MAX PLANCK INSTITUTE
FOR DEMOGRAPHIC RESEARCH

Konrad-Zuse-Strasse 1 · D-18057 Rostock · Germany · Tel +49 (0) 3 81 20 81 - 0 · Fax +49 (0) 3 81 20 81 - 202 · www.demogr.mpg.de

MPIDR Working Paper WP 2022-011 | March 2022
https://doi.org/10.4054/MPIDR-WP-2022-011

Indirect estimation of the timing of first union dissolution with incomplete marriage histories

Ben Malinga John | john@demogr.mpg.de
Natalie Nitsche | nitsche@demogr.mpg.de

Indirect estimation of the timing of first union dissolution with incomplete marriage histories

Ben Malinga John[1,2,3] & Natalie Nitsche[1]

1. Max Planck Institute for Demographic Research (MPIDR)
2. Stockholm University
3. Department of Population Studies, University of Malawi

Contact details

Ben Malinga John

Email: john@demogr.mpg.de / bjohn@unima.ac.mw

Natalie Nitsche

Email: nitsche@demogr.mpg.de

Acknowledgements

We are grateful for the feedback on an earlier version of this paper from Christian Dudel and Angela Carollo of Max Planck Institute for Demographic Research. Their insights helped us tremendously to refine the message of this manuscript. Nonetheless, none of them is responsible for how we decided to use their comments and suggestions.

Abstract

Lack of nationally representative data with detailed marriage histories in developing countries impedes comprehensive understanding of essential aspects of union dissolution in these countries – for example, the timing of first union dissolution. This research note proposes a method, denoted as 'Indirect Life Table of first Union Dissolution' (ILTUD), for estimating quantum adjusted measures of the timing of first union dissolution from incomplete marriage histories. ILTUD estimates a survival function of first union from a simple tabulation of ever-married women by duration since first union, classified by union dissolution status (intact vs dissolved first union). It then uses the relationships between life table functions to generate the distribution of marriages ending each year (θ_t) for a given marriage cohort. Using the distribution of θ_t, ILTUD generates quantum adjusted first union survival rates from which the percentiles of first union dissolution are calculated. ILTUD estimates are consistent with estimates produced using traditional statistical methods such as the Kaplan Meier estimator. In addition, ILTUD is simple to implement and has simple data requirements – available in most nationally representative surveys, e.g. the Demographic Health Surveys. Thus, ILTUD presents an opportunity for broadening our understanding of union dissolution dynamics in developing countries.

Keywords: Marriage, Union dissolution, Demographic methods, Survival analysis, Indirect estimation

Introduction

Experiencing a union dissolution may impact the life course in several ways. For example, women who experience a union dissolution typically have lower fertility than those who remain in their first union (John and Adjiwanou 2021; Uddin and Hosain 2013; Van Bavel et al. 2012). Moreover, not only whether but also when a union dissolution occurs is likely consequential (Meggiolaro and Ongaro 2010; Thomson et al. 2012). Thus, assessing how long individuals have been partnered before experiencing the first union dissolution is relevant for studying marriage and fertility dynamics.

However, nationally representative information on the timing of union dissolution is lacking in several developing countries, mainly due to limited availability of detailed marriage histories in available data sources. The most widely used source of nuptiality information in these countries (e.g. Demographic Health Surveys (DHS)) only show whether a first union ended but not when (date or age) the event occurred. Therefore, traditional statistical techniques such as the Kaplan Meier estimator (Kaplan and Meier 1958) cannot be used to analyze the time to first union dissolution. A growing interest in understanding union dissolution variation and change in these countries (Cherlin 2017; Clark and Brauner-Otto 2015) underscores the need to develop approaches that can produce reasonable measures of union dissolution timing from incomplete marriage histories. We, therefore, propose a technique for estimating quantum adjusted first union survival rates, which we call 'Indirect Life Table of first Union Dissolution' (ILTUD). ILTUD calculates the first union survival rates under some assumptions from incomplete marriage histories. It then uses relationships between life table functions to calculate the distribution of marriages ending each year for a given marriage cohort. We discuss and demonstrate the application and robustness of ILTUD using World Fertility Survey data.

