

Family Complexity and Diversity across the World

Andrés Castro (University of Pennsylvania), Joan García-Roman (Centre d'Estudis Demogràfics), Albert Esteve-Palos (Centre d'Estudis Demogràfics)

Abstract

Across the world, families adopt multiple forms due to context-dependent interactions between social norms and individual behavior regarding union formation (dissolution), reproduction and co-residence. However, we lack global cross-national analyses about the family based on comparable indicators. Much of the classical theorizing on global family change occurred without the benefit of empirical data. It is only in the last few years that researchers have gained access to large-scale representative datasets. In this paper, we leverage IPUMS-I, DHS and EU-LFS microdata to analyze for the first-time variations in family forms in 110 countries over the last four decades. We summarize trends in 12 key family indicators concerning union formation, household size and complexity and the position of women within households. Results show quite distinct patterns both within and across continental regions and a mixture of common trends and persistent singularities.

Introduction

Across the world, families adopt multiple forms. Family diversity arises from a wide spectrum of context-dependent interactions between social norms and behaviors regarding union formation (and dissolution), reproduction and co-residence. As a result, union formation and dissolution are not evenly spread across the world, men and women marry and have children at different ages, the size and composition of households differs across regions, and the role of women in these households is quite distinct (Bianchi, 2014; Goode, 1963; Lesthaeghe, 2014; Therborn, 2004; Thornton, 2013; Todd, 1985). However, we lack global cross-national analyses about the family based on comparable indicators. Much of the classical theorizing on the family at a global scale occurred without the benefit of empirical data, and it is only in the last few years that researchers have gained access to large-scale representative databases which allow the observation of social changes on a worldwide scale (Carmichael & Rijpma, 2017; Rijpma & Carmichael, 2016; Ruggles, 2014).

This is even more the case when it comes to studies on the gendered nature of family dynamics. No global comparison based on demographic indicators have been conducted in this area. Recent demographic accounts of the role of gender in family change—mostly fertility—have focused on a handful set of countries, typically of high income (Anderson & Kohler, 2015; McDonald, 2000). This is despite of the overwhelming qualitative and historical evidence on key importance of gender on family dynamics (Alesina, Giuliano, & Nunn, 2013; England, 2010). Our study is then

also innovative by appropriating a quantitative approach, to a hitherto distinctively qualitative goal.

Hence, in this paper, we leverage newly available data to analyze for the first time, variation in family forms on several geographic scales over a five-decade period. Our ultimate aim is to make informed inferences about present and future changes in the family, for which a detailed and in-depth description of time trends of family indicators constitutes the first necessary step.

More specifically, we summarize trends in 12 key family life indicators in 110 countries, spanning from 1970 to 2015. These indicators measure dimensions such as union formation and dissolution, fertility, household size and complexity, and the position of women within the couple. We rely on the analysis of census and survey microdata samples, gathered and harmonized from three major sources, the Integrated Public Use of Microdata Series (IPUMS-I) (Minnesota Population Center, 2015), Demographic Health Surveys (DHS), and Eurostat's Labor Force Survey data (EU-LFS). We examine recent patterns and trends in the 12 indicators and run a cluster analysis to identify regions that share similar characteristics.

Our approach is inductive and comparative in nature. At this stage, we do not aim to prove any global theory but to provide the foundations for theory confirmation and development. By comparing many countries and their changes over time we can better grasp the large heterogeneity in family forms across the world. Any global theory on family change should be able to explain sustained heterogeneity as evidence supporting convergence on family forms seems weak, if not absent (Kohler & Anderson, 2016; McLanahan, 2000; Pesando & GFC-team, 2018)

Data and methods: country coverage and temporal frame

We use data from 110 countries spanning from 1970 to 2015. These data come from nationally representative household microdata samples from censuses and surveys. We use all available datasets in IPUMS international. IPUMS-I offers access to 365 census microdata samples from 94 countries. In the vast majority of cases, IPUMS samples provide both household and individual level microdata. Sample densities in census microdata are usually in the range of 1 to 10% of the total population and, thus, they usually provide the most robust and coherent estimates. For this reason, we give priority to the census over the other sources when two or more datasets are available for the same country and year. However, IPUMS-I is clearly under-represented in some areas and outdated in others. To fill these gaps, we use all available DHS and EU-LFS microdata. The former allows us to include many African and low-income countries not present in IPUMS and the latter allow us to obtain more recent estimates for Europe.

