

Application des modèles hiérarchiques en démographie**Farida Laoudj Chekraoui****Université de Jijel****Claire Kersuzan,****Université de Bordeaux, France**

Abstract. Following to the other social sciences, demography is increasingly using hierarchical models. Indeed, the standard regression models, called "Uni-level", assume that the individual observations are independent of each other conditionally to the variables X_i introduced in the model. This hypothesis may lead to erroneous statistical inferences in the presence of a dependence of observations in which the resorts remain unobserved through the variables collected in the survey (Goldstein, 1995; Sjniders and Bosker, 1999). Unobserved heterogeneity can be linked to the sample design of a survey (in particular clusters) or to the sharing of contextual, spatial and / or family influences by members of the same group. If unobserved heterogeneity in modeling is to be taken into account in order to obtain an unbiased estimate of the parameters of a model, it may also be of interest in itself, by offering the possibility of decomposing the total variability of a phenomenon between a compositional effect (level 1 people's profile) and a context effect (residual variance between level 2 units after compositional effects were controlled). Using two different application examples, the purpose of this paper is to demonstrate the contribution of hierarchical methods in demography:

- Either to correct the estimation of the parameters of a model of the dependence problem of observations and thus limit the risk of erroneous inference (applied here to the case of the influence of early parental death on the trajectory of children in Burundi);
- Or to measure the share of contextual effects and composition effects in the total variability of a phenomenon (applied here to cantonal mortality in metropolitan France from the 1980s).

Through these two examples, we are interested in two methods where the dependence of the observations in the groups is an endogenous element to the model (the correlation between the observations is explicit in the model and influences the estimates of the parameters through the introduction of a new factor in the series of linear predictors): multilevel models (or random effects model, conditional or mixed model) and fixed-effect models. Moreover, and taking into consideration the nature of the variables modeled: binary in the study of the influence of early parental death on the schooling of children (never having attended school) and quantitative in the case of the analysis of the territorial variability of mortality (mortality rate), we apply logistic models in the first case and linear models in the second case.

Keywords. Hierarchical models, multilevel models, fixed effects models, demographics

1 Introduction

De plus en plus utilisés en démographie, les modèles hiérarchiques permettent de répondre aux problèmes désormais classiques que posent l'analyse des observations qui peuvent être regroupées au sein d'unités plus larges et partagent ainsi caractéristiques communes inobservées [1-2]. Appliqués à une série de problématiques socio-démographiques, ces modèles sont généralement mobilisés dans deux directions distinctes : 1) l'estimation sans biais de l'impact propre d'une variable d'intérêt sur un phénomène [3-5] 2) la mesure de la variabilité totale d'un phénomène imputable à des effets contextuels [6-8]. A partir de deux exemples différents, l'objectif de cet article est d'illustrer la portée des modèles hiérarchiques pour répondre à ces deux types de questionnement. Parmi l'ensemble des méthodes pour données corrélées aujourd'hui disponibles, deux modèles sont ici considérés : les modèles multiniveaux et les modèles à effets fixes. Après avoir montré l'intérêt de recourir à une modélisation hiérarchique pour dégager l'influence propre de l'expérience précoce du décès parental sur la scolarisation des enfants au Burundi, nous nous intéressons à un cas où les effets de contexte sont considérés non plus comme une nuisance mais comme ayant un intérêt en soi. Cette perspective est celle adoptée pour mesurer la part des effets contextuels et des effets de

APPLICATION DES MODELES HIERARCHIQUES EN DEMOGRAPHIE

composition dans la variabilité totale de la mortalité cantonale en France métropolitaine en 2005-2007.

2 L'apport des modèles hiérarchiques pour la recherche « causale » : l'exemple de la mesure du lien entre l'expérience précoce du décès parental et la scolarisation des enfants au Burundi.

L'objectif initial de cette recherche est d'élargir les travaux portant sur les conséquences de l'expérience précoce du décès parental sur la scolarisation des enfants dans les pays du SUD à un contexte où la principale cause de décès chez l'adulte n'est pas le VIH/SIDA mais une grave crise socio-politique [9-11]. Enclavé dans l'est de l'Afrique centrale, le Burundi est l'un des épicentres du conflit traversé dans les années 1990 par la région des Grands Lacs. La « crise » burundaise, caractérisée par des massacres de masse, a entraîné d'importantes pertes humaines et a augmenté le risque pour les enfants de connaître tôt le décès d'au moins un parent.