Method description

Estimation of survival rates of first union dissolution ($S(t)$) from incomplete marriage histories

For cross-sectional data with complete marriage histories, the survival rate of first union can be derived using traditional statistical methods. The calculated survival rate of first union at time t $S(t)$ from such methods refers to the probability of the first union surviving t years following its onset. Therefore, if first union survival rates have remained reasonably stable across different marriage cohorts, $S(t)$ can equally be approximated by calculating the proportion of women in intact first unions among women who first married t years before the survey (ρ_t). This implies that we can estimate $S(t)$ from a tabulation of ever-married women by duration since first union, classified by first union dissolution status (intact vs dissolved first union). Such tabulation does not require knowledge of when a first union ended. Thus, it can be produced from cross-section data with incomplete marriage histories.

Our suggestion is to use ρ_t as a measure of first union survival rates $S(t)$ for data with incomplete marriage histories. However, first union survival rates may likely vary across marriage cohorts, in which case the $\rho_t's$ values cannot equal the underlying first union survival rates. Thus, the $S(t)$ estimates we propose to derive using $\rho_t's$ generally relate to a synthetic cohort. For example, $S(15)$ would be interpreted as the proportion of women of a given marriage cohort whose first unions would be expected to end during the first 15 years of first union if the observed duration specific first union survival rates continue to apply throughout their life. This interpretation resembles other hypothetical measures used in demographic analysis, such as the Total Fertility Rate (TFR).

Plotting the values of $\rho_t's$ against t should thus yield a monotonically decreasing function (since $\rho_t \equiv S(t)$). However, $\rho_t's$ estimates may not decrease monotonically because of limited sample size, data reporting errors or random variation. We can address this issue by smoothing the distribution of $\rho_t's$ to yield monotonically decreasing $S(t)$ estimates. We suggest smoothing the ρ_t values by fitting a polynomial function. One could also consider smoothing ρ_t estimates using parametric models like the log-normal function.

Estimating quantum adjusted first union survival rates

The ultimate goal of ILTUD is to measure how long, on average, individuals are partnered before experiencing a first union dissolution from incomplete marriage histories. Because not every married woman experiences a union dissolution and the levels of union dissolution vary

substantially over space and time, the timing measures of first union dissolution that are fully adjusted for level (quantum) of union dissolution are more meaningful. ILTUD uses $S(t)$ estimates derived as described above and the relationships between life table functions to produce such estimates.

The life table is one of the traditional demographic techniques for estimating survival rates. It involves estimating survival rates from the distribution of events and censored observations at predefined time points. Table 1 shows the typical data matrix required for this estimation. The first column, t, represents the time intervals. Column n_t depicts the number of observations at the beginning of interval t. ω_t and θ_t denote the number of censored observations and events during interval t, respectively. $S(t)$ indicates the survival rate during interval t.

Table 1: Representation of data matrix for calculation of survival rates using a life table method

t	interval	n_t	ω_t	θ_t	S(t)
1	[0, 1)	n_1	ω_1	θ_1	S(1)
2	[1, 2)	n_2	ω_2	θ_2	S(2)
.	.	.	.	.	.
.	.	.	.	.	.
t	[t-1, t)	n_t	ω_t	θ_t	S(t)

Notes

In Table 1, the values of n_t, ω_t and θ_t are related as

$$n_t = n_{t-1} - (\omega_{t-1} + \theta_{t-1}) \qquad t > 1 \tag{1}$$

The survival rate at interval t, $S(t)$ is estimated as

$$S(t) = \prod_{j=1}^{t} \frac{\left(n_j - \frac{\omega_j}{2}\right) - \theta_j}{\left(n_j - \frac{\omega_j}{2}\right)} \tag{2}$$

