Research

journal homepage: www.elsevier.com/locate/alcr

Mechanisms of family formation: an application of Hidden Markov Models to a life course process

Sapphire Yu Han[a,b,*], Aart C. Liefbroer[a,c,d], Cees H. Elzinga[a,d]

[a] Netherlands Interdisciplinary Demographic Institute (NIDI/KNAW), The Hague, the Netherlands
[b] University of Groningen, the Netherlands
[c] University Medical Centre Groningen, University of Groningen, The Netherlands
[d] Vrije Universiteit Amsterdam, the Netherlands

ABSTRACT

Life courses consist of complex patterns of correlated events and spells. The nature and strength of these correlations is known to depend on both micro- and macro-covariates. Life-course models such as event-history analysis and sequence analysis are not well equipped to deal with the processual and latent character of the decision- making process. We argue that Hidden Markov Models satisfy the requirements of a life course model. To illustrate their usefulness, this study will use Hidden Markov chains to model trajectories of family formation. We used data from the Generations and Gender Programme to estimate Hidden Markov Models. The results show the potential of this approach to unravel the mechanisms underlying life-course decision making and how these processes differ both by gender and education.

1. Introduction

A life course is an individual's narrative of all events or states, ordered in time, that this individual has experienced. In the course of the last decades, a series of important holistic[1] and non-holistic methods have been developed that increase our ability to understand the life course and its correlates. Event history analysis (EHA) uses regression models to predict the occurrence of particular (combinations of) life-course events (Blossfeld, Golsch, & Rohwer, 2007). Sequence analysis (SA) (Cornwell, 2015) is mostly used to find the characteristic patterns in a set of life courses and thereto, SA creates a distance-based representation of this set of life courses. Eventually, this spatial configuration can be used to examine the association between a set of covariates and the final classification (Studer, Ritschard, Gabadinho, & Müller, 2011). Clearly, methods like EHA and SA are indispensable. What is lacking, though, are statistical models that allow us to understand how the observed life-course patterns are generated. In this article, we discuss Hidden Markov Models (HMMs) as a promising class of models that can be used to understand how life-course patterns are generated and vary according to personal or contextual characteristics. In a nutshell, an HMM describes a life course as a sequence of observed states that are outcomes of a latent, unobservable decision making process. HMMs have a number of characteristics that make them very attractive to study how life-course sequences come about. At a practical level, an advantage of using a model that examines transitions between latent states is that the number of states and the number of transitions to be modelled is much smaller than if one would have to model all possible transitions between all possible observed states. For example, if one has an observed state-space of 10 positions, the number of possible transitions (including self-transitions) is 100, which is clearly unmanageable. If this system of transitions can be represented by four latent states, the number of possible transitions reduces to 16, which is much more manageable. However, using HMMs is also very attractive because its characteristics nicely mirror key theoretical assumptions of the life-course paradigm. A first key assumption of the life-course approach is that individuals exert life-course agency, and thus reflect on their available options and consciously make decisions on appropriate action sequences (Hitlin et al., 2007). Life courses can be conceptualized as the result of individual decisions about, for instance, whether and when to leave the parental home, to finish education before entering parenthood, to change jobs or to move to another residence (Elder et al., 2003). Such individual decisions are the result of unobservable mental processes. Thus, any model that is to explain the individual life course has to contain a mechanism that represents these unobservable mental processes. Second, it is generally held that events and choices made early in life may affect the (non-)occurrence of other

* Corresponding author at: Netherlands Interdisciplinary Demographic Institute (NIDI/KNAW), The Hague, the Netherlands.
E-mail address: han@nidi.nl (S.Y. Han).

[1] We say that a method is 'holistic' if it allows for the derivation of sentences pertaining to any data-instance whose truth is decidable, given the procedures of the method or the assumptions of the model. According to this definition, classifiers like Sequence Analysis and Latent Class Analysis are holistic and Hazard Models are not.

https://doi.org/10.1016/j.alcr.2019.03.001
Received 30 April 2018; Received in revised form 17 February 2019; Accepted 2 March 2019
Available online 29 July 2019

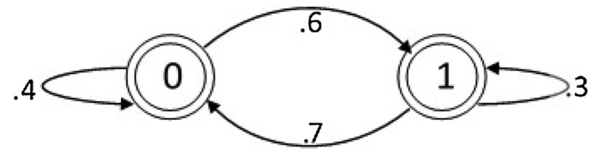
events or their outcomes in later life (Bissell, 2000; Gangl, 2004; Mayer, 2009). Therefore, a holistic model of the life course should have a memory in the sense that it contains a mechanism that makes earlier outcomes affect later life. Third, we know that the life course and its outcomes depend on micro- and macro-level covariates that are not affected by individual mental processes (Blossfeld, Klijzing, Mills, & Kurz, 2006; Specht, Egloff, & Schmukle, 2011). A life-course model should therefore also allow the inclusion of relevant time-constant and time-varying micro- and macro-level covariates. HMMs, formulated on the basis of a latent, hidden, random process over a finite set of states (a Markov chain), nicely satisfy all three of these assumptions of the life-course paradigm. Latent states can be viewed as key nodes in the mental decision-making process, emphasizing the agentic nature of the life-course process. Furthermore, HMMs have a memory in the sense as intended, and they allow for the inclusion of time-constant and time-varying micro- and macro-level covariates. The parameters of these models can be estimated (Bartolucci, Farcomeni, & Pennoni, 2012) and easily allow for causal analysis once formulated as a log-linear regression model (Paas, Vermunt, & Bijmolt, 2007). In particular, the Latent Class (LC)-model is a special case of the HMM in the sense that there is no state-switching possible within an LC-model. Precisely for this reason, the LC-model cannot be used to represent a latent dynamic decision process; HMMs can. HMMs belong to a larger family of latent structure models that has been amply described (e.g. Vermunt, 1997). HMMs hold great theoretical promise, but are, at the same time, not always easy to use. Using them asks for a large number of decisions to be made, for instance about the number of latent states to retain and their interpretation and how to link covariates to the model. These decisions have to be based on a combination of theoretical and statistical considerations. This paper aims to demonstrate how HMMs can be used, by applying them to modeling the family formation process. Next to the school-to-work transition, family formation is a key aspect of the transition to adulthood (Buchmann & Kriesi, 2011). Specifically, we focus on the family-life trajectories of French men and women born between 1956 and 1965, using data from the French Generations and Gender Survey (GGS). The paper is structured as follows. Given that many life-course researchers may be quite unfamiliar with Hidden Markov Models, we first discuss their main concepts. We try to do so in a relatively non-technical manner and with a focus on their application in life-course research. However, the use of technical language cannot completely be avoided. In section 3 we discuss our data and the methods used for our illustrative example. In section 4 we discuss the main results and in section 5 we summarize our results, try to draw both substantive and methodological conclusions and make suggestions for further research.

2. Hidden Markov Models

In this section, we discuss Hidden Markov Models and their application in life-course research. In section 2.1 we briefly introduce the basic aspects of HMMs. In section 2.2, we discuss some key issues if one wants to apply these models in life-course analysis. In section 2.3, we discuss some practical aspects of estimating these models in life-course research.

2.1. The basic HMM

A Markov-model or Markov-chain is a random process over a set of states such that the probability of being in a particular state at the next observation only depends on the state-history of the process. If the relevant state history just consists of the present state, such a chain is called "first-order". Fig. 1 shows a graphical representation of a first-order 2-state Markov-chain and its matrix of transition probabilities. Let us denote the k distinct states of a Markov chain as $Q = \{q_1, \ldots, q_k\}$ and let S_t denote the state that the system is in at time t, i.e. S_t could have any of the "values" or labels from the set Q. Then we say that a



$$A = \{a_{ij}\} = \begin{pmatrix} .4 & .6 \\ .7 & .3 \end{pmatrix}$$

Fig. 1. A graph, showing a first-order, 2-state Markov chain and its transition probability matrix A. The states are labeled as "0" and "1" and the arrows represent the transition probabilities.

random process over Q is a first-order Markov-chain, precisely when

$$Prob(S_t = q_j | S_0 \ldots S_{t-1}) = Prob(S_t = q_j | S_{t-1}) = a_{ij} \tag{1}$$

If we now define the initial state probabilities as $Prob(S_0 = q_i) = \pi_i$, the Markov-chain λ is fully defined by the k -vector $\boldsymbol{\pi} = (\pi_1, \ldots, \pi_k)$ of initial state probabilities and the $k \times k$ -matrix of transition probabilities $\boldsymbol{A} = \{a_{ij}\}: \lambda = (\pi, \boldsymbol{A})$.

