

Indices au niveau individuel de fécondité, de mortalité des enfants et de nuptialité

Dominique TABUTIN

9

Université catholique de Louvain

Département des Sciences de la Population
et du Développement

Indices au niveau individuel de fécondité, de mortalité des enfants et de nuptialité

Dominique TABUTIN

Document de Travail n° 9
Mai 2000

Texte publié sous la responsabilité de l'auteur



Université catholique de Louvain
Département des sciences de la population et du développement
1, Place Montesquieu, bte 4 B-1348 Louvain-la-Neuve
Tel. : (32 10) 47 40 41 Fax : (32 10) 47 29 52
E-mail : letocart@sped.ucl.ac.be

INDICES AU NIVEAU INDIVIDUEL DE FECONDITE, DE MORTALITE DES ENFANTS ET DE NUPTIALITE

Dominique TABUTIN

Institut de Démographie

Avec notamment la multiplication des enquêtes, la recherche démographique depuis une trentaine d'années s'est de plus en plus orientée vers l'analyse des données individuelles. Sur le plan explicatif, on est passé du niveau macro et agrégé, longtemps privilégié, au niveau micro¹, avec les divers outils statistiques disponibles en sciences sociales (classification, régressions...). Dans une enquête ou un recensement, chaque individu est bien caractérisé par son âge, son sexe, son état matrimonial, son éducation, sa profession, parfois son ethnie ou sa religion... qui constituent les variables dites explicatives ou indépendantes. Mais comment définir simplement les variables dépendantes (ou à expliquer) que peuvent être la fécondité d'une femme (ou d'un homme), l'histoire de la femme (ou d'un couple) en matière de mortalité des enfants ou encore l'intensité de la mobilité conjugale ?

Le problème est de créer des indices de synthèse au niveau individuel, résumant une vie ou une expérience et contrôlant les variables de confusion que sont souvent en démographie l'âge de l'individu, l'âge et la durée de mariage ou encore le nombre d'enfants nés. On positionnera le plus

Je remercie B. Schoumaker et G. Wunsch de leurs commentaires sur ce texte. Cet article paraîtra en 2001 dans l'un des volumes de l'ouvrage *Démographie : Analyse et synthèse*, coordonné par G. Caselli, J. Vallin et G. Wunsch.

1 Avec aujourd'hui une tendance à essayer de relier les divers niveaux (approches multi-niveaux).

souvent chaque individu par rapport à un standard ou une référence. Ces indicateurs permettront de comparer directement les fécondités individuelles de femmes (ou d'hommes) et de les intégrer dans des analyses statistiques individuelles, ou encore – si besoin est – de construire des groupes, selon les niveaux de sur ou sous-fécondité (par rapport au standard) ou de sur ou sous-mortalité des enfants.

Bien sûr, on peut aussi procéder à des analyses statistiques individuelles à partir de l'occurrence (ou pas) d'événements dans une période donnée : par exemple, naissance ou pas au cours des cinq dernières années ? décès d'enfants ou pas ? recours ou non à la contraception ? ... Ou encore à partir du nombre d'événements survenus sur une période récente bien précise (les cinq dernières années par exemple). Mais cela n'est point ici notre propos principal, comme nous n'aborderons pas non plus le problème, vaste et complexe, des outils statistiques requis pour ces analyses de déterminants individuels.

Nous nous centrerons essentiellement sur la fécondité et la mortalité, de loin les plus traités dans le domaine des indices individuels, avec néanmoins quelques exemples sur ce que l'on fait ou pourrait faire en matière de nuptialité. Pour chaque phénomène, nous présenterons les indices les plus courants, en partant des principes, en dégagant aussi leurs intérêts et limites. Précisons de suite que cette approche a surtout été utilisée dans les études « explicatives » sur les pays du Sud, basées pour une grande partie d'entre elles sur les données des enquêtes de type EMF (Enquête Mondiale de Fécondité) ou EDS (Enquête Démographique et de Santé).

1. Indicateurs individuels de fécondité

L'indice de fécondité cumulée le plus classique et le plus immédiat au niveau individuel, résumant l'histoire féconde d'une femme (ou d'un homme), est *la parité* (le nombre d'enfants rés-vivants que déclare une femme à un moment donné). Pour contrôler l'effet de l'âge de la femme, de son âge au mariage et de sa durée de vie conjugale, on peut calculer *une parité corrigée* de la durée d'exposition au risque de fécondité, ou encore construire *un indice standardisé* (comme le DRAT) qui positionne la fécondité de la femme par rapport à un standard en contrôlant l'âge ou la durée de mariage.

1.1. La parité déclarée

Nombre d'études sur les relations au niveau individuel entre fécondité et variables socio-économiques ont utilisé – et utilisent encore – le nombre d'enfant nés-vivants (ou parité), c'est-à-dire le nombre cumulé de naissances vivantes qu'une femme a eu du début de sa vie génésique jusqu'au moment de l'enquête. Il s'agit le plus souvent des femmes mariées.

C'est une information a priori simple (pas de datation, pas de période de référence, pas de problèmes donc de localisation de l'événement dans le temps), mais qui n'est pas non plus dénuée de risques d'omissions (notamment des enfants décédés rapidement après leur naissance) ou de confusions (entre mort-nés et naissances vivantes par exemple). Ces risques sont en général d'autant plus grands que la femme est âgée, qu'elle a eu une descendance élevée et que la mortalité du contexte est forte. Il faut aussi que l'âge des femmes soit correctement déclaré².

Dans les analyses statistiques utilisant la parité, on se doit dès lors de contrôler l'âge et/ou la durée de mariage. Il y a deux grandes alternatives :

La première est de prendre la parité de chaque femme comme variable dépendante, en introduisant l'âge et la durée de mariage parmi les variables explicatives. Mais dans ce cas de figure, une grande partie de « l'explication » se focalisera sur ces deux variables, et surtout on fait l'hypothèse que les relations entre la fécondité et les variables explicatives ne varient pas selon l'âge ou la génération de femmes.

La seconde est de travailler sur certains groupes d'âges (ou de générations) de femmes, de stratifier en quelque sorte l'échantillon. Là encore, divers choix sont possibles :

- *ne travailler que sur les femmes les plus âgées, celles qui ont quasiment achevé leur vie reproductive (les 40-49 ou 45-49 ans), en contrôlant cependant la durée de mariage.* Mais il est évident que sur le plan explicatif les comportements de ces femmes relèvent du passé, reflétant peu les pratiques actuelles, dans les situations au moins où la fécondité recule. Par ailleurs, les caractéristiques de ces femmes à l'enquête peuvent différer de ce qu'elles étaient à l'époque de leurs grandes décisions en matière de fécondité : variable à expliquer et variables explicatives ne sont plus alors dans le même champ temporel. Cela peut conduire à des biais et erreurs non négligeables dans l'interprétation.
- *centrer l'analyse sur un groupe d'âges intermédiaires, en pleine vie féconde et d'effectifs suffisamment importants : cela limite « le mélange » des générations et des histoires reproductives.* C'est ce que

2 On assiste fréquemment, en Afrique par exemple, à un vieillissement des jeunes femmes et à un rajeunissement des femmes les plus âgées, conduisant selon le cas à des sous-estimations ou sur-estimations fictives du nombre d'enfants.

B. Schoumaker (2000) a fait dans une étude récente des déterminants individuels et contextuels (analyse multi-niveaux) de la fécondité du Maroc rural : il a pris le nombre d'enfants nés-vivants des femmes de 25-39 ans comme variable dépendante et introduit les groupes d'âges 30-34 ans et 35-39 ans ainsi que la durée de mariage parmi les vingt variables explicatives. Avec d'autres variables comme l'accès aux médias ou le niveau contextuel de mortalité, l'âge mais surtout la durée de mariage ont montré un effet significatif sur la fécondité.

- stratifier l'échantillon des mères en quelques groupes d'âges pour procéder à des analyses séparées sur, par exemple, les 25-34 ans, 35-44 ans et 45 ans et plus, en introduisant bien sûr comme variable de contrôle la durée de mariage. Cela peut permettre de dégager des facteurs explicatifs propres à ces trois groupes de générations, à histoires et vécus différents. Mais cela conduit aussi à travailler sur des effectifs moins importants et ainsi à une perte de précision statistique.

Chacune de ces formules a ses avantages et inconvénients. Le choix final dépendra de la taille de l'échantillon dont on dispose, de la précision recherchée, du contexte dans lequel on est (changements socio-démographiques plus ou moins rapides et récents) et aussi bien sûr de ce que l'on cherche.

1.2. Des parités ajustées

On peut aussi contrôler l'âge et la durée de mariage d'une autre façon : en divisant le nombre d'enfants nés-vivants déclaré par chaque femme par sa durée d'exposition au risque de fécondité, on crée *un indicateur standardisé*, l'équivalent au niveau individuel d'un taux.

Si P_i est la parité d'une femme i

n_i le nombre d'années d'exposition au risque

$$P_i = \frac{P_i}{n_i}$$

Pour une femme avec 4 enfants en 8 ans, $P_i = 0,50$; pour une autre avec 4 enfants en 12 ans, $P_i = 0,33$. Cet indicateur peut être utilisé dans des analyses statistiques individuelles ou agrégées (groupes sociaux, groupes de générations...), comme il permet de classer les femmes ou les groupes en sur ou sous-fécondité par rapport à une moyenne.

En ne considérant que les femmes mariées, n peut être la durée écoulée depuis le premier mariage, ou mieux, si l'on veut aussi contrôler les périodes de divorces et de veuvage, le nombre total d'années passées en

union(s)³. Dans les sociétés où assez fréquemment des naissances surviennent avant le mariage, le début de la période d'exposition d'une femme pourrait être fixé comme étant l'année de sa première naissance moins une année⁴.

On peut calculer cet indicateur de « performance » individuelle soit pour les femmes mariées à l'enquête, soit pour les femmes mariées ou l'ayant été. Nous conseillons d'exclure les jeunes femmes (les 15-19 ans) et de le calculer sur les femmes mariées depuis 4 ou 5 ans (ou l'ayant été au moins 4 ou 5 ans). Par ailleurs, sa construction, simple dans son principe, requiert une information correcte sur l'âge au premier mariage (ou la durée écoulée depuis) et/ou la durée de chaque union, ce qui est loin d'être toujours le cas.

Ces parités que nous qualifions d'ajustées (par la durée d'exposition au risque) n'ont pas souvent été utilisées dans les analyses statistiques des déterminants ou des prédicteurs de la fécondité⁵. On les trouve néanmoins dans quelques travaux liés à l'Enquête Mondiale de fécondité des années 1970 et 1980 (voir par exemple R. Little, 1977 ; Th. Pullum, 1987).

Dans le même ordre d'idées, on peut travailler avec *des naissances survenues sur une période ou entre des âges donnés*. Un exemple, déjà ancien, avec l'étude des déterminants socio-économiques de la fécondité légitime au Sri Lanka de R. Little et S. Perera (1981), menée à partir de l'enquête EMF de 1975. Les auteurs ont analysé la fécondité de trois grandes cohortes de femmes définies par leur durée de mariage⁶ : moins de 10 ans de mariage, 10 à 19 ans et 20 ans et plus. Dans chacun de ces groupes, les femmes ont des durées d'exposition différentes. Prenons par exemple celles qui ont entre 10 et 19 ans de mariage : les durées d'exposition sont complètes pour les 10 premières années de mariage, mais pas pour les suivantes (certaines femmes en sont à 11 ans et d'autres à 19 ans de mariage). Pour contrôler ces durées d'exposition et pouvoir ainsi comparer des fécondités individuelles, les auteurs ont construit par femme un indicateur du nombre estimé des naissances sur 10 ans, en d'autres termes la descendance constituée entre 10 et 20 ans de mariage. Pour les cohortes avec 10 à 20 ans de mariage :

Si N (10-19) sont les naissances d'une femme qui est entre 10 et 20 ans de mariage

M (10-19) le nombre de mois vécus par cette femme depuis son dixième anniversaire de mariage

3 Si l'on a bien sûr les données requises pour le calculer (durée des différentes unions, ou dates de début et fin de chaque union).

4 Compte tenu des neuf mois de la grossesse.

5 Sans que nous sachions réellement pourquoi, on leur préfère souvent le nombre de naissances survenues au cours des cinq années précédents l'enquête, qui constitue, il est vrai, un meilleur indicateur du moment.

6 Les auteurs ont préféré à juste titre la durée de mariage à l'âge compte tenu de l'augmentation rapide de l'âge au mariage au Sri Lanka et du recul de la fécondité dès les années 1970.

le nombre estimé de naissances sur 10 ans = $120 \frac{N(10-19)}{M(10-19)}$

On le calcule de la même façon dans les autres cohortes.

On peut utiliser cet indicateur de fécondité dans des analyses individuelles (régression...) comme dans des analyses agrégées (par cohorte de mariages ou groupe socio-économique).

1.3. Une variante sur la fécondité masculine

Unions de diverses natures, forte mobilité conjugale, polygamie élevée, parfois fécondité hors mariage sont les caractéristiques de nombre de sociétés en Afrique sub-saharienne. Rares sont les vraies enquêtes masculines de fécondité, retraçant les vies « reproductives » et matrimoniales des hommes, union par union, femme par femme, enfant par enfant. Ce fut le cas néanmoins dans une enquête menée au Sud-Bénin en 1989 sur 1650 hommes mariés de 20 ans et plus (F. Donadje et D. Tabutin, 1994). La descendance finale des hommes mariés y était de 11,4 enfants, la polygamie touchant à peu près la moitié d'entre eux. Mais comment dans un tel contexte mesurer la fécondité individuelle des hommes ?

Nous aurions pu construire un indicateur du type DRAT contrôlant l'âge et la durée de mariage (nous verrons un peu plus loin ce qu'il est pour les femmes et ce qu'il aurait pu être pour les hommes)⁷. Pour tenir compte de la grande variabilité de l'âge des répondants et de leur histoire, nous devons bien sûr contrôler la durée passée en union. Pour chaque homme, nous avons simplement calculé *un nombre d'enfants par femme-année*, en divisant le nombre total d'enfants nés-vivants qu'un homme a eu par la somme totale des années passées en union avec chacune de ses femmes⁸. Pour éliminer les très jeunes époux, les cas de stérilité et les divorces rapides, il a été calculé sur les hommes ayant passé au moins cinq années en union et ayant au moins un enfant.

Cet indicateur a permis de confirmer la forte fécondité masculine (il est de 0,30 en moyenne générale, soit un enfant tous les trois ans), d'en étudier les disparités sociales ou spatiales et les déterminants individuels, après contrôle donc des durées d'union.

Mais venons-en à une autre approche cherchant à positionner la parité de chaque femme par rapport à un modèle de référence. Nous en étendrons la procédure aux hommes.

⁷ Nous ne l'avons pas fait à l'époque faute d'y avoir pensé.

⁸ Plus précisément du mariage jusqu'à l'âge de 49 ans pour les femmes qui ont à l'enquête plus de 50 ans ou qui sont sorties de l'union après 50 ans.

1.4. Le DRAT (Duration Ratio)

Avec le double souci d'éviter la segmentation des échantillons due à la nécessité de contrôler à la fois l'âge et la durée de mariage et d'améliorer « l'explication », B. Boulier et M. Rosenzweig vont en 1978 proposer une mesure individuelle standardisée de fécondité cumulée. Il s'agit en fait d'un indicateur de niveau relatif de la fécondité, défini comme étant le rapport d'une fécondité observée à une fécondité théorique. Pour chaque femme mariée, on rapporte le nombre d'enfants qu'elle a eu au nombre théorique d'enfants qu'avec son âge et sa durée d'union elle aurait dans un schéma de fécondité naturelle. Le DRAT d'une femme d'âge a est ainsi égal à :

$$\text{DRAT}(a) = \frac{C(a)}{\int_m^a f(a) da}$$

où $C(a)$ est le nombre d'enfants de la femme d'âge a
 $f(a)$ le taux de fécondité naturelle par âge
 m l'âge d'entrée en union de la femme

Il s'agit bien d'une *standardisation indirecte*, à l'image au niveau individuel de l'indice I_g de fécondité légitime que A. Coale avait construit au niveau agrégé⁹. Les DRAT sont des indices de fécondité qui permettent de comparer des femmes (ou des groupes de femmes) qui ont été exposées au risque de procréer pendant des durées différentes. On conseille de les calculer pour les femmes mariées depuis au moins 4 ou 5 ans.

Le DRAT contrôle l'effet biologique de l'âge sur la fécondité, mais dans une certaine mesure seulement, car on fait l'hypothèse que la fécondité théorique dépend uniquement de l'âge, excluant donc le rôle de l'âge au mariage sur la fécondité, souvent démontré dans les réalités contemporaines ou historiques.

Autre hypothèse en quelque sorte dans ce modèle : les fécondités observées et théoriques ont le même calendrier, même si leurs niveaux diffèrent. Cela peut conduire à des distorsions. *Le choix de la fécondité de référence est donc d'importance*. On utilise couramment la série des taux de fécondité naturelle calculés par A. Coale et al. (1975)¹⁰. Mais comme B. Boulier et M. Rosenzweig eux-mêmes le soulignaient déjà, on peut choisir

9 Défini, rappelons-le, comme étant dans une population le nombre de naissances observées au nombre de naissances que l'on aurait si les femmes de cette population suivaient le schéma de fécondité naturelle des Huttérites, une des fécondités naturelles légitimes les plus élevées dans l'histoire (indice synthétique de 12,4 enfants par femme mariée).

10 Une série-type calculée à partir de treize populations à fécondité naturelle pour la plupart européennes, conduisant à un indice synthétique légitime de 11,8 enfants par femme mariée. Publiée par groupes d'âges quinquennaux, cette série doit être décomposée par année d'âge (par interpolation par exemple) pour le calcul des DRAT.

d'autres modèles de référence, «collant» mieux à la population étudiée. Plutôt qu'un modèle théorique ou bâti sur des données anciennes de l'Europe, rien n'interdit de se construire sa propre fécondité de référence, reposant sur les réalités du pays ou de la région étudiée. C'est par exemple ce qu'a fait A. Ilinigumugabo (1989) dans son étude multivariée des déterminants de la fécondité au Rwanda ; les deux standards (le sien et celui de A. Coale et al.) aboutissent aux mêmes relations, mais le «standard rwandais » qu'il s'est construit à partir des données de fécondité du Rwanda double la variance expliquée. Dans un autre ordre d'idées, on peut très bien prendre aussi la fécondité moyenne par âge du pays comme référence : les femmes (ou les groupes) en sous-fécondité auront un DRAT plus ou moins inférieur à 1, celles en surfécondité seront plus ou moins au-delà de 1.

Cela dit, un indice comme le DRAT réduit l'hétéroscédasticité, un des problèmes classiques que l'on a avec les régressions du nombre d'enfants en fonction de l'âge, le fait autrement dit que la variance de la variable dépendante est fonction d'une variable explicative et qu'elle n'est pas constante. C'est bien le cas avec le nombre d'enfants dont la variance augmente sensiblement avec l'âge des femmes.

Sans faire fureur et supplanter le nombre d'enfants par femme, ces DRAT ont néanmoins été utilisés dans un certain nombre d'analyses des déterminants de la fécondité des pays du Sud. Par exemple, R. Lesthaeghe et al. (1985) les ont pris comme variable dépendante dans leur étude des effets individuels et contextuels de l'éducation sur les variables intermédiaires de la fécondité au Kenya. De même plus récemment, B. Schoumaker et D. Tabutin (1999) dans des analyses multi-niveaux de la fécondité rurale au Maroc.

1.5. Un DRAT masculin ?

Dans le même ordre d'idées pour un contexte polygamique ou de forte mobilité conjugale, on pourrait très bien comparer le nombre total d'enfants qu'un homme a eu avec ses différentes femmes (successives ou simultanées) au nombre théorique qu'il aurait eu si chacune de ses femmes avait suivi un schéma de fécondité naturelle (ou théorique) de référence durant le temps passé avec cet homme.

Si $C_1, C_2, C_3 \dots$ sont les nombres d'enfants nés-vivants qu'un homme a eu avec ses femmes 1, 2, 3, ..., n
 $m_1, m_2, m_3 \dots$ l'âge d'entrée en union de chaque femme
 $a_1, a_2, a_3 \dots$ l'âge à l'enquête de chaque femme (toujours en union) ou l'âge qu'elle avait à la fin d'union en cas de rupture de l'union

$$\begin{aligned} \text{DRAT masculin} &= \frac{C_1 + C_2 + C_3 + \dots}{\int_{m_1}^{a_1} f(a)da + \int_{m_2}^{a_2} f(a)da + \int_{m_3}^{a_3} f(a)da + \dots} \\ &= \frac{\sum_1^n C_n}{\sum_1^n \int_{m_n}^{a_n} f(a)da} \end{aligned}$$

On doit faire les mêmes hypothèses que pour le modèle féminin. On supposerait aussi - hypothèse peut-être forte - que les comportements de l'homme (sa sexualité par exemple) sont indépendants de son âge (il est jeune à la première union, nettement plus vieux à la nième) et de l'âge des différentes épouses en cas de polygamie (pas de comportements discriminatoires extrêmes). Nous conseillerions de le calculer pour les hommes de 30 à 55 ans.

Quand deux conjoints sont toujours ensemble dans leur première union, le DRAT de l'homme serait équivalent à celui de sa femme.

1.6. Des mesures individuelles de la fécondité récente

Si l'on veut véritablement comprendre et expliquer ce qui se passe actuellement, on ne peut se contenter de travailler avec des indices de fécondité cumulée - comme ceux que l'on vient d'examiner - qui reposent sur un « stock d'enfants » constitué sur de nombreuses années (G. Farooq, 1985). Dans l'étude des changements de la fécondité, on doit essayer de mesurer des comportements récents. D'autant plus que, dans la plupart des cas, les variables explicatives dont on dispose (éducation, résidence, niveau de vie...) concernent la période de l'enquête.

L'approche générale revient à utiliser *le nombre de naissances déclarées par les femmes (mariées, ou mariées et l'ayant été) au cours d'une période récente*. On pourrait prendre les naissances des 12 derniers mois, une information d'ailleurs souvent disponible dans les recensements, mais parfois de qualité douteuse (mauvaise perception de la période de référence, non-déclaration de ces naissances récentes décédées) ; dans les enquêtes rétrospectives, cela conduirait aussi à des nombres assez faibles d'événements et donc à de grandes erreurs de sondage. Mieux vaut élargir la période de référence.

Le plus courant est dès lors d'utiliser *le nombre de naissances survenues dans les cinq années précédant l'enquête*¹¹, une sorte de compromis entre le désir d'étudier les changements actuels et le besoin de minimiser les erreurs de déclaration ou de sondage. Cinq années peut même permettre d'étudier les intervalles fermés entre naissances pour les femmes qui auront eu plus d'un enfant dans la période.

On devra là aussi contrôler ce nombre de naissances par la durée d'exposition au mariage dans cette période : de 5 ans pour un grand nombre de femmes, de moins pour d'autres si elles se sont mariées dans l'intervalle ou ont divorcé. Le rapport du nombre de naissances dans l'intervalle (N) à la durée de mariage (d) conduit à un indicateur :

$$f = \frac{N}{d}.$$

Il est au niveau individuel l'équivalent d'un taux de fécondité légitime. C'est ce type d'indicateur qui sera utilisé comme variable dépendante dans nombre d'études des déterminants individuels de la fécondité dans les pays du Sud.

2. Indices de mortalité des enfants par femme

Quelques mots sur les principes et les données requises avant une brève présentation des indices possibles.

2.1. Principes et données

L'étude de la mortalité en démographie est restée longtemps très descriptive, basée sur les outils classiques (taux ou quotients) ou les approches indirectes (les fameuses questions Brass). On se contentait de quelques approches différentielles agrégées en termes de niveau d'instruction, groupe social, milieu d'habitat de la mère ou des parents... C'était et c'est toujours utile, mais tout à fait insuffisant pour des approches plus explicatives, en termes par exemple de risques individuels ou de population à risques. Dès lors, comme pour la fécondité mais avec un peu de retard, se développeront dans les années 1980 les approches dites individuelles de la mortalité des enfants, qui, en simplifiant, consistent à mesurer pour chaque femme un indice de sur- ou de sous-mortalité de sa descendance par rapport à une moyenne ou un standard. Comme toute nouvelle méthodologie, cette appro-

¹¹ Ou parfois aussi le temps écoulé depuis le premier mariage quand celui-ci a moins de cinq ans.

che a ses avantages et ses inconvénients. Se multiplieront aussi, mais nous ne les aborderons point ici, les analyses « explicatives » partant de la survie ou non des enfants nés dans une période récente, analyses qui utilisent les techniques de la statistique multivariée.

Ces approches, en termes d'indices par femme, développés par J. Trussell et S. Preston au début des années 1980, puis reprises notamment par A. Noubissi dans les années 1990, répondent à un double souci :

- a) améliorer l'analyse différentielle et la recherche des déterminants de la mortalité aux jeunes âges en intégrant l'hétérogénéité individuelle dans l'étude des différences entre groupes. A un âge donné, tous les enfants d'un même groupe ne sont pas exposés à des risques identiques de décéder¹². On récuse en quelque sorte l'hypothèse d'homogénéité du groupe propre à l'analyse agrégée, basée sur des moyennes censées représenter les risques de chaque individu ;
- b) utiliser de façon optimale la masse d'informations individuelles sur les mères et leurs enfants incluses dans les gros échantillons des enquêtes de type EMF ou EDS¹³. En définitive, il s'agit d'aller au-delà des analyses classiques des déterminants de la mortalité des enfants.

Le principe de construction de ces indices de mortalité des enfants¹⁴ par femme (IME) est simple : on cherche à mesurer « l'expérience relative » de chaque mère (i) en matière de mortalité infanto-juvénile en rapportant le nombre de décès que chacune a connu (Di) au nombre estimé qu'elle aurait eu (Ei) si la mortalité était celle d'un standard (théorique ou empirique) :

$$\text{IME (i)} = \frac{D_i}{E_i}$$

Tout repose sur l'idée que la proportion d'enfants décédés d'une femme (ou d'un groupe) peut être une mesure de mortalité des enfants après bien sûr ajustement de sa durée de mariage ou de son âge¹⁵. Une idée déjà ancienne¹⁶ à la base de toutes les techniques indirectes d'estimation (au niveau agrégé) de la mortalité des enfants à partir des proportions d'enfants décédés. Les différentes variantes de l'approche relèvent de la méthode retenue d'estimation des décès théoriques.

12 Tout enfant naît avec un « capital santé » personnel qui se modifie plus ou moins rapidement au cours de sa vie en fonction des agressions extérieures qu'il subit, des comportements sanitaires ou alimentaires de ses parents...

13 Echantillons souvent autour de 5 000 femmes et de 15 à 20 000 naissances vivantes.

14 Une excellente présentation, critique et illustrée, en est faite dans l'ouvrage d'A. Noubissi sur les *Méthodologies d'analyse de la mortalité des enfants* (1996).

15 Un ajustement nécessaire dans la mesure où plus les femmes sont âgées ou ont une durée longue de mariage, plus leurs enfants sont âgés et ont une durée croissante d'exposition au risque de mortalité. Plus leur proportion d'enfants décédés sera élevée.

16 Développée par W. Brass dès 1964, reprise par lui-même et par nombre d'auteurs dans les années 1970 et 1980.

2.2. Les indices de Trussell et Preston

Ces deux auteurs (1982, 1984) vont développer trois versions de leur modèle qui diffèrent dans la façon d'estimer les décès théoriques et dans les données requises¹⁷. Celles-ci pour une femme vont des simples nombres d'enfants nés-vivants et décédés (questions du type Brass) dans une version à la date de naissance de chaque enfant né-vivant dans une autre. Nous nous en tiendrons ici à la méthode la plus simple et la plus courante qui utilise le nombre total d'enfants (nés-vivants et décédés) - une information simple et disponible dans toutes les enquêtes et dans de nombreux recensements des pays du Sud - et l'âge ou la durée de mariage de chaque femme.

L'indice pour une femme i d'âge j (ou de durée de mariage j) sera égal au nombre de ses enfants décédés (D_i) divisé par un nombre attendu d'enfants décédés (E_i) :

$$\text{IME}_{ij} = \frac{D_i}{E_i}$$

Le nombre de décès observés pour une femme est le produit de ses naissances (N_i) par la proportion réelle d'enfants décédés PD_j ; le nombre attendu d'enfants décédés sera le produit de ses naissances par une proportion attendue (on pourrait dire théorique) de décès PD'_j :

$$\text{soit } \text{IME}_{ij} = \frac{D_i}{E_i} = \frac{N_i \cdot PD_j}{N_i \cdot PD'_j} = \frac{PD_j}{PD'_j}$$

L'estimation de la proportion attendue se fait à partir de la technique de Brass qui, rappelons-le, permet d'estimer la probabilité de décès des enfants (de la naissance à un âge a) comme étant le produit de la proportion – $d(j)$ – de décès d'enfants chez les mères d'un âge donné j par un coefficient multiplicateur $k(j)$, reflétant la plus ou moins grande précocité de la fécondité :

$$q(a) = d(j) \cdot k(j) \quad \text{ou} \quad d(j) = \frac{q(a)}{k(j)}$$

Ces coefficients $k(j)$, qui sont fonction des schémas de fécondité, sont à calculer à partir des parités déclarées selon l'âge (15-19, 20-24, 25-29 ans...) ou selon la durée de mariage (0-4, 5-9, 10-14 ans...). Plutôt que la technique originelle de Brass (qui n'envisage que l'âge), on utilisera souvent

¹⁷ Voir les auteurs ou A. Noubissi (1986, pp. 91-114) pour leur présentation et discussion critique.

la variante de Trussell (1980 ; Nations Unies, 1984, pp. 76-86) basée sur les tables types de mortalité de Coale et Demeny¹⁸.

Pour l'estimation de la proportion attendue d'enfants décédés PD', il suffit donc de se choisir une mortalité standard, $q_s(a)$, et de la convertir en proportions d'enfants décédés :

$$PD'_j = \frac{q_s(a)}{k(j)}$$

La mortalité standard sera choisie comme étant celle qui représente le mieux la situation réelle. Chez Trussell (1980), on a le choix entre les familles Nord, Sud, Est et Ouest des tables de Coale et Demeny ; chez Palloni-Heligman (1986), entre cinq modèles des tables des Nations Unies¹⁹. On veillera aussi à aboutir à ce que pour l'ensemble de la population le nombre de décès actuels soient à peu près égaux aux décès attendus.

En définitive, l'indice de mortalité des enfants pour une femme i d'âge j se calcule comme suit :

$$IME_{ij} = \frac{D_i}{N_i \cdot q_s(a) / k(j)}$$

C'est un indice standardisé de mortalité puisqu'il contrôle l'âge ou la durée de mariage des mères. Comme le DRAT, on peut l'utiliser dans des analyses individuelles ou agrégées (où il représentera alors la moyenne d'un groupe). Il est relativement simple à mettre en œuvre.

Cela dit, rappelons que cet indice ne permet de mesurer que les expériences relatives des mères en matière de mortalité, et en aucune façon des niveaux de mortalité. De même, il ne se rapporte à aucune période, ni à aucun type bien précis de mortalité : plus les femmes sont âgées, plus la période en moyenne est ancienne, plus aussi il y a mélange de types de mortalité (infantile, juvénile, adolescente, parfois même adulte). Il y a aussi bien sûr le problème de choix du standard, un problème constant dans les procédures de standardisation indirecte, sur lequel nous reviendrons en présentant l'indice proposé par A. Noubissi.

Cet indice de Trussell et Preston a été utilisé – entre autres – dans une vaste étude des Nations Unies (1985) sur les différentiels socio-économiques de la mortalité des enfants dans 15 pays en développement²⁰. Basée sur les données d'enquêtes EMF ou d'échantillons de recensements (ayant

18 Mais rien à notre avis n'empêche d'utiliser une autre version de la technique, comme celle de A. Palloni et L. Heligman (1986) basée sur les tables types de mortalité des Nations Unies pour les pays en développement.

19 Amérique Latine, Chili, Asie du Sud, Extrême Orient et général (la moyenne des quatre autres).

20 Sept d'Afrique, cinq d'Asie et trois d'Amérique Latine.

inclus les questions Brass), elle prendra l'IME comme variable à expliquer et, selon les pays, de 10 à 17 variables explicatives (dont certaines contextuelles). Pour la douzaine de variables privilégiées²¹ que sont l'instruction de la mère et du père, le milieu et la région de résidence, l'ethnie, la religion, le revenu, la structure familiale..., l'analyse sera systématiquement menée en univarié, bivarié et multivarié, combinant donc analyses agrégées et analyses individuelles. Cette étude pionnière de par sa méthodologie démontrera tout l'intérêt de l'utilisation d'un indice comme l'IME de Trussell et Preston pour « l'explication » de la mortalité des enfants.

2.3. L'indice de Noubissi

Nous l'avons dit, A. Noubissi reprendra au début des années 1990 la problématique générale de la mesure par femme de la mortalité des enfants avec le souci de mieux comprendre les inégalités sociales en intégrant le problème de l'hétérogénéité individuelle.

L'auteur conserve l'approche et les hypothèses générales de Trussell et Preston, cherchant donc aussi à mesurer l'expérience relative des femmes en rapportant des décès observés à des décès de référence. Mais il propose une simplification du calcul de l'indice en prenant au dénominateur non pas les décès tirés d'un standard de mortalité, mais les décès observés dans l'ensemble de la population.

Pour une femme i d'âge j (ou de durée de mariage j) :

$$\text{IMF}_{ij} = \frac{D_{ij}}{E_{ij}} = \frac{D_{ij}}{N_{ij} \cdot \text{PD}_{tj}} = \frac{\text{PD}_{ij}}{\text{PD}_{tj}}$$

avec N_{ij} = les naissances de la femme i d'âge j

PD_{ij} = la proportion d'enfants décédés chez la femme i

PD_{tj} = la proportion d'enfants décédés chez l'ensemble des femmes d'âge j .

La formulation reste donc la même, mais la procédure de calcul est simplifiée en rapportant la mortalité observée chez une femme à celle de la moyenne du groupe de femmes de même âge (ou de même durée de mariage). Chacune sera ici en sur ou sous-mortalité par rapport à ses congénères du même âge. Cela résout le problème éventuel du standard. En fait, A. Noubissi (1996, pp. 98-102) a clairement démontré que les trois types d'indices proposés par Trussell et Preston (nous n'en n'avons présenté qu'un) ainsi que le sien n'étaient pas sensibles au choix du standard. En d'autres termes, les divers standards possibles conduisent à peu de choses près aux

21 Qui feront chacune l'objet d'un chapitre dans l'ouvrage.

mêmes résultats. Dès lors, il nous semble logique de recommander le plus simple, celui de Noubissi, en le calculant pour les femmes de 20 à 45 ans ou de 25 à 40 ans²².

A. Noubissi (1996) appliquera ces diverses techniques sur les données des deux enquêtes de fécondité du Cameroun de 1978 et de 1991. Il fera de l'analyse univariée et de l'analyse multivariée, en multipliant les approches et les tests. Il dégagera le rôle des déterminants socio-économiques et culturels, plus fort que celui des déterminants bio-démographiques propres à chaque enfant, l'importance de l'hétérogénéité individuelle et les caractéristiques des femmes à hauts risques que l'on retrouve dans chaque groupe social.

3. Quelques possibilités en matière de nuptialité

Dans les enquêtes des pays du Sud, la nuptialité est en général beaucoup moins bien documentée que la fécondité ou la mortalité des enfants. Les enquêtes EMF des années 1970 et 1980 envisageaient cependant divers éléments de la nuptialité de chaque femme, avec dans le questionnaire-type huit questions sur l'histoire du mariage actuel et quatre sur chaque mariage antérieur éventuel. On pouvait ainsi reconstituer l'histoire matrimoniale précise de chaque femme (durée des unions, causes des ruptures...). Dans les EDS en revanche, on ne trouve que quelques données (date de la première union, nombre de co-épouses), rien d'autre permettant de reconstituer l'histoire de la nuptialité²³.

La stabilité des unions, le nombre de mariage ou encore les durées totales de vie conjugale (ou de rupture), permettant de bien mesurer les durées d'exposition, sont des déterminants cruciaux de la fécondité d'une femme à un âge donné. Toutes choses égales par ailleurs, plus la vie matrimoniale d'une femme est stable, plus forte sera sa fécondité. Par ailleurs, le degré de mobilité conjugale peut être considéré aussi comme une variable psycho-sociale d'identité de vie personnelle.

Donnons simplement quelques exemples d'indicateurs de nuptialité que l'on peut créer à partir du moment où l'on dispose des durées passées dans chaque union.

22 Nous recommandons d'exclure les femmes les plus jeunes et les plus âgées en raison notamment de la surmortalité des enfants propre aux âges extrêmes. On gagnera en qualité, mais on perdra en précision puisque la taille de l'échantillon diminue.

23 Une erreur à notre sens, et maintes fois dénoncée, des concepteurs de ces enquêtes, qui n'a point été corrigée dans les phases récentes de ce programme d'enquêtes.

3.1. Des exemples pour les femmes

Un indice de stabilité conjugale, mesurant la durée totale passée en union(s) par une femme par rapport à ce que cette durée aurait pu être compte tenu de son âge au mariage, peut être calculée en divisant le nombre total d'années passées en union par la différence entre son âge actuel et son âge au premier mariage, autrement dit la durée écoulée depuis son premier mariage. Pour une femme :

$$ISC = \frac{D_1 + D_2 + D_3 + \dots}{x_a - x_m} = \frac{\sum_{m=1}^n D_m}{x_a - x_m}$$

D_m étant la durée de chaque union, de la première à la nième

x_a l'âge actuel de la femme

x_m son âge au premier mariage

Cet indice mesure pour une femme d'un âge donné la proportion de temps passé en union depuis son premier mariage. Pour une femme sans rupture, l'indice est égal à 1 ; il est sinon inférieur à 1, et d'autant plus que la (ou les) durées des unions ont été courtes. Il serait à calculer pour les femmes mariées (ou l'ayant été) pendant quelques années, disons cinq ans.

Inversement, on peut imaginer *un indice d'instabilité conjugale*, en rapportant l'ensemble des durées passées hors mariage au nombre d'années écoulées depuis le premier mariage :

$$IIC = \frac{H_1 + H_2 + H_3 + \dots}{x_a - x_m} = \frac{\sum_{m=1}^n H_m}{x_a - x_m} = 1 - ISC$$

H_m étant chacune des durées de veuvage ou de divorce, de la première à la nième

x_a et x_m les âges actuels et au premier mariage.

Cet indice mesure pour une femme mariée (ou l'ayant été) la proportion de temps qu'elle a passée hors mariage, en d'autres termes en l'état de veuve ou divorcée. Il va de 0 (si la première union est toujours là) à près de 1 dans le cas d'un (ou de plusieurs) mariage(s) rapidement rompu(s).

Dans le même ordre d'idées, on pourrait *calculer un indice individuel relatif de stabilité conjugale* en rapportant l'indice d'une femme d'un âge donné (a) à l'indice moyen des femmes du même âge a (moyenne des divers indices individuels) :

$$\text{IRSC}_{aj} = \frac{\text{ISC}_{aj}}{\text{ISC}_{aT}}$$

ISC_{aT} étant l'indice moyen de l'ensemble des femmes d'âge a.

Une femme a une mobilité moins forte que la moyenne si son indice est supérieur à 1, plus élevée dans le cas contraire. On peut aussi regrouper les valeurs de l'indice en quelques grandes classes (< 0,70, 0,90-0,85, 0,85-1,95, 0,95-1,05, 1,05-1,15, > 1,15 par exemple) et attribuer à chaque femme une modalité correspondante. Chacune est alors classée dans un type de mobilité relative (du genre : très forte, forte, moyenne, faible).

3.2. Un exemple masculin sur l'Afrique

Dans les sociétés monogames, on peut utiliser pour les hommes des indices voisins de ceux que l'on vient de présenter pour les femmes. En revanche, le problème sera différent dans les sociétés polygamiques. Dans ces systèmes matrimoniaux qui combinent souvent union (plus ou moins stable), mariage monogame et mariage polygame, le cycle de vie matrimoniale d'un homme peut être extrêmement diversifié : de la monogamie à la bigamie, de la bigamie à la « trigamie », ou encore retour à la monogamie après rupture d'un des deux mariages... Tout cela avec des durées et causes de ruptures entre unions ou états tout à fait variables selon son statut social ou son âge.

Dans l'étude sur la fécondité masculine au Bénin dont nous avons déjà parlé, nous avons cherché à étudier les différents modèles matrimoniaux des béninois et à mesurer l'impact de ces modèles sur la fécondité des hommes (F. Donadje, 1992 ; F. Donadje et D. Tabutin, 1994). Disposant des histoires fécondes et matrimoniales précises de chaque homme (dates, durées et causes de rupture éventuelle ; fécondité dans chacun des mariages), nous avons finalement retenu²⁴ quatre grands types de profil matrimonial combinant stabilité et type de mariage :

- le monogame stable : l'homme qui n'a connu qu'une union dans sa vie, au moins jusqu'à l'enquête
- le monogame instable : celui qui a eu plusieurs épouses, mais toujours successivement
- le polygame stable : celui qui, après un premier mariage, a pris une seconde puis éventuellement une troisième ou quatrième femme, mais sans jamais revenir à la monogamie

24 Après divers autres essais plus détaillés, mais qui se heurtaient surtout à un problème d'effectifs dans la mesure où par la suite nous devions dans les analyses contrôler les groupes de générations (ou d'âges).

- la polygame instable : l'homme à la vie la plus «mouvementée », qui passe de la monogamie à la polygamie, pour revenir à la monogamie avant éventuellement de reprendre une autre femme.

Ce type d'indicateur individuel, aussi simple soit-il²⁵ et à objectif différent des indices précédemment examinés, a permis de préciser pour les béninois les relations entre leur nuptialité et leur fécondité, mesurée par la descendance finale après 55 ans, ou mieux le nombre d'enfants par femme-année (voir section 1) qui contrôle la durée totale du (ou des) mariage(s). Avec ces deux indices individuels de nuptialité et de fécondité, les monogames apparaissent jusqu'à 50 ans comme un peu plus féconds que les polygames (0,35 enfant par femme-année contre 0,29)²⁶, tandis que chez les uns comme chez les autres, le degré de stabilité de leur(s) union(s) n'a pas d'impact significatif sur la fécondité. On voit là le rôle du remariage, intense et rapide, des veufs et divorcés.

4. Intérêts et limites

Nous ne reviendrons pas longuement sur l'intérêt de ces indices individuels, qui permettent d'aller bien au-delà - d'une certaine façon - de l'analyse classique agrégée, de contrôler dès le départ les grandes variables de confusion incontournables en démographie comme l'âge des individus, la durée de mariage ou même la fécondité, de synthétiser parfois divers éléments d'information, de mieux préciser les facteurs de risque ou les facteurs explicatifs qui relèvent de l'individu.

Mais rappelons-en aussi brièvement quelques limites ou problèmes intrinsèques.

- *le risque de sélection* : par définition, la plupart de ces indices sont construits à partir des déclarations faites par des femmes (ou des hommes) survivantes au décès ou à la migration. On a toujours un risque inhérent de sélection, de non-représentativité parfaite, dans les contextes notamment de forte mortalité et de grande mobilité interrégionale ou internationale.
- *la nature rétrospective des données* : toutes les informations sur la fécondité et la mortalité des enfants relèvent par définition du passé, mais d'un passé plus ou moins lointain selon le type de questions et selon l'âge des individus. Il y aura le plus souvent décalage chronologique entre ces indices (à expliquer) et les variables explicatives qui, dans la plupart des

²⁵ Mais qui requiert des données adéquates et de bonne qualité.

²⁶ Ce n'est plus le cas après 50 ans, à ces âges où les femmes des monogames n'ont plus guère de fécondité, tandis que les polygames – notamment instables – continuent à avoir des naissances avec leur(s) plus jeune(s) femme(s).

- cas, concernent la période de l'enquête. Un problème insoluble en l'absence de biographies complètes en matière sociale et économique.
- *la qualité de l'information* : elle est liée au caractère rétrospectif des données. Tant pour la fécondité que pour la mortalité des enfants ou la nuptialité, on fait appel à la mémoire des individus. Le démographe est toujours exigeant, avec nombre de questions sur les âges, les dates ou les durées. Dès lors, on n'est jamais à l'abri d'omissions, de confusions, de mauvaises déclarations. Et un indice calculé au niveau individuel est toujours plus sensible à ces problèmes qu'un indice agrégé ou qu'une moyenne, sur lesquels des effets de compensation peuvent jouer.
 - *les fluctuations aléatoires* : le faible nombre parfois de naissances et de décès d'enfants par femme (et même par groupe de femmes) peut conduire à une grande variabilité des indices individuels. Pour contrôler ce problème, on peut éliminer les femmes « exceptionnelles » (très jeunes, peu de durée de mariage stable, peu fécondes,...) ; dans des analyses différentielles entre groupes de femmes, on calculera les intervalles de confiance statistique. Une chose certaine : ce type d'indice n'est pas à utiliser dans des contextes de faibles fécondité et mortalité.
 - enfin, *le choix des standards ou des références* : un problème constant dès que l'on procède à des standardisations indirectes. Nous en avons déjà parlé.

En définitive, ces approches analytiques reposant sur des indices individuels sont d'un apport incontestable pour la richesse de « l'explication » au niveau individuel. Cependant elles ne sont pas non plus sans problèmes divers, dont la nature peut varier d'un système d'informations à un autre.

Bibliographie

- BOULIER B. et ROSENZWEIG M., 1978, «Age, biological factors and socioeconomic determinants of fertility : a new measure of cumulative fertility for use in the empirical analysis of family size », *Demography*, vol. 15, n° 4, pp. 487-497.
- COALE A., HILL A.G. et TRUSSELL J., 1975, « A new method of estimating standard fertility measures from incomplete data », *Population Index*, vol. 41, n° 1, pp. 203-213.
- DONADJE F., 1992, *Nuptialité et fécondité des hommes au Sud-Bénin. Pour une approche des stratégies de reproduction au Bénin*, Academia, Louvain-la-Neuve, 222 p.
- et TABUTIN D., 1994, «Male nuptiality and fertility in Southern Benin », in Th. Locoh et V. Hertrich (eds), *The Onset of Fertility Transition in Sub-Saharan Africa*, Derouaux, Ordina Editions, Liège, pp. 135-162.

- FAROOQ G., 1985, « The definition of fertility : measurement effects », in G. Farooq et G. Simmons (eds), *Fertility in Developing Countries*, MacMillan, pp. 125-148.
- ILINIGUMUGABO A., 1989, *L'espacement des naissances au Rwanda : niveaux, causes et conséquences*, Ciaco Editeur, Institut de Démographie, Louvain-la-Neuve, 243 p.
- LESTHAEGHE R. et al., 1985, « Individual and contextual effects of education on proximate fertility determinants and on life-time fertility in Kenya », in J.B. Casterline (ed), *The Collection and Analysis of Community Data*, International Statistical Institute, pp. 31-59.
- LITTLE R.J.A., 1977, « Regression models for differentials in fertility », *WFS Technical Paper*, 635.
- et PERERA S., 1981, « Illustrative analysis : socio-economic differentials in cumulative fertility in Sri Lanka. A marriage cohort approach », *WFS, Scientific Reports*, n° 12, 77 p.
- Nations Unies, 1984, *Manuel X. Techniques indirectes d'estimation démographique*, Etudes Démographiques, n° 81, New York, 324 p.
- NOUMBISSI A., 1996, *Méthodologies d'analyse de la mortalité des enfants. Applications au Cameroun*, L'Harmattan/Academia, Paris / Louvain-la-Neuve, 304 p.
- PALLONI A. et HELIGMAN L., 1986, « Re-estimation of structural parameters to obtain estimates of mortality in developing countries », *Population Bulletin of the United Nations*, n° 18.
- PULLUM Th., 1987, « Analytical methodology », in J. Cleland and C. Scott (eds), *The World Fertility Survey. An Assessment*, Oxford University Press, pp. 644-676.
- SCHOUMAKER B., 2000 (à paraître), « Déterminants de la fécondité et contexte local au Maroc rural : une application des modèles multi-niveaux », in F. Gendreau (ed.), *Les transitions démographiques dans les pays du Sud*, AUPELF-UREF/Estern, Paris, 29 p.
- et TABUTIN D., 1999, « Analyse multi-niveaux des déterminants de la fécondité : problématique, modèles et applications au Maroc rural », in *La population africaine au 21^e siècle*, Actes de la troisième conférence africaine de population, Durban, UEPA/NPU, vol. 1, pp. 299-331.
- TRUSSELL T. et HILL K., 1980, « Fertility and mortality estimation from the Panama retrospective demographic survey », *Population Studies*, vol. 34, n° 3, pp. 551-563.
- TRUSSELL J. et PRESTON S., 1982, « Estimating the covariates of childhood mortality from retrospective reports of mothers », *Health Policy and Education*, vol. 3, n° 1.

-----, 1984, « Estimating the covariates of childhood mortality from retrospective reports of mothers », in J. Vallin, J. Pollard et L. Heligman (eds), *Methodologies for the Collection and Analysis of Mortality Data*, Ordina Editions, Liège, pp. 331-364.

United Nations, 1985, *Socio-economic Differentials in Child Mortality in Developing Countries*, New York, ST/ESA/SER.A/97, 319 p.