For estimation of first union survival rates, t in Table 1 represents the time since the onset of the first union. n_t is equivalent to the number of ever-married women at the beginning of interval t. ω_t denotes the number of individuals in intact first unions. θ_t corresponds to the number of marriages ending during interval t. Thus, the $S(t)$ in equation (2) estimates the proportion of first unions that survive marriage interval t. Essentially, it is the distribution of θ_t we need to understand how long, on average, individuals are partnered before experiencing a first union dissolution. We can compute the values of $\theta_t's$ by rearranging equation (2) as

$$S(t) = \left(\prod_{j=1}^{t-1} \frac{\left(n_j - \frac{\omega_j}{2}\right) - \theta_j}{n_j - \frac{\omega_j}{2}}\right) \times \left(\frac{\left(n_t - \frac{\omega_t}{2}\right) - \theta_t}{n_t - \frac{\omega_t}{2}}\right) \tag{3}$$

Equation (3) reduced to

$$S(t) = S(t-1) \times \left(\frac{\left(n_t - \frac{\omega_t}{2}\right) - \theta_t}{n_t - \frac{\omega_t}{2}}\right) \tag{4}$$

Hence solving for θ_t, in equation (4), we have

$$\theta_t = \left(n_t - \frac{\omega_t}{2}\right)\left(1 - \frac{S(t)}{S(t-1)}\right) \tag{5}$$

Since ω_t denotes the number of individuals in intact first unions during interval t, we can directly extract the values of $\omega_t's$ from a tabulation of ever-married women by duration since first union, classified by union dissolution status. We can also directly observe the value of $n_{1=[0,1)}$ since it indicates the number of ever-married individuals in the dataset. Thus, equation (1) and equation (5) suggest that if $S(t)$ estimates are available, one can obtain the values of $\theta_t's$ and $n_t's$ $(t > 1)$ iteratively. ILTUD uses the smoothed $S(t)'s$ described earlier to produce the distribution of $\theta_t's$.

Once the distribution of marriages ending each year for a given marriage cohort ($\theta_t's$) is produced using equation (5), the *quantum adjusted* survival rate of the first union at duration t ($S^{\star}(t)$) can be estimated using an equivalent of equation (2) with $\omega_t's = 0$. Denoted as

$$S^{\star}(t) = \prod_{j=1}^{t} \frac{n^{\star}_j - \theta_j}{n^{\star}_j} \tag{6}$$

Where

$$n^{\star}_t = \sum_{j=t}^{\tau} \theta_j \qquad \tau = \text{maxmum marriage duration} \tag{7}$$

Using linear interpolation, one can use $S^{\star}(t)$ estimates to calculate summary measures (e.g. median) of the timing of first union dissolution that are fully adjusted for the quantum of union dissolution.

Method application

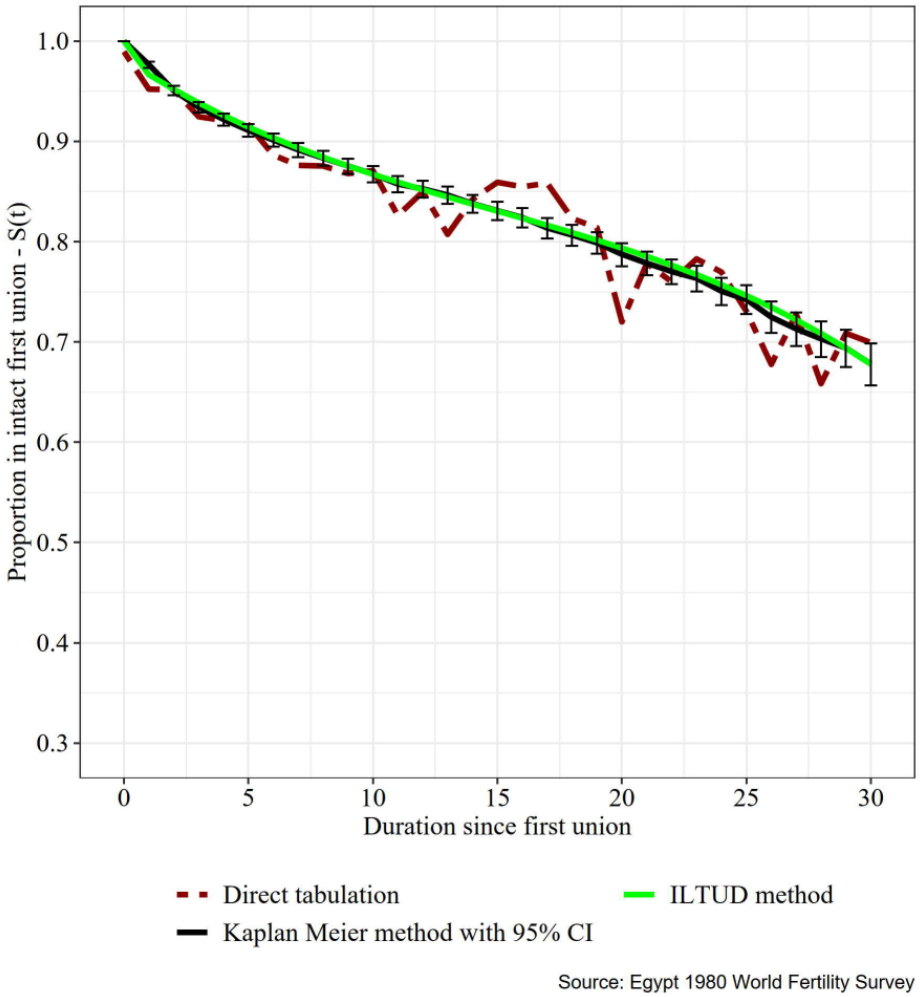
We now illustrate the application and robustness of ILTUD using data from World Fertility Surveys collected in 12 countries in Africa (Table 2). These datasets contain detailed marriage histories. Thus, they allow us to derive first union survival rates using traditional statistical methods, which we use as a benchmark to evaluate ILTUD estimates. The purpose of this illustration is twofold. First, we intend to demonstrate the step by step application of ILTUD. Second, we underscore the idea that estimates derived using ILTUD relate to a synthetic cohort and are robust. We have used data from Egypt for the first objective. We then applied the method to each country for the second objective.

The Egypt dataset has a sample of 8,788 ever-married women, of whom 1,438 (16.4.6%) had a dissolved first union. The first step of ILTUD is to tabulate ever-married women by the number of years since the first union, according to first union dissolution status. We use the output of this tabulation to calculate the proportion of women in intact first unions ($\rho_t{}'s$) at each union duration. For example, 372 women first married [2, 3) years before the survey, of whom 354 were still in first unions. Thus,

$$\rho_{3=[2,3)} = \left(\frac{352}{372}\right) = 0.952$$

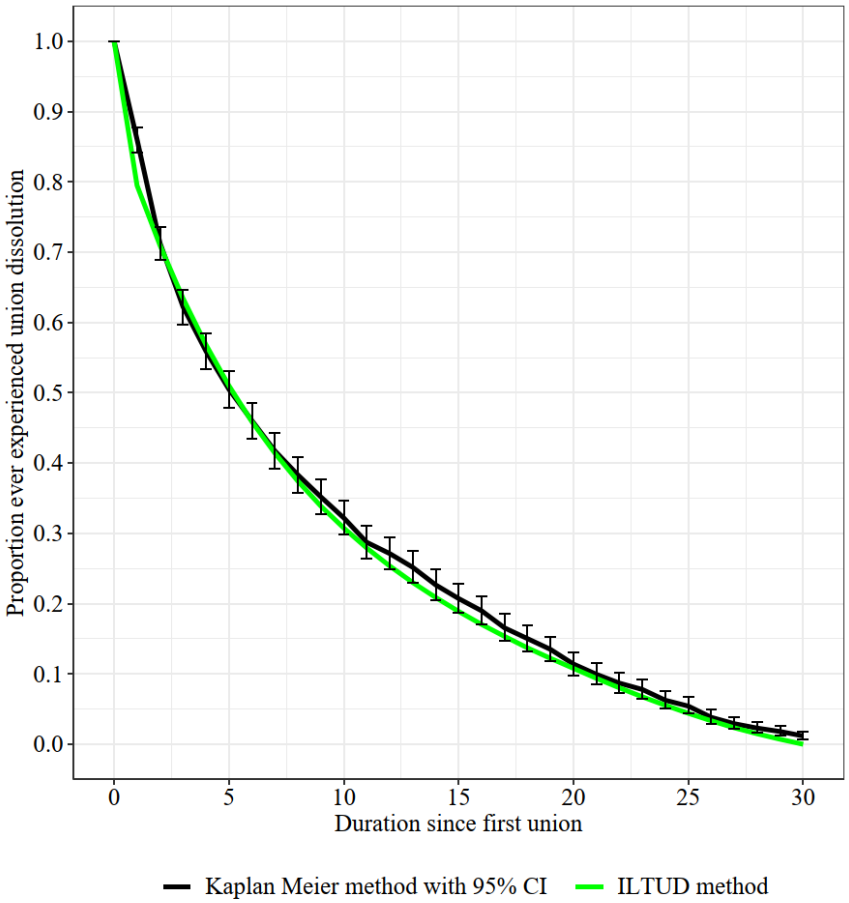
The second step involves smoothing the $\rho_t's$ values to produce monotonically decreasing first union survival rates. This analysis smoothed the ρ_t values corresponding to marriage duration 0 to 30 years using a third-degree polynomial function. Figure 1 shows the observed and smoothed estimates of ρ_t values alongside $S(t)$ estimates derived using the Kaplan Meier method. The results in Figure 1 reveal that the smoothed ρ_t estimates are closely identical to those derived using the Kaplan Meier approach. This pattern implies that union dissolution rates of younger marriage cohorts reasonably reflect the experience of older cohorts (see Figure 3).

Figure 1: Survival rates of the first union by duration since first union, calculated using (1) complete marriage histories and Kaplan Meier method (black - as a benchmark) and (2) incomplete marriage histories (red = observed $\rho_t's$, green = ILTUD method (polynomial smooth of $\rho_t's$))



The third step involves using the smoothed $\rho_t's$ estimates and the distribution of women in intact first unions to calculate the number of marriages ending each year ($\theta_t's$) for a given marriage cohort. We perform this calculation by first using equation (5) to produce $\theta_t's$ and then using equation (1) to calculate the number of ever-married women at the beginning of each interval ($n_t's$). For example, we produced the number of unions expected to end during the first year of union (θ_1) by setting $n_1 = 8{,}788$. The number of women in intact unions during this interval is 375. The fitted $S(1)$ is 0.9712. Thus, using equation (5) yields 248 as the number of first unions that dissolved during the first year. Subtracting 375 and 248 from 8,788 gives the number of ever-married women at the beginning of the second year of union ($n_2 = 8165$). We then use equation (5) again to produce θ_2 and equation (1) to produce n_3 and so forth.

Figure 2: Quantum adjusted survival rates of first union by duration since the first union calculated using (1) complete marriage histories and Kaplan Meier method (black - as a benchmark) and (2) incomplete marriage histories using ILTUD method (green)



Finally, we estimate the survival rates of the first union among women whose first union ended using equations (6) and (7). Figure 2 shows these estimates based on the ILTUD and Kaplan Meier methods.

Figure 2 shows that ILTUD estimates are identical to those derived using the Kaplan Meier method. Using ILTUD, we found that half of the women whose first union ended experienced the event within 5.2 (4.4 – 6.3 95% CI) years of the first union. This estimate is 5.2 (4.6 – 5.8 95% CI) for the Kaplan Meier method.

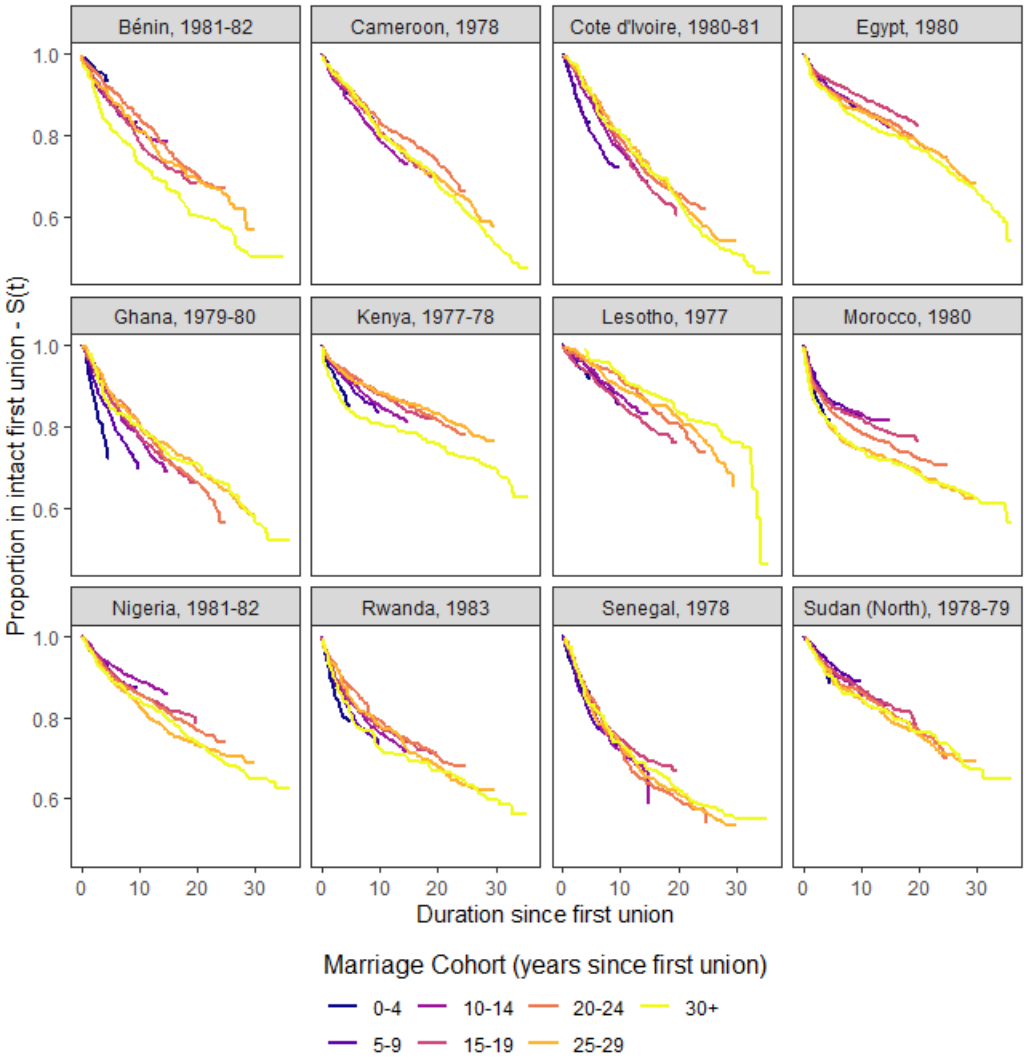
Table 2 summarizes the results of the application of ILTUD to each of the 12 countries we considered. It shows the duration at which 25%, 50%, and 75% of first unions were dissolved among individuals whose first union ended, derived using the ILTUD and Kaplan Meier methods. Table 2 illustrate that the estimates from these two methods are relatively similar in Benin, Cameroon, Egypt, Lesotho, Nigeria, Senegal and Sudan. These countries have first union survival rates comparable across marriage cohorts (Figure 3). Generally, Table 2 and Figure 3 reveal that significant discrepancies between estimates derived using ILTUD and Kaplan Meier methods exist in countries where first union survival rates for younger cohorts differ substantially from older marriage cohorts (for example, in Ghana and Morocco). In these countries, the smoothed ρ_t values at a shorter union duration depict the union dissolution experience of the younger marriage cohorts and not all women. These findings underscore that estimates derived using ILTUD should be interpreted as synthetic cohort measures.

Table 2: Summary measures of quantum adjusted timing of first union dissolution (expressed in years), comparing estimates calculated using (1) complete marriage histories and Kaplan Meier method (as a benchmark) and (2) incomplete marriage histories using ILTUD method

			25 Percentile		Median		75 Percentile	
County and survey year	*Sample size*	*% of dissolved first unions*	*ILTUD method (95% CI*	*Kaplan Meier method (95% CI)*	*ILTUD method (95% CI)*	*Kaplan Meier method (95% CI)*	*ILTUD method (95% CI*	*Kaplan Meier method (95% CI)*
Benin, 1981-82	3577	19.9	3.1 (2.6, 3.6)	2.2 (1.8, 2.5)	5.7 (4.9, 6.8)	5.4 (4.8, 6.1)	10.6 (8.7, 12.4)	10.7 (9.8, 11.6)
Cameroon, 1978	7088	22.8	2.8 (2.5, 3.2)	2.8 (2.4, 3.1)	5.7 (5.1, 6.5)	6.8 (6.5, 7.3)	11.7 (10, 13.3)	13.1 (12.1, 13.9)
Cote d'Ivoire, 1980-81	4990	26.8	1.4 (0.9, 1.9)	2.4 (2.3, 2.7)	3.9 (3.4, 4.4)	5.6 (5.2, 6.0)	8.0 (6.9, 9.1)	10.8 (10.0, 11.6)
Egypt, 1980	8788	16.4	1.5 (0.9, 2.2)	1.8 (1.6, 1.9)	5.2 (4.4, 6.3)	5.2 (4.6, 5.8)	12.2 (10.3, 14)	13.1 (11.9, 14.0)
Ghana, 1979-80	4912	27.3	0.7 (0.5, 0.8)	2.3 (2.1, 2.4)	2.5 (1.5, 3.2)	4.9 (4.5, 5.3)	7.2 (6.0, 8.8)	10.2 (9.8, 10.8)
Kenya, 1977-78	6241	16.3	0.6 (0.5, 0.9)	1.4 (1.3. 1.6)	2.1 (1.0, 3.0)	4.3 (3.8, 4.8)	6.2 (4.9, 8.0)	9.8 (8.9, 10.7)
Lesotho, 1977	3603	15	2.5 (1.0, 3.3)	3.5 (2.8, 3.9)	6.1 (4.9, 7.6)	7.3 (6.4, 8.4)	11.3 (9.2, 12.9)	13.9 (12.9, 15.1)
Morocco, 1980	4105	23.6	0.6 (0.5, 0.8)	1.1 (1.0, 1.2)	2.9 (0.9, 5.1)	2.6 (2.3, 3.0)	12.7 (9.7, 15.5)	7.4 (6.5, 8.8)
Nigeria, 1981-82	8188	15.5	2.3 (1.5, 2.9)	2.2 (2.1, 2.4)	5.8 (4.9, 7.1)	5.2 (4.8, 5.5)	12.6 (10.7, 14.4)	10.7 (9.8, 11.3)
Rwanda, 1983	4071	25.2	0.6 (0.5, 0.9)	1.3 (1.2, 1.4)	1.8 (0.9, 2.7)	3.3 (3.0, 3.5)	5.3 (4.2, 6.8)	7.7 (7.0, 8.3)
Senegal, 1978	3472	28.7	1.2 (0.8, 1.9)	2.2 (2.0, 2.3)	4.6 (3.6, 5.5)	4.5 (4.1, 4.9)	9.7 (8.1, 11.2)	9.3 (8.3, 10.1)
Sudan (1978-79)	3115	17.3	1.9 (0.7, 3.3)	2.3 (2.0, 2.7)	6.7 (4.9, 8.6)	5.3 (3.8, 6.0)	12.8 (10.4, 14.7)	11.6 (10.7, 13.4)

Source: Authors calculation. ILTUD CIs are based on 1000 bootstrap samples, and they are calculated using a percentile approach

Figure 3: Survival rates of first union by marriage cohorts



Summary and conclusion

Estimating the quantum adjusted timing of first union dissolutions when such information is not directly available is a necessary next step for broadening our understanding of country and period differences in the occurrence of first union dissolutions. This article proposed and demonstrated an Indirect Life Table of first Union Dissolution (ILTUD) method for producing such estimates from incomplete marriage histories. ILTUD can advance scholarship on family dynamics in developing countries, where representative data on union dissolution timing is often lacking while incomplete marriage histories are available. Measures derived using ILTUD generally relate to a

synthetic cohort. They represent a life-course union dissolution trajectory that a given marriage cohort would experience if the observed duration-specific first union survival rates prevail, similar to other demographic synthetic measures like the TFR. We indeed demonstrated that ILTUD yields estimates comparable to those derived using the Kaplan Meier method, particularly when first union survival rates across marriage cohorts have remained reasonably stable.

The key data input for ILTUD is the distribution of ever-married individuals by duration since the first union, classified by first union dissolution status. While such data is often available when data on dissolution timing is lacking, its quality could be affected by reporting errors such as the omission of early unions (leading to misclassification of union dissolution status) or misreporting of date of first marriage (leading to misclassification of duration since the first union). Evaluation of marriage histories in developing countries suggests that these reporting errors are pronounced with increasing age or marriage duration (Chae 2016; Gage-Brandon 1995, Mensch et al. 2006). Therefore, we recommend estimating the first union survival rates based on data from younger marriage cohorts. In this paper, we considered women who were married for 0 to 30 years.

Furthermore, due to data errors or limited sample size, the distribution of the proportion of individuals in intact first unions by marriage duration may require smoothing to produce reasonable estimates. We suggested the use of polynomial regression for such smoothing. However, other parametric or semi-parametric models that adequately fit the data could be used. Finally, the required data for ILTUD can easily be extracted from representative surveys available in most developing countries (e.g. DHSs). Thus, ILTUD presents an opportunity to assess country and period differences in the timing of first union dissolutions in these countries.

References

Chae, S. (2016). Forgotten marriages? Measuring the reliability of marriage histories. *Demographic Research, 34(19),* 525-562. Doi: 10.4054/DemRes.2016.34.19

Cherlin, A. J. (2017). Introduction to the special collection on separation, divorce, repartnering, and remarriage around the world. *Demographic Research, 37(38),* 1275-1296. Doi: https://dx.doi.org/10.4054/DemRes.2017.37.38

Clark, S., & Brauner-Otto, S. (2015). Divorce in sub-Saharan Africa: Are unions becoming less stable? *Population and Development Review*, *41(4)*, 583-605. Doi: https://doi.org/10.1111/j.1728-4457.2015.00086.x

Gage-Brandon, A. J. (1995). An assessment of the quality of data on age at first union, first birth, and first sexual intercourse for Phase II of the Demographic and Health Surveys Program. Macro International Incorporated.

John, B. M., & Adjiwanou, V. (2021). Fertility decline in sub-Saharan Africa: Does remarriage matter? *Population Studies*, 1-21. Doi: https://doi.org/10.1080/00324728.2021.1933148

Kaplan, E. L., & Meier, P. (1958). Nonparametric estimation from incomplete observations. *Journal of the American statistical association*, *53(282)*, 457-481.

Meggiolaro, S., & Ongaro, F. (2010). The implications of marital instability for a woman's fertility: Empirical evidence from Italy. *Demographic Research, 23(34)*, 963-996. Doi https://dx.doi.org/10.4054/DemRes.2010.23.34

Mensch, B. S., Grant, M. J., & Blanc, A. K. (2006). The changing context of sexual initiation in sub-Saharan Africa. *Population and Development Review,32(4)*, 699-727. Doi: https://doi.org/10.1111/j.1728-4457.2006.00147.x

Thomson, E., Winkler-Dworak, M., Spielauer, M., & Prskawetz, A. (2012). Union instability as an engine of fertility? A microsimulation model for France. *Demography*, *49(1)*, 175-195. Doi: https://doi.org/10.1007/s13524-011-0085-5

Uddin, M. S. G., & Hosain, M. M. (2013). Factors affecting marital instability and its impact on fertility in Bangladesh. *ASA University Review*, *7(2)*, 35-41. http://www.asaub.edu.bd/data/asaubreview/v7n2sl4.pdf.

Van Bavel, J., Jansen, M., & Wijckmans, B. (2012). Has divorce become a pro-natal force in Europe at the turn of the 21st century? *Population Research and Policy Review*, *31(5)*, 751-775. http://dx.doi.org/10.1007/s11113-012-9237-6