Modeling Age-Specific Fertility Curves to Understand Demographic Transition Theory

Athena Pantazis[1,*] and Samuel J. Clark[1,2,3,4]

[1]Department of Sociology, University of Washington
[2]MRC/Wits Rural Public Health and Health Transitions Research Unit (Agincourt), School of Public Health, Faculty of Health Sciences, University of the Witwatersrand
[3]ALPHA Network, London School of Hygience and Tropical Medicine, London, UK
[4]INDEPTH Network, Accra, Ghana
[*]Correspondence to: apantazi@u.washington.edu

April 6, 2015

1 Introduction

Persistent high fertility in sub-Saharan Africa (SSA) has been of concern to demographers and global health practitioners for decades. Demographic transition theory posits that once begun fertility decline should be relatively steady and irreversible, and most demographers agree that fertility decline began in almost all of SSA by the late 1980s (see Bongaarts and Casterline, 2013) or the 1990s (see Caldwell et al., 1992). However, the most recent population projections projected an increase in world population, reaching 9.6 billion by 2050, largely due to high SSA fertility (United Nations, 2013). Analysis of the convergence of fertility has demonstrated that high fertility in SSA has been the outlier in global fertility convergence for decades (Dorius, 2008). Theoretical and empirical work seeking to understand the pace of fertility decline in SSA and whether or not fertility decline in some countries has stalled has not yet provided a definitive explanation that situates SSA fertility trends within global fertility trends and demographic transition theory satisfactorily. This study aims to understand trends in SSA fertility from 1950-2010 in the context of global fertility trends though the analysis of age-specific fertility rates, from the UN World Population Prospects, as curves.

Sub-Sahara Africa (SSA) is the last world region to undergo fertility transition and has seen the slowest decline. The transition in SSA is the most recent and SSA began its transition at, and currently has, the highest fertility. Figure 1 shows trends in total fertility rates (TFR) for the world and world regions, and SSA regions from 1950-55 until 2005-10. SSA, including East Africa, West Africa and Central Africa have curves that, despite some evidence of decline, remain at high levels of fertility, while all other regions of the world have seen a steady decline through the last half of the 20th Century, with TFRs well below 4 by 2005-10. Only the Southern African transition, led by South Africa, resembles the transitions seen elsewhere in the developing world (Shapiro, 2012). According to Bongaarts and Casterline (2013), most of SSA is still early in the transition with high fertility, citing the different reproductive behaviors in African countries, compared to non-African countries. The transition did not begin in SSA until at least the 1980s, and the fertility levels at the beginning of the transition were higher in SSA than they were elsewhere at the onset of fertility decline; the average total fertility rate at the beginning of decline was 6.5 in SSA compared to 5.8 elsewhere (Bongaarts and Casterline, 2013). Additionally, the recent pace of the fertility decline in SSA is much slower than the pace of decline in Asia and Latin America in the 1970s. Ideal family size in SSA pre-transition was higher, by about 1 child, than ideal family size in pre-transition Asia, and while ideal family size and TFR are closely correlated and have declined in almost all countries, at any TFR ideal family size is higher in SSA than in Asia or Latin America and ideal family size in SSA remains higher (at about 4.6) than it was in Asia pre-transition (Bongaarts and Casterline, 2013). The persistent high fertility in SSA, in addition to being a subject of importance for population projections, has posed challenging to incorporate into demographic transition theory, resulting in a few debates about the mechanics of fertility decline.

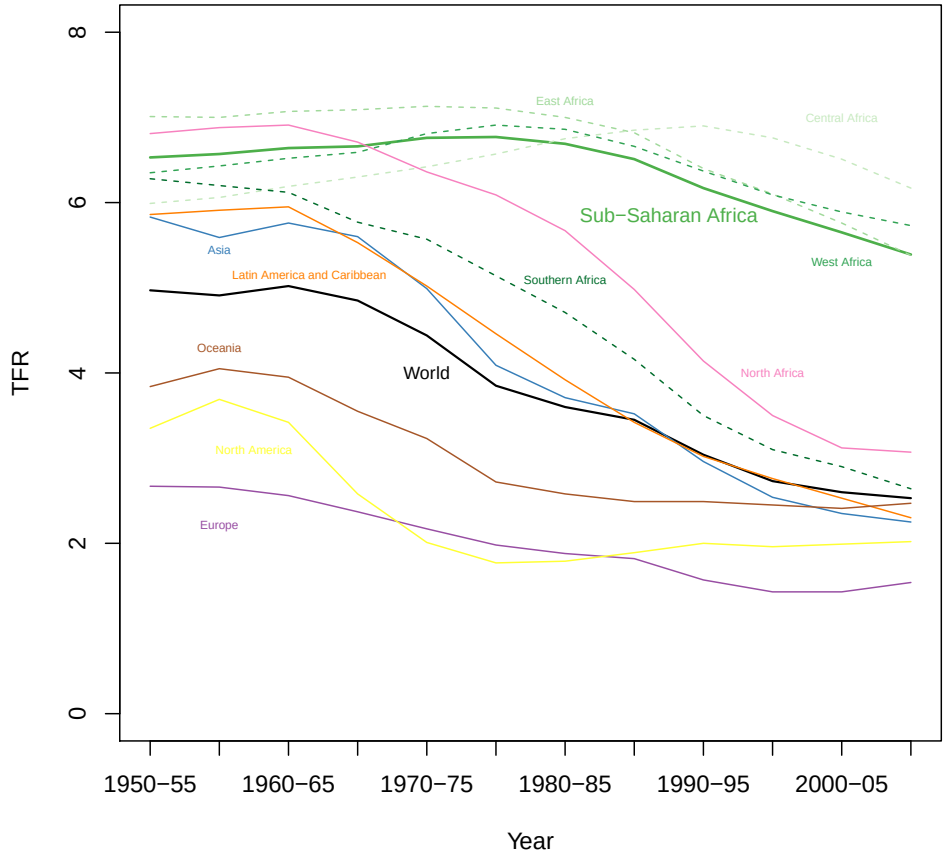

Figure 1: Trends in TFR by world region, with sub-regions for sub-Saharan Africa. Data from United Nations (2013).

The slow pace of fertility decline in SSA, much slower than the earlier declines in Latin America and Asia, has resulted in theorizing about the slow decline and analysis of fertility trends. Looking at fertility trends in real time has led to the identification of countries for whom there has been a failure to decline between two time points, or even an increase in fertility between two time points; it has been suggested that these situations indicate a stall in fertility decline. Substantial effort has been made to identify and characterize SSA fertility stalls (see Machiyama, 2010, for an overview of some of this work), although some researchers believe the putative stalls are due to data quality issues and measurement errors or simply artifacts of the data (e.g. Schoumaker, 2009; Machiyama, 2010; Bongaarts and Casterline, 2013). While there has been a fair amount of work dedicated to analyzing the fertility stalls seen in SSA in the 1990s and 2000s, there is no consensus on the causes (Schoumaker, 2009; Westoff and Cross, 2006).

While most literature that seeks to characterize and understand fertility transition focuses on the TFR, a single number for a place and time, the TFR is built from age-specific fertility rates (ASFR). ASFRs compose the fertility regime that is under analysis when one looks at the TFR. However, very different age patterns of fertility can produce the same TFRs, and the age pattern of fertility, and how it changes, is important for understanding the transition of fertility in a population. Generally, fertility decline created through the control of fertility has been hypothesized to decline at older ages first, followed by a decline at youngest ages. Knodel (1977) found evidence of this pattern in Europe and in most of Asia. However, Caldwell et al. (1992) argue that SSA decline will occur much differently from that seen in Asia and the West, due to different constraints on premarital and extramarital sexuality, differences in marital stability, and different emphases on the need and reasons for birth spacing; the authors hypothesize that in SSA the fertility decline will occur at all ages. van de Walle and Foster (1990) applied the characteristics of Western and Asian fertility decline, including marked decline at oldest ages and/or highest parities, to SSA and found "considerable uncertainty about the causes and permanence of these trends", indicating, at least in part, that SSA was not exhibiting the patterns associated with early decline seen by Knodel (1977) or others. More recent work by Moultrie et al. (2012) has found widening birth intervals at all ages and parities associated with fertility decline in SSA, supporting Caldwell et al. (1992)'s claims of differences in the SSA decline at least in part.

To create a quantitative characterization of fertility schedules for comparison through time and across countries, we consider the ASFRs rather than the TFR and treat the schedules as functional data. This approach acknowledges the different ASFRs that may aggregate to similar TFRs but mean something different about fertility transition or decline. This approach is similar to methods used in various previous work by the authors and their colleagues (e.g. INDEPTH Network [Prepared by Samuel J. Clark], 2002; Clark et al., 2009; Clark, 2014; Sharrow et al., 2014).

2 Data

Data used for this analysis come from the UN's World Population Prospects (WPP) 2012 Revision data (United Nations, 2013). ASFRs are provided for all countries in five year intervals from 1950-55 to 2005-10 for five year age groups (ages 15-19 to ages 45-49). Countries with a population of at least one million were included in the analysis, and three countries with unique TFR trajectories over the time under study were excluded (Yemen, Gabon and Timor-Leste); 154 countries were included in the final analysis. Six age-specific fertility rates are used (age 15-19 years, 20-24 years, 25-29 years, 30-34 years, 35-39 years, and 40-44 years) for the 12 time periods. The last age group, ages 45-49, was omitted from this analysis because the fertility rates for that age group were so close to zero, or were zero, and inclusion greatly influenced the results of the analysis at the expense of details related to fertility at earlier and higher fertility ages. Logged ASFRs are used for the analysis.

3 Methods

This section summarizes the material in Clark (2014).

The singular value decomposition (SVD) (e.g. Strang, 2009) factorizes a matrix $\mathbf{X}$ such that

$$\mathbf{X} = \mathbf{U}\mathbf{S}\mathbf{V}^{\mathrm{T}} \ , \tag{1}$$

where $\mathbf{U}$ contains the orthogonal (independent) 'left singular vectors' $\mathbf{u}_i$ (columns of $\mathbf{U}$), $\mathbf{V}$ contains the orthogonal (independent) 'right singular vectors' $\mathbf{v}_i$ (columns of $\mathbf{V}$), and $\mathbf{S}$ is a diagonal matrix containing the 'singular values'.

The right singular vectors are a new set of orthonormal dimensions for the points defined by the rows of $\mathbf{X}$. The product of the left singular vectors[1] and their corresponding singular values are the projections of the points defined by the rows of $\mathbf{X}$ along the new dimensions defined by the right singular vectors.

The SVD is estimated by minimizing the distance between the actual points (rows of $\mathbf{X}$) and the best approximations of those points using successively more of the new dimensions defined by $\mathbf{V}$. The singular values correspond to the fraction of the overall squared distance from the origin to the points along the new dimensions $\mathbf{V}$ that is captured by each individual new dimension $\mathbf{v}_i$. The first new dimension is oriented to capture as much of this squared distance as possible, and each successive new dimension captures the most possible of what remains.

The product of the SVD factors can be algebraically rearranged to yield another equivalent expression called the *Eckart-Young-Mirsky formula* (Golub et al., 1987)

$$\mathbf{X} = \sum_{i=1}^{\rho} s_i \mathbf{u}_i \mathbf{v}_i^{\mathrm{T}} \ . \tag{2}$$

Equation (2) expresses $\mathbf{X}$ as a sum of rank-1 matrices, where ρ is the rank of $\mathbf{X}$. By construction (above) the first term in this sum captures or explains the bulk of the variation in the original data (rows of $\mathbf{X}$), and each subsequent term explains less and less. The expression for $\mathbf{X}$ in Equation (2) can be further rearranged to express each column vector $\mathbf{x}_\ell$ in $\mathbf{X}$ as

$$\mathbf{x}_\ell = \sum_{i=1}^{\rho} s_i v_{\ell,i} \mathbf{u}_i \ . \tag{3}$$

Equation (4) says that we can write all the columns in $\mathbf{X}$ as weighted sums of the left singular vectors scaled by their corresponding singular values. The weights are the ℓ^{th} elements of each corresponding right singular vector. Moreover, the Eckart-Young-Mirsky matrix

[1] The left singular vectors are also orthonormal.

approximation theorem (Golub et al., 1987) reveals that these sums have the property of concentrating most of the variation in the first few terms, and consequently we only need the first few terms to produce approximate values for $\mathbf{x}_\ell$ that are very close to the actual values. This allows us to closely approximate the columns of $\mathbf{X}$ with (potentially very) few effective parameters – just the first few weights.

Using the SVD, the $6 \times 1,848$ (age $\times$ country, time) matrix of ASFRs is factored into $i \in \{1, \ldots, 6\}$ orthogonal age-varying components $s_i \cdot \mathbf{u}_i$ (the left singular vectors scaled by their corresponding singular values, six elements each, one for each age group) and country-time-varying weights associated with those components, the 1,848 elements of each $\mathbf{v}_i$.

Following Clark (2014), the first component $s_1 \cdot \mathbf{u}_1$ is the underlying 'shape' of the age-specific fertility schedules, and the remaining components define increasingly subtle refinements to that underlying shape. As the Eckart-Young-Mirsky matrix approximation theorem suggests, adding additional components, or terms to the sum in Equation (4), adds increasingly more refined but less consequential nuances to the reconstructed fertility schedule, until when all all components are included, the reconstruction is equal to the original.

In what follows we focus on the first three components in the Eckart-Young-Mirsky approximation of the fertility schedules. These use the first three left and right singular vectors and singular values and closely approximate the majority of the 1,848 empirical fertility schedules included in the matrix originally factored using the SVD. Consequently, the model of age-specific fertility that we manipulate is

$$\mathbf{f}_{c,t} = \sum_{i=1}^{3} v_{i,c,t} \cdot s_i \mathbf{u}_i \ , \tag{4}$$

where $\mathbf{f}_{c,t}$ is the age-specific fertility schedule for country c in time period t, and i indexes the three SVD-derived age-specific components that we retain. The columns of the original data matrix $\mathbf{X}$ are country and time-specific; hence each column is identified by a unique combination of c and t.

We describe how the weights (elements of the right singular vectors) change through time by country and region, and using the `mclust`[2] model-based clustering method, we group the three-element 'weight vectors' associated with each age-specific fertility schedule into clusters of similar weight vectors, and hence similar fertility schedules. Each resulting cluster has its own characteristic age pattern of fertility.

The SVD factorization allows us to work with a high quality, three-parameter approximation of the full age-specific fertility schedules. Beyond having fewer dimensions, the SVD factorization produces effective parameters, the v weights, that are independent and interpretable because they are associated with fixed age-specific components whose age profiles are meaningful.