Données

Les données utilisées sont issues d'une enquête rétrospective réalisée au Burundi en 2002 avec la collaboration de l'Institut d'Etudes Démographiques de l'Université Montesquieu Bordeaux IV : l'Enquête socio-démographique et de santé de la reproduction (ESDSR). L'échantillon de l'ESDSR est composé de 7 119 ménages (32 085 personnes ont été enquêtées). Le plan de sondage de l'enquête est complexe (population répartie en différentes strates puis sélection de certains groupes d'individus à travers le tirage de grappe dans chaque strate) [12]. Contrairement à la majorité des enquêtes rétrospectives disponibles pour les pays du SUD, l'ESDSR fournit la date de décès éventuel des parents. Elle permet ainsi de vérifier si l'impact de cet événement sur la scolarisation varie selon l'âge de l'enfant au moment de son occurrence. La principale limite de l'ESDSR est que les caractéristiques des ménages des enfants interrogés sont enregistrées à la date de l'enquête et ne reflètent donc pas nécessairement celles de leur ménage d'origine, en particulier en cas de décès d'au moins un parent. Néanmoins, la plupart des travaux menés à partir de données de panels concluent à l'existence de différences modérées dans les conclusions qui auraient été tirées en exploitant ces panels comme des données d'enquêtes rétrospectives [13-14]

Méthodologie d'analyse

L'analyse porte sur le groupe des enfants âgés de 7 à 17 ans en 2002. La scolarisation distingue les enquêtés selon qu'ils aient ou non déjà été scolarisés au moment de l'enquête. La population des enfants est décomposée en 4 sous-groupes (mutuellement exclusifs) en fonction de l'expérience du décès d'un parent avant différents âges (2 ans, 5 ans, 10 ans, 15 ans et 18 ans) : les orphelins partiels de père (enfants dont seul le père est décédé), les orphelins partiels de mère (enfants dont seule la mère est décédée), les orphelins totaux (enfants dont les deux parents sont décédés) et les enfants dont les deux parents sont en vie. Des modèles multivariés sont appliqués dans la perspective de recréer, à posteriori, les conditions d'une expérience aléatoire en contrôlant statistiquement l'ensemble des facteurs parasites supposés perturber l'inférence de l'impact du décès précoce d'un parent sur la scolarisation des enfants. La variable modélisée étant une variable dichotomique : « avoir déjà été à l'école » (1/0), le modèle de régression logistique qui exprime sous forme de probabilité la relation entre une variable Y dichotomique et une ou plusieurs variables Xi est considéré. La forme de la modélisation doit néanmoins tenir compte de la structure hiérarchique des observations disponibles tout en s'adaptant à l'objectif de la recherche, ie l'estimation sans biais de l'expérience précoce du décès parental sur la probabilité d'être entré à l'école.

L'inadaptation du modèle logistique standard uni-niveau

Dans sa forme standard uni-niveau, le modèle de régression logistique permet d'estimer l'effet de l'expérience précoce du décès parental sur la scolarisation des enfants une fois contrôlé l'influence d'une série de variables de contrôle et de leur possible corrélation. Les variables de contrôle retenues sont celles qui, selon les recherches existantes, sont liées à la scolarisation des enfants et qui sont

APPLICATION DES MODELES HIERARCHIQUES EN DEMOGRAPHIE

inégalement distribuées selon la survie des parents. Trois grands groupes de variables sont ainsi considérés et introduits séquentiellement dans l'analyse afin d'observer d'éventuels effets de médiation. Le premier renvoie aux caractéristiques socio-démographiques de l'enfant : âge et âge au carré, sexe et appartenance religieuse (Modèle I). Le second décrit les facteurs socioéconomiques et culturels du milieu « d'origine » de l'enfant : niveau d'instruction du chef de ménage, quartile de niveau de vie relatif du ménage, accès à l'aide humanitaire, évolution de la possession de vache(s) du ménage entre 1992 et 2002 et milieu de résidence (Modèle II). Enfin le dernier groupe caractérise l'environnement social de l'enfant à l'enquête : sexe du chef de ménage, présence dans le ménage d'au moins un des membres de la famille nucléaire, nombre de femmes de 18 ans et plus du ménage et présence dans le ménage d'autre(s) enfant(s) biologique(s) du chef de ménage (Modèle III). Les paramètres du modèle logistique sont estimés à l'aide de la méthode du maximum de vraisemblance. Elle suppose que les observations individuelles sont, conditionnellement aux variables X_i introduites dans les modèles, indépendantes entre elles. Cette hypothèse pose problème pour l'analyse envisagée. En effet, les données de l'ESDSR ont été collectées via un plan de sondage complexe, les enregistrements individuels sont donc organisés selon une structure hiérarchique. La dépendance des observations apparaît de façon nette à l'échelle des 168 grappes composant le plan de sondage de l'ESDSR. En conséquence, les individus des différentes grappes de sondage sont susceptibles de se ressembler sur tout un ensemble de caractéristiques dont certaines vont demeurer inobservées. Il peut en résulter un biais au niveau de l'estimation des écarts-types du modèle qui vont, en général, se trouver sous-estimés [1-2, 15]. L'absence de prise en compte de ces corrélations intra-grappes peut conduire à des inférences erronées quant aux conséquences de l'expérience précoce du décès parental sur la scolarisation des enfants. Il est donc indispensable d'en tenir compte pour estimer sans biais la liaison recherchée.

Le modèle hiérarchique multiniveaux et ses limites

A cause du problème des paramètres incidents, il est impossible d'ajouter les variables indicatrices des 168 grappes dans l'analyse. La prise en compte de ces possibles effets de grappes requiert donc des stratégies de modélisation adaptées. Dans le cadre d'un modèle à 2 niveaux appliqué au cas où la fonction qui décrit la relation entre la variable dépendante et les variables explicatives est la fonction logit, la formule générique du modèle de régression hiérarchique peut s'écrire de la manière suivante :

$$\log\left(\frac{P_{ij}}{1 - P_{ij}}\right) = \beta_0 + \beta_1 X_{pij} + \dots + \beta_p X_{pij} + u_j + e_{ij} = \beta_1 X'_{ij} + \beta_2 X'_j + u_j$$

Avec

β_0 : le logarithme du rapport de chance moyen

$\beta_1 X'_{1j}$: le vecteur de paramètres de régression β_i associés au vecteur de variables explicatives X_j qui varient essentiellement entre les macro-unités de niveau 2, j

u_j : le résidu associé à chaque unité de niveau 2, j qui représente l'effet des unités de niveau 2 sur le logarithme du rapport de chances c'est-à-dire l'hétérogénéité inobservée entre les unités de niveau 2. Ce paramètre peut être vu comme un facteur d'exposition à la probabilité d'avoir déjà été scolarisé Y qui affecte de façon identique tous les individus d'une même unité de niveau 2.

Il faut ensuite choisir la distribution des résidus u_j , hétérogénéité inobservée entre les macro-unités de niveau 2, qui peuvent être considérés soit comme des effets fixes, soit comme des effets aléatoires. Les modèles à effets aléatoires (modèles multiniveaux) sont ceux qui sont le plus couramment appliqués en démographie pour traiter le problème de l'hétérogénéité inobservée. Ce modèle est ici appliqué pour dégager l'influence de l'expérience précoce du décès parental sur la scolarisation des enfants en tenant compte des effets contextuels partagées par les membres d'une