Selection of indicators

The number of family-life indicators that can be constructed is obviously larger than the ones that can be effectively obtained from cross-sectional censuses and surveys. However, for the sake of completeness in geographic coverage, we opted for few, simple and comparable indicators across countries. We began with a battery of more than 30 indicators, which, after a careful selection based on quality concerns and correlation analysis, was reduced to 12. These indicators reflect the main family dimensions on which family demographers and sociologists are interested: union formation and dissolution, reproduction, household structures and complexity and the role of women. We are aware some of these dimension and indicators overlap, however, we sorted them and classified them by dimension to ease interpretation of trend and differences. Table 1 displays the list of indicators along with the short label we use for them in the graphs.

Regarding union formation, we measure the proportion of women (Ind1) and men (Ind2) in union at the ages of 20-24. We indirectly measure gender differences in union formation through the age difference between spouses (Ind3). Unfortunately, the distinction between marriage and cohabitation was not possible in all countries and, therefore, we did not include it. Regarding household size and composition, we estimated the percentage of households with 5 or more individuals (Ind4), the percentage of unipersonal households (Ind5), and the average number of children 15 or younger per household (Ind6). For internal household complexity, we included the average age range between the youngest and eldest household member (Ind7), the percentage of households including with other relatives (beyond parents, children and spouses) and non-relatives (Ind8), and the percentage of 0 – 14 children residing with both parents (Ind9). The fourth and last dimension refers to the position of women within households. In this dimension we include the percentage of women 35-39 that are: separated or divorced (Ind10), not in union but with children (Ind11), and household heads (Ind12).

Cluster analysis

We conduct a cluster analysis to identify groups of countries with similar family traits across the 12 indicators. For this analysis we focus on the most recent sample by country collected after 2000. This gives us an analytical sample of 99 countries. We standardized and identified the main factorial dimensions across the 12 indicators using factorial analysis techniques. Further we use factorial dimensions to compute a pair-wise (99 x 99) dissimilarity matrix that measures the cross-country dissimilarities in family characteristics. The generic term of this matrix, d_{ij} measures the dissimilarity in family characteristics between countries i and j . The greater this number the more dissimilar are families between these two countries.

Further, we use this matrix as the input for a hierarchical clustering following the Ward method (Kaufman & Rousseeuw, 1990; Pardo & Del Campo, 2007). This strategy is suitable to our aim for two reasons. First, it minimizes the within-cluster variance, thus providing a partition that

groups countries with similar family features and separates countries with divergent family traits. Second, it does not require to anticipate the number of clusters, instead, the clustering structure can be analyzed to better understand discrepancies across countries and how they group together. This second condition is aligned with our inductive approach and favors theory development rather than theory testing.

Results

Trends in family life in 110 countries

Figure 1 shows patterns in family life indicators based on the most recent available data for each country since year 2000. Boxplots summarize patterns of countries. Color indicates continent. Differences across countries exist for all indicators. Africa is the most diverse continent, followed, in this order, by Asia, Europe and North America and Latin America.

African women enter their unions at the youngest ages and show the largest age gap between spouses. European women are at the opposite end. African and Asian households are larger than in another region of the world. Europe and North America have the smallest households and largest percentage of unipersonal households. Unipersonal households are rare in Asia, Africa and Latin America. Concerning household complexity, Africa, Asia and Latin America have similar levels of intergenerational co-residence as shown by the age range between the eldest and youngest household member, which in Europe is ten years lower than in the other continents. The presence of other relatives or non-relatives in European is household is systematically lower than in Africa, Latin America and Asia. As for the family context of children, the percentage of children living with both parents in Africa and Latin America is lower 15 percent points lower (on average) than in Europe and Asia. Finally, Europe and Latin America have the highest rates of separated and divorced women. In Latin America, we find the highest proportions of women not in union raising children. Finally, Africa, Latin America and Europe show similar proportions of women heading households, but less in Asia.

Regarding trends, Figure 2 summarizes trends over time in the 12 indicators. There is one panel per indicator. All panels have the same structure: time in the horizontal axis and the variable of interest in the vertical one. The thin lines depict country trends. Thick lines indicate average continental trends. Results show that women and men are postponing union formation except in Latin America. Age differences between spouses are shirking. Household size is declining worldwide, especially in Latin America, and the percentage of unipersonal is growing everywhere except in Asia. A common trend is the declining presence of children in households, which is not necessarily reducing complexity within households. The age range within households and the presence of other relatives and non-relatives is quite stable. But the percentage of children co-residing with both parents is declining everywhere except in Asia. Finally, regarding the position

of women, Figure 2 shows that divorce and separation are increasing in Europe and in Latin America. A similar trend applies to women not in union with children at home. Finally, more women 35-39 are heading their households. Again, Asia is an exception.

Clusters solution

Five clusters were identified based on classification dendrogram in Figure 3. This cluster solution is adequate as it explains a large proportion of the total variance in the distance matrix, 73%. Moreover, this solution yields an Average Silhouette Width (ASW) of 0.41 which indicates strong cluster consistency (Studer, 2013). All other measures for the goodness of fit of the partition were aligned, peaking with a five clusters solution.

The most distinct group appears at the bottom of the dendrogram and it comprises mostly European and North American countries. Continuing from bottom to top, the second cluster comprises mostly Latin American countries, being Kyrgyzstan, Ethiopia and Thailand the only three exceptions. The following cluster includes mostly South African countries, along with Bolivia and some Caribbean countries. North African, Middle eastern and Asian countries constitute the majority in group fourth. Finally, the group one is composed mostly by Central and Western African countries, with some Asian countries as exceptions.

Further steps

1. Further explore the geographical distribution of clusters paying attention to the exceptions identified above
2. Explore cross-clusters differences across the 12 indicators as displayed in Figure 4
3. Replicate the analysis with older samples and assess clusters consistency over time

References

- Alesina, A., Giuliano, P., & Nunn, N. (2013). On the Origins of Gender Roles: Women and the Plough. *The Quarterly Journal of Economics*, 128(May), 469–530. <https://doi.org/10.1093/qje/qjt005>. Advance
- Anderson, T., & Kohler, H. P. (2015). Low Fertility, Socioeconomic Development, and Gender Equity. *Population and Development Review*, 41(3), 381–407.
- Bianchi, S. M. (2014). A Demographic Perspective on Family Change. *Journal of Family Theory & Review*, 6(1), 35–44. <https://doi.org/10.1111/jftr.12029>
- Carmichael, S., & Rijpma, A. (2017). Blood is thicker than water: Geography and the dispersal of family characteristics across the globe. *Cross-Cultural Research*, 51(2), 142–171. <https://doi.org/10.1177/1069397117691025>
- England, P. (2010). The Gender Revolution: Uneven and Stalled. *Gender & Society*, 24(2), 149–166. <https://doi.org/10.1177/0891243210361475>
- Goode, W. J. (1963). *World Revolution and Family Patterns*. Free Press of Glencoe.
- Kaufman, L., & Rousseeuw, P. J. (1990). *Finding Groups in Data: An Introduction to Cluster Analysis*. New York: Wiley.
- Kohler, H., & Anderson, T. (2016). Divergent Demographic Destinies. Population Association of America.
- Lesthaeghe, R. (2014). The second demographic transition: A concise overview of its development. *Proceedings of the National Academy of Sciences of the United States of America*, 111(51), 18112–18115. <https://doi.org/10.1073/pnas.1420441111>
- McDonald, P. (2000). Gender Equity in Theories of Fertility Transition. *Population and Development Review*, 26(3), 427–439.
- McLanahan, S. (2000). Family, state and child well-being. *Annual Review of Sociology*, 26, 703–706.
- Minnesota Population Center. (2015). Integrated Public Use Microdata Series, International: Version 6.4.
- Pardo, C. E., & Del Campo, P. C. (2007). Combinación de métodos factoriales y de análisis de conglomerados en R: El paquete FactoClass. *Revista Colombiana de Estadística*, 30(2), 231–245.
- Pesando, L. M., & GFC-team. (2018). Global Family Change: Persistent Diversity with Development. *Population and Development Review*.
- Rijpma, A., & Carmichael, S. G. (2016). Testing Todd and Matching Murdock: Global Data on Historical Family Characteristics. *Economic History of Developing Regions*, 31(1), 10–46. <https://doi.org/10.1080/20780389.2015.1114415>
- Ruggles, S. (2014). Big Microdata for Population Research. *Demography*, 51(1), 287–297. <https://doi.org/10.1007/s13524-013-0240-2>
- Studer, M. (2013). WeightedCluster Library Manual: A practical guide to creating typologies of trajectories in the social sciences with R. *LIVES Working Papers*, 24. <https://doi.org/http://dx.doi.org/10.12682/lives.2296-1658>
- Therborn, G. (2004). *Between Sex and Power: Family in the World 1900–2000*. Routledge.
- Thornton, A. (2013). *Reading History Sideways. The fallacy and enduring impact of the*

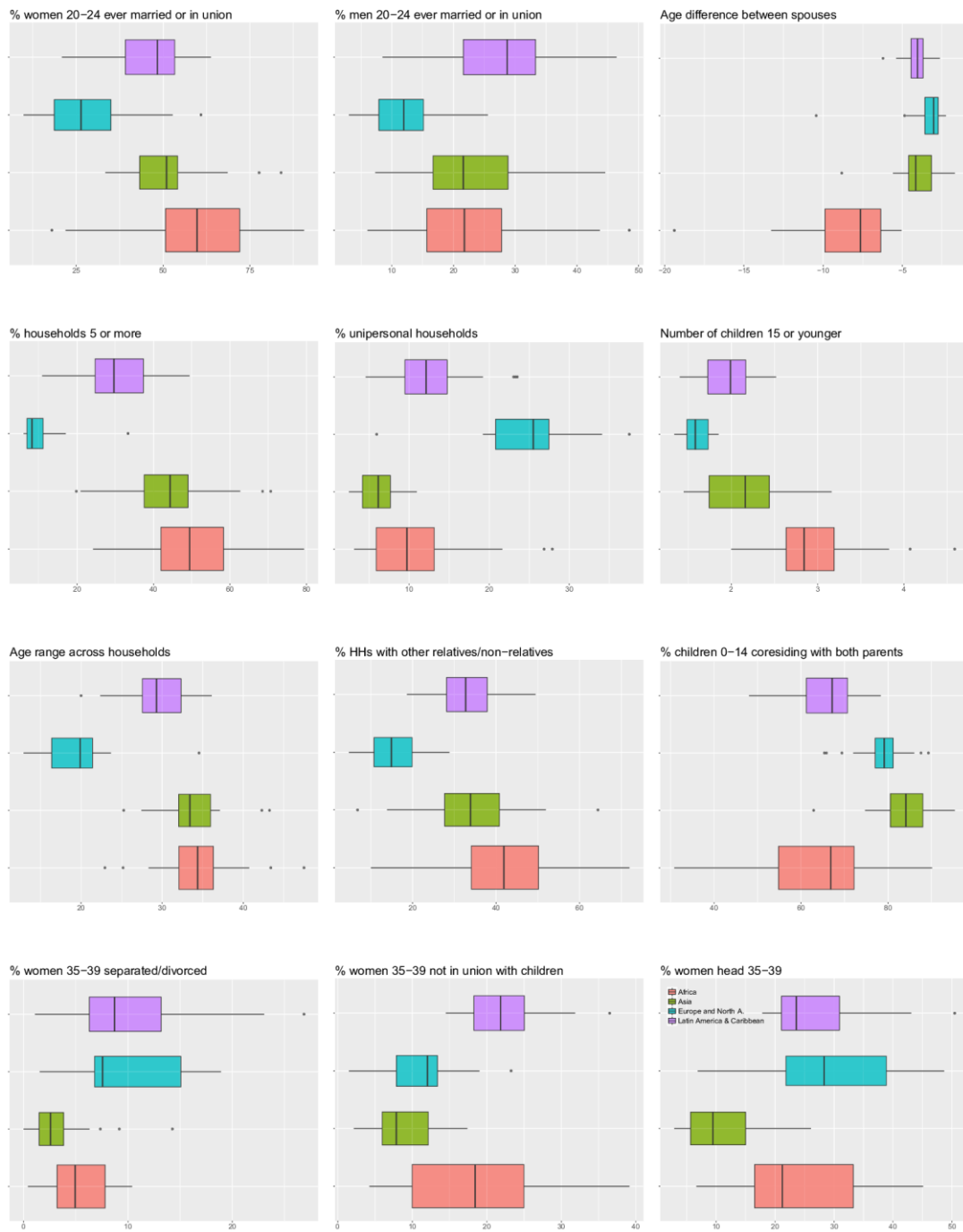
developmental paradigm on family life. Chicago: The University of Chicago Press.

Todd, E. (1985). *The Explanation of Ideology: Family Structures and Social Systems*. Oxford: Blackwell.

Table 1. Family dimensions and family life indicators by dimension

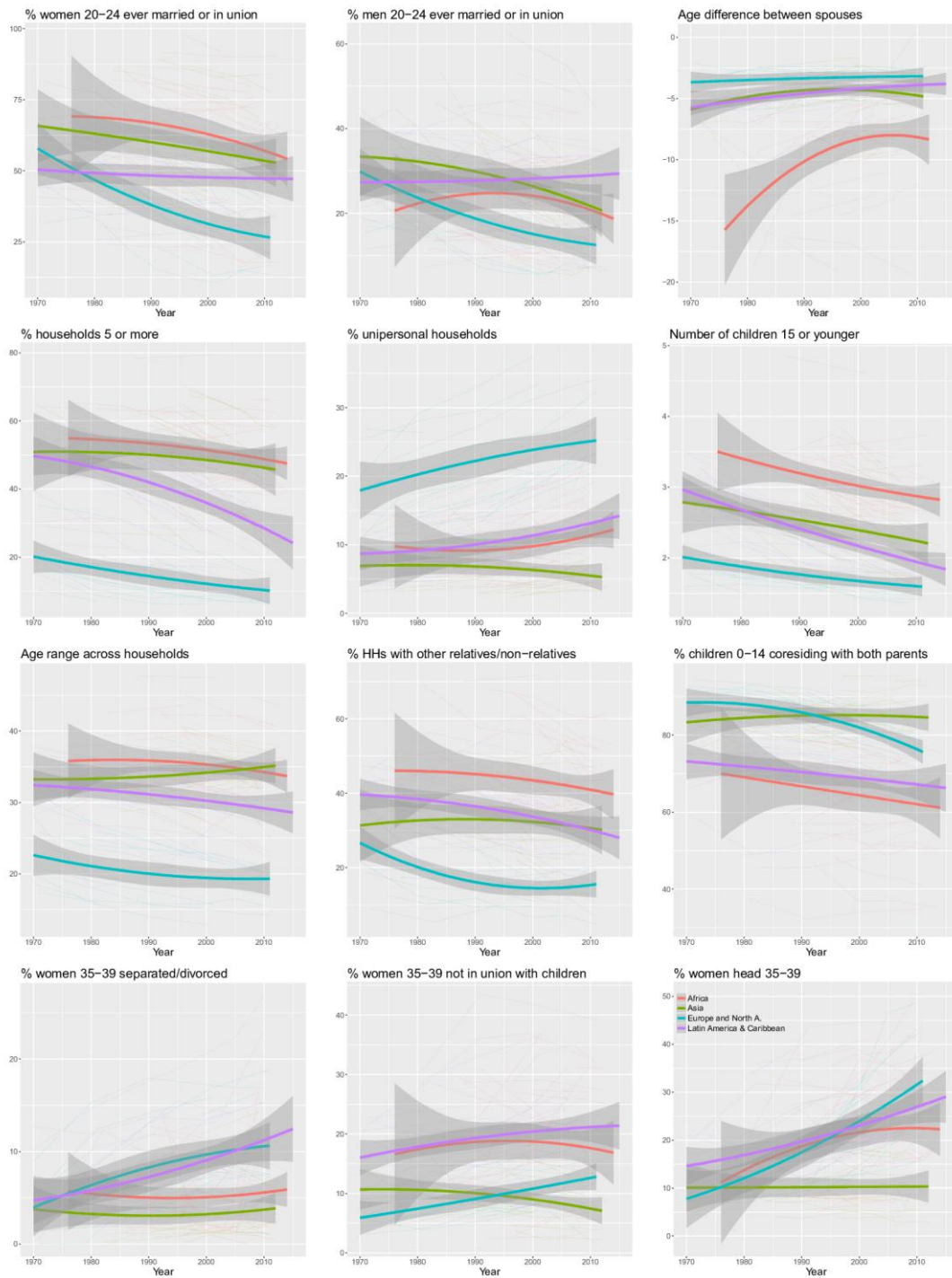
Dimension	Indicator	Short label
Union formation and dissolution	Females 20-24 ever married or in union	Ind1
	Males 20-24 ever married or in union	Ind2
	Average age difference between spouses (women 30-34)	Ind3
Reproduction	% households 5 or more	Ind4
	% unipersonal households	Ind5
	Average number of children 15 or younger	Ind6
Household structure and complexity	Average age range across households	Ind7
	% households with other relatives and non-relatives	Ind8
	Children 0-14 co-resident living with both parents	Ind9
Women's position within the family	% 35-39 women separated or divorced	Ind10
	% Females 35-39 not in union with children	Ind11
	% 35-39 female heads	Ind12

Figure 1. Patterns in 12 family-life indicators in 110 countries by continent (most recent available data since 2000)



Source: Own calculations based on IPUMS-I census microdata and DHS and EU-LFS survey microdata.

Figure 2. Trends in 12 family-life indicators in 110 countries and continental average trends



Source: Own calculations based on IPUMS-I census microdata and DHS and EU-LFS survey microdata.

Figure 3. Classification dendrogram for 99 countries

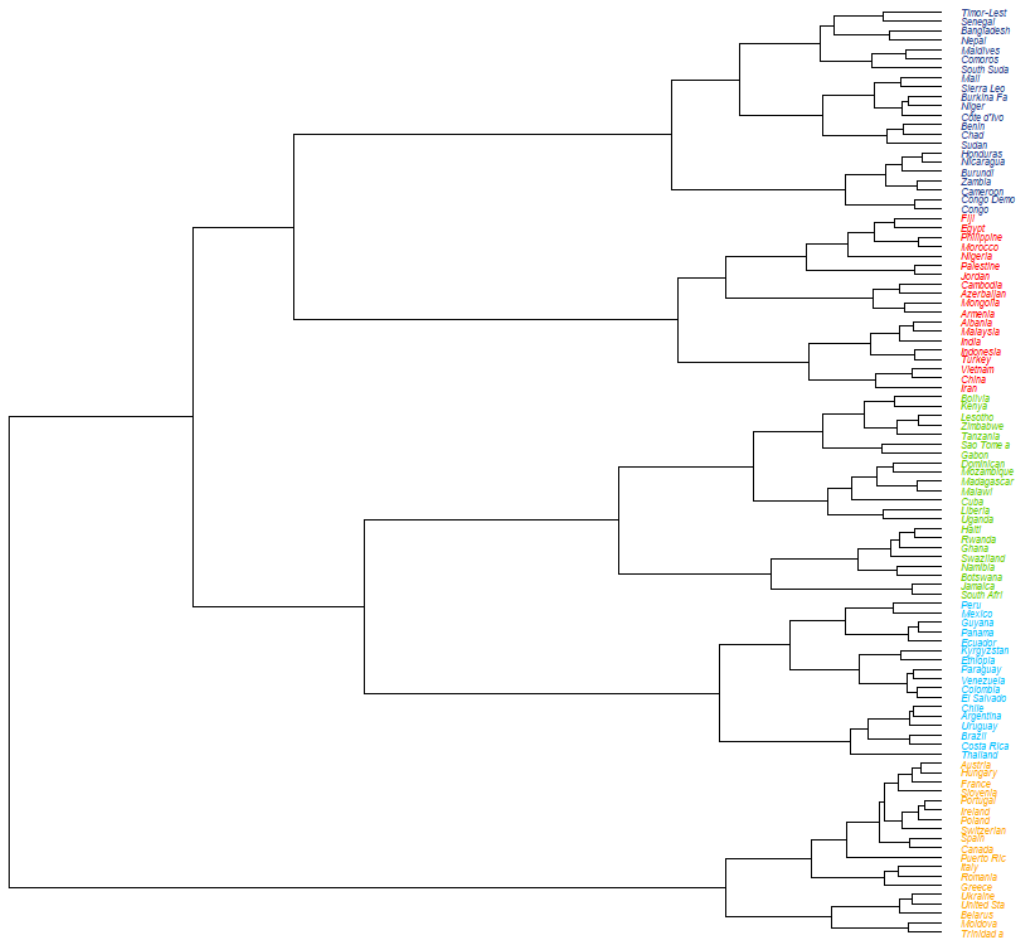


Figure 4. Mean and confidence intervals for the twelve standardized family indicators by cluster and geographical distribution