[2]Model-based clustering is conducted using the `mclust` package in R (Fraley et al., 2012) on the first three weights selected from the SVD results. Bayesian Information Criteria (BIC) is used to select the optimal number of clusters

4 Results

4.1 Results from SVD

Fertility schedules can be reconstructed perfectly using all six components and the country and period associated weights. However, the bulk of information is captured in the first few components and with this reduced set of data reasonable approximations of fertility schedules can be recreated. Figure 2 displays a scree plot of the singular values produced by the SVD of the age-specific fertility schedules. We can see that the first component captures the vast majority of the variation among the fertility schedules. Traditionally, the elbow would likely be drawn at the second component, however, given the importance of variations in age-specific fertility curves for understanding how fertility changes, we chose to look at the first three components as they all seemed to be contributing a fair amount of information and those contributions were both distinct and interpretable.

Figure 3 plots the shapes of the age-specific curves of the first three components, denoted as u_i. The first component captures the overall shape of the curve of fertility with age, rising steeply from the first age to a peak in the second through fourth age groups, and steadily declining again. The weights on u_1 are all positive. Larger weights on u_1 result in lower fertility, pulling the curve down, and smaller weights result in higher fertility, pushing the curve up.

The second component adjusts the earliest and oldest age groups, to accommodate higher or lower early or late fertility, while the third component adjusts the peak fertility age, in the 20s and early 30s, either flattening the curve or intensifying the peak. One can see that positive weights for the u_2 curve would drive down the fertility in the earliest age while simultaneously pulling up fertility at later ages, leaving fertility levels at the second age group unchanged (the opposite would occur with a negative weight on u_2, pulling up earliest fertility while suppressing later fertility).

Similarly, depending on the sign of the weight on u_3, combining this curve with u_1 would accentuate the peak seen at ages 25-29, pushing up the fertility at these ages while suppressing fertility at the extreme ages. Subsequent components made further, more complex and more subtle adjustments to the basic curve given by u_1.

4.1.1 SVD Weights over Time

The SVD gives weights for each country and at each period for the three components, denoted as $v_{i,c,t}$, which are used to capture the fertility curves of each country in each period. Figure 4 shows the weights for the first component for each country over time, with the lines colored by world region. Higher values of $v_{1,c,t}$ are associated with lower overall fertility and the median curve shown in black shows that values of $v_{1,c,t}$ have been steadily increasing globally since the 1960s. Though world regions do overlap, we can see distinctly that SSA countries are

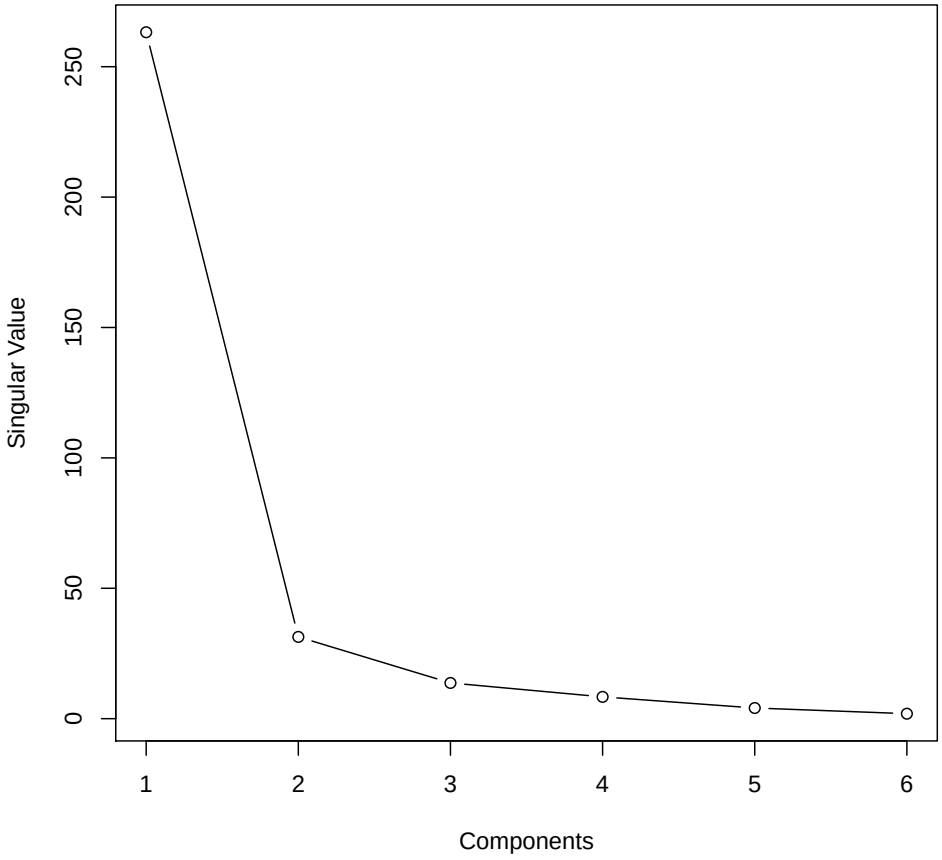

Figure 2: Scree plot of singular values from SVD factorization of age-specific fertility schedules.

predominantly at the bottom of the graph, with the lowest values of $v_{1,c,t}$ through time and with the smallest increase over time. The West, comprised of the US, Canada, Australia, New Zealand and Europe, had some of the highest values for this weight in the earliest years and have maintained, after a rise in the 1960s and 1970s, high values for this weight. For East Asia we can see a dramatic increase from low levels to the highest values of $v_{1,c,t}$.

Weights on the second component are largely centered around zero and the variance is much smaller than seen for $v_{1,c,t}$ (shown in Figure 5. Looking at the way $v_{2,c,t}$ changes over time for each country, colored by world region, we do not see the regional patterns that is present in the relationship between $v_{1,c,t}$ and time. However, SSA countries seem highly concentrated close to 0, in contrast to other regions that experience more variation in values of $v_{2,c,t}$ over time. *This clearly reveals the lack of age-specific fertility change in SSA compared to other world regions.* Tracking individual country lines through time, there appear to be substantial

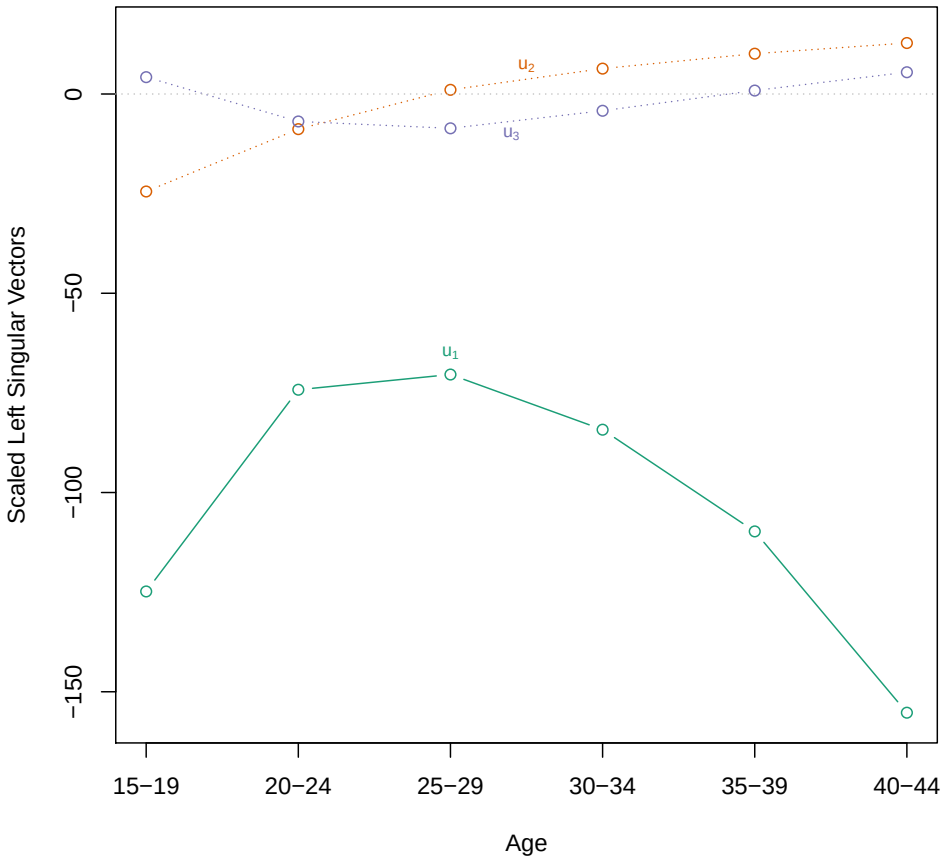

Figure 3: First three components (scaled left singular vectors) from SVD factorization of age-specific fertility schedules.

changes in values of $v_{2,c,t}$ period to period as well as a wide variety of patterns for overall time trends.

Weights on the third component over time for each country are shown in Figure 6. Like the second component, these weights are largely concentrated around 0 and regional patterns for changes in $v_{3,c,t}$ over time do not seem particularly pronounced, though SSA countries seem concentrated close to 0 through time, though almost exclusively positive, while East Asia and the West seem to have much higher volatility in values and predominantly negative values for $v_{3,c,t}$. As the third component is responsible for how peaked the fertility curves are, we see that SSA values are associated with flatter curves, or curves with peaks that span over the 20s and early 30s, while the West and especially East Asia have curves that are quite peaked with highest fertility occurring in only one age group.

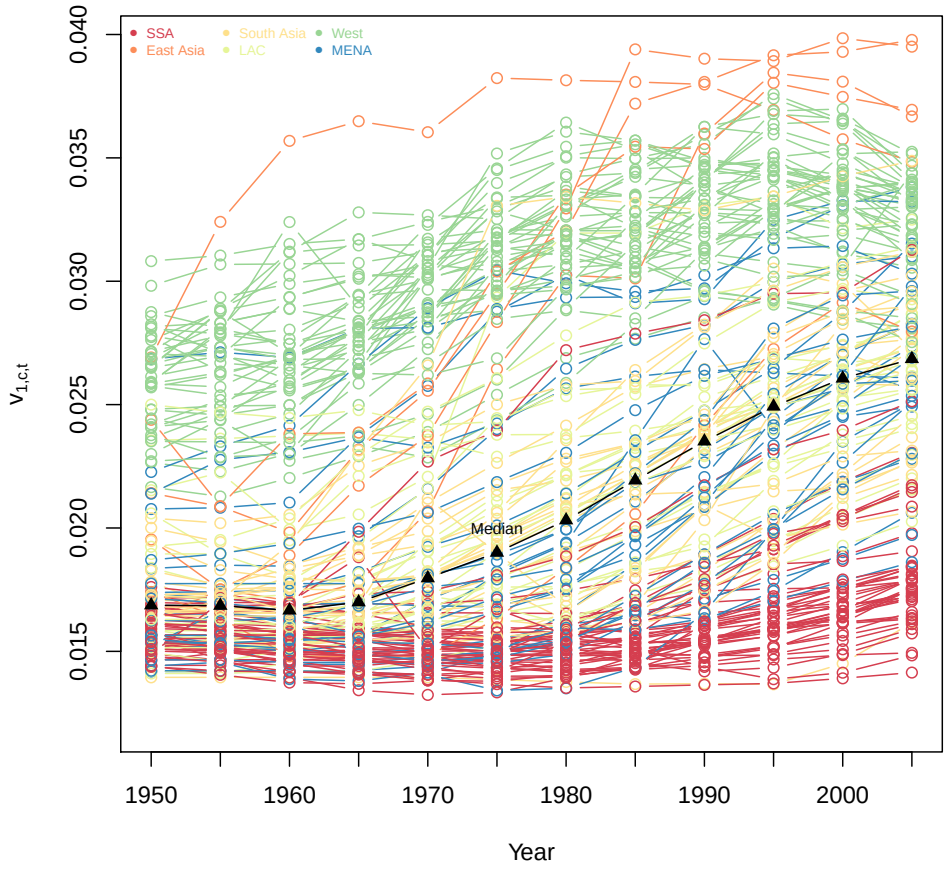

Figure 4: Weight on first component by country and time, colored by world region.

4.1.2 SVD Weights and Total Fertility Rates

The first component, which corresponds to the general shape of fertility curves and the overall fertility level, is closely correlated with the TFR. Plotting the weights for the first component against TFR for each country and period (see Figure 7) there is a pronounced curvilinear relationship, with lower values of $v_{1,c,t}$ associated with high levels of TFR and curving through medium values to the highest values of $v_{1,c,t}$, which are exclusively associated with below replacement TFRs. There is distinct regional patterning here, with each region forming a fairly pronounced clump of points and taking on bounded ranges of both TFR and $v_{1,c,t}$, which is in-line with regional patterns of fertility levels.

The second component adjusts early and late fertility, raising one while decreasing the other. Plotting the weights for the second component against TFR, the distinct pattern seen for

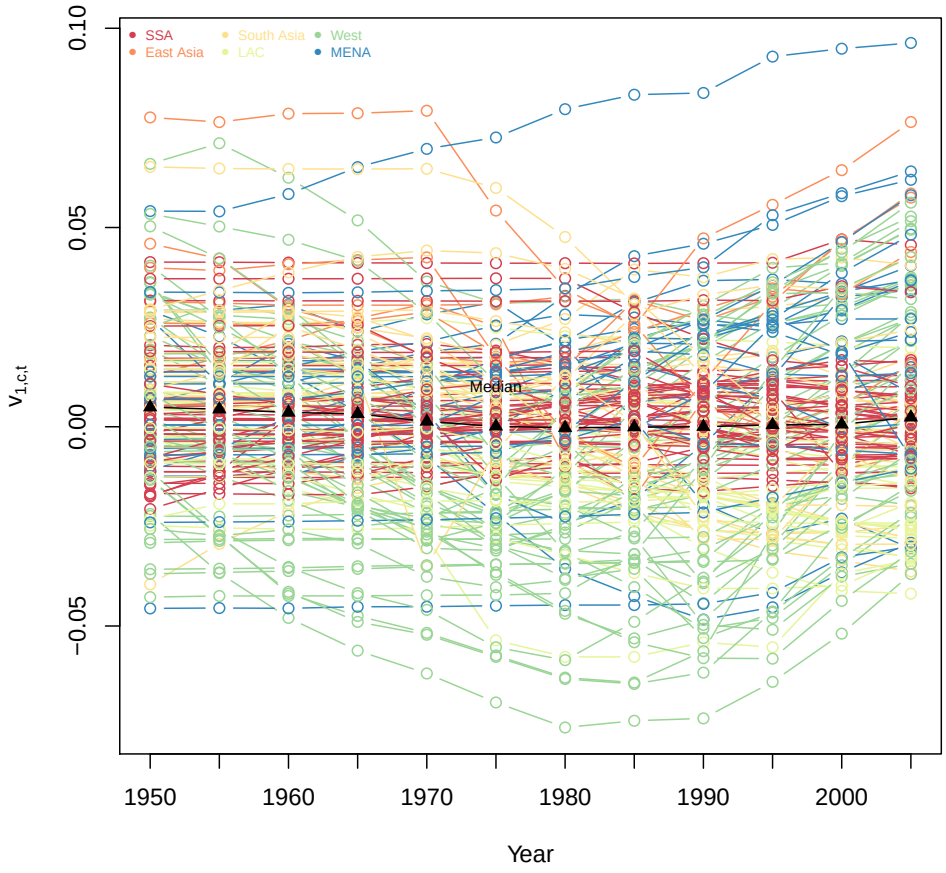

Figure 5: Weight on second component by country and time, colored by world region.

$v_{1,c,t}$ is not nearly so pronounced. Instead there is a large cloud, with many values close to 0. There is some regional clumping, notably for SSA and the West. However, the West, and lower TFR levels generally, have much higher variance in the values of $v_{2,c,t}$, while there is much less variation in $v_{2,c,t}$ at lower levels of TFR. This is consistent with general theories of fertility decline, which would suggest a standard curve not unlike our base first component (u_1), with a wide fertility peak across the 20s and 30s, followed by a slow decline in fertility levels at older ages, while at lower levels of fertility (i.e. mid and post-decline) the fertility curves are more varied as fertility is changing at different rates at different ages. However, excluding SSA, most other regions do still see a fair amount of variability in values of $v_{2,c,t}$ at all levels of TFR.

Overall, the weights for the third component do not appear to form a distinctive pattern with TFR. These weights, which determine whether the fertility curve is very peaked or

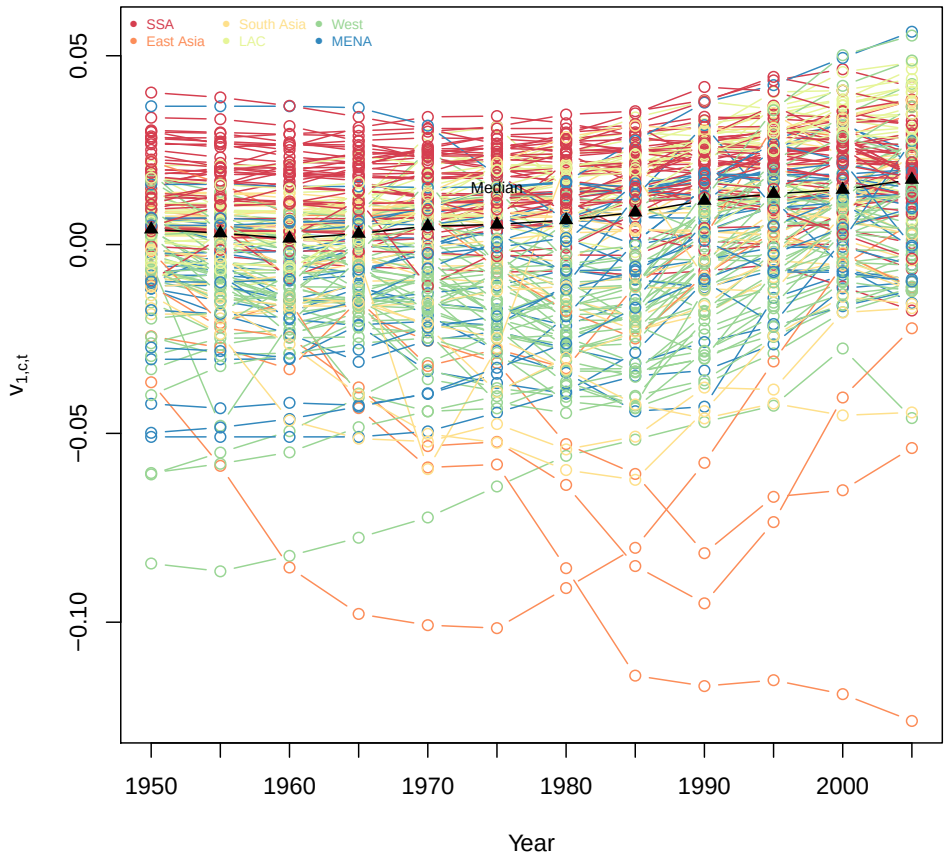

Figure 6: Weight on third component by country and time, colored by world region.

relatively flat, do display regional relationships with TFR. SSA, in red, is highly clumped in this plot, with a suggested increase in values of $v_{3,c,t}$ at lower levels of TFR. The West also forms a distinct cloud. East Asia forms a band at the lowest values of TFR and through the lowest values of $v_{3,c,t}$. Latin America and the Caribbean form a band in between SSA and the West, showing a tendency towards higher values of $v_{3,c,t}$ at the lower levels of TFR. It is more difficult to find a distinct cloud for the Middle East and North Africa countries or the South Asia countries. The distinct clouds here and for $v_{1,c,t}$ and TFR, as well as the more subtle clouds seen in for the relationship between $v_{2,c,t}$ and TFR suggest that there are likely to be distinctive characteristics associated between the fertility curves and TFR, perhaps within regions.

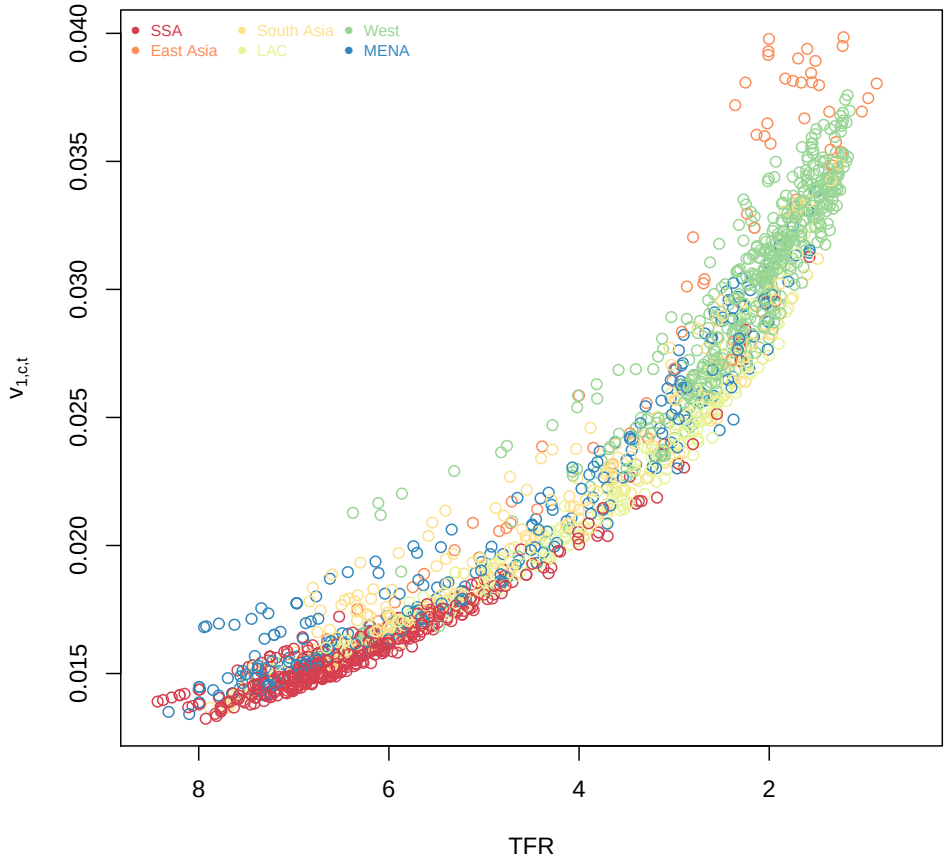

Figure 7: First component weight by TFR, all countries and periods.

4.2 Clustering Results

Seven clusters are obtained from the analysis: three high, early transition clusters; two middle transition, medium fertility clusters; one low fertility cluster near replacement and one low fertility, late or post-transition cluster. Median fertility regime curves are shown for each cluster in Figure 10.

Cluster 1, with 131 schedules, has the lowest levels of fertility, capturing the lowest-low, or below replacement, schedules. The median TFR was 1.7. Generally, the ASFRs are low at all age groups, with a distinct peak in ages 25-29, though some schedules categorized here had high very high peaks at ages 25-29. Only more recent schedules were placed in this cluster, from Europe, Canada, and East Asia.

Cluster 2 had a mean TFR of 2.2 and included 434 schedules. Most TFRs for schedules in this cluster were between 1 and 3. This cluster includes schedules from Latin America, Asia,

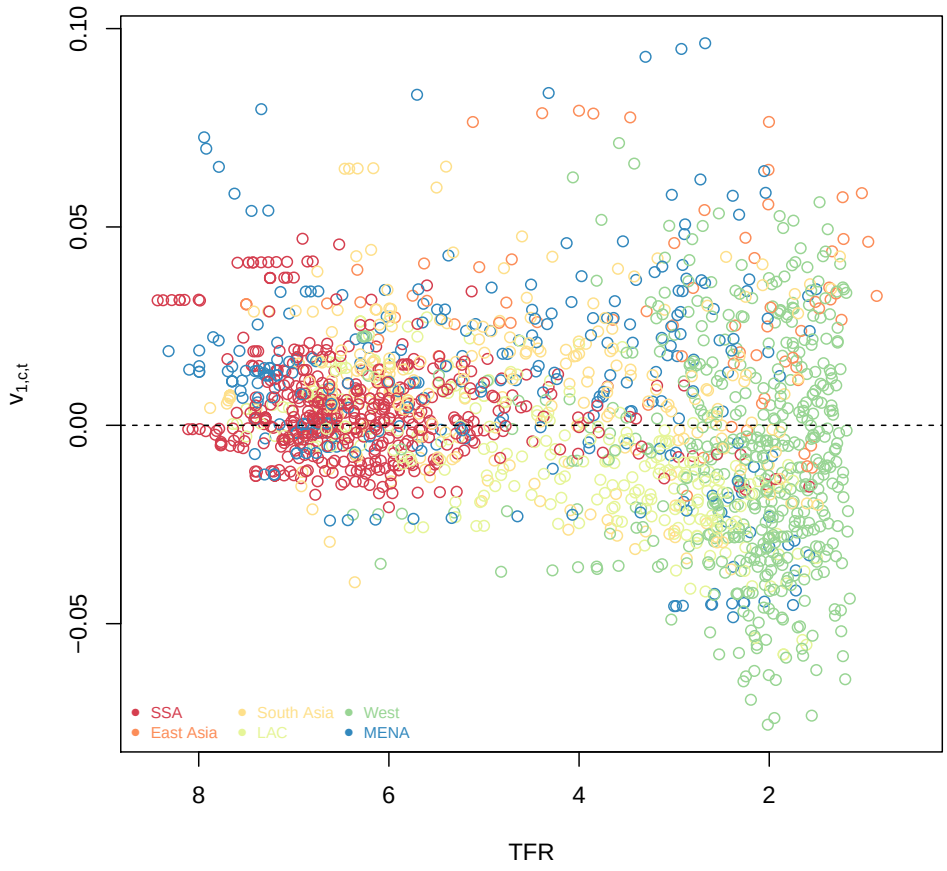

Figure 8: Second component weight by TFR, all countries and periods.

former Soviet Republics, Europe, the US, Canada, and Australia. Cluster 2 has a higher peak fertility that occurs at younger ages than Cluster 1.

Cluster 3, with 233 members, captures a pattern of medium fertility. The median TFR in this group is 3.3, and the TFRs are concentrated between 2 and 4, with some schedules with higher fertility. This cluster includes earlier schedules from Europe, North African and Middle Eastern schedules, and early Asian schedules through the 1980s and 1990s. Curves in this cluster were more varied than other clusters, though generally with a high and distinct peak at ages 20-24 or ages 25-29.

Cluster 4, comprising 211 schedules has a broadest range of TFRs, though most of the TFRs from schedules in this cluster fall between 2 and 4, with a median TFR of 3.6. A wide variety of fertility curve shapes are also possible in this cluster. The unifying characteristic of this group is medium fertility levels and higher fertility at ages 15-19 than curves in Cluster 3, as well as a tendency for fertility to peak at ages 20-24 instead of ages 25-29. This cluster is

Figure 9: Third component weight by TFR, all countries and periods.

largely comprised of schedules from Latin America and the Caribbean and from Asia.

Cluster 5 is characterized by markedly higher fertility, with a median TFR of 5.6. 285 schedules were placed in this cluster, largely earlier Latin America and the Caribbean schedules, schedules from the Middle East and North Africa, and more recent schedules from SSA. Distinctive for these curves is very low fertility at ages 15-19 compared to high fertility through ages 20-24 and 30-34, peaking at ages 25-29. The fertility is declining pretty steadily by the late 30s in schedules included in this group.

Cluster 6 is characterized by high fertility, with a median TFR of 6.4 for the group. This cluster includes 257 schedules, and is dominated by high fertility SSA schedules, with some schedules from Eurasia and the Middle East. In this group, fertility is high at all ages, peaking around ages 20-24 but declining slowly with age.

Cluster 7 includes 297 schedules and has the highest median TFR, 6.9. TFRs for this cluster

are even more highly concentrated at the higher levels, with the bulk of the distribution having TFRs above 6. Curves in this cluster are distinct from those in Cluster 6 by having lower fertility at the earliest ages but higher fertility at the peak reproductive ages (20-24 and 25-29). The decline in fertility is also more rapid at older ages in Cluster 7 than in Cluster 6. This cluster is also dominated by SSA but includes more Latin American and Caribbean, Asian and Middle East schedules.

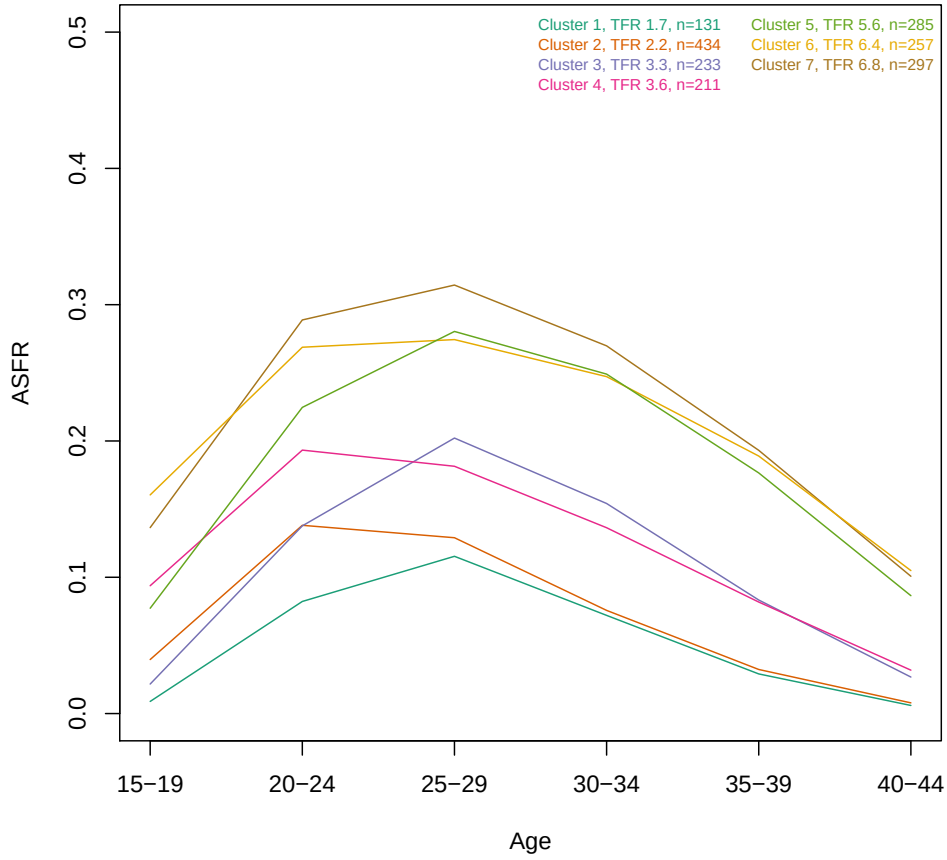

Figure 10: Median age-specific fertility schedules by cluster

A country's demographic transition can be traced through time by membership in the different clusters. Figure 11 shows the transitions of SSA schedules over time through the clustering schema. SSA schedules are predominant in the high fertility, early transition clusters, though almost all schedules have transitioned out of the pre-transition, highest fertility cluster. For example, South Africa transitioned from Cluster 5, with TFR ranging from 5.8-6.6 to Cluster 4. South Africa saw a steady decline in TFR over the period under study though only minimal changes in the age structure of its fertility. Kenya was in Cluster 7, the

highest fertility levels with pre-transition age-specific fertility pattens until the 1980s when fertility patterns change to include Kenya in Cluster 5. In the most recent data, Kenya's fertility had declined to the 4.2-5.0 range and the regime had shifted into Cluster 4, indicating a shift in age-specific fertility more in line with mid-transition patterns has accompanied the fertility decline.

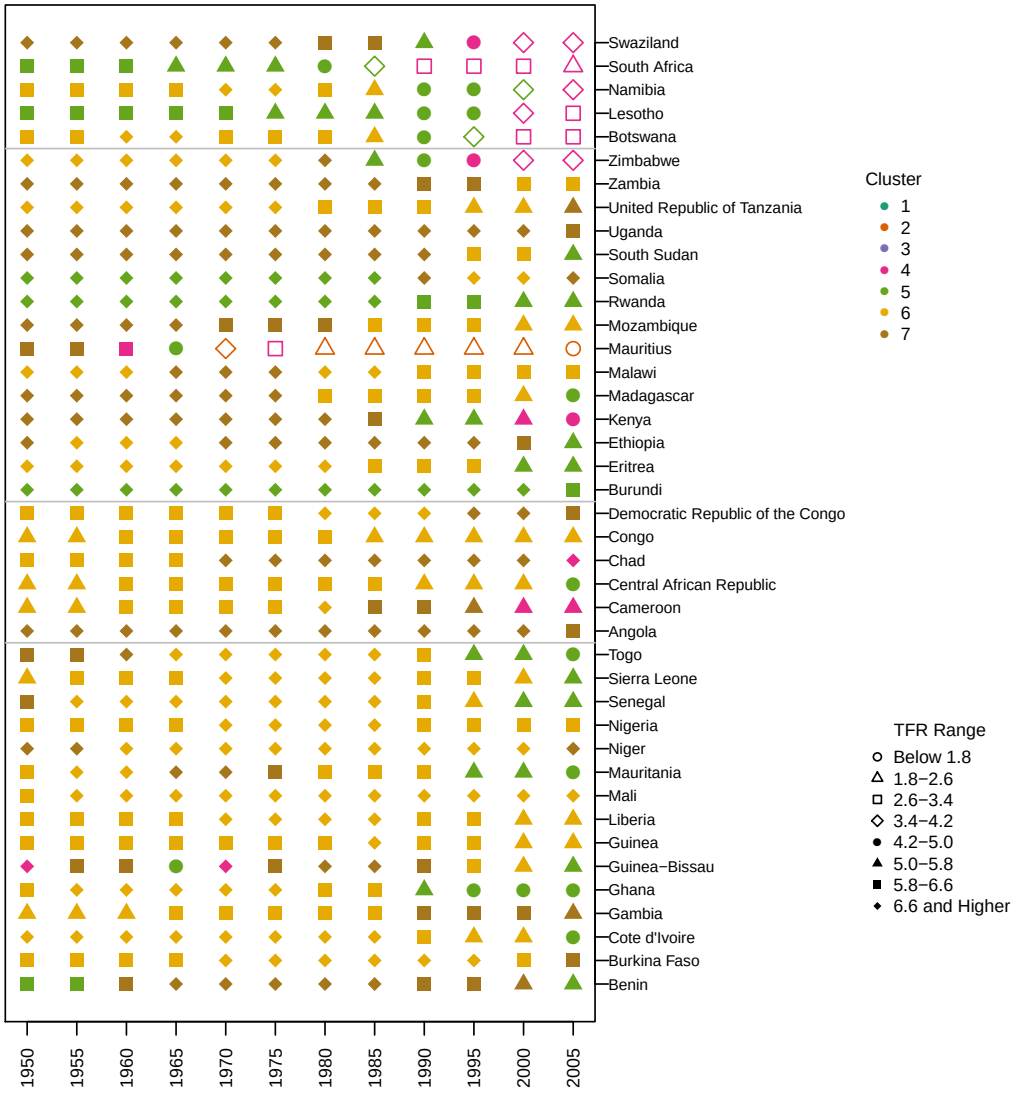

Figure 11: Cluster transitions and associated TFR for sub-Saharan African countries by time.

Figure 12 shows the TFR trends for Kenya and South Africa from 1950-55 through 2005-10 with the associated cluster membership. For Kenya, there was some decline in fertility level while retaining an age-specific fertility schedule compatible with Cluster 7, the highest fertility cluster. After having dropping to a TFR below 6, Kenya's age profile shifted to

that of Cluster 5, suggesting a more pronounced fertility peak in the late 20s and relative reductions in fertility at the youngest ages while older fertility remained fairly high. While TFR seemed relatively stable around 5, the age profile of fertility shifted again to Cluster 4, indicating a fertility peak at younger ages and a reduction of fertility in late reproductive age. Meanwhile, South Africa's age profile was categorized as Cluster 5 as early as 1950-55 and remained in that cluster through a substantial decline in TFR. Only at TFR levels below 4 did South Africa transition into Cluster 4, where it has remained as it's neared replacement fertility. For comparison, Sweden and South Korea have been included. South Korea underwent fertility decline to below replacement levels during this period. Even though initial TFRs for South Korea were high, and increased for a period, South Korea's fertility regimes were consistently categorized as Cluster 3 (excepting the peak TFR of 1955-60, whose associated fertility regime as categorized as Cluster 5). These curves have low early and later fertility, with a pronounced peak around ages 25-29. Nearing replacement South Korea transitioned to Cluster 1, the lowest-low fertility regime group, shaped similarly to Cluster 3, but with much lower fertility at all age groups. Sweden had attained replacement fertility by the start of the data used here, but transitioned from Cluster 2, with fertility peaking at younger ages, to Cluster 1, even as the TFR rose (and then declined and rose again).

5 Conclusion

Incorporating age patterns of fertility explicitly in modeling provides more depth to the analysis of global fertility trends by making explicit how changes in fertility may occur within stable age patterns of fertility and also how changing age patterns of fertility may result in only small TFR changes in the short-term. This analysis makes explicit the idea that many different patterns of fertility can result in similar levels of fertility, and the variety present in the mid-transition clusters indicates the many different pathways a high fertility regime may take to arrive at lower fertility (Hirschman, 1994). By reducing the fertility curves using the SVD, we are able to look at these patterns with fewer parameters, without losing detail, which simplifies analysis. By categorizing the fertility regimes into clusters we can more clearly see that changes in the age structure of fertility are occurring differently across fertility declines. Preliminary work suggests regional patterning, which is further supported by analyzing the SVD weights directly. Considering SSA fertility trends in their global context this way, we can anticipate that given the relative stability until most recent decades in age patterns of fertility a substantial amount of change is still to come, but that does not preclude rapid decline in fertility, as seen in some parts of Asia and Latin America and the Caribbean, once age-specific fertility patterns change. Investigating not only how the TFR is changing over time globally, but also by understanding how the age patterns of fertility change during the decline, can be used to potentially predict subsequent age patterns of fertility, based on overall fertility level, and to explore relationships between factors that influence fertility, such as economic development, women's education, infant and

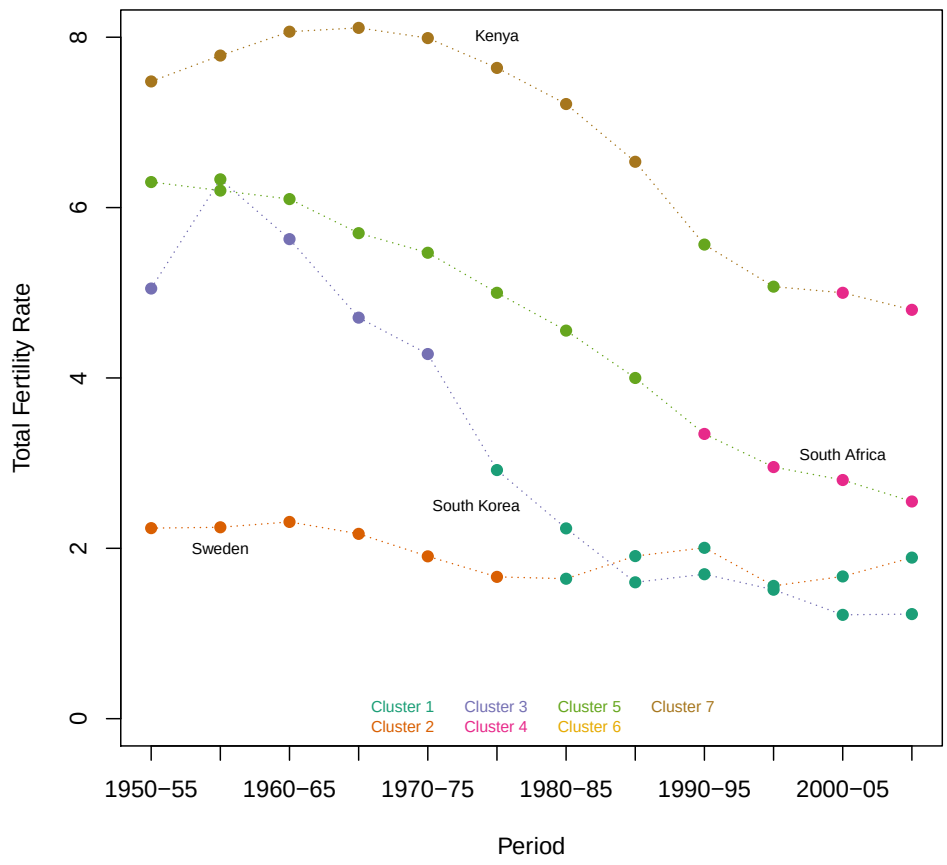

Figure 12: Trends in Total Fertility Rates with Cluster Assignments for Select Countries

child mortality, HIV/AIDS, etc.

References

Bongaarts, J. and J. Casterline (2013). Fertility transition: Is sub-saharan africa different? *Population and development review 38*, 153–168.

Caldwell, J. C., I. O. Orubuloye, and P. Caldwell (1992). Fertility decline in africa: A new type of transition? *Population and development review 18*(2), 211–242.

Clark, S. J. (2014). A singular value decomposition-based factorization and parsimonious component model of demographic quantites correlated by age. Working Paper 143, University of Washington Center for Statistics and Social Sciences.

Clark, S. J., M. Jasseh, S. Punpuing, E. Zulu, A. Bawah, and O. Sankoh (2009, May). Indepth model life tables 2.0. In *Annual Conference of the Population Association of America (PAA)*. Population Association of America (PAA).

Dorius, S. F. (2008). Global demographic convergence? a reconsideration of changing intercountry inequality in fertility. *Population and development review 34*(3), 519–537.

Fraley, C., A. E. Raftery, T. B. Murphy, and L. Scrucca (2012). mclust version 4 for r: Normal mixture modeling for model-based clustering, classification, and density estimation. Technical Report 597, Department of Statistics, University of Washington.

Golub, G. H., A. Hoffman, and G. W. Stewart (1987). A generalization of the eckart-young-mirsky matrix approximation theorem. *Linear Algebra and Its Applications 88*, 317–327.

Hirschman, C. (1994). Why fertility changes. *Annual Review of Sociology*, 203–233.

INDEPTH Network [Prepared by Samuel J. Clark] (2002). *INDEPTH Mortality Patterns for Africa*, Volume 1 of *Population and Health in Developing Countries*, Chapter 7, pp. 83–128. Ottawa: IDRC Press.

Knodel, J. (1977). Family limitation and the fertility transition: Evidence from the age patterns of fertility in europe and asia. *Population Studies 31*(2), 219–249.

Machiyama, K. (2010). A re-examination of recent fertility declines in sub-saharan africa. DHS Working Paper 68, MEASURE DHS.

Moultrie, T. A., T. S. Sayi, and I. M. Timæus (2012). Birth intervals, postponement, and fertility decline in africa: A new type of transition? *Population studies 66*(3), 241–258.

Schoumaker, B. (2009). *Stalls in Fertility Transitions in Sub-Saharan Africa: Real or Spurious?* Département des sciences de la population et du développement, Université catholique de Louvain.

Shapiro, D. (2012). Women's education and fertility transition in sub-saharan africa. *Vienna Yearbook of Population Research 10*, 9–30.

Sharrow, D. J., S. J. Clark, and A. E. Raftery (2014). Modeling age-specific mortality for countries with generalized hiv epidemics. *PloS one 9*(5), e96447.

Strang, G. (2009). *Introduction to Linear Algebra 4e*. Wellesley-Cambridge Press.

United Nations (2013). World population prospects: the 2012 revision. http://esa.un.org/wpp/index.htm.

van de Walle, E. and A. Foster (1990). Fertility decline in africa: Assesment and prospects. Technical Paper 125, World Bank.

Westoff, C. F. and A. R. Cross (2006). The stall in the fertility transition in kenya. DHS Analytical Studies 9, MEASURE DHS.