APPLICATION DES MODELES HIERARCHIQUES EN DEMOGRAPHIE

même grappe (distance des établissements scolaires, insécurité, mode de vie et traditions culturelles des populations). En effet, grâce aux composants aléatoires des unités de niveau 2, le modèle multiniveaux est capable de corriger les écarts-types des paramètres des effets fixes de la dépendance des observations au sein des grappes [1,16]. Ce modèle présente toutefois un inconvénient majeur pour l'objectif poursuivi : il suppose l'absence de corrélation entre les variables explicatives et les effets contextuels de niveau 2. Or, il n'est pas impossible que le fait d'être un enfant orphelin soit corrélé avec les caractéristiques des grappes de sondage via un mécanisme qui échappe à l'observation. Les jeunes orphelins, en particulier de père, ne se répartissent pas de façon uniforme entre les grappes de l'enquête et il n'est pas exclu que cette distribution s'effectue en fonction de facteurs inobservés liés, directement ou indirectement, à la scolarisation des enfants (par exemple via des mécanismes d'assistance aux populations ou une appartenance ethnique particulière).

Le modèle à effets fixes pour tester la robustesse des résultats obtenus

Afin de s'assurer de la robustesse des résultats obtenus, les estimations tirées des modèles multiniveaux sont comparées à celles déduites d'un modèle à effets fixes ménage. A l'instar des modèles multiniveaux, ces modèles tiennent compte de la structure hiérarchique des observations mais à l'inverse de ces derniers, ils n'impliquent pas de supposer l'indépendance entre les variables explicatives et les facteurs inobservés associés aux différents niveaux de l'analyse. Les effets contextuels (les u_j) sont supposés être un ensemble de paramètres fixes propre à chaque entité qui ne sont pas modélisés [17]. La méthode utilisée pour l'estimation des paramètres est celle du maximum de vraisemblance conditionnelle (Chamberlain, 1980). Ce modèle est moins efficace statistiquement que le modèle multiniveaux puisque il n'utilise pas l'information des groupes qui comprennent uniquement une observation ou qui ne comportent que des observations homogènes. Les modèles à effets fixes ménage sont ici utilisés pour comparer au sein du même ménage la probabilité de scolarisation des enfants orphelins à celles des non orphelins d'âge et de sexe similaires.

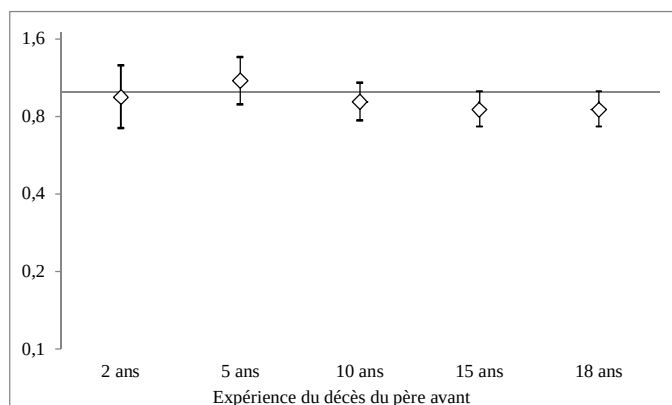
Résultats

Les coefficients (OR et IC 95%) estimés pour la variable d'intérêt principal (l'expérience précoce du décès parental avant différents âges) à partir des modèles complets (Modèle III) à effets « grappe de sondage » aléatoires sont présentés par type de perte parentale (père, mère ou deux parents) sur les figures 1. Ces coefficients permettent d'évaluer à chaque âge le risque supplémentaire des enfants orphelins par rapport à ceux dont les deux parents sont en vie de ne pas être scolarisés à l'enquête une fois les caractéristiques individuelles observées et contextuelles inobservées liées aux grappes de sondage contrôlées dans l'analyse. Toutes autres caractéristiques fixées par ailleurs, les orphelins de père ont une probabilité similaire à celles des enfants dont les deux parents sont en vie d'être scolarisés en 2002 (figure 1a). C'est principalement la pauvreté socioéconomique et culturelle de leur ménage qui explique la légère sous-scolarisation des orphelins de père observée à l'enquête. En revanche, le décès de la mère et des deux parents exercent des effets néfastes sur la chance des enfants d'être entrés à l'école. L'expérience d'une perte maternelle ou double avant le début officiel d'entrée à l'école ou autour des âges entourant cette période (7-10 ans) est plus discriminante pour l'entrée dans le système scolaire que la survenue du même événement à un âge plus tardif (Figures 1b 1c).

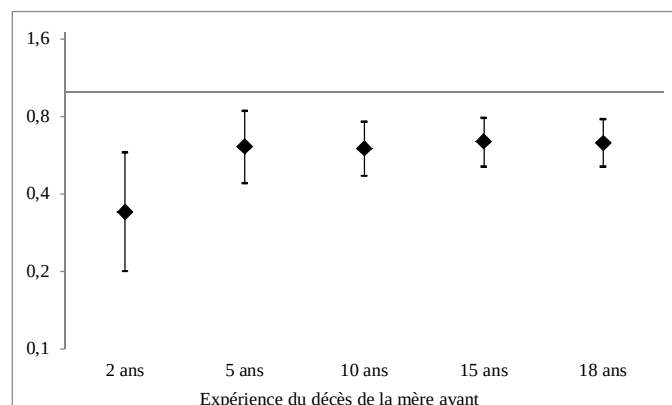
Figures 1 : Paramètres estimés (OR et IC 95%) par les modèles multiniveaux à effets « grappe de sondage » aléatoires appliqués à la probabilité de ne jamais avoir été scolarisé, selon le type de perte et l'âge de l'enfant au moment du décès (enfants âgés de 7-17 ans en 2002)

APPLICATION DES MODELES HIERARCHIQUES EN DEMOGRAPHIE

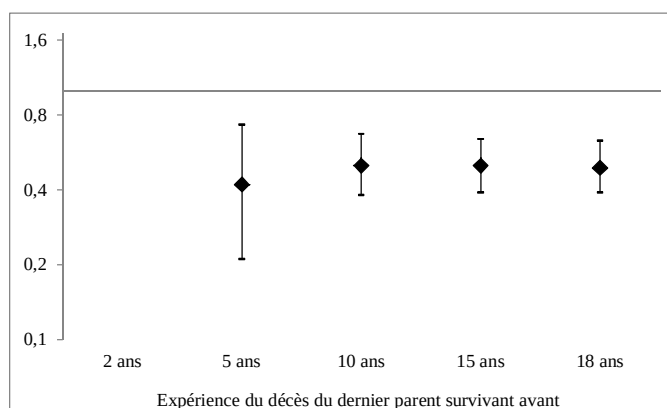
Orphelin de père



Orphelin de mère



Orphelin absolu



Significativité : ◆ $p < 0,0001$; ◇ Non significatif

Les résultats tirés des modèles multiniveaux sont confirmés par ceux obtenus à partir des modèles à effets fixes ménage qui montre une sous-scolarisation des orphelins de mère et de deux parents par rapport aux non orphelins avec lesquelles ils vivent (Tableau 1). Il demeure toutefois difficile de déterminer à l'aide de ces résultats ce qui tient à un effet direct de l'expérience précoce du décès parental sur la scolarisation des enfants (réaction psychologique de l'enfant notamment) de ce qui est liée à une conséquence indirecte de cet événement. Il est par exemple probable que l'accueil fréquent des orphelins de mère et de deux parents dans des ménages de la famille élargie se traduit par des discriminations scolaires intra-ménage à l'encontre des enfants orphelins. Dans un contexte caractérisé par une grande précarité des conditions de vie et où la scolarité est payante, les membres du ménage peuvent être amenés à réaliser des choix dans la répartition des rares ressources disponibles entre les enfants du ménage. Ces choix peuvent s'exercer à l'encontre des orphelins avec lesquels ils ne partagent pas de lien biologique direct (Hamilton, 1964).

Tableau 1 : Paramètres estimés (OR et écart-type) par le modèle à effets fixes « ménage » appliqué à la probabilité de ne jamais avoir été scolarisé, selon le type de perte parentale (enfants âgés de 7-17 ans en 2002)

Père décédé	0,69 (0,28)
Mère décédée	0,43*** (0,32)
2 parents décédés	0,23*** (0,27)
n=3 949 / *** : $p < 0,0001$	

3 L'apport des modèles multi niveaux pour l'étude de la variabilité territoriale de la mortalité

APPLICATION DES MODELES HIERARCHIQUES EN DEMOGRAPHIE

L'approche la plus spontanément pour mