Social Research Institute

UCL

# A Sequence-Analysis Approach to the Study of the Transition to Adulthood in Low- and Middle-Income Countries

**Maria Sironi**

**Nicola Barban**

**Luca Maria Pesando**

**Frank F. Furstenberg**

**Quantitative Social Science**

Working Paper No. 20-17

October 2020

## Disclaimer



Quantitative Social Science

UCL Social Research Institute

University College London

55-59 Gordon Square

London WC1H 0NU

Social Research Institute

UCL

# A Sequence-Analysis Approach to the Study of the Transition to Adulthood in Low- and Middle-Income Countries

**Maria Sironi**[1]
**Nicola Barban**[2]
**Luca Maria Pesando**[3]
**Frank F. Furstenberg**[4]

**Abstract**

This study investigates whether young people in low- and middle-income countries (LMICs) have experienced processes of de-standardization of the life course similar to those observed in high-income societies. We provide two contributions to the relevant literature. First, we use data from 263 Demographic and Health Surveys (DHS) across 69 LMICs, offering the richest comparative account of women's transition to adulthood (TTA) patterns in the developing world existing to date. In so doing, we shift the focus from individual life-course transitions towards a holistic approach that allows us to characterize the life-course complexity by detailed sequences of events, namely first sexual intercourse, first union, and first birth. Second, using a clustering algorithm based on optimal-matching distances of life-course sequences, we identify clusters of TTA and explore their changes across cohorts by region and urban/rural location of residence. Results stress the importance of investigating cross-regional differences in partnership and fertility trajectories by looking at the interrelation and complexity of status combinations. Summarizing the ensuing heterogeneity through four clusters, we document significant differences by macro-regions yet relative stability across cohorts. We interpret the latter as suggestive of cultural specificities that make the TTA resistant to change and slow to converge across regions, if converging at all.

**Keywords** Transition to adulthood, union formation, fertility, sequence analysis, cluster analysis, Demographic and Health Surveys, low- and middle-income countries

**Codes:** J12, J13, O57

[1] Social Research Institute, University College London
[2] Institute for Social and Economic Research, University of Essex
[3] Department of Sociology and Centre on Population Dynamics, McGill University
[4] Department of Sociology and Population Studies Center, University of Pennsylvania

## Introduction

Adolescence and early adulthood are critical periods of human development that set young people on trajectories that shape their future roles as adults, alongside their capacity to make informed decisions about the timing of key life events (National Research Council and Institute of Medicine, 2005). During adolescence, important decisions are made relating to transitions out of school, into work, into sexual relations, marriage, and parenthood. Although these transitions are not homogenous across contexts, it is indisputable that they shape the subsequent life course and – in an intergenerational perspective – the context in which children are born and raised (Bongaarts et al., 2017).

Over the past century, there have been significant changes in the prevalence, timing, density, and complexity of transitions to adulthood (TTA) in the United States and Europe. These have been extensively documented, with the majority of studies providing evidence of increasing "de-standardization" (or disorder) of family-formation trajectories (e.g., Billari and Liefbroer 2010; Rindfuss, Swicegood, and Rosenfeld 1987; Shanahan 2000). Family-life trajectories of young adults in high-income societies (HICs) seem to have become less similar to one another, and the variation in the types of family-trajectories to have increased (Elzinga & Liefbroer, 2007; Van Winkle, 2018). Some other evidence suggests that, although no convergence of trends is observed yet, these diversified pathways to adulthood are changing in the same direction across most parts of Europe towards a late, protracted, and complex TTA (Billari & Liefbroer, 2010).

In low- and middle-income countries (LMICs), the literature on the TTA is still in its infancy, yet interest has been growing steadily in light of the significant socio-economic and demographic transformations these countries have been undergoing (Juárez & Gayet, 2014; Pesando & GFC Team, 2019). Inserted within a framework of rooted poverty and persistent economic and social inequalities, factors such as massive educational expansion, changes in family forms and behaviors, high unemployment rates, multiple health vulnerabilities, and shifting preferences regarding marriage types are likely to manifest in a world in which fewer certainties lead to new ways of experiencing the transition to adulthood (Beguy et al., 2011). Drawing also on the idea that individual life courses are strongly standardized by institutional demands – and the presence and nature of institutions in LMICs is heterogeneous and often lacking – existing (single-country) research in LMICs has demonstrated a vast diversity of situations in different regions of the world, alongside transitions to adulthood occurring at different ages and modalities within the same society (Juárez & Gayet, 2014; National Research Council and Institute of Medicine, 2005).

This study seeks to contribute knowledge on the nature of the transition to adulthood in low- and middle-income countries. We provide two contributions to the relevant literature. First, we offer – to the best of our knowledge – the first comparative account of women's TTA patterns in LMICs using data from 263 Demographic and Health Surveys (DHS) across 69 countries. In so doing, we shift the focus from the study of independent life-course transitions towards a holistic approach that

characterizes the life course as a complex combination of events. The literature on the TTA is currently organized along three lines. First are studies that describe the complexity of life-course transitions focusing on one or two countries only. Examples from HICs are Aassve, Billari, and Piccarreta (2007) on the UK, Oris and Ritschard (2014) on Belgium, Ravanera, Rajulton, and Burch (1998) on Canada, Robette (2010) on France, Salmela-Aro et al. (2011) on Finland, Sironi, Barban, and Impicciatore (2015) on Italy and the US, etc. Conversely, examples from LMICs are Goldberg (2013) on South Africa, Echarri Cánovas and Pérez Amador (2007) and Fussell (2005) on Mexico, Beguy et al. (2011) on Kenya, Mensch et al. (2014) on Malawi, and Tian (2016) on China. Second are cross-national comparative analyses focused on industrialized societies, such as Billari and Liefbroer (2010), Elzinga and Liefbroer (2007), Schwanitz (2017), and Van Winkle (2018) covering, respectively, 26, 19, eight, and 14 countries and the institutional differences (e.g., welfare states) thereof. Third are cross-national comparative studies focused on LMICs, such as Lloyd and Mensch (2008), Bongaarts, Mensch, and Blanc (2017), and Pesando and GFC Team (2019). These provide good bases for the understanding of the TTA in LMICs, yet they treat transitions such as entering first marriage or first birth as independent – rather than interrelated – events, thus missing the holistic approach we adopt here.

Building on a similar observation that analyses of TTA in less developed countries largely focus on one or two transitions in isolation, this paper is closest in aim and ambition to Grant and Furstenberg (2007), who examined whether increases in female school participation, the timing of marriage, and socio-economic changes in Africa and Latin America influenced the timing and pattern of the transition to adulthood using DHS data from six countries. We here capitalize on Grant and Furstenberg (2007) by enriching their analyses in three directions. First and most importantly, we do not only rely on one combined index of heterogeneity to describe the complexity of the life course. Second, we provide updated estimates drawing on all surveys collected up until the year 2016. Third, we expand the geographical coverage from six to 69 countries, looking at cross-regional heterogeneity and changes across cohorts and rural and urban areas.

Our second contribution is methodological. We use sequence analysis to characterize the life-course complexity through detailed sequences of events defined by first sexual intercourse, first union, and first birth. To maximize the number of survey waves and countries covered, we focus on women only, and restrict to events occurring up to age 30. The latter decision, together with our focus on first birth – rather than the whole fertility history – reflects an explicit choice to characterize the early part of the life course and highlight as much heterogeneity as possible "independently" of higher-order fertility, which would be a dominant and driving feature, for instance, in sub-Saharan Africa (SSA).[1] Drawing on information from the sequences, we compute new descriptive measures of *mean age at events* and *time between events* and explore their changes across cohorts. Additionally, we use a clustering algorithm based on Optimal Matching (OM) distances of life-course sequences to summarize the emerging heterogeneity through clusters of TTA and explore their changes overtime by region and urban/rural location of residence.

Conceptually, our chice of sequence analysis rests on the idea that increases in women's mean or median ages at key life-course events (e.g., median age at first marriage) are by now well-documented in LMICs (see, for instance, Bongaarts et al. 2017). Through this methodological approach, we are able to understand how transitions to first sex, first union, and first birth are interrelated and influence each other – and we intend to convey that background forces such as educational expansion, urbanization, and cultural, societal, and institutional changes do exert an influence not only on mean ages, but on the whole "package" of TTA status combinations that young adults go through. In the context of LMICs, sequence analysis has been employed for single-country cases to study family planning trajectories in Malawi (Furnas, 2016), relationship quality and well-being in Malawi (Frye & Trinitapoli, 2015), and the use of time among the elderly in South Africa (Grapsa & Posel, 2016). This is among the first studies to use sequence analysis and clustering methodologies based on regression-tree techniques to study the TTA using large-scale comparative data from LMICs.

**Background: The Transition to Adulthood in LMICs**

Similarly to more industrialized societies, individuals in low-income societies have been witnessing a transfer of responsibility from family to other societal agencies, and emerging new institutions with the potential to alter the timetable for growing up in nations where significant educational expansion has occurred (Grant & Furstenberg, 2007). Yet the extent to which the dynamics and the contexts in which youth make transitions to adulthood in low-income countries resemble those of more advanced economies is still an open question. Focusing on independent transitions, Behrman and Sengupta (2005) undertook an investigation of this kind, finding that low-income countries have converged – with East Asia and the Pacific converging the most and sub-Saharan Africa converging the least – towards characteristics of high-income countries in a number of respects, such as the higher dependence on markets rather than family enterprises, formal schooling rather than working, more awareness of options and lifestyles from contexts broader than the local community, smaller gender gaps favoring males, and much more mobility in several dimensions. Conversely, more recent scholarly reviews (Juárez & Gayet, 2014) point towards a vast diversity of situations in different regions of the world, with tensions arising from globalization leading to imitate Western lifestyles in some respects, and circumstances of social inequality, poverty, and exclusion unique to the low-income world influencing the timeframe in which transitions occur, in other respects. Although globalization could have contributed to a homogenization of the transitions on a global scale, it seemed to result instead in even greater diversification of life paths – some of which fully unique to LMICs.

As the expansion of universal school attendance to the secondary level is a key force in standardizing the ages at which young people begin the transition to adulthood – and in some (mostly) sub-Saharan African countries universal primary enrolment has not been achieved yet (Hewett & Lloyd, 2005; Psaki et al., 2018) – we expect to observe widespread heterogeneity in the experience of early life-

course transitions, and life paths that are rather unstandardized in late adolescence. Given the weak role of institutions and the wide variation in institutional contexts across countries – both by region and by level of economic development – we anticipate marked differences across regions. As for changes overtime, our expectations are less clear-cut, yet recent studies on global family change have provided evidence of slow changes in family domains in LMICs (Castro et al., 2019; Cherlin, 2016; Pesando & GFC Team, 2019). Marriage, union formation, divorce, and union dissolution practices have proven less responsive to socioeconomic changes in LMICs because they are tied to elements of the social structure that are harder to change. These structural features of societies include religious beliefs, marriage-related-laws and prohibitions, inheritance rights, shortages in the housing market, and persistent disparities in gender roles and dynamics (Coontz, 2014). This inertia overtime would be consistent with what observed by Fussell (2005) in Mexico and by Grant and Furstenberg (2007) who – putting aside concerns related to the validity of a single heterogeneity index for summarizing large volumes of information – observed relatively small changes in the age-specific index of heterogeneity from the first to the most recent survey in each country.

Yet, whether the transition to adulthood conceptualized as a holistic package of status transitions in LMICs is (un)structured or (dis)ordered and whether it resembles the transition to adulthood in HICs is ultimately an empirical question. We hence turn to a more detailed description of our data, methodological approach, and empirical analyses.

## Data and Methods

### *Data*

We use data from 69 low- and middle-income countries drawing upon 263 Demographic and Health Survey (DHS) waves. Each survey is representative of the national population of women aged 15-49 and includes information on key variables of interest, i.e., age at first sexual intercourse, age at first marriage, and fertility histories. Information from DHS data used in this study has been collected between 1985 and 2016. We keep all respondents (women) with complete information on age at first sexual intercourse, first union, and first birth. Although DHS data are also collected for men in some countries, we here focus on women only to maximize the number of countries covered and survey waves included. Due to the complexity of disentangling formal and informal marriages, we refer to unions as the combination of marriage and cohabitation (coded as "living together" in the survey), as typically done in studies of this kind (Casterline et al., 1986; Clark & Brauner-Otto, 2015). The number of survey waves and the availability of data for a specific year vary by country. Middle Eastern and North-African (MENA) countries – except for Egypt – are excluded from the investigation due to missing information on age at first sex, typically not asked for reasons of cultural appropriateness. Also, selected surveys from Asia were omitted because the DHS only interviews ever-married women, thus impairing our ability to observe early life-course transitions. DHS sample weights have been used in the analysis throughout.

As we conducted analyses at the regional level, countries were grouped in the following macro-regions following the classification used in previous studies of global family change (Pesando & GFC Team, 2019): Americas, Asia, Former Union of Soviet Socialist Republics (USSR), and sub-Saharan Africa (SSA). Among these, we relied on a further classification by subregions, namely Central America, South America, Central Asia, South-East Asia, West Asia, Eastern Europe, East Africa, West Africa, South Africa, and Central Africa. Table 1 below shows the regional and sub-regional classifications adopted in the analysis, and the sample size of women with complete sequences by subregion. For the number of waves per region, and the detailed regional and sub-regional classification of countries included in the analysis, see Appendix Table A1 and Figure A1.

**Table 1**: Regional and sub-regional classification adopted in the analysis, and sample of women with complete sequences

| **Region** | **Subregion** | **Observations** |
|---|---|---|
| Americas | Central America | 101,272 |
| | South America | 218,590 |
| Asia | South-East Asia | 201,643 |
| Former USSR | Central Asia | 15,979 |
| | West Asia | 17,556 |
| | Eastern Europe | 12,151 |
| SSA | East Africa | 265,390 |
| | West Africa | 217,405 |
| | Central Africa | 65,492 |
| | South Africa | 28,575 |

*Methods*

This analysis looks at a combination of family life-course events that are closely interrelated, namely first sexual intercourse, first union, and first birth. As already mentioned above, we focus on first births only – rather than the whole fertility history – to better characterize the early part of the life course and highlight as much heterogeneity as possible "independently" of higher-order fertility. Rather than analyzing the median ages at different life-course events (e.g., age at marriage and children progression), we use sequence analysis to describe TTA trajectories. By focusing on the analysis of trajectories rather than the occurrence of single events, sequence analysis overcomes the limitations of event history analysis in studying multiple interrelated events. Rather than studying the age at each event, using sequence analysis we are able to explore the occurrence and the timing of events together with the intervals between events, and this allows us to identify patterns and create TTA typologies that would

not emerge otherwise (Barban & Sironi, 2019). We are not just interested in the timing of events, but also in the time intervals between events, and how experiencing (or not) one event might have consequences for subsequent markers in the TTA.

In sequence analysis, each life-course trajectory is represented by a string of characters, which resembles the one used to code DNA molecules in biological sciences. Hence, every trajectory is made up of a number of values that correspond to the number of years (or months) each individual is observed for. Trajectories can be analyzed by representing the data as a categorical time series. Each individual $i$ can be associated to a variable $s_{it}$ indicating the individual's life-course status at time $t$. In other terms, trajectories can be represented as strings or sequences of characters, with each character denoting one particular state that describes a specific family role.

Sequence analysis is associated to a family of algorithms used to quantify dissimilarities between life-course trajectories. *Optimal Matching* algorithm (OM) is the most known technique that has been applied to social science. The development of OM started in the seventies and the technique has been described in details by Kruskal (1983). OM expresses distances between sequences in terms of the minimal amount of effort, measured in terms of edit operations (insertion; deletion, and substitution), that is required to change two sequences such that they become identical. Abbott (1995) adapted OM to social sciences assigning to three elementary operations different costs, based on the 'social' differences between states (Lesnard, 2006). The choice of the operations' costs determines the matching procedure and influences the results obtained. Sequence analysis algorithms identify differences in trajectories due to changes in *timing* (when events happen), *quantum* (what and how many transitions), and *ordering* (in what order) of life-course events (Billari et al., 2006; Billari & Piccarreta, 2005). We construct trajectories using a state-space composed of five different states: 'no sexual intercourse – not in a union', 'sexual intercourse – not in a union', 'no sexual intercourse – in a union', 'sexual intercourse – in a union, no children', 'sexual intercourse – in a union, one child or more.'

After constructing sequences of events, we obtain a series of descriptive statistics computed from the sequences themselves, such as the *mean age at specific events* – specifically, age at first sexual intercourse, first union, and first child – and the *time between events* – specifically, sex to first union, sex to first child, and union to first child. All summary measures are computed conditionally on the event occurring by age 30.

After building partnership and fertility trajectories and computing summary measures from the sequences, we build clusters to identify patterns in the data and highlight "typical" life-course trajectories (Abbott & Tsay, 2000; Aisenbrey & Fasang, 2010; Barban, 2013; Barban & Billari, 2012). In this study, we identify clusters using a clustering algorithm (*Agglomerative Hierarchical Clustering*) based on OM distances of life-course sequences. Differently from other studies that use the entire dissimilarity matrix to classify observations in different groups representing typologies of trajectories, we used a subsample of the data to estimate a classification of respondents in different groups, and we assigned the remaining

sample to a cluster based on their similarity to a representative sequence of each cluster. This method is appropriate for large samples, for instance when it is computationally impractical to estimate a complete pairwise dissimilarity matrix. Furthermore, the partition rule obtained can be applied to different datasets. The method proceeds as follows.

First, we randomly selected a subsample of 20,000 respondents that is used as training sample. We calculated the pairwise dissimilarity matrix for this subsample based on OM metric, with substitution costs based on transition rates. Second, we applied a clustering algorithm and identified four groups of trajectories. We decided on a four-cluster solution based on the visual inspection of the dendrogram (Figure A2 in the Appendix) and the comparison of different measures of cluster performance provided by the R package `WeightedCluster` described in (Studer, 2013). A complete description of goodness of classification of different cluster solutions is included in Table A2 in the Appendix. Third, for each cluster we calculated a representative trajectory based on the medoid sequences. A medoid is the observation with the minimum distance from all other individuals in a cluster (Aassve et al., 2007). Fourth, we assigned each observation of the remaining sample to the cluster with the minimum distance with respect to the medoid sequence. In this way, we reduced the computational burden and memory requirements by calculating only distances from four representative sequences. As input of analysis, we used the entire set of variables included to construct the sequences, that is, the time-dependent indicators of life-course states from age 12 to age 30. As sensitivity analysis, we compared the summary statistics in the subsample with the entire sample (Appendix Table A3). Furthermore, since the subsample is randomly chosen, we compared the cluster solutions obtained in ten independent random draws by looking at measures of classification concordance. We can safely conclude that the approximation obtained with this methodology does not affect the quality of results.

Once we classify individuals into typologies, we are able to explore cluster distributions by regions and birth cohorts, in order to ascertain differences and changes in typical transitions over time. Moreover, to understand potential drivers of the TTA over time and across countries, we investigate differences between urban and rural settings and the relationship between the level of GDP per capita and the probability of belonging to each cluster.

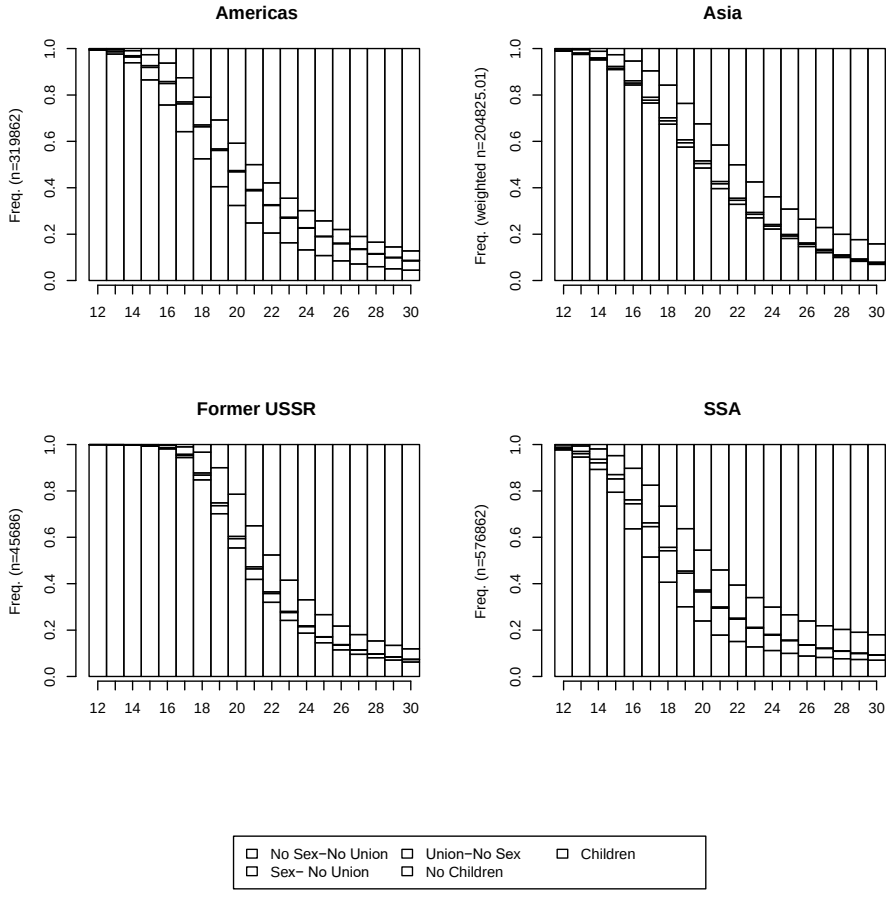
## Findings

### *Sequences*

Figure 1 presents the distribution of states across the four macro-regions – Americas, Asia, Former USSR, and sub-Saharan Africa – between ages 12 and 30. The four panels show an important degree of heterogeneity across areas. First, women in SSA and the Americas experience transitions to the first event earlier than in the Former USSR and Asia. As a matter of fact, in South and Central America, half of the women in the sample experienced their first event in the TTA, i.e., first sexual intercourse, by age

19, and in SSA by age 18. The median age is higher in Asia and former USSR, 20 and 21 years of age, respectively. Also, the state 'sex-no union' is very distinct in the Americas and SSA, it is less widespread in the Former USSR, and almost non-existent in Asia, suggesting a closer connection between first sexual intercourse and being in a union in the latter regions. Lastly, the transition from 'sexual intercourse - in a union (no children)' to having at least one child is fastest in the Americas and slowest in the Former USSR. By age 30, more than 80% of women in all macro-regions have completed the transition to motherhood, but slightly more so in South and Central America and Former USSR, relative to Asia and SSA.

**Figure 1**: Distribution of states across macro-regions




*Descriptive summary measures based on sequences*

Drawing information directly from the sequences, we obtain some descriptive summary measures to provide a *prima-facie* characterization of the cross-context heterogeneity. Mean ages at events and mean times between events are reported in Figures 2 and 3, respectively. Starting from Figure 2, three trends are evident. First, the three variables – mean age at first sex, first union, and first child – follow very similar trends overtime in each subregion. Second, within each panel there is heterogeneity across regions both in terms of levels and – to a lesser extent – in terms of trends. In terms of levels, mean

ages at first sex, union, and childbearing are lowest in Central, East and West Africa and highest in West Asia, Central Asia, and Eastern Europe. The overall trend is one of delayed transitions to adulthood with increasing mean ages at each event. For example, age at first sexual intercourse increases from 13.5 to 14 in Central Africa, from 14.7 to 15.6 in East Africa, from 17 to 18 in Southeast Asia across women born between 1940 and 1980. Age at first union is increasing in most regions as well: from 14.4 to 15.5 in Central Africa, from 15.1 to 15.9 in West Africa, from 14.7 to 16.8 in East Africa, from 16.2 to 16.8 in Central America, and from 16.4 to 18.5 in Southeast Asia. Age at first birth has increased the least over time across subregions out of the three events, but we can still observe an increase from 16.9 to 18 in East Africa, from 17.7 to 18.1 in Central America, from 19.5 to 19.9 in Central Asia, and from 18.4 to 19.8 in Southeast Asia.

Despite this general trend, some subregions stand out as having experienced more marked – or different – changes overtime. South-East Asia is among these, as it shows the steepest increases in the three variables over time (around a 1-year increase in age at first sex, a 2-year increase in age at first union, and 1.5-year increase in age at first child). Conversely, in line with the literature, South America stands out as the only subregion exhibiting declining mean ages at events, particularly for sexual intercourse – in line with Bongaarts et al. (2017). In South America, age at first sexual intercourse declines from 17.1 among women born in the 1940s to 15.3 among those born in the 1980s. Age at first union declines approximately by half a year (from 17.8 to 17.4), and age at first birth from 18.7 to 18.4. Third, two panels – the one on age at first sex and the one on age at first child – suggest the emergence of two groups of subregions that seem to converge towards similar mean ages, namely Central Asia, South-East Asia, West Asia, and Eastern Europe (group a), and Central America, South America, South Africa, West Africa, Central Africa and East Africa (group b). This distinction reflects the 'speed' of the transition to adulthood that is apparent also in Figure 1. Conversely, the panel on age at first union does not show a distinction of this kind.

Differently from panels in Figure 2, panels in Figure 3 show heterogeneous trends across variables, mostly increasing for 'time sex to union' (panel a, top left), flat for 'time sex to child' (panel b, top right) and decreasing for 'time union to child' (panel c, bottom left) – a set of variables and findings which are, to the best of our knowledge, new in this literature. Within each panel, sub-regional trends are also heterogeneous. For instance, in South America – the only subregion with declining mean ages at first sexual intercourse – we observe the most marked increases in the intervals between first sexual intercourse and first union, suggesting delayed progressions to subsequent events: from 9 months for women born in the 1940s to 27 months for women born in the 1980s. In the four subregions of SSA and in Central America we also observe a large increase: from half a year to a year in West Africa, from 4 months to 16 months in East Africa, from 11 to 19 months in Central Africa, from 14 to 26 months in South Africa, and from 4 months to a year in Central America. The interval between first sexual intercourse and first union is instead relatively flat in Central Asia, Southeast Asia, West Asia and Eastern Europe.

As far as 'time sex to child' is concerned, most subregions show rather flat patterns, suggesting that even if transitions to first sexual intercourse occur later, there's no evidence that childbearing is also shifted ahead once women have engaged in first sexual intercourse. In SSA (particularly in West, East, and Central Africa), we observe narrowing intervals between first sexual intercourse and childbearing: from 3.3 to 2.6 years, from 2.8 to 2.4 years, and from 3.6 to 3 years respectively. South America, also in this case, is the subregion that shows the largest increase, from 2.2 to 3.4 years.

When looking at the transition from first union to first child, we observe a general declining trend. SSA (except for South Africa) is the region where time from union to child has declined most rapidly across birth cohorts. The decline is evident in West (from 3 to 1.75 years), East (from 2.8 to 1.7 years) and Central Africa (from 2.5 to 1.3 years), and in Southeast Asia (from 2.25 to 1.7 years). The trend is flatter in other subregions. These results combined suggest that there is a general delay in entering into first union after experiencing first sexual intercourse, but that once women get married or start a coresidential union, the transition to motherhood is relatively quick.

**Figure 2**: Mean age at specific events: first sexual intercourse (panel a, top left), first union (panel b, top right), and first child (panel c, bottom left)

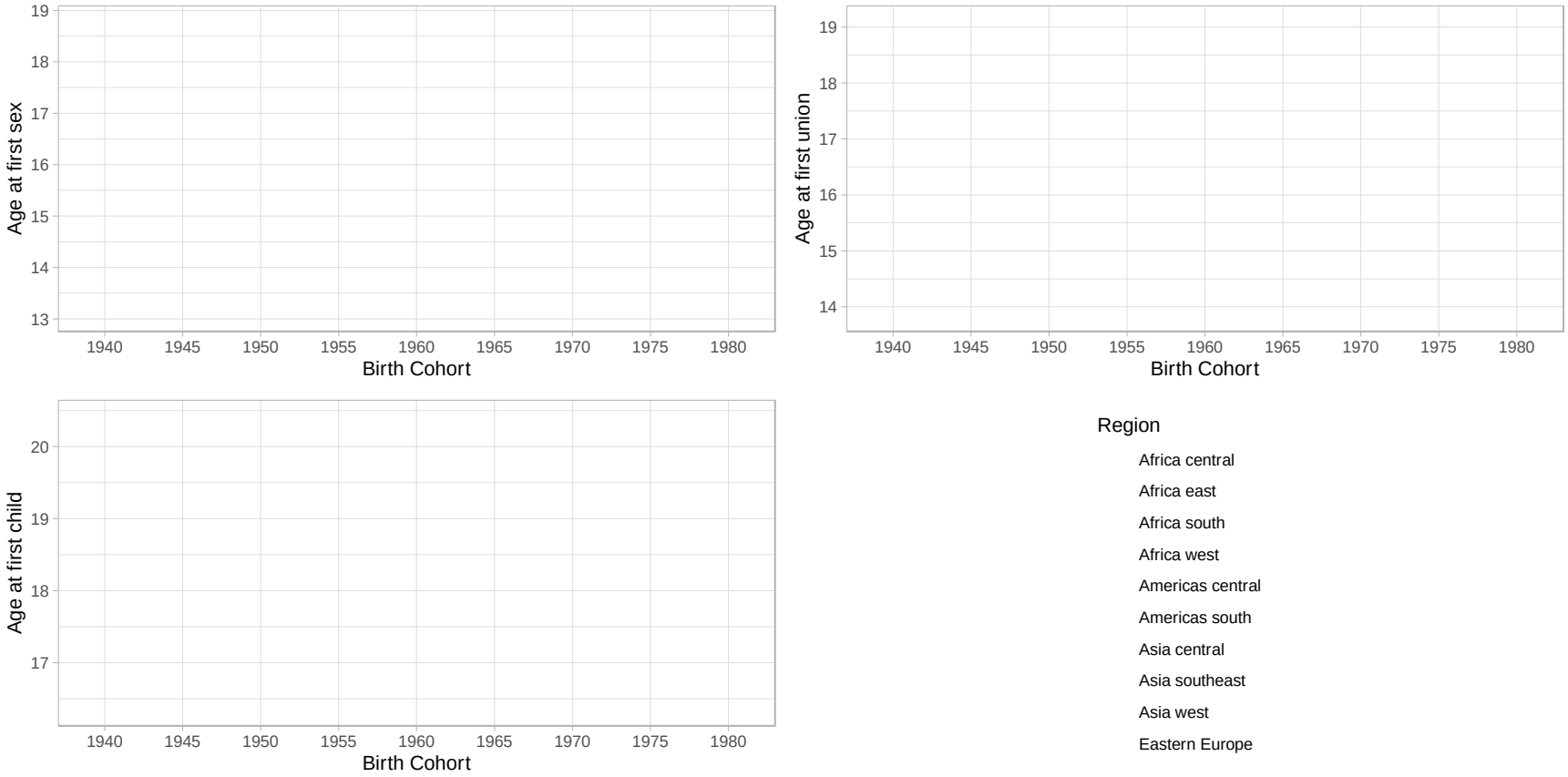
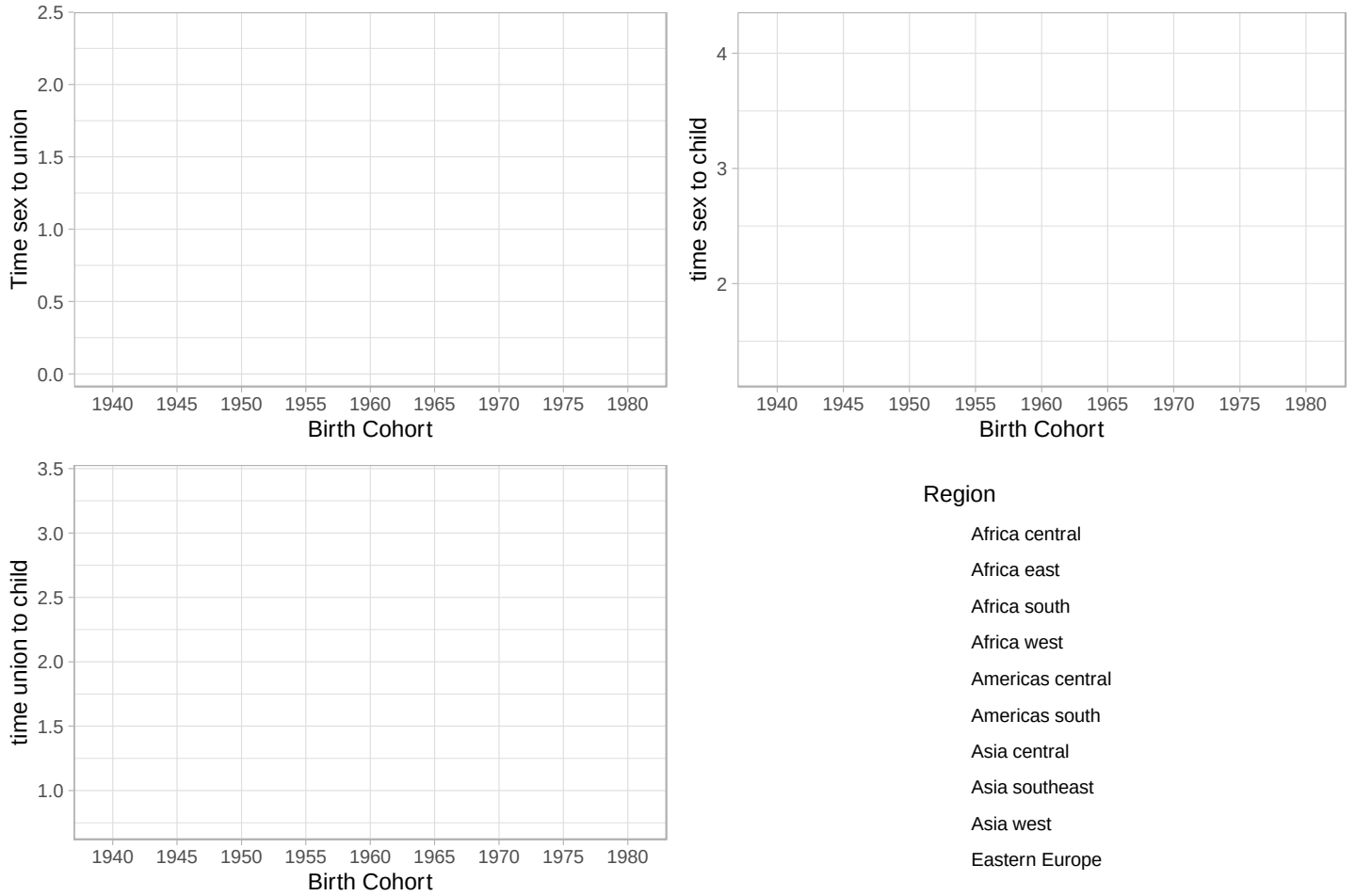


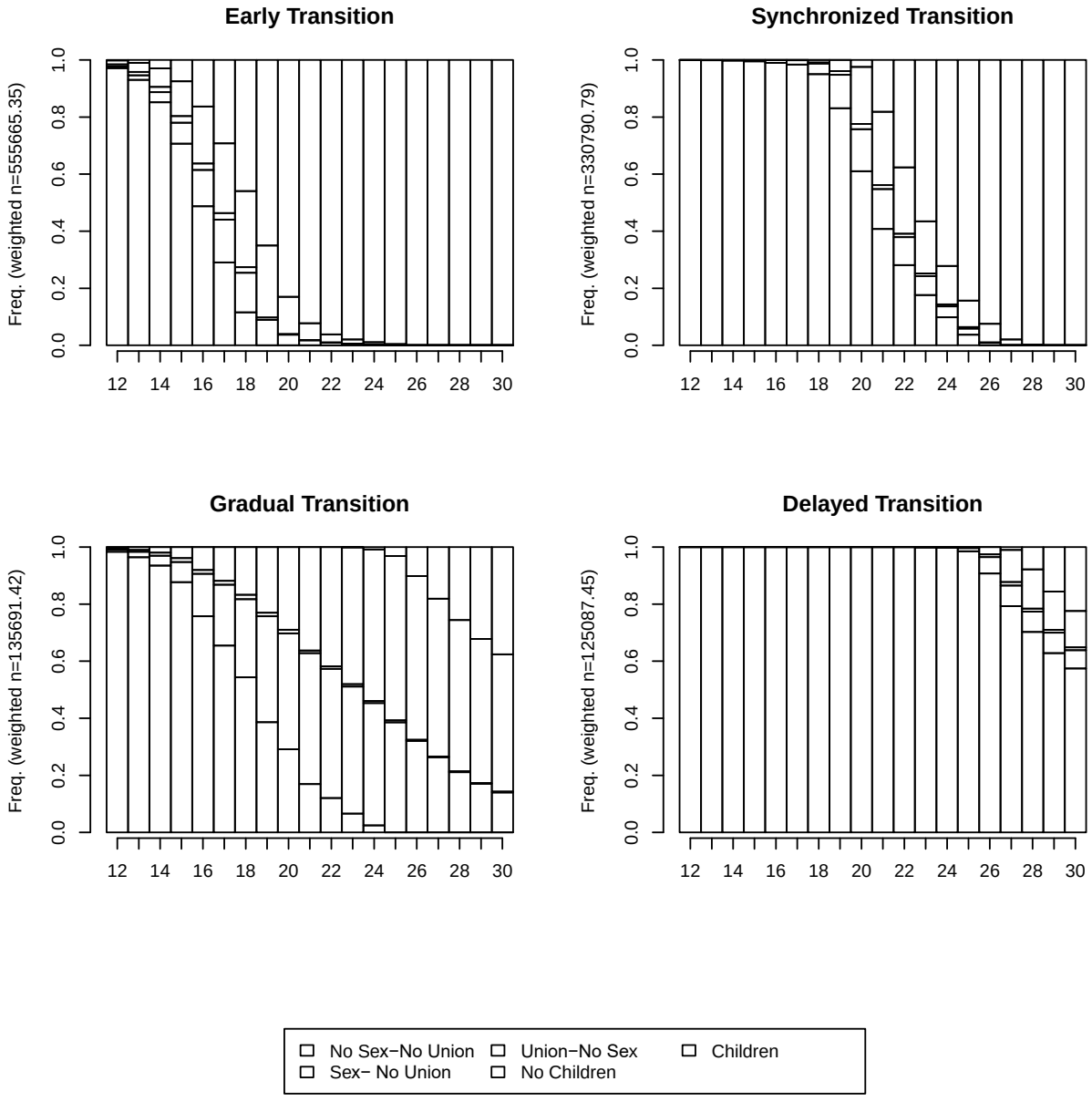
**Figure 3**: Time between specific events: sex to union (panel a, top left), sex to child (panel b, top right), union to child (panel c, bottom left)



*Clusters*

The analysis proceeds by obtaining clusters to describe typologies of transition to adulthood. The cluster analysis suggests the emergence of four distinct groups, labeled for convenience and ease of reference according to their respective characteristics. Figure 4 describes the importance of the five states described above in differentiating the typologies of clusters.

**Figure 4:** Cluster analysis: features and typologies




Cluster 1, labeled the *'Early Transition'* cluster, is characterized by early transitions to first sexual intercourse outside of a union: by age 16, half of the women in this cluster have experienced this event,

followed by a relatively short period in a union with no children, and then a quick transition to first birth. By age 17, all women in this cluster have experienced sexual intercourse, and by age 23 almost every woman had her first child. The average interval between sexual intercourse and first child is 2 years. The second cluster, labeled the '*Synchronized Transition'* cluster, is characterized by short interval times between transitions. The main feature of this cluster is that once the transition starts (around age 18), most events happen in a very short span of time, on average less than 2 years. The third cluster, '*Gradual Transition'*, shows a slower and more gradual transition to first sexual intercourse, as well as a slower transition into a union and into childbearing. The average age at transition to first sexual intercourse is 16, followed by an average interval of 4 years before entering a union and 4 additional years before childbearing. Also, 38% of women in this typology have not had children by age 30. The last cluster, *'Delayed or Late Transition'*, is characterized by extremely delayed TTA patterns where until age 25 no one in the cluster has experienced childbearing, and by age 30 less than half of the sample has started the transition. In Table 2, we report the sample sizes – frequencies and weighted percentages – in each typology. The 'Early Transition' cluster is the most frequent one with 48.4% of women belonging to this group, followed by 28.8% in the 'Synchronized Transition' cluster, 11.8% in the 'Gradual Transition' typology, and 10.9% in the 'Delayed Transition' one.

**Table 2**: Sample sizes and labels of the clusters

| Cluster | Label | Frequencies | % (weighted) |
|---|---|---|---|
| 1 | Early Transition | 554,228 | 48.44 |
| 2 | Synchronized Transition | 329,535 | 28.83 |
| 3 | Gradual Transition | 135,540 | 11.83 |
| 4 | Delayed Transition | 124,750 | 10.90 |

Each typology of TTA has its distinctive features, further outlined in Table 3. In the 'Early Transition' cluster, the mean age at each event is lower than 16. On average, women in this group start their TTA when they are just 13 years old and have their first child by age 16. Also, the time between events is relatively short: 8 months between first sex and first union, and 1.56 years between first union and first child. The 'Synchronized Transition' cluster is also characterized by short intervals between events (8 months between first sex and first union, and 1.31 years between first union and first child), but the TTA of these women starts later, given that they experience their first sexual intercourse when they are 18.4 years old and they have their first child when they are 20.4, on average. Women in the 'Gradual Transition' typology starts their TTA after those in the 'Early Transition' group, but earlier than those in the 'Synchronized Transition' one, since the mean age at first sex is 16. The transition in this third cluster is also slower, since the mean age at first child is 24.9, and almost 4 years elapse in between first sex and first union and in between first union and first child. Finally, the 'Delayed Transition' cluster

is characterized by a late start, with the mean age at first sexual intercourse being 24.9 (the same age as when women in the 'Gradual Transition' group has their first child) and the mean age at motherhood being 26.3. Even though the beginning of the TTA is postponed, the transition from one event to the other is relatively quick: 5 months between first sex and first union, and 1.2 years between first union and first child). Figure A3 in the Appendix complements Table 3 by providing a graphical depiction of the average time spent in each state, by cluster of sequences.

**Table 3**: Mean age and time between events, by cluster

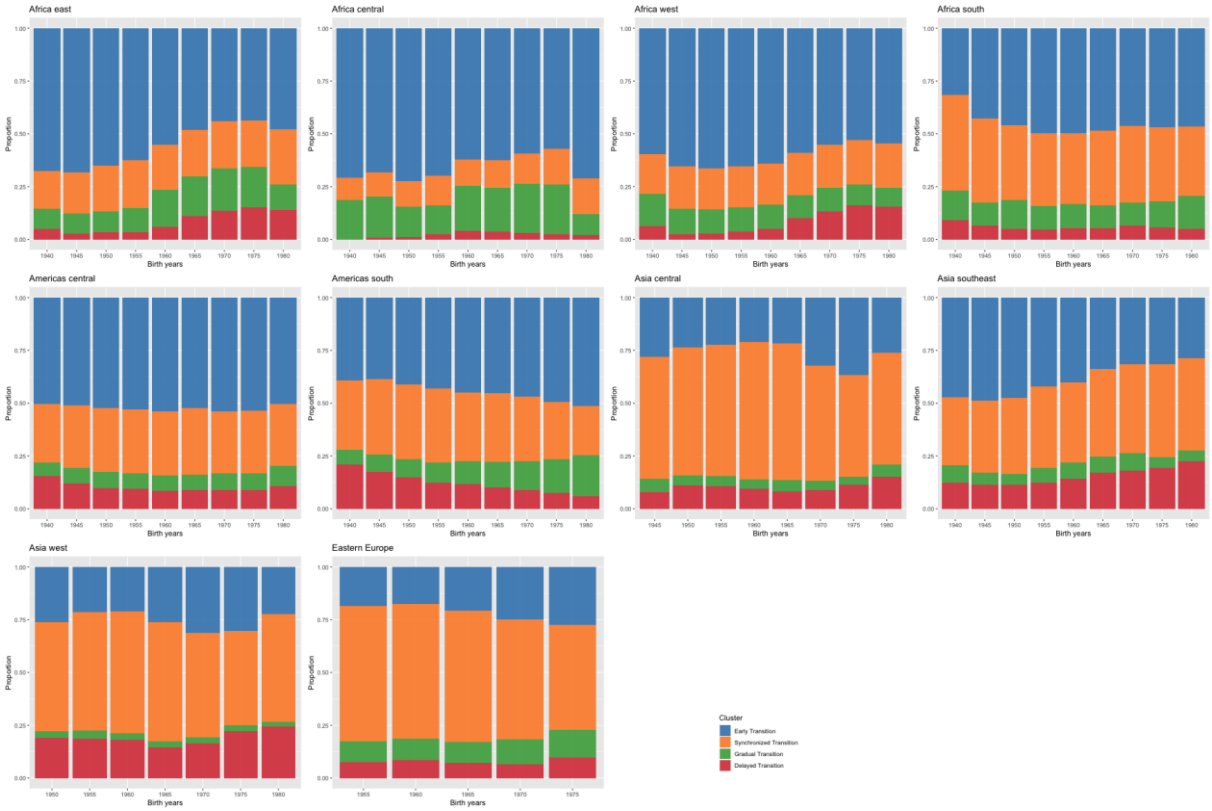
| | | **Early Transition** | **Synchronized Transition** | **Gradual Transition** | **Delayed Transition** | **All sample** |
|---|---|---|---|---|---|---|
| Age at first transition | Mean | 13.33 | 18.34 | 15.77 | 24.99 | 15.80 |
| | (SD) | (1.77) | (2.13) | (2.92) | (1.65) | (3.73) |
| Age at first sexual intercourse | Mean | 13.48 | 18.43 | 16.02 | 24.93 | 15.80 |
| | (SD) | (1.82) | (2.15) | (2.86) | (1.53) | (3.60) |
| Age at first union | Mean | 14.03 | 19.04 | 19.61 | 25.39 | 16.80 |
| | (SD) | (2.13) | (1.98) | (4.55) | (1.54) | (4.02) |
| Age at first child | Mean | 15.59 | 20.35 | 24.93 | 26.29 | 18.12 |
| | (SD) | (2.23) | (1.96) | (1.81) | (1.17) | (3.84) |
| Time between sex to union (yrs) | Mean | 0.70 | 0.70 | 3.94 | 0.38 | 1.05 |
| | (SD) | (1.46) | (1.42) | (4.16) | (0.98) | (2.20) |
| Time between sex to child (yrs) | Mean | 2.13 | 1.93 | 9.13 | 1.48 | 2.45 |
| | (SD) | (2.02) | (1.65) | (2.72) | (1.15) | (2.54) |
| Time between union to child (yrs) | Mean | 1.56 | 1.31 | 4.22 | 1.18 | 1.61 |
| | (SD) | (1.82) | (1.24) | (4.13) | (0.99) | (1.96) |

*Cluster trends over time and place*

Figure 5 reports ten panels, one per subregion, and shows which clusters characterize each region, and the extent to which the prevalence of the clusters has changed over time by birth cohort within each subregion.[2] Figure 6 reports the overall prevalence of each cluster in every country included in the analysis (while the most common cluster in each country by birth cohort is reported in Appendix Figure A4). At first glance, we observe that the 'Early Transition' cluster is still the most prevalent in Africa and South and Central America, with between 40% and 60% of women in this cluster across all birth cohorts. The prevalence of this typology was also high in southeast Asia, but the percentage of women in this group decreased from 47% for those born in 1940s to 29% for those born in 1980s. Moreover, this cluster is

not as common in Central Asia, West Asia, and Eastern Europe, with a prevalence of between 20% and 30% across cohorts.

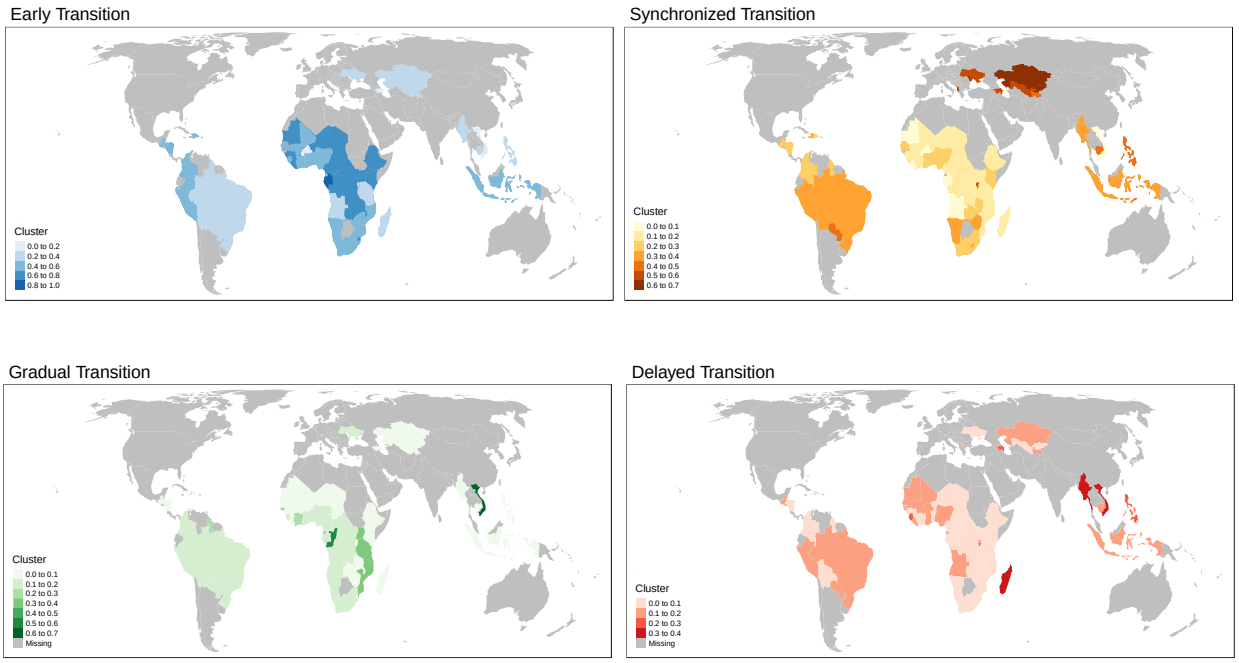
**Figure 5:** Evolution over birth cohorts of cluster typologies for each subregion




In most subregions, the proportion of women belonging to this typology hasn't fluctuated substantially over time, with the exception of a decreasing trend in East Africa and South East Asia, and an increasing prevalence in South America. This may be suggestive of a process of demographic transition which is complete in subregions where the cluster is almost absent (e.g., Central Asia, West Asia and Eastern Europe), more advanced in subregions where the prevalence has narrowed across cohorts (e.g., East Africa and South-East Asia), and still unfolding or stalling in subregions where the prevalence of an early transition is still very common and stable across cohorts (e.g., Central and West Africa). The stable shares of women undergoing early transitions in Central and West Africa are also consistent with cultural specificities of these regions, such as rooted patriarchal norms, tighter control over age at marriage, wider age differences between spouses, and very high prevalence of child marriage (Odimegwu, 2020; Tabutin & Schoumaker, 2004). The second cluster, 'Synchronized Transition' doesn't show particular changes in trends over time. It has become slightly more prevalent in South East Asia, and less prevalent in South America. This cluster is the most prevalent in Asia and Europe: more than half of women in Central Asia and Eastern Europe fall into this typology. Also, the 'Gradual Transition' typology doesn't present substantial fluctuations by birth cohorts in most regions, except for a gradual increase across cohorts in South America and East Africa, from about 10% to 20% in both contexts. The 'Late Transition' cluster is the least prevalent one in sub-Saharan Africa, yet one that shows relevant changes over time in most

contexts. There has been a substantial increase in its prevalence in East and West Africa, from almost non-existent to between 10% and 15%, and a steady high prevalence in East Asia, followed by South East Asia. Although perhaps speculative and outside the scope of this paper, the increasing importance of the 'Gradual Transition' and 'Delayed Transition' typologies in contexts such as Asia and specific parts of sub-Saharan Africa is a trend that well fits with the idea of emerging Second Demographic Transition (SDT) pillars in some low- and middle-income countries.

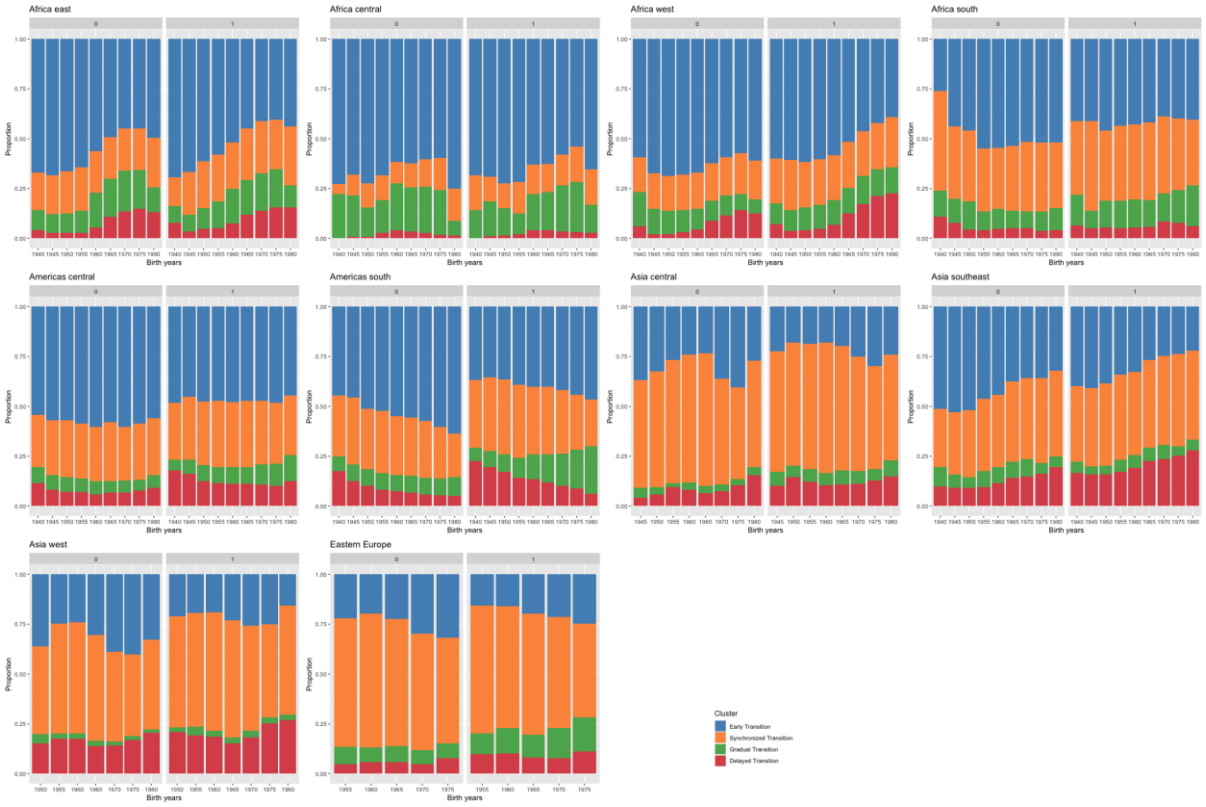
**Figure 6**: Cluster prevalence, by country




*Differences by urban and rural settings and relationship with GDP per capita*

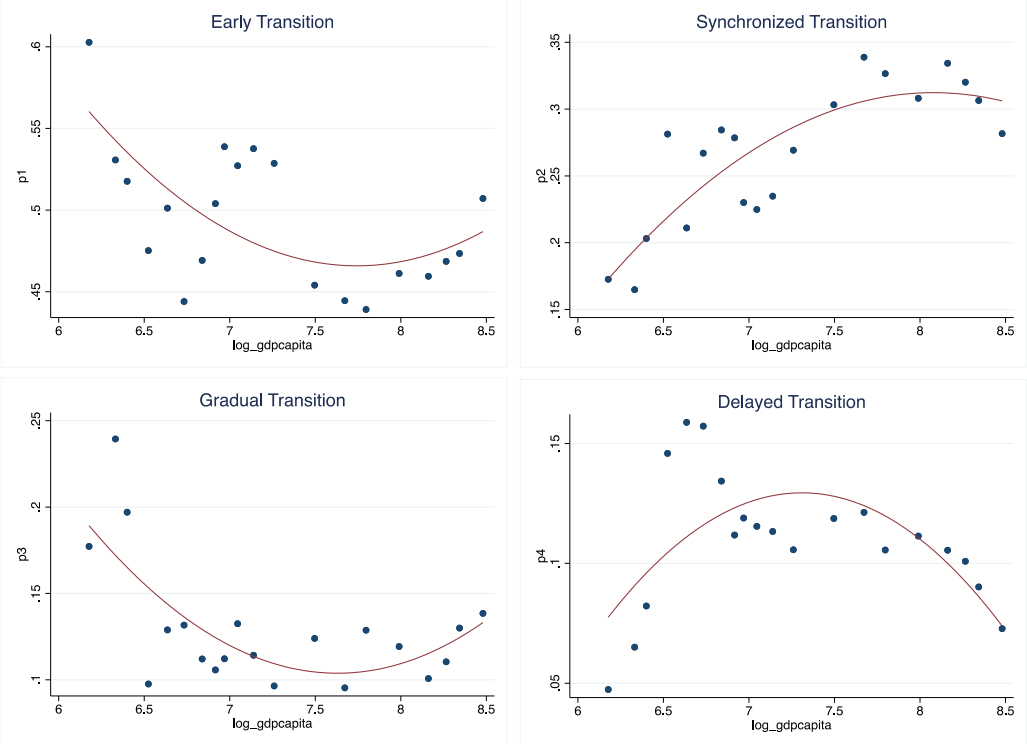
The data further allow us to distinguish between urban and rural settings. We exploit this information to indirectly get at whether changes in life-course trajectories vary by level of development and urbanization. We hence replicate Figure 5 having distinct panels for urban and rural locations within each subregion (Figure 7). Results shows that, on the whole, there are no marked differences within subregions among those who live in urban versus rural areas. There are small differences in levels, but barely any difference in trends across birth cohorts. The only visible exception is South America, where the 'Early Transition' cluster became relatively more prevalent in rural areas compared to urban areas: its prevalence increased from 35% to almost 45% in urban areas, while it increased from 45% to 65% in rural areas. Conversely, the 'Gradual Transition' cluster has become more common in urban areas (from 10% to almost 25%), but not in rural areas, where the prevalence has remained between 10% and 15% across cohorts. Also, in Latin America (both South and Central), West Africa, and Asia (both South East and West), the 'Late Transition' cluster has expanded proportionally more in urban areas.

**Figure 7:** Evolution over cohort of cluster typologies by subregion, by urban (1-right)/rural (0-left)




The picture presented in Figure 7 is suggestive of the idea that urbanization is only one possible driver – and even a rather minor one – underlying changes in the transition to adulthood in some of the subregions considered. We thus conclude this analysis by investigating the association between GDP per capita at the country-cohort level and the probability of belonging to each TTA cluster.[3] Figure 8 reports the predicted probability of being in each typology of the TTA based on the logarithm of GDP per capita.[4] Results presented are from a multinomial logistic regression with the urban/rural dummy as control variable and with clustered standard errors. As we can see from Figure 8, the probability of being in the 'Early Transition' group is highest at low levels of GDP per capita, and it decreases - at a diminishing rate - as GDP per capita increases. We do observe a very similar trend in the 'Gradual Transition' typology, even though the associated probabilities are on average lower than in the 'Early Transition' cluster. An opposite trend is visible in the 'Synchronized Transition' typology. In fact, the probability of belonging to this group is lowest for low levels of GDP per capita and it increases – at a diminishing rate – as GDP increases. Finally, there is an inverted U-shape association between the log(GDP per capita) and the probability of being in the 'Delayed Transition' cluster. The probability is low for low and high levels of GDP per capita, and highest in the middle of the GDP distribution.

**Figure 8**: Association between GDP per capita and the probability of belonging to each cluster




## Conclusions and Discussion

This study has provided a rich account of women's transition to adulthood patterns across 69 LMICs using all available Demographic and Health Surveys. In so doing, we contributed to the relevant literature in two directions. Theoretically, we shifted the focus from the study of independent life-course transitions towards a holistic approach that characterizes the life course as a complex combination of events, thus reflecting the idea that background forces – such as educational expansion, urbanization, and cultural, societal, and institutional changes – do exert an influence on the whole "package" of TTA status combinations that young adults go through, rather than on single events. Methodologically, we built partnership and fertility trajectories using sequence analysis; we computed novel summary measures directly from the sequences, and we identified clusters of typical life-course trajectories using a robust clustering algorithm (*Agglomerative Hierarchical Clustering*) based on optimal-matching distances of sequences.

In line with other studies focusing on independent life-course transitions in LMICs (e.g., Bongaarts et al., 2017; Pesando & GFC Team, 2019), we found that – on the whole – the transition to adulthood has been delayed in LMICs. However, our analyses reveal a high degree of cross-regional idiosyncrasies such as declining mean ages at first sexual intercourse in South America, declining time intervals between first sexual intercourse and first child across East and West Africa, and declining time

intervals between first union and first child across all sub-regions considered, but particularly so across West and Central Africa. These results combined suggest that there is a general delay in entering into first union after experiencing first sexual intercourse, but that once women get married or start a coresidential union, the transition to motherhood is relatively quick. These are simple findings, yet some that add nuances to the somehow "unidirectional" idea that the transition to adulthood has simply been delayed, and also some that – to the best of our knowledge – have not yet been documented using analogous measures drawn directly from sequences.

We then identified clusters of TTA in an innovative data-driven manner and summarized the ensuing heterogeneity through four cluster solutions, labeled 'Early Transition,' 'Synchronized Transition,' 'Gradual Transition,' and 'Delayed Transition.' These are four very distinct typologies that concisely summarize the high volume of existing heterogeneity. 'Early Transition' – characterized by quick transitions to first sexual intercourse outside of unions, followed by a short period in a union with no children, and then a quick transition to first birth – remains the most common typology for all LMICs combined, with almost 50% of women belonging to it (especially across sub-Saharan Africa) and significant declines across cohorts in East Africa and South East Asia, while less so in South Africa, where prevalence of this typology was lower to start with. Conversely, the opposite cluster typology, 'Delayed Transition' – characterized by a late start, with a mean age at first sexual intercourse around 25 and a relatively quick transition to motherhood – is the least common typology for all LMICs combined, yet still about 10% of women belong to it. This typology is the least prevalent in sub-Saharan Africa, yet one that has shown relevant changes over time in most contexts. For instance, there has been a substantial increase in its prevalence in East and West Africa, from almost non-existent to between 10% and 15%, and a steady high prevalence in East Asia, followed by South East Asia. We believe the existence of such cluster of TTA is a demographic phenomenon that went quite unnoticed in the family literature existing to date. While its prevalence sounds less surprising in the Asian context, its existence and significant increase across cohorts in East and West Africa is a novel finding which will deserve an in-depth analysis of its own. Overall, our results point to marked differences across subregions and "novel" cluster typologies, thus underscoring the importance of investigating cross-regional differences in partnership and fertility trajectories looking at the interrelation among different events in a holistic way.

Focusing specifically on changes overtime, when looking at all LMICs combined we did not observe massive changes across cohorts. Rather, there seems to be relative stability of cluster typologies over time. Indeed, specific subregions underwent more apparent changes away from an 'Early Transition' and towards a 'Gradual and Delayed Transition', such as East Africa and South East Asia, yet in other regions stability has been the norm, such as Central Africa (with high and stable prevalence of 'Early Transition'), West Asia (with high and stable prevalence of 'Synchronized' and 'Delayed Transition'), and Eastern Europe (with high and stable prevalence of 'Synchronized' and 'Gradual Transition'). Although not much literature exists on changes in the transition to adulthood in LMICs, these findings align with recent studies that have provided evidence of slow changes and little convergence in family domains in

LMICs (Castro et al., 2019; Cherlin, 2016; Pesando & GFC Team, 2019). In this sense, our findings confirm the idea that union formation and fertility practices might be less responsive to socioeconomic changes in LMICs because they are tied to elements of the social structure that are more resistant to change. These structural features of societies include religious beliefs, marriage-related-laws and prohibitions, inheritance rights, shortages in the housing market, and persistent disparities in gender roles and dynamics (Coontz, 2014).

Although the analysis was not set up to explicitly test for cross-regional convergence in TTA overtime, our results are suggestive of a reality that is far from convergence and likely supportive of persistent differences across regions. Going back to the specific literature that inspired this study, our results align with findings from Grant and Furstenberg (2007), who observed relatively small changes – or high inertia – in the age-specific index of heterogeneity from the first to the most recent survey in each country. We see this high level of consistency as valuable given the far broader country coverage of our study, the larger sample sizes, and the different and varied methodological approaches.

Indeed, given this set of findings, the next logical step is to understand and identify at least some drivers of these changes (or lack thereof). Our analyses by rural/urban location of residence and relationship with GDP per capita at the country level were initial attempts to get at underlying drivers. Except for South America, where differences between rural and urban areas were apparent due a more marked increase in the 'Gradual Transition' typology and higher levels of the 'Delayed Transition' typology in urban areas, differences were quite minor elsewhere. We took this evidence as pointing towards the idea that urbanization was only one minor driver of changes in the TTA in LMICs. Similarly, the relationships with GDP per capita, albeit interesting, do not point towards a unidirectional interpretation whereby more "secularized" clusters ('Gradual Transition' and 'Delayed Transition') are more prevalent where GDP per capita is higher, thus complicating further long-standing debates on the relationship between family change and socio-economic development and pointing, once again, towards a scenario that best conforms to the summary label of 'persistent diversity with development' (Pesando & GFC Team, 2019). The lack of explanatory power of household location of residence, alongside the complex relationship identified with GDP per capita, call for the need to look for additional factors that might explain the heterogeneity in the TTA across LMICs, possibly the expansion of education, cultural and institutional variables, or information on different socioeconomic strata within countries. As a matter of fact, albeit indirectly, our results point towards a reality whereby homogeneity in TTA patterns seems higher within regions than across them.

Returning to the original puzzle in this study, the question of whether young people in LMICs have experienced processes of de-standardization of the life course similar to those observed in high-income societies does not lend itself to a unique answer. In contrast to what Billari & Liefbroer documented in Europe (2010), pathways to adulthood have not changed in the same direction across all

LMICs towards a late, protracted, and complex TTA. To the contrary, excepting a few generalized and common trends, each sub-region seems to have followed very diversified pathways.

Our data-driven methodological approach delivers four clearly distinct typologies which, in some respect, impose order and structure to the existing heterogeneity – a finding which would suggest that the transition to adulthood in LMICs is instead ordered and structured. If we took our clusters as a starting point, we could summarize this whole scenario by stating that the TTA in LMICs is ordered and structured, yet it has changed in ways that do not necessarily resemble changes observed in high-income societies. Ultimately, an assessment of this kind hinges upon theoretical perspectives that have been adopted to understand changes in families across the world with socio-economic development. The reader might have noticed that we explicitly shied away from dealing with the applicability of the SDT in this study, given the heated debates surrounding it (Zaidi & Morgan, 2017). We did so because we could not explicitly test for it in great detail. We will simply state that the emergence of the 'Gradual Transition' cluster – and the 'Delayed Transition' cluster, to a lesser extent – resembles some of the pillars of the SDT. The fact that these clusters are more apparent in South East Asia, West Asia, and Eastern Europe would also support the idea that some SDT forces are at play. Yet our findings combined – especially the complex relationships with GDP per capita – seem most supportive of a scenario in which the SDT has limited applicability to selected subregions, rather than to all LMICs combined.

The study has limitations that lay the ground for subsequent research. First, our analysis only focuses on women, mostly to prioritize breadth and country coverage. An alternative study using DHS could be conducted focusing on the same life-course transitions for men, though sample size would be significantly reduced. We nonetheless see this as a fruitful avenue for future research. Second, although not a limitation *per se*, we focus on first child rather than the whole fertility history to better characterize heterogeneity in the earlier part of the life course. Third, although we look at changes overtime for groups of countries (rather than countries individually), when evaluating changes across cohorts we run the risk that some trends might be masked by the different composition of countries by birth cohorts. Fourth, although it would be ideal to predict the four cluster solutions as a function of multiple country-specific characteristics, our analyses by birth cohort – with women born as early as the 1930s – are such that it is hardly impossible to obtain time series for variables such as the Human Development Index or Gender Inequality Indicators that date back to such a period. Our use of GDP per capita was an attempt in this direction, yet we acknowledge that the scope and ambition of this study remains predominantly descriptive.

**References**


Aassve, A., Billari, F. C., & Piccarreta, R. (2007). Strings of Adulthood: A Sequence Analysis of Young British Women's Work-Family Trajectories. *European Journal of Population*, *23*, 369–388.

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

Abbott, A., & Tsay, A. (2000). Sequence Analysis and Optimal Matching Methods in Sociology: Review and Prospect. *Sociological Methods & Research*, *29*(1), 3–33.

Aisenbrey, S., & Fasang, A. F. (2010). New Life for Old Ideas: The "Second Wave" of Sequence Analysis Bringing the 'Course' Back Into the Life Course. *Sociological Methods & Research*, *38*(3), 420–462.

Barban, N. (2013). Family Trajectories and Health: A Life Course Perspective. *European Journal of Population*, *29*(4), 357–385.

Barban, N., & Billari, F. C. (2012). Classifying life course trajectories: A comparison of latent class and sequence analysis. *Journal of the Royal Statistical Society. Series C: Applied Statistics*, *61*(5), 765–784.

Barban, N., & Sironi, M. (2019). Sequence Analysis as a Tool for Family Demography. In R. Schoen (Ed.), *Analytical Family Demography* (The Spring, pp. 101–123). Springer, Cham.

Beguy, D., Kabiru, C. W., Zulu, E. M., & Ezeh, A. C. (2011). Timing and Sequencing of Events Marking the Transition to Adulthood in Two Informal Settlements in Nairobi, Kenya. *Journal of Urban Health*, *88*(SUPPL. 2), 318–340.

Behrman, J. R., & Sengupta, P. (2005). Changing Contexts in Which Youth are Transitioning to Adulthood in Developing Countries: Converging Toward Developed Economies? In C. B. Lloyd (Ed.), *The Changing Transitions to Adulthood in Developing Countries: Selected Studies* (pp. 13–55). National Academies Press.

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*(1), 37–65.

Billari, F. C., & Liefbroer, A. C. (2010). Towards a New Pattern of Transition to Adulthood? *Advances in Life Course Research*, *15*(2–3), 59–75.

Billari, F. C., & Piccarreta, R. (2005). Analyzing Demographic Life Courses through Sequence Analysis. *Mathematical Population Studies*, *12*(2), 81–106.

Bongaarts, J., Mensch, B. S., & Blanc, A. K. (2017). Trends in the age at reproductive transitions in the developing world: The role of education. *Population Studies*, *71*(2), 139–154.

Casterline, J. B., Williams, L., & McDonald, P. (1986). The age difference between spouses: Variations among developing countries. *Population Studies*, *40*(3), 353–374.

Castro, A. F., Pesando, L. M., Kohler, H.-P., & Furstenberg, F. F. (2019). Family Change and Variation Through the Lens of Family Configurations in Low- and Middle-Income Countries. In *University of Pennsylvania Population Studies Center Working Papers (PSC/PARC)* (No. 2019–31; Population Studies Center Working Papers (PSC/PARC)).

Cherlin, A. J. (2016). A Happy Ending to a Half-Century of Family Change? *Population and Development Review*, *42*(1), 121–129.

Clark, S., & Brauner-Otto, S. (2015). Divorce in sub-Saharan Africa: Are Unions Becoming Less Stable? *Population and Development Review*, *41*(4), 583–605.

Coontz, S. (2014). *Marriage, A History: How Love Conquered Marriage.* (1st ed.). Penguin Books.

Echarri Cánovas, C. J., & Pérez Amador, J. (2007). En Tránsito Hacia la Adultez: Eventos en el Curso de Vida de los Jóvenes en México. *Estudios Demográficos y Urbanos*, *22*(1), 43–77.

Elzinga, C. H., & Liefbroer, A. C. (2007). De-standardization of Family-Life Trajectories of Young Adults: A Cross-National Comparison using Sequence Analysis. *European Journal of Population*, *23*(3–4), 225–250.

Frye, M., & Trinitapoli, J. (2015). Ideals as Anchors for Relationship Experiences. *American Sociological Review*, *80*(3), 496–525.

Furnas, H. E. (2016). Capturing Complexities of Relationship-Level Family Planning Trajectories in Malawi. *Studies in Family Planning*, *47*(3), 205–221.

Fussell, E. (2005). Measuring the Early Adult Life Course in Mexico: An Application of the Entropy Index. *Advances in Life Course Research*, *9*(04), 91–122.

Goldberg, R. E. (2013). Family Instability and Pathways to Adulthood in Cape Town, South Africa. *Population and Development Review*, *39*(2), 231–256.

Grant, M. J., & Furstenberg, F. F. (2007). Changes in the transition to adulthood in less developed countries. *European Journal of Population*, *23*(3–4), 415–428.

Grapsa, E., & Posel, D. (2016). Sequencing the Real Time of the Elderly: Evidence from South Africa. *Demographic Research*, *35*(1), 711–744.

Hewett, P. C., & Lloyd, C. B. (2005). Progress Toward Education for All: Trends and Current Challenges for sub-Saharan Africa. In *The Changing Transitions to Adulthood in Developing Countries: Selected Studies* (pp. 84–115). National Academies Press.

Juárez, F., & Gayet, C. (2014). Transitions to Adulthood in Developing Countries. *Annual Review of Sociology*, *40*(1), 521–538.

Kruskal, J. B. (1983). An Overview of Sequence Comparison: Time Warps, String Edits, and Macromolecules. *SIAM Review*, *25*(2), 201–237.

Lesnard, L. (2006). Optimal Matching and Social Sciences. *Centre de Recherche En Économie et Statistique Working Paper Series*, *2006*(1).

Lloyd, C. B., & Mensch, B. S. (2008). Marriage and Childbirth as Factors in Dropping out from School: An Analysis of DHS Data from Sub-Saharan Africa. *Population Studies*, *62*(1), 1–13.

Mensch, B. S., Soler-Hampejsek, E., Kelly, C. A., Hewett, P. C., & Grant, M. J. (2014). Challenges in Measuring the Sequencing of Life Events Among Adolescents in Malawi: A Cautionary Note. *Demography*, *51*(1), 277–285.

National Research Council and Institute of Medicine. (2005). *Growing Up Global: The Changing Transitions to Adulthood in Developing Countries* (C. Lloyd (ed.)). National Academies Press.

Odimegwu, C. (2020). *Family Demography and Post-2015 Development Agenda in Afrca*. Springer.

Oris, M., & Ritschard, G. (2014). Sequence Analysis and Transition to Adulthood: An Exploration of the Access to Reproduction in Nineteenth-Century East Belgium. In P. Blanchard, F. Buhlmann, & J. A. Gauthier (Eds.), *Advances in Sequence Analysis: Theory, Method, Applications* (pp. 151–167). Springer, Cham.

Pesando, L. M., & GFC Team. (2019). Global Family Change: Persistent Diversity with Development. *Population and Development Review*, *45*(1), 133–168.

Psaki, S. R., McCarthy, K. J., & Mensch, B. S. (2018). Measuring Gender Equality in Education: Lessons from Trends in 43 Countries. *Population and Development Review*, *44*(1), 117–142.

Ravanera, Z. R., Rajulton, F., & Burch, T. K. (1998). Early Life Transitions of Canadian Women: A Cohort Analysis of Timing, Sequences, and Variations. *European Journal of Population*, *14*(2), 179–204.

Rindfuss, R. R., Swicegood, C. G., & Rosenfeld, R. A. (1987). Disorder in the Life Course: How Common and Does It Matter? *American Sociological Review*, *52*(6), 785.

Robette, N. (2010). The Diversity of Pathways to Adulthood in France: Evidence from a Holistic Approach. *Advances in Life Course Research*, *15*(2–3), 89–96.

Salmela-Aro, K., Kiuru, N., Nurmi, J. E., & Eerola, M. (2011). Mapping Pathways to Adulthood among Finnish University Students: Sequences, Patterns, Variations in Family- and Work-Related Roles. *Advances in Life Course Research*, *16*(1), 25–41.

Schwanitz, K. (2017). The Transition to Adulthood and Pathways Out of the Parental home: A Cross-National Analysis. *Advances in Life Course Research*, *32*, 21–34.

Shanahan, M. J. (2000). Pathways to Adulthood in Changing Societies: Variability and Mechanisms in Life Course Perspective. *Annual Review of Sociology*, *26*(1), 667–692.

Sironi, M., Barban, N., & Impicciatore, R. (2015). Parental Social Class and the Transition to Adulthood in Italy and the United States. *Advances in Life Course Research*, *26*(1), 89–104.

Studer, M. (2013). *WeightedCluster library manual: A practical guide to creating typologies of trajectories in the social sciences with R* (No. 24; LIVES Working Papers).

Tabutin, D., & Schoumaker, B. (2004). The demography of sub-Saharan Africa from the 1950s to the 2000s: A survey of changes and a statistical assesment. *Population*, *59*(3–4), 457–556.

Tian, F. F. (2016). Transition to Adulthood in China in 1982-2005: A Structural View. *Demographic Research*, *34*(1), 451–466.

Van Winkle, Z. (2018). Family Trajectories Across Time and Space: Increasing Complexity in Family Life Courses in Europe? *Demography*, *55*(1), 135–164.

Zaidi, B., & Morgan, S. P. (2017). The Second Demographic Transition Theory: A Review and Appraisal. *Annual Review of Sociology*, *43*, 139-473–492.

## Appendix

### Tables

**Table A1:** Regional and sub-regional classification of countries, together with number of waves

| **Americas (48)** | **Asia (39)** | **Former USSR (14)** | **SSA (162)** |
|---|---|---|---|
| ***Americas Central (21)*** | ***Asia South (17)*** | ***Asia Central (6)*** | ***Africa Central (19)*** |
| Dominican Republic | Afghanistan | Kazakhstan | Angola |
| Guatemala | Bangladesh | Kyrgyzstan | Cameroon |
| Haiti | India | Tajikistan | Central Africa |
| Honduras | Maldives | Uzbekistan | Chad |
| Nicaragua | Nepal | ***Asia West (5)*** | Congo |
| ***Americas South (27)*** | ***Asia Southeast (22)*** | Armenia | DR of Congo |
| Bolivia | Cambodia | Azerbaijan | Gabon |
| Brazil | Indonesia | ***Eastern Europe (3)*** | Sao Tome and Principe |
| Colombia | Myanmar | Albania | ***Africa East (71)*** |
| Guyana | Philippines | Moldova | Burundi |
| Paraguay | Timor-Leste | Ukraine | Comoros |
| Peru | Vietnam | | Ethiopia |
| | | | Kenya |
| | | | Madagascar |
| | | | Malawi |
| | | | Mozambique |
| | | | Rwanda |
| | | | Tanzania |
| | | | Uganda |
| | | | Zambia |
| | | | Zimbabwe |
| | | | ***Africa South (9)*** |
| | | | Lesotho |
| | | | Namibia |
| | | | South Africa |
| | | | Swaziland |
| | | | ***Africa West (63)*** |
| | | | Benin |
| | | | Burkina Faso |
| | | | Cote d'Ivoire |
| | | | Gambia |
| | | | Ghana |
| | | | Guinea |
| | | | Liberia |
| | | | Mali |
| | | | Mauritania |
| | | | Niger |
| | | | Nigeria |
| | | | Senegal |
| | | | Sierra Leone |
| | | | Togo |

**Table A2:** Best number of cluster solutions according to several measures of goodness of fit available in the `weightedCluster` package (Studer, 2013).

| Index | Abbrv. | Best number of groups | Stat |
|---|---|---|---|
| Point Biserial Correlation | PBC | 4 | 0.59 |
| Hubert's Gamma | HG | 10 | 0.86 |
| Average Silouette Width | ASW | 4 | 0.38 |
| Calinksi- Harabasz index | CH | 2 | 6849 |
| Pseudo $R^2$ | R2 | 10 | 0.64 |
| Hubert's $C$ | HC | 10 | 0.06 |

**Table A3:** Summary statistics from subsample used to create cluster solution with entire sample.

| | | **Subsample** | **Whole sample** |
|---|---|---|---|
| Age at first transition | Mean | 15.79 | 15.80 |
| | (SD) | (3.69) | (3.73) |
| Age at first sexual intercourse | Mean | 16.58 | 15.80 |
| | (SD) | (4.52) | (3.60) |
| Age at first union | Mean | 16.76 | 16.80 |
| | (SD) | (3.98) | (4.02) |
| Age at first child | Mean | 18.12 | 18.12 |
| | (SD) | (3.84) | (3.84) |
| Time between sex to union | Mean | 1.05 | 1.05 |
| | (SD) | (2.52) | (2.20) |
| Time between sex to child | Mean | 2.18 | 2.45 |
| | (SD) | (2.44) | (2.54) |
| Time between union to child | Mean | 1.61 | 1.61 |
| | (SD) | (1.94) | (1.96) |
| Sample size | | 20,000 | 1,144,053 |

## Figures

**Figure A1:** Countries included in the analysis

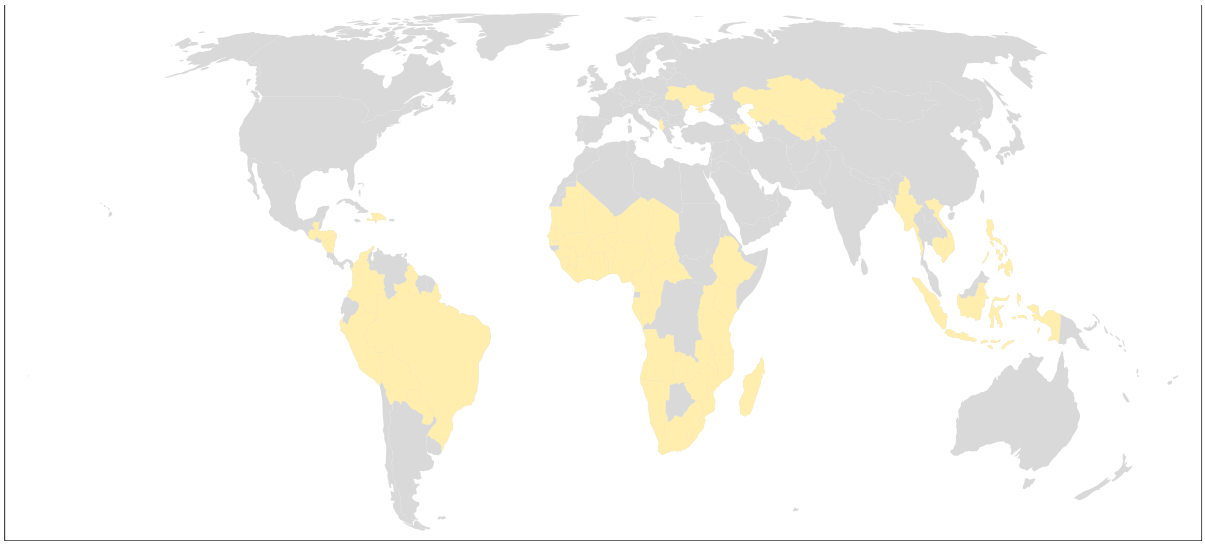


**Figure A2:** Cluster Dendrogram

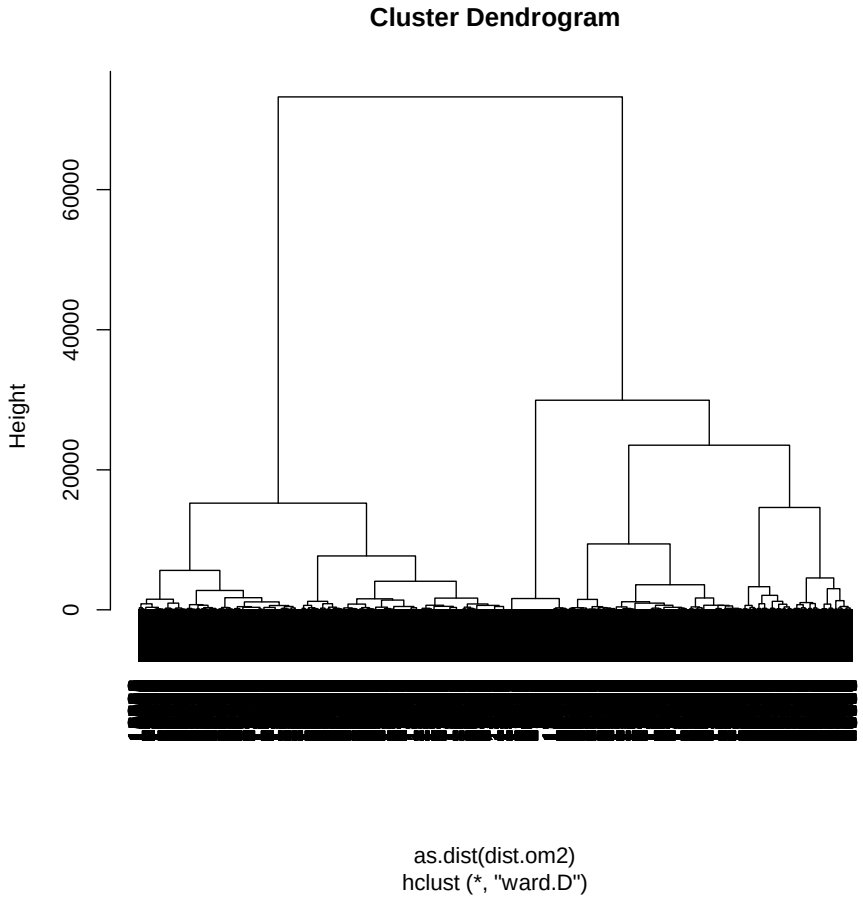


**Figure A3:** Average time spent in each state, by cluster of sequences

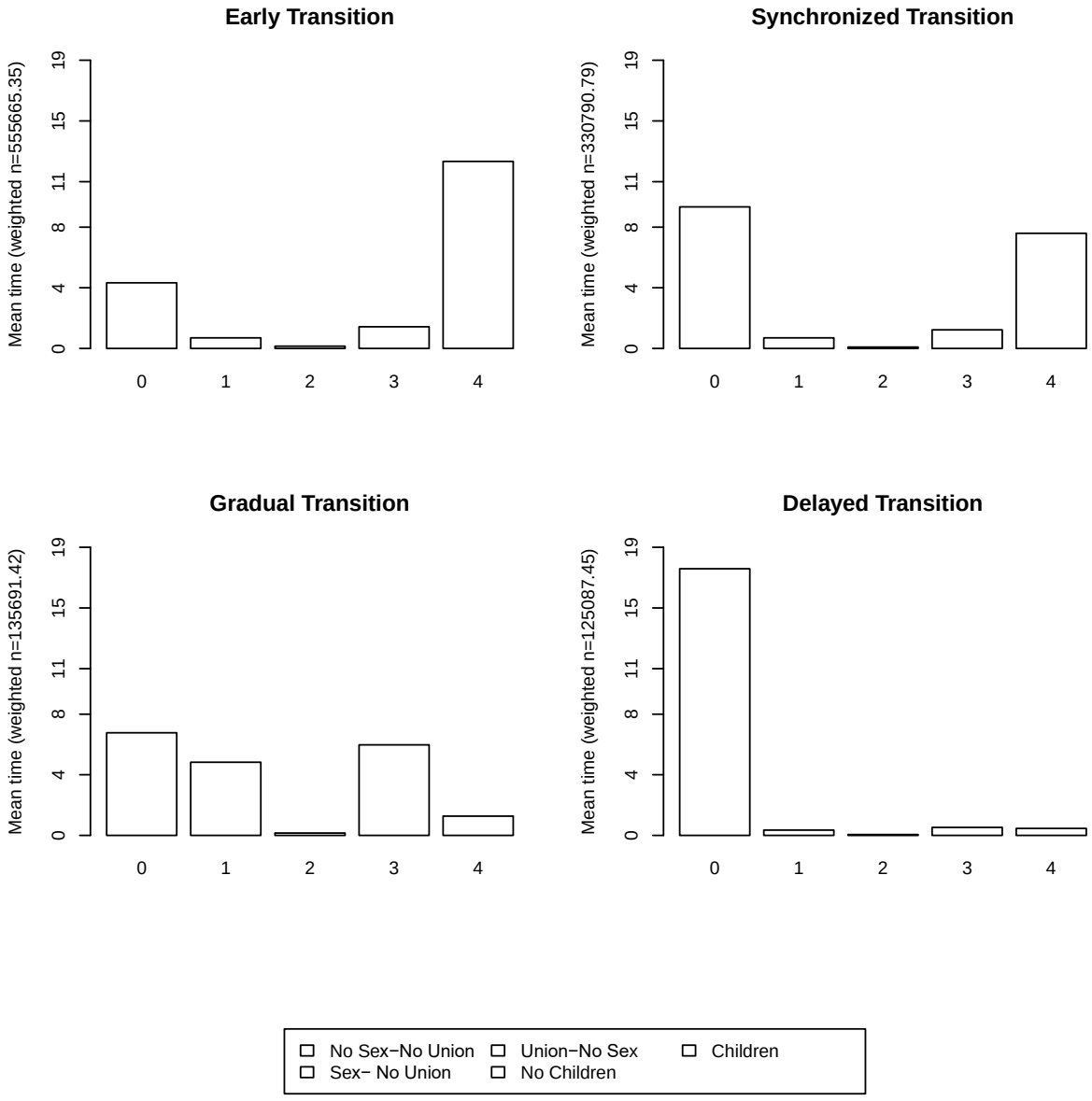
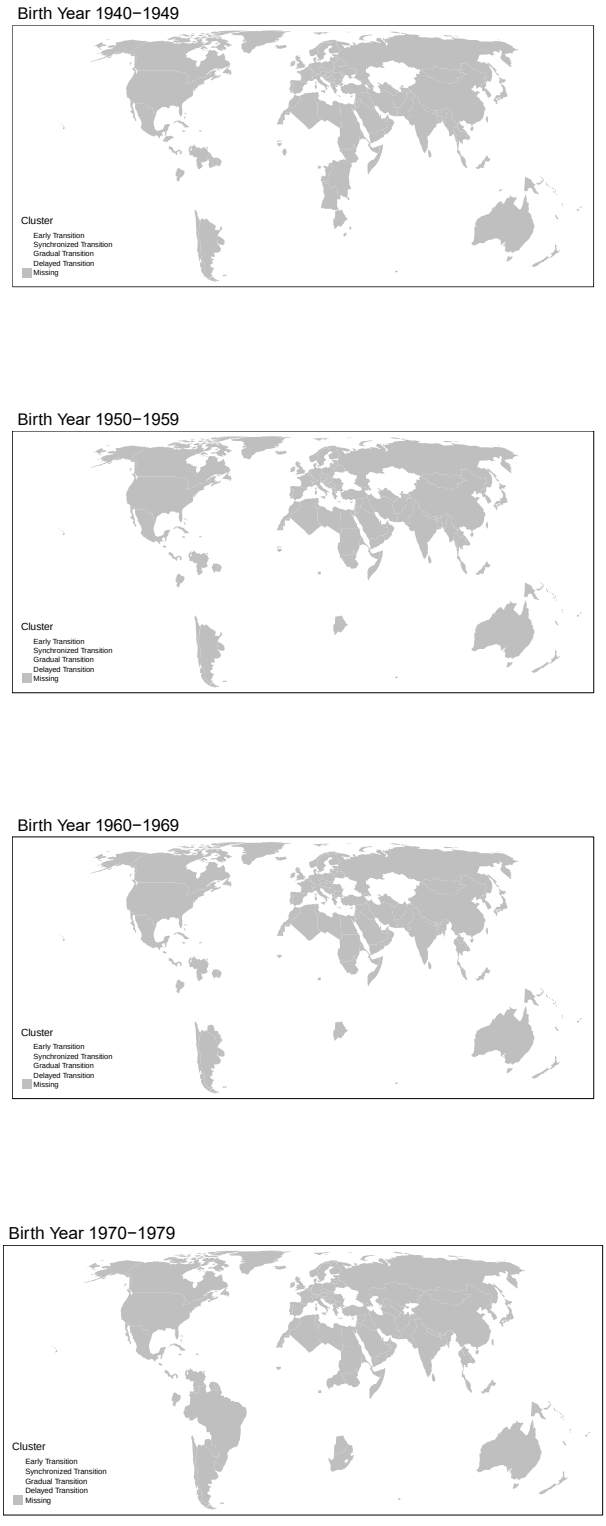


**Figure A4:** Most common cluster in each country by birth cohort



**Endnotes**

[1] We are currently undertaking two related projects adopting a similar methodological approach yet also (i) focusing on men, and (ii) focusing on complete fertility histories with attention to sex composition of births.

[2] We acknowledge that in this analysis some trends might be partly masked by the different composition of countries by birth cohorts. By looking at groups of countries (subregions), we are rather confident these compositional effects net out.

[3] We obtained time series on GDP per capita and merged these by birth cohort considering the year in which the woman was 12 years old (i.e., birth year plus 12). Note that information on GDP per capita before the year 1940 is hardly available.

[4] We also included the quadratic term of log(GDP per capita) in the multinomial logistic regression, to take into account the fact that the association might be non-linear.