In a Hidden Markov Model, the Markov chain is defined over a set of latent, unobservable states. Furthermore, it is supposed that, at each state, the process "emits" an observable according to a state-specific probability distribution over the full set of observables, in the present context the observable states of a life course. Thus, in a k -state HMM with a set of observables $O = \{o_1, \ldots, o_n\}$, there must be a set B of k state-specific probability distributions $\boldsymbol{b}_j = (b_{j1}, \ldots, b_{jn})$. We write the likelihood of observation o_t at time t as

$$b_{ji} = Prob(o_t = o_i | S_t = q_j). \tag{2}$$

This allows us to represent the set B as a $(k \times n)$ -matrix

$$\boldsymbol{B} = \begin{pmatrix} b_{11} & \cdots & b_{n1} \\ \vdots & \ddots & \vdots \\ b_{1k} & \cdots & b_{nk} \end{pmatrix} = \begin{pmatrix} \boldsymbol{b}_1 \\ \vdots \\ \boldsymbol{b}_k \end{pmatrix} \tag{3}$$

where each row is a distinct probability distribution over the observables and the complete HMM $\lambda = (\boldsymbol{\pi}, \boldsymbol{A}, \boldsymbol{B})$ is specified by the initial state distribution $\boldsymbol{\pi}$, the $(k \times k)$ - matrix $\boldsymbol{A}$ of transition probabilities and the $(n \times k)$ -matrix $\boldsymbol{B}$ of emission probabilities.

In Fig. 2, we show a graph of the HMM-generated events in a time-window $(t - 1, t + 1)$: at $t - 1$, the system arrives in state S_{t-1} and emits observable o_{t-1} (governed by $\boldsymbol{B}$) and then switches to state S_t (governed by $\boldsymbol{A}$) and again emits an observable, etc..

When we apply HMMs to model life courses, we know that some states are almost irreversible: for example, once parenthood is entered, it is a lasting state except for the rare cases of child loss. Models in which a return to a previously occupied state is impossible or highly unlikely are called "left-to-right" models.

The theory of log-linear models in relation to HMM's has been extensively dealt with in e.g Bartolucci et al. (2012). Under such a model, we then have that, for example

$$log\frac{Prob(S_t = q_j | S_{t-1} = q_k, \boldsymbol{v})}{Prob(S_t = q_r | S_{t-1} = q_k, \boldsymbol{v})} = \alpha_{jr} + \boldsymbol{\beta}'_{jr}\boldsymbol{v}_i + \varepsilon_i \tag{4}$$

wherein q_r is the reference state. Similar models can be obtained for the log-odds of the initial state probabilities and the emission distributions. So, by expressing an HMM as a log-linear model, the regression-weights estimate the effect of the covariates on the log-odds. Parameter estimation is usually carried out by an EM-algorithm implemented in the form of the forward-backward algorithm. Once parameter estimates have been computed, standard errors are commonly associated to these estimates. For more information on the estimation of standard errors, the reader is referred to Bartolucci et al. (2012).

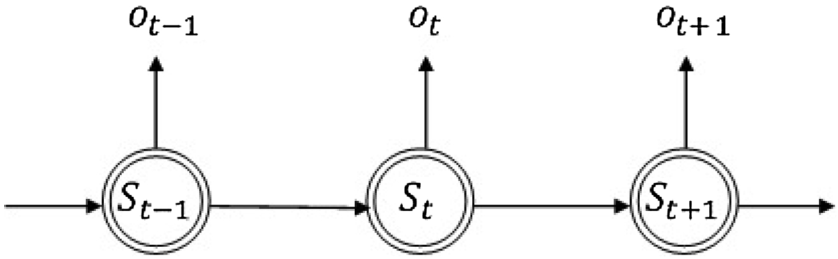


Fig. 2. A graph showing the time-window ($t-1,\ t+1$) of a Hidden Markov process. At each time t, the system is at some latent state S_t and emits an observable. Note that the hidden state S_t is not necessarily different from S_{t+1}. The observable is a random sample from the set of observables, according to a probability distribution that is specific for each state q_i, $i = 1, \ldots, k$

2.2. HMM-assumptions and life course research

A number of key decisions have to be taken if one wants to apply HMMs to life-course research. In this section, we highlight four of these. First, we discuss the extent to which a focus on first-order Markov chains (that is a situation where transition probabilities between latent states only depend on the previous latent state) limits the use of HMMs. Second, many processes in young adulthood are age-dependent, and we discuss ways in which age-dependency can be accounted for in HMMs. Third, deciding on the number of latent states to be included in a HMM is discussed. Finally, we discuss ways in which covariates can be linked within HMMs.

The assumption that the process is first-order is a gross simplification: we know that some events early in the life course may have lasting effects long after their occurrence (see e.g. Gangl, 2004, on the scarring effects of unemployment). However, extending the memory to $m \geq 1$ previous states will result in a transition probability matrix of size $k \times k^m$ since there will be distinct transition probabilities for each of the k^m possible state-histories $S_{t-1} \ldots S_{t-m-1}$. Several solutions have been proposed to reduce the size of this parameter space. Berchtold and Raftery (2002) proposed the so-called Mixture Transition Model that reduces the number of state parameters by considering each of ℓ lags separately and approximating their effects by ℓ lag-weights. This approach has been applied to estimating higher-order latent Markov chains by Berchtold (2002). Another approach (Eggeling, Gohr, Bourguignon, Wingender, & Grosse, 2013; Mächler & Bühlmann, 2004) retains only specific higher-order paths of possibly variable length.

So, in principle, estimable extensions of HMMs to higher-order latent chains are available but still, the computational burden is far from trivial. However, more important is that we do not know how to use such models in a theory-driven way: at present, we do not have specific ideas about how various autocorrelations come about, and thus modeling them through a Markovian model would amount to quite a data-mining operation. Therefore, we choose to limit our modeling to first-order Markov chains and focus on other aspects of the theory of life course generation; such first-order chains at least allow for some propagation of long-term effects through the system.

A second important aspect of a basic HMM is the fact that age does not play a role in the parametrization of the model: the transition-rates in the transition-matrix are constant over time whereas we know that certain transitions are highly age-dependent (Fasang, 2012). However, when an HMM is used to study the life courses of a cohort that is not too wide, the transition parameters in fact estimate rates for roughly the same age-groups: the waiting-times for transition are, in a basic left-to-right HMM, geometric distributions that sharply taper off over time. So, age dependency is implicit in such models. However, it is possible to specify other distributions of the state durations (e.g. see Dewar, Wiggins, & Wood, 2012). We did not use such models because, again, we do not have a clear hypothesis on duration dependence in our application. Another option is simply to add age as a covariate to the model and examine whether it influences the transition rates between latent states and whether it changes the estimates of other parameters of interest. For our application presented below, we estimated models that include age and age squared as covariates. Including these covariates did not change the estimates of any of the other covariates of interest in any substantive way. Estimates for this alternative model are available upon request.

A third important aspect of an HMM is determining the number k of latent states. This number has to be fixed by the researcher; it is not a free, estimable parameter. So, when we believe that an HMM is a valid model, the next step is to compare HMMs with different numbers of latent states and use the BIC to select the most parsimonious model (see Burnham & Anderson, 2002, Section 6.5.3). Unfortunately, the BIC-curve may not show a clear knee or will keep decreasing for ever bigger numbers of latent states. In such cases, the best alternative is to fix k at the optimal value that yields a substantively sound interpretation.

Finally, when using covariates in modeling with HMMs, one has to decide how these covariates affect the behavior of the stochastic system as a whole. This system consists of two main parts: the Markov chain over the latent states and the mechanism through which it expresses itself, i.e. the set of emission probability distributions, one for each of the latent states. Covariates may affect either or both of these components. For example, we know (Manning, Brown, & Payne, 2014) that both the timing and the type of first union are education dependent. Within the present context, timing of first union formation is a result of state transition, i.e. of demographic decision making, and type is the result of an emission distribution over observable relation types. Hence, one might postulate that education affects both components of the stochastic process. Other covariates might affect transition but not emission probabilities or vice versa. Unravelling all these intricacies is beyond the space limitations and scope of the present paper. Instead, we confine ourselves to demonstrating the effect of covariates on state transitions. The reason for this preference is that we know that most life courses in developed countries mainly differ in the timing and duration of the various stages on the route to adulthood (Billari, Fürnkranz, & Prskawetz, 2006; Halpin, 2010). Due to the relative stability of sequence order, in most countries, the behavioral alternatives and the order in which they are expressed are roughly the same for most people.

Of course, in modeling (facets of) the life course, the choice of observables does matter. Modeling family formation will change when "cohabitation" and perhaps other partnering relations are introduced and when modeling labor market careers, using finely grained scales of being (un-)employed or pensioned will affect the resulting models. We consider this to be a strength of HMM-modeling since choosing different observables reflects a theoretical position on the importance and possible impact of these observables.

2.3. Modeling and model-selection with HMM's: Some practical considerations

Let $O_i = o_{i1} \ldots o_{iT}$ denote an observed sequence from a set $O = \{O_1, \ldots, O_N\}$ of such sequences and let $Prob(O_i|\lambda)$ denote the likelihood of that sequence, given the model.

Furthermore, let $Q_i^* = q_{i1}^* \ldots q_{iT}^*$ denote the path along the latent states that maximizes $Prob(Q_i|O_i, \lambda)$, i.e. the latent sequence that "best accounts" for the observations, given the model.

Being able to calculate the likelihood of the observations given the model is a precondition for EM-estimation of the parameters of the model and calculating Q_i^*, the most probable latent sequence, is a precondition for a substantive interpretation of the model.

Both problems, evaluating $Prob(O|\lambda)$ (Baum, Petrie, Soules, & Weiss, 1970), and calculating Q_i^* (Viterbi, 1967) were already solved in the sixties of the previous century. Here, we will not deal with the intricacies of these methods. Instead, we will discuss some practical issues that are related to these methods and their output.

One should be aware that evaluating an HMM involves the estimation of quite some parameters: with k postulated latent states, we have to estimate $k - 1$ estimated initial state parameters $\hat{\pi}_i$; $k - 1$ since we must have that $\sum_i^k \hat{\pi}_i = 1$. Likewise, we have to estimate $k(k - 1)$ parameters to obtain the matrix of estimated state transition probabilities $\hat{\boldsymbol{A}}$ and $k(n - 1)$ parameters to get the matrix of estimated state-specific probability distributions $\hat{\boldsymbol{B}}$. Consequently, in total, we have to estimate $k^2 - 1 + k(n - 1)$ parameters for a model with k latent states and n observables. When estimating a model with covariates, this number is multiplied with the number of covariate-value combinations. There are two problems associated with this big number of parameters. First, it implies that testing the adequacy of a (Hidden) Markov Model is practically impossible (Eggar, 2002): almost always the parameters can be chosen such that the model fits the data. All that can be tested is the relative efficiency of different models for the same data using likelihood-ratio tests (Giudici, Rydèn, & Vandekerkhove, 2000). Second, the surface of the likelihood function $Prob(O|\lambda)$ is quite irregular and therefore, attempts to find its maximum will most often converge to a local instead of the global maximum. Extending the HMM to incorporate covariates will only aggravate this problem. Therefore, the estimation of an HMM should be repeated many times to find a configuration $(\hat{\boldsymbol{\pi}}, \hat{\boldsymbol{A}}, \hat{\boldsymbol{B}})$ that (probably) comes close to the maximum sought for. For example, we display the density of the likelihood maxima as obtained over 1000 repetitions of estimating a 4-state model in Fig. 3.

Clearly, these values are quite different, as are the underlying configurations $(\hat{\boldsymbol{\pi}}, \hat{\boldsymbol{A}}, \hat{\boldsymbol{B}})$. Obtaining this curve took almost two hours of computation time and quite some memory. Increasing the number of states and the number of trials soon requires unfeasible computation time and memory. However, in life course modeling, this should not pose a problem as the number of postulated latent states is small when these states are interpreted as pertaining to demographic choices - there are only few of these (but see our remarks on data encoding).

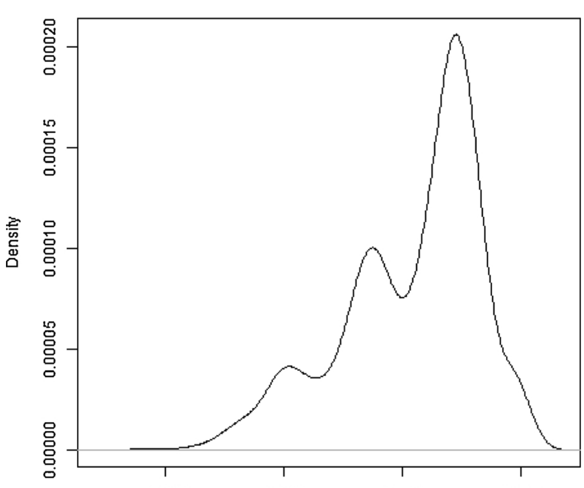


Fig. 3. Likelihood-density plot as obtained from repeating the estimation of a 4-state HMM 1000 times with random initial values. The horizontal axis shows the estimated likelihood $\hat{\mathscr{L}}$.

2.3.1. Model selection and validation

With mixture models, model fit is to be judged through evaluating the likelihood $\mathscr{L} = Prob(O|\lambda)$ of the data O, given the model λ. However, by allowing for an ever bigger number of parameters k, $\mathscr{L}$ can be made arbitrarily close to 1. So, most criteria[2] to judge model-fit penalize with some function of k, like for example in $BIC = 2\,log + k\,ln\,N$, wherein, in our application, N denotes the number of sequences. So, in practice, selecting a model amounts to balancing $\mathscr{L}$ and model parsimony. We then hope that this "balance" unveils itself through a clear knee in the plot of BIC vs k. If such a knee is absent, unclear or occurs at a value of k that is beyond interpretation, such criteria are not sufficient. In such cases, one has to select the best-fitting model within the range of k that can be sensibly interpreted. The next step then is logical inference: to see if the model fits in with other relevant facts that do not belong to the data used to estimate the model and it is here that covariates come in.

Covariates like gender, SES and religion and macro-variables like welfare regime and the occurrence of natural disasters, economic crises or political instability are generally held to affect demographic decisions in specific, well researched ways (see e.g. Neels, Theunynck, & Wood, 2015; Härkönen & Dronkers, 2006; Sobotka & Toulemon, 2008; Studer, Liefbroer, & Mooyaart, 2018) Thus one should test whether or not the selected k-state model allows for logical inference: i.e. whether or not the model is able to reproduce these effects. If such tests fail, the model should be rejected. If such tests do not fail, i.e. when the model is capable of reproducing known effects, the model has become one of the nodes in a nomological network (Han, Liefbroer, & Elzinga, 2017; Torgerson, 1958) about life courses and related phenomena. Perhaps, hopefully, such a model will be succeeded by a model that allows for statistical inference too.

2.3.2. Interpreting HMM's

How do we interpret the latent states? Interpreting the latent states always implies combining information from all three transition matrices within HMMs. The emission probability matrix tells us which observed states are most clearly linked to a latent state, the initial state matrix tells us whether there is a clear starting state or not, and the transition probabilities matrix tells us which transitions between latent states are likely and which are not. Again peeking around the corner of our analysis yet to be presented, we show a plot of the shift in the distribution of latent states over time for a 4-state HMM in Fig. 6: this plot shows something that is not immediate from the estimated transition probabilities: most subjects start in the latent state labeled as LS1, so LS1 most probably is to be associated with a decision about leaving the parental home. Indeed, the emission probabilities in Table 4 suggest that LS1 is characterized by young adults being in the parental home without partner and/or child, whereas the emission probability of being in the parental home is close to zero for each of the other latent states. So, the marginal state occupancies over time and the emission distributions will help us to interpret the latent states. However, these considerations do not suffice for a credible interpretation.

A credible interpretation can only arise in the light of the way covariates affect the parameters of the model: do the estimated effects of covariates corroborate, or at least are not at variance with, the knowledge that we already have about the effects of these covariates on the occurrence and timing of life-course events. For example, we may expect that low-educated will enter parenthood earlier than high-

[2] We observe that BIC heavily penalizes likelihood for over-fitting in big datasets (N) with too many parameters (k), i.e. with too complex models. An alternative to BIC is AIC, which is more lenient with respect to complex models. For a detailed comparison of AIC and BIC, the reader is referred to (Burnham & Anderson, 2002, especially Ch. 6). The conclusions to be presented in this paper are not affected in any substantive way by using AIC instead of BIC (the pertaining results are obtainable from the first author).

Table 1
Information on the distribution across key dependent and independent variables, and mean age at the time of key observable family formation events (N = 1900)

Observables	Category	%	Mean ages
Fertility	0	19	
(# children)	1	22	27
	2	37	30
	> 2	22	32
Partnership	Single		
	Married		25
	Cohabiting		24
Left home	yes		21
	no		
Covariates			
Education	low	81	
	high	19	
Gender	male	44	
	female	56	

educated, which is the case when the transition probabilities for lower-educated are higher across the whole age range than the transition probabilities of the higher-educated. Therefore, it is not enough to only evaluate a HMM as such: we need to enrich the model with relevant covariates in order to decide on the credibility of the interpreted model.

3. Data and Method

3.1. Data

The Generations and Gender Programme (GGP) is a longitudinal survey of 18-79 year olds in nineteen countries that examines the relationships between generations and genders, by collecting representative data in all participating countries. Fokkema, Kveder, Hiekel, Emery, and Liefbroer, (2016) provide extensive information on the design and representativeness of the GGS. From this study, we selected males and females from the French GGP, 1900 in total, born between 1956 and 1965. For this subset of the data, information is available for ages between 15 and 40, on annual fertility, partnering and leaving the parental home as well as background information on gender and the level of education.

From these data, we constructed the family formation history as a convolution of three trajectories: we reconstructed fertility histories (four categories: no, one, two or more than two children), partnership histories (three categories: single, cohabiting or married) and a binary trajectory for having or not having left the parental home. Such convolutions are known as "multichannel sequences" (Gauthier, Widmer, Bucher, & Notredame, 2010). The two background variables were categorical too: male or female and two categories for educational level: high or low. We tabulated these data characteristics in Table 1.

In Fig. 4, we show the sequence index plots of the three channels, each plot sorted by the final stage[3] .

3.2. Method

In all analyses, we estimated HMMs for the data as described above, using EM in the form of the Baum-Welch algorithm (see e.g. Bartolucci, Pandolfi, & Pennoni, 2017). To reduce the risk of getting trapped on a local maximum of the likelihood function, each estimated model was picked as the best solution out of 1000 trials, each starting with a

[3] We could have produced plots where in each plot, the subjects are sorted according to one and the same criterion (Helske & Helske, 2017). We do not show such plots because they give a less clear picture of the distribution of patterns in the sample.

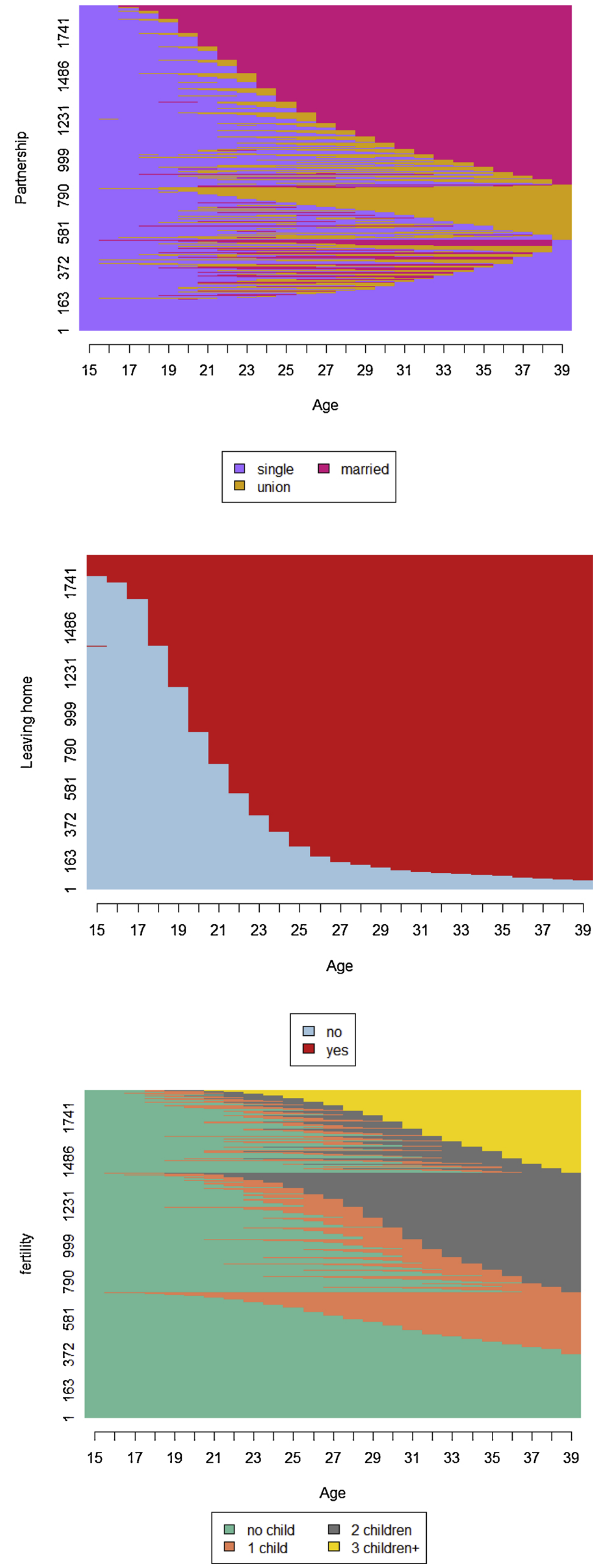


Fig. 4. Sequence index plots of the three channels of family formation: leaving home (upper panel), union formation (middle panel) and fertility (lower panel).

Table 2
Estimated initial probability distribution of a 4-state HMM

Latent State	1	2	3	4
	.94	.06	0	0

randomly chosen set of initial parameter-values. The covariates were only used in the estimation of the 5-state model.

The algorithms used were implemented in the R-based LMest (Bartolucci et al., 2017) and markovchain (Spedicato, Kang, Yalamanchi, & Yadav, 2014). We used TraMineR (Gabadinho, Ritschard, Müller, & Studer, 2011) for visualization.

4. Results

In this section, we present results of HMMs that differ in their number of latent states and discuss their interpretation. Below, we will denote an HMM with k latent states as a "k-state HMM". An estimated k-state HMM will consist of an estimated k-vector of initial state occupancy probabilities, an estimated $k \times k$ -matrix of state transition probabilities and k estimated emission probability distributions. For each HMM with k states, we discuss the model with the lowest BIC value selected from a set of 1000 estimated k-state models.

We fitted models with k varying from 3 to 8, but for reasons to be explained later, we will focus on models with 4 and 5 states and, very concisely, on a model with 6 states. We amply discuss the 4-state model although its BIC is appreciably higher than the BIC for the 5-state model. We do this to show what substantive reasons also make us favor the 5-state model.

4.1. A 4-state HMM

Key parameter estimates of 4-state HMM are shown in Tables 2–4 and graphically depicted in Fig. 5. As the initial probability distribution in Table 2 shows, only two of the four postulated states are initially occupied and the vast majority, 94% of the respondents is estimated to be in Latent State 1 (LS1). On its diagonal, the state-transition matrix (Table 3) shows the parameter a_{ii} of the postulated geometric waiting time distribution for a transition to another state. For example, it is estimated that $a_{11} = .85$ and hence the probability that someone will stay in LS1 for precisely 6 years equals $.85^6(1.85) = .057$. From Table 3, one observes that the estimated model is a "forward"-model: if a transition from a particular state i occurs, it is almost invariably to state $i + 1$ or, rarely, to a state $i + 2$. This is visualized in Fig. 5: almost all respondents move across the states in the order LS1-LS2-LS3-LS4 and a small minority skips LS2.

Table 4 should be read column-wise: each column presents the emission probabilities of the observed states, given the latent states. It shows that, for example, the vast majority of those estimated to be in LS1 have no children, and those in LS3 are almost certain to have one child. On the other hand, all of those estimated to be in LS4 have at least two children. Similarly, of those in LS4, 76% is married and 99% has left the parental home.

To interpret this HMM and its latent states, one has to simultaneously examine the different types of probabilities and take into account that one has to interpret the latent states as cognitive states or processes in which demographic decisions are considered. Most respondents in LS1 still are in the parental home without a partner and without children, and given the transition probabilities they mainly move to LS2, a state characterized by "having left the parental home", having no children and a mix of partnership states. Thus the latent state interpretation of LS1 is that it seems to be the state wherein most youngsters consider or decide whether, and if so, when, to leave the parental home. For respondents in LS2, reproduction seems to be the key decision under consideration: none of those estimated to be in LS2

Table 3
Estimated transition probability distributions of a 4-state HMM

State	1	2	3	4
1	.85	.14	.01	0
2	0	.91	.09	0
3	0	0	.85	.15
4	0	0	0	1

Table 4
Estimated emission probability distributions of a 4-state HMM

State	1	2	3	4
# Children				
0	.99	1	0	0
1	.01	0	1	0
2	0	0	0	.68
> 2	0	0	0	0
Partnership				
Single	.97	.53	.17	.10
Cohabiting	.02	.27	.24	.14
Married	.01	.20	.59	.76
Left home				
No	1	0	.02	.01
Yes	0	1	.98	.99

is a parent yet, but once they transition to LS3 (and this is very likely to happen), they all have their first child and almost 85% of these respondents have partnered - in LS2, the latter percentage was still less than 50.

It seems that in LS3, the key mental decision process is about the kind of partnership and the number of children one wants. Thus, the key decision process is about family extension. Those in LS3 can only move to LS4, and if they do so, they expand their household with at least one more child. In LS4, almost 80% is married while only 10% has no partner. The observed demographic events and decisions one has to make at each latent state are shown in Table 5.

In summary, the 4-state HMM model suggests that the transition to adulthood is driven by a chain of demographic decisions pertaining to leaving the parental home (LS1), reproduction (LS2) and expanding the family and choosing an appropriate partnership (LS3). LS4 acts as a kind of absorbing state where no further (observed) demographic decisions are considered.

Fig. 6 shows the estimated fraction of each age group that occupies each of the four latent states. In Fig. 7, we show the most probable paths (Bartolucci et al., 2012, Section 7.5.2) along the latent states, given the estimated model and the observed categorical variables.

Fig. 7 shows that almost everyone has left LS1 by their late twenties, but that considerable fractions are still in LS2 and LS3 at the end of their thirties. Thus, some respondents have not started having children or have not decided on family expansion by their late thirties, indicating that they may end up with no or just one child.

This 4-state HMM models the transition to adulthood as a process that is mainly driven by fertility-related decisions, as there is no explicit room for decisions pertaining to the mode of partnering. According to Second Demographic Transition Theory (e.g. Lesthaeghe, 1995), one might expect a more prominent role for decisions concerning the type of partner relationships young adults engage in. Therefore, and because it has a much smaller BIC, we closely inspect the results of an estimated 5-state HMM in the next subsection.

4.2. A 5-state HMM

The initial state distribution (Table 6) is the same as that of the 4-state model and again, the model is almost perfectly "left-to-right" as all the transition probabilities below the diagonal of Table 7 are (very close

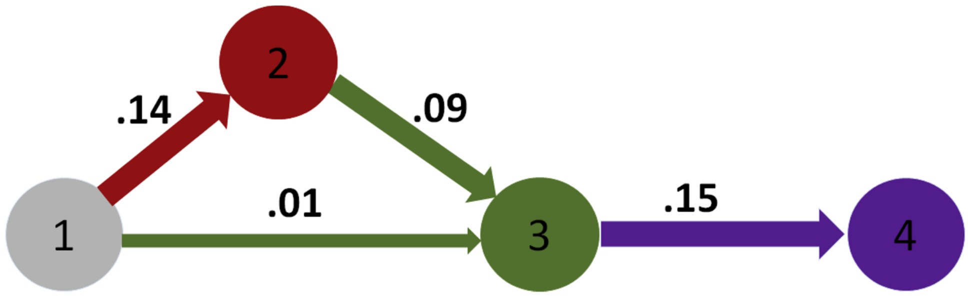


Fig. 5. State transition plot of an estimated 4-state HMM. The numbered circles represent identically numbered latent states and the arrows represent transition probabilities. To help interpret the figures in the sequel, we tagged the states with a color: grey ⟶ LS1, red ⟶ LS2, green ⟶ LS3, purple ⟶ LS4. The arrows have been colored according to the "target-state". The thickness of the arrows and the numbers above the arrows reflect the transition probabilities (self-transition probabilities not shown).

Table 5
Interpretation of latent states in terms of observed demographic events and mental decision processes in a 4-state HMM.

State	Observation	Decision Process
LS1	In parental home	Leaving parental home
LS2	Residential-independent	Family formation
LS3	One-child family	Family extension
LS4	Multi-child family	(absorbing state)

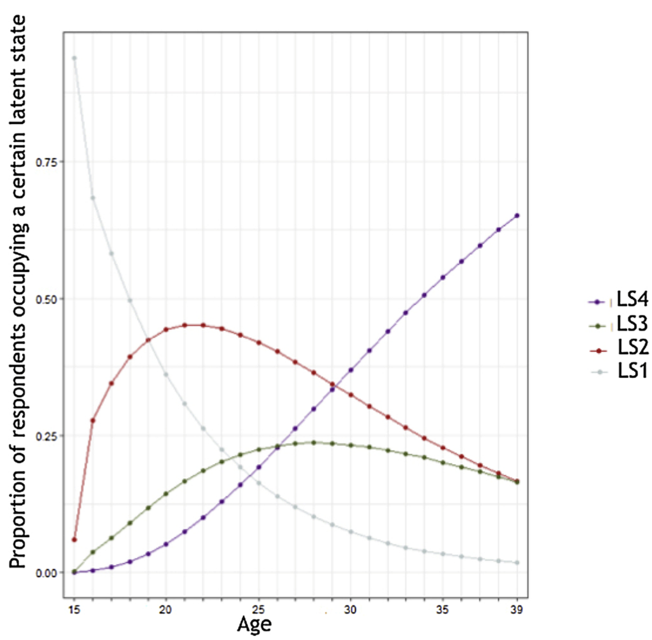


Fig. 6. Latent state occupancy fraction plot of an estimated 4-state HMM.

to) zero while the bigger off-diagonal probabilities are in the upper-diagonal part of the estimated matrix.

Following the same reasoning as when interpreting the 4-state model, here LS1 is a state wherein people consider when and how to leave the parental home. Unlike the 4-state model, though, two different options seem to be available, as LS2 is dominated by people either single or in cohabitation, whereas LS3 is dominated by people who are married.

People in LS2 again face a clear choice about how to continue their family formation process. Either they move to LS3, where marriage is the dominant living arrangement, or they move to LS4 where they have children but do not marry. This latter decision path that leads to becoming an unmarried parent, runs via LS2 to LS4, and is taken by 31.2% of the sample. According to this model, there are two decision paths to becoming a married parent. The first path, taken by roughly 35.1% of the sample, involves the decision to (first) cohabit and arrive in LS2, then to marry and arrive in LS3, and finally to decide to complete the family formation process with either one or two more children and arrive in LS5. A second, almost equally likely path starts with the decision to marry directly after leaving home and thus runs from LS1 directly to LS3, and next to LS5. According to this model, all other decision paths are quite unlikely.

Unlike in the 4-state model, in the 5-state HMM, there are two absorbing states: LS4 and LS5. Transitioning between both states is (in these data) a rare phenomenon. Unmarried parents end up in LS4 and then only rarely consider marrying, whereas married parents mostly end up with a completed family in LS5 and then rarely consider divorce or lose their partner. The observed demographic events and the underlying mental decision process at each latent state are shown in Table 9.

In Fig. 8, we visualize the decision paths by arrows whose thickness reflects the probability of moving between the respective latent states.

We show the estimated fraction of the respondents that occupies each of the latent states by age for this 5-state HMM in Fig. 9 and the most probable state paths in Fig. 10. As already stated in Section 2.3, BIC- or entropy-based model selection of HMMs for big social-demographic data sets should not be expected to be feasible. Indeed, we estimated HMMs with the number of latent states ranging from 3 to 8 and each successive addition of an extra latent state appeared to be associated with a smaller BIC-value, as shown in Table 10.

Reliably estimating models with even more latent states proved to be practically unfeasible. Here, we only discuss the estimated 4- and 5-state models since these models allow for an interpretation of the latent process that is sensible within the context of modern, more general social-demographic theoretical frameworks. Discussing HMMs with a bigger latent state space does not lead to new insights. To illustrate this, we present the transition plot of a 6-state model in Fig. 11 and the observed demographic events and decision processes in Table 11. In our view, this solution only adds complexity without offering new insights in the transition to adulthood. The main difference is that two latent states where married people have children are distinguished, one with one child and one with 2+ children. This does not add much to our understanding of the underlying family formation process.

In our assessment, the 4-state model does not adequately reflect modern demographic decision notions about the decline of traditional family values and the increased importance of autonomy, whereas the 5-state model does fit in with these ideas and has a much lower BIC as well. Therefore, in the next section, we confine to discussing the results obtained when one adds gender and educational level as covariates to the 5-state model.

4.3. Testing model-validity: the 5-state model with gender and education

It is known that demographic decision processes differ by gender

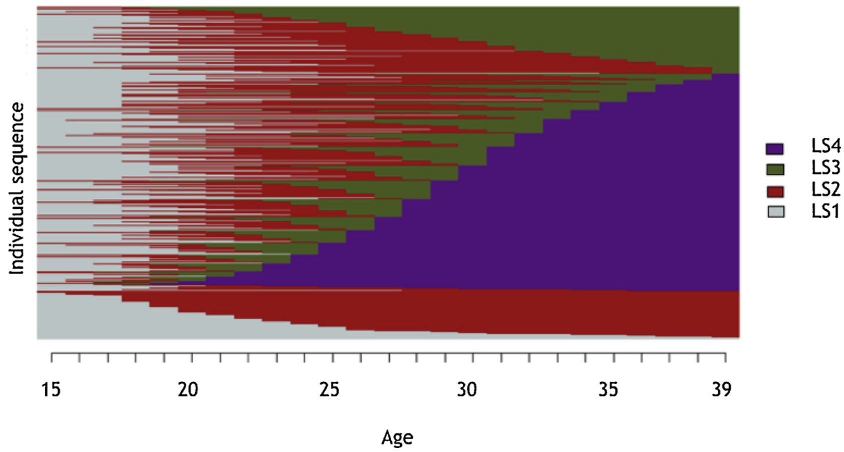


Fig. 7. Sequence index plot of the most probable latent state sequences, according to an estimated 4-state HMM.

Table 6
Estimated initial probability distribution of a 5-state HMM

State	1	2	3	4	5
	.94	.06	0	0	0

Table 7
Estimated transition probability distributions of a 5-state HMM

State	1	2	3	4	5
1	.86	.11	.03	0	0
2	0	.90	.06	.04	0
3	0	.01	.86	.02	.12
4	0	0	.02	.96	.02
5	0	0	0	.02	.98

Table 8
. Estimated emission probability distributions of a 5-state HMM

# Children\State	1	2	3	4	5
0	1	1	.40	0	0
1	0	0	.60	.53	0
2	0	0	0	.32	.67
> 2	0	0	0	.15	.33
Partnership\State	1	2	3	4	5
Single	.98	.67	0	.42	0
Cohabiting	.02	.33	0	.58	0
Married	0	0	1	0	1
Left home	1	2	3	4	5
No	1	0	.02	.03	.01
Yes	0	1	.98	.97	.99

Table 9
Interpretation of latent states in terms of observed demographic events and mental decision processes in a 5-state HMM

State	Observation	Decision processes
LS1	Living in parental home	Why and how to leave parental home
LS2	Childless, unmarried, residential-independent	Mode of partnership, entering parenthood
LS3	Married with 0 or 1 child	Family extension
LS4	Non-marital family	Marriage
LS5	Multi-child family	Staying together

and educational level. Men experience many of the events in the family formation process, in particular marriage and parenthood, later than women (Aassve, Billari, Mazucco, & Ongaro, 2002; Andersson & Philipov, 2002), thus transition rates between latent states may be expected to generally be lower for men than for women. If indeed these transition rates are lower for men, we may expect odds that are smaller than 1 for men, when compared to women. Educational differences in the family formation process usually are a bit more complex, and have partly been found to be gender-specific. The higher educated generally delay key family formation events like marriage and parenthood (Kravdal & Rindfuss, 2008; Liefbroer & Corijn, 1999), but in many countries they are very reluctant to enter unmarried cohabitation as well (Perelli-Harris et al., 2010). In some societies, having children out of wedlock is viewed as a low-class experience, and thus it is likely that the higher educated will be less prone to move to latent states characterized by births outside marriage. If the 5-state HMM validly describes individual demographic decision making, we may expect that the estimated effects of these covariates on the latent state transitions fit in with this knowledge. So, fitting a 5-state model with covariates amounts to testing the validity of the (interpretation of the) 5-state HMM.

As explained in Section 2.2, we assume that covariates affect the transition distribution. To evaluate this effect, we estimated the odds of these binary covariates and their interaction through logistic regression.

In Table 12 we only report those effects that have a p -value below .10.

Homogeneous models imply that age does not play a role in the timing of transitions. However, gender and education do play a role in the timing of crucial stages/events in family formation. A gendered effect cannot show up when gender is pooled, whether or not the model is time-homogeneous. However, even a time-homogeneous model should be able to reproduce such gendered effects. Therefore, testing for the presence of such gendered effects on timing is a good way to test for the validity of the model.

Men have a lower rate to experience the transition from LS1 to LS3 (marrying directly from the parental home), but do not differ in their rate of experiencing the transition from LS1 to LS2 (leaving the parental home by living single or in a cohabitation). Given that the transition from LS1 to LS2 is more likely than the transition from LS1 to LS3, this implies that men are somewhat slower than women in making their decision about leaving the parental home. Men are also slower than women to decide on further steps in the family formation after leaving home. In particular, they delay moving from LS2 to either LS3 (taking

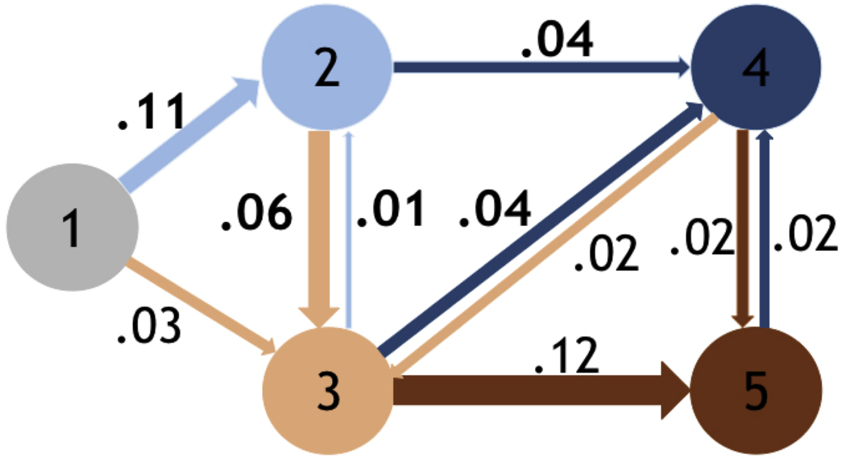


Fig. 8. State transition plot of a 5-state HMM. The thickness of the arrows and the numbers above the arrows reflect these transition probabilities (self-transitions not shown). Here, compared to the states of the 4-state HMM, the interpretation of LS2-LS5 has been changed. Therefore, the coloring has been changed: LS2 ⟶ light blue, LS3 ⟶ light brown, LS4 ⟶ dark blue, LS5 ⟶ dark brown. The probabilities of transitions that are part of the path leading to LS5 have been printed in a thicker font.

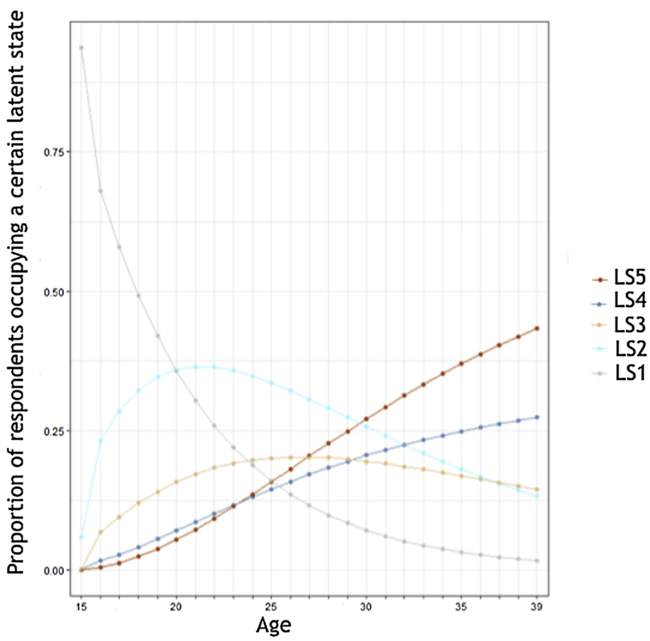


Fig. 9. Latent state occupancy fraction plot of a 5-state HMM.

the more traditional family formation path via marriage) or LS4 (taking the path via out-of-wedlock parenthood). Very clear educational differences can be observed as well. High-educated respondents are much more likely to transition from LS1 to LS2 than from LS1 to LS3, implying that the high educated delay traditional family formation patterns, but are more likely to first live on their own or to cohabit unmarried. In a next step, those in LS2 can either go to LS3 or to LS4. We observe a statistically significant interaction between level of education and gender for the LS2 LS3 transition. For women, higher education delays this transition ($b = .74$), but for men higher education slightly speeds up the transition ($b = .74 \times 1.56$). Thus higher educated women are a bit more reluctant to move into a traditional, marriage-like family pathway than higher educated men. The higher educated are clearly more reluctant to make the transition between LS2 and LS4 than the low educated, suggesting that out-of-wedlock parenthood among this cohort of French young adults is still mainly a low-class phenomenon. High education, finally, speeds up the transition between LS3 and LS5, suggesting that although the highly educated delay embarking on the traditional family process, they are more willing to speed up once they have entered the process. This is partly due to higher educated respondents usually having higher financial security to have more children when they already have one, and maybe the delay in the earlier life course triggers them to catch-up. The effect of these covariates is visualized in Fig. 12.

The roles played by the covariates in affecting the transition probabilities agree with what we know about these effects from other studies and thus confirm the choice of the 5-state HMM.

5. Conclusion and Discussion

Most life courses are made up of a multitude of changes in multiple life domains. A key challenge of life-course research is to make sense of this complexity by searching for fundamental processes that drive these observable transitions and by examining which factors influence them. In this paper, we claim that Hidden Markov modeling holds great promise in unraveling these processes, and we provide a relatively simple example of its potential by applying it to the family transition into adulthood among French men and women born between 1956 and 1965.

We feel that our results reveal a number of interesting viewpoints on the family formation process. The 4-state solution represents the transition to adulthood in the family domain as a process that is mainly driven by fertility. The first challenge that young adults face is about when and how to leave parental home. The next steps in this intergenerational reproductive process are about the initiation of a family (entry into parenthood), followed by successive phases of family expansion and family completion. Thus, the 4-state HMM suggests a model of the full family cycle starting as a child in a family of origin and ending up as an adult in a next generation family.

The 5-state HMM provides another interesting view on the family transition into adulthood. Rather than viewing this transition as a linear trajectory where young adults only differ in the likelihood and speed of moving to successive stages as is central to the 4-state HMM, the 5-state HMM distinguishes between two alternative family pathways into adulthood. As in the 4-state HMM, the first challenge every young adult faces is when to leave the parental home. One pathway strongly resembles the traditional pathway where young adults first establish a "traditional" family, characterized by marriage and possibly a child, followed by a subsequent stage of family expansion. However, a second pathway is distinguished as well, where young adults opt for a more autonomous lifestyle, characterized by single living and/or unmarried cohabitation. After this stage, these young adults are confronted by

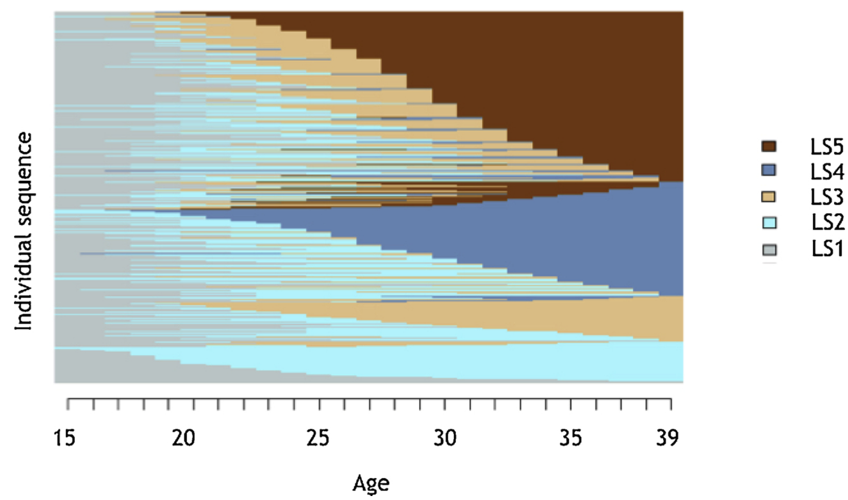


Fig. 10. Sequence index plots of the most probable latent state sequences, according to an estimated 5-state HMM.

Table 10
BIC obtained in estimating various HMM's

# states	BIC
3	136816.8
4	110833.3
5	100913.9
6	91321.3
7	82831.8
8	76296.6

Table 11
Interpretation of latent states in terms of observed demographic events and mental decision processes in a 6-state HMM

State	Observation	Decision Processes
LS1	Living in parental home	When and how to leave parental home
LS2	Childless, unmarried, residential independent	Mode of partnership and entering parenthood
LS3	Married with 0 or 1 child	Family extension
LS4	Non-marital family	Marriage
LS5	Multi-child marital family	Staying together, family extension
LS6	Multi-child marital family	-

another fundamental choice, either to continue this alternative lifestyle track and opt for children outside marriage, or to align themselves into the traditional pattern by moving "back" into the traditional family pathway. Our analysis also reveals clear differences in the speed and likelihood of transitions when linking covariates to the structural part of the 5-state HMM. For instance, high educated respondents are more likely to start off on the alternative track of independent living than low educated respondents, but once they enter the traditional pattern that emphasizes marital fertility, they complete this process faster than lower educated respondents. The interaction of gender and education also offers interesting insights in the switching between the alternative track and the traditional pattern. High educated males are faster in transiting whereas high-educated women are much more likely to delay switching to the traditional family formation pathway.

Whether one interprets the data on the basis of the 4-state or 5-state HMM solution at least partly depends on one's theoretical interests. The 4-state HMM offers a succinct interpretation of the traditional family

Table 12
Estimated odds of transitions in a 5-state HMM with 3 covariates in the form of weights in a logistic regression equation. Only odds that are significantly ($p < .01$) different from 1 are shown. The states are numbered in accordance with Fig. 8 and Table 8.

transition	male	high education	Inter-action
LS1→LS2	-	1.41	-
LS1→LS3	.40	.40	-
LS2→LS3	.78	.74	1.56
LS2→LS4	.72	.50	-
LS3→LS5	-	1.26	-
LS4→LS3	1.46	-	-
LS5→LS4	.64	-	-

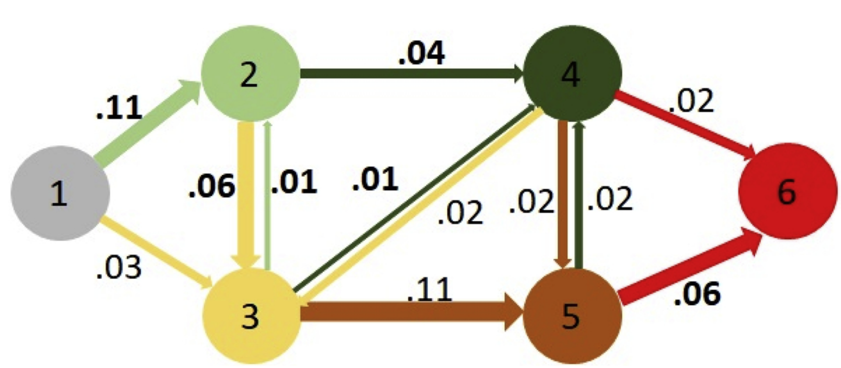


Fig. 11. Estimated state transition plot of a 6-state HMM. The colorings of some of the states and transitions have been changed, since the interpretations have been changed).

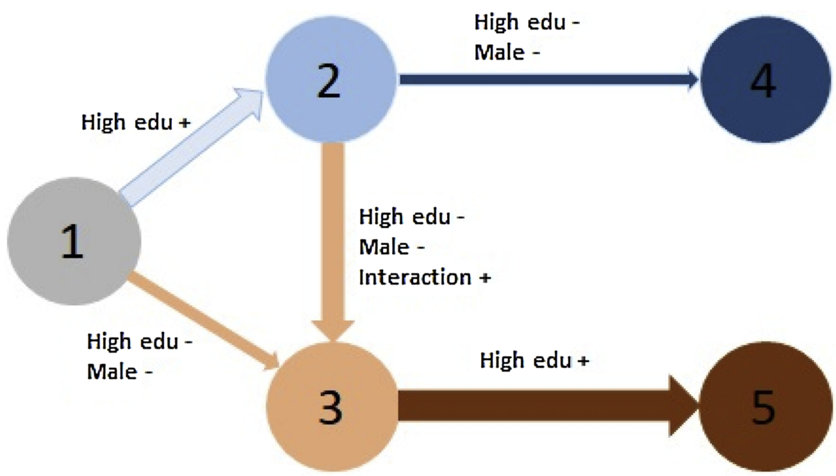


Fig. 12. Covariate effects in the estimated 5-state HMM. The coloring scheme is the same as that used in that of Fig. 8.

life pattern, pointing at three major decisions to be taken in the course of the family-life cycle (Glick, 1955). The 5-state HMM incorporates more heterogeneity into this family life cycle (Glick, 1989), and offers interesting opportunities to study the process of family change that is often captured under the heading of the Second Demographic Transition (Lesthaeghe, 1995). Furthermore, the analysis with covariates underscores the validity of the 5-state model.

A major advantage of both of these models is that they greatly limit the complexity of the process of transition into adulthood, by reducing the large number of transitions between observable states to a small number of transitions between unobservable, latent states. To quantify the potential of an HMM, we compare the estimation of an HMM with the estimation of a transition system of w parallel channels with c_j, $j = 1, . . . , w$, observational categories. Such a system has $\prod_{i=1}^{w} c_i = C$ distinct observable states.

In our application, we have three channels: the binary "living-at-home" channel, the 3-valued partnering channel and the 4-valued reproduction channel and thus this system has 24 observational states. So, the system has $C - 1$ initial state probabilities and $C(C - 1)$ transition probabilities, totalling to $(C + 1)(C - 1) = C^2 - 1$ states - in our application, this amounts to 575 parameters to estimate. In comparison, a k-state HMM with w channels and c_j states per channel has $k^2 - 1 + k\sum_j^w (c_j - 1)$ parameters; our 5-state HMM has only $24 + 45 = 69$ parameters! This potential could be even more useful if the number of potential observational states and transitions becomes even larger, for instance if one wants to study both family transitions and career-related transitions in one model.

The inclusion of covariates in a HMM serves two purposes. It offers the opportunity to study the influence of covariates at different stages of the life course and to compare their relative importance at these stages. However, if one already has extensive knowledge about the relationships between covariates and the processes under study, that knowledge could be used to test the validity of the HMM. If the observed relationships do not strongly resemble those based on prior knowledge, this suggest that the validity of the HMM solution is questionable.

The models introduced in this paper have clear merit for life-course research. Several extensions of the Hidden Markov Model can be envisaged, for example, constrained HMM (see e.g. Roweis, 2000). Constrained HMM is useful when one has a clear idea about the structure of the transition pattern, and wants to test the hypothesized transition probability distribution. This paper did not elaborate on this type of topic yet, but it can be of great interest for future research.

Furthermore, it is generally held that the lives of spouses and parents and children are linked (Elder, Johnson, & Crosnoe, 2003). Such linkages might well be modelled using multiple HMMs whose emissions affect each other's latent processes. Such mutually affecting models have been proposed in e.g. Elzinga, Hoogendoorn, and Dijkstra, (2007).

Generally, in applying these models to life-course data, researchers have to be aware of both theoretical and practical restrictions on the analyses. Models should not become too complex in order for them to be mathematically feasible to estimate and to be theoretically interpretable. Our paper suggests a number of guidelines in this respect that may prove useful to future users. Finally, we want to stress the point that applying mixture models is possible only when building on the results obtained from (hazard-based) regression models and classifiers like SA or LCA.

Acknowledgement

The research leading to these results has received funding from the European Research Council under Grant Agreement No 324178FP7/ERC (Project: Contexts of Opportunity, PL Aart C. Liefbroer)

The paper has benefitted from the comments of two anonymous reviewers.

References

Aassve, A., Billari, F. C., Mazucco, S., & Ongaro, F. (2002). Leaving home: a comparative analysis of ECHP data. *Journal of European Social Policy, 12*, 259–275.
Andersson, G., & Philipov, D. (2002). Life-table representations of family dynamics in Sweden, Hungary and 14 other FFS countries: A project of descriptions of demographic behavior. *Demographic Research, 7*, 67–144.
Bartolucci, F., Farcomeni, A., & Pennoni, F. (2012). *Latent Markov models for longitudinal data. Statistics in the Social and Behavioral Sciences*. Boca Raton: CRC Press.
Bartolucci, F., Pandolfi, S., & Pennoni, F. (2017). LMest: an R package for latent Markov models for categorical longitudinal data. *Journal of Statistical Software, 81*, 2–38.
Baum, L. E., Petrie, T., Soules, G., & Weiss, N. (1970). A maximization technique occurring in the statistical analysis of probabilistic functions of Markov chains. *Annals of Mathematical Statistics, 41*, 164–171.
Berchtold, A. (2002). High-order extensions of the double chain Markov model. *Stochastic Models, 18*, 193–227.
Berchtold, A., & Raftery, A. E. (2002). The mixture transition distribution model for high-order Markov chains and non-Gaussian time series. *Statistical Science,* 328–356.
Billari, F. C., Fürnkranz, J., & Prskawetz, A. (2006). Timing, sequencing, and quantum of life course events: A machine learning approach. *European Journal of Population, 22*, 37–65.
Bissell, M. (2000). Socio-economic outcomes of teen pregnancy and parenthood: A review of the literature. *The Canadian Journal of Human Sexuality, 9*, 191.
Blossfeld, H.-P., Golsch, K., & Rohwer, G. (2007). *Event History Analysis with STATA*. Mahwah, NJ: Lawrence Erlbaum.
Blossfeld, H.-P., Klijzing, E., Mills, M., & Kurz, K. (2006). *Globalization, uncertainty and youth in society: The losers in a globalizing world.* New York: Routledge.
Buchmann, M. C., & Kriesi, I. (2011). Transition to adulthood in Europe. *Annual Review of Sociology, 37*, 481–503.
Burnham, K. P., & Anderson, D. R. (2002). *Model Selection and Multimodel Inference. A*

Practical Information-Theoretic Approach (2nd edn). Heidelberg: Spinger.
Cornwell, B. (2015). *Social Sequence Analysis: Methods and Applications. No. 37 in Structural Analysis in the Social Sciences.* New York: Cambridge University Press.
Dewar, M., Wiggins, C., & Wood, F. (2012). Inference in Hidden Markov Models with Explicit State Duration Distributions. *IEEE Signal Processing Letters, 19*, 235–238.
Eggar, M. H. (2002). Validity of fitting a first-order Markov chain model to data. *Journal of the Royal Statistical Society, Series D (The Statistician), 51*, 259–265.
Eggeling, R., Gohr, A., Bourguignon, P.-Y., Wingender, E., & Grosse, I. (2013). *Inhomogeneous parsimonious Markov models. Joint European Conference on Machine Learning and Knowledge Discovery in Databases*. Springer321–336.
Elder, G. H., Jr, Johnson, M. K., & Crosnoe, R. (2003). The emergence and development of life course theory. In J. T. Mortimer, & M. J. Shanahan (Eds.). *Handbook of the life course* (pp. 3–19). Boston: Springer Handbooks of Sociology and Social Research, chap. 1.
Elzinga, C. H., Hoogendoorn, A. W., & Dijkstra, W. (2007). Linked Markov Sources: Modeling outcome-dependent social processes. *Sociological Methods & Research, 36*, 26–47.
Fasang, A. E. (2012). Retirement patterns and income inequality. *Social Forces, 90*, 685–711.
Fokkema, T., Kveder, A., Hiekel, N., Emery, T., & Liefbroer, A. C. (2016). Generations and Gender Programme Wave 1 Data Collection: An overview and assessment of sampling and fieldwork methods, weighting procedures, and cross-sectional representative-ness. *Demographic Research, 34*, 499–524.
Gabadinho, A., Ritschard, G., Müller, N. S., & Studer, M. (2011). Analyzing and visua-lizing state sequences in R with TraMineR. *Journal of Statistical Software, 40*, 1–37.
Gangl, M. (2004). Welfare states and the scar effects of unemployment: A comparative analysis of the United States and West Germany. *American Journal of Sociology, 109*, 1319–1364.
Gauthier, J.-A., Widmer, E. D., Bucher, P., & Notredame, C. (2010). Multichannel se-quence analysis applied to social science data. *Sociological methodology, 40*, 1–38.
Giudici, P., Rydèn, T., & Vandekerkhove, P. (2000). Likelihood-Ratio Tests for Hidden Markov Models. *Biometrics, 56*, 742–747.
Glick, P. C. (1955). The life cycle of the family. *Marriage and Family Living, 17*, 3–9.
Glick, P. C. (1989). The family life cycle and social change. *Family relations, 38*, 123–129.
Halpin, B. (2010). Optimal matching analysis and life-course data: The importance of duration. *Sociological Methods & Research, 38*, 365–388.
Han, S. Y., Liefbroer, A. C., & Elzinga, C. H. (2017). Comparing methods of classifying life courses: Sequence Analysis and Latent Class Analysis. *Longitudinal and Life Course Studies, 8*, 319–341.
Härkönen, J., & Dronkers, J. (2006). Stability and Change in the Educational Gradient of Divorce. A Comparison of Seventeen Countries. *European Sociological Review, 22*, 501–517.
Helske, S., & Helske, J. (2017). *Mixture hidden Markov models for sequence data: the seqHMM package in R.* arXive:1704.00543[stat.CO]https://arxiv.org/abs/1704.00543.
Hitlin, S., Elder, G. H., Jr., & Elder, G. H., Jr. (2007). Time, self and the curiously abstract concept of agency. *Sociological Theory, 25*(2), 170–191.
Kravdal, Ø., & Rindfuss, R. R. (2008). Changing relationships between education and fertility: A study of women and men born 1940 to 1964. *American Sociological Review, 73*, 854–873.
Lesthaeghe, R. (1995). The Second Demographic Transition in Western countries: An interpretation. In K. Oppenheim Mason, & A.-M. Jensen (Eds.). *Gender and Family Change in Industrial Countries* (pp. 17–62). Wotton- under-Edge, UK: Clarendon Press International Studies in Demography.
Liefbroer, A. C., & Corijn, M. (1999). Who, what, where an when? Specifying the impact of educational attainment and labour force participation on family formation. *European Journal of Population, 15*, 45–75.
Mächler, M., & Bühlmann, P. (2004). Variable Length Markov Chains: methodology, computing, and software. *Journal of Computational and Graphical Statistics, 13*, 435–455.
Manning, W. D., Brown, S. L., & Payne, K. K. (2014). Two decades of stability and change in age at first union formation. *Journal of Marriage and Family, 76*, 247–260.
Mayer, K. U. (2009). New directions in life course research. *Annual Review of Sociology, 35*, 413–433.
Neels, K., Theunynck, Z., & Wood, J. (2015). Economic Recession and First Births in Europe: Recession-Induced Postponement and Recuperation of Fertility in 14 European Countries Between 1970 and 2005. *International Journal of Public Health, 58*, 43–55.
Paas, L. J., Vermunt, J. K., & Bijmolt, T. H. (2007). Discrete time, discrete state latent Markov modelling for assessing and predicting household acquisitions of financial products. *Journal of the Royal Statistical Society: Series A (Statistics in Society), 170*, 955–974.
Perelli-Harris, B., Sigle-Rushton, W., Krayenfeld, M., Låppegard, T., Keizer, R., & Berghammer, C. (2010). The educational gradient of childbearing with cohabitation in Europe. *Population and Development Review, 36*, 775–801.
Roweis, S. T. (2000). Constrained Hidden Markov Models. In S. A. Solla, T. K. Leen, & K. Müller (Eds.). *Advances in Neural Information Processing Systems 12* (pp. 782–788). MIT Press.
Sobotka, T., & Toulemon, L. (2008). Changing Family and Partnership Behaviour: Common Trends and Persistent Diversity across Europe. *Demographic Research, 19*, 85–138.
Specht, J., Egloff, B., & Schmukle, S. C. (2011). Stability and change of personality across the life course: the impact of age and major life events on mean-level and rank- order stability of the Big Five. *Journal of Personality and Social Psychology, 101*(4), 862–882.
Spedicato, G. A., Kang, T. S., Yalamanchi, S. B., & Yadav, D. (2014). *The markovchain package: A package for easily handling discrete Markov chains in R.* ftp://mirrors.xmu.edu.cn/CRAN/web/packages/markovchain/vignettes/an_introduction_to_markovchain_package.pdf.
Studer, M., Liefbroer, A. C., & Mooyaart, J. E. (2018). Understanding Trends in Family Formation Trajectories: An Application of Competing Trajectories Analysis. *Advances in Life Course Research, 36*, 1–12.
Studer, M., Ritschard, G., Gabadinho, A., & Müller, N. S. (2011). Discrepancy analysis of state sequences. *Sociological Methods & Research, 40*, 471–510.
Torgerson, W. S. (1958). *Theory and methods of scaling.* New York: J. Wiley.
Vermunt, J. K. (1997). *Log-Linear Models for Event Histories. Vol. 8 of Advanced Quantitative Techniques in the Social Sciences*. Thousand Oaks: Sage.
Viterbi, A. J. (1967). Error bounds for convolutional codes and an asymptotically optimal decoding algorithm. *IEEE Transactions on Information Theory, 13*(2), 260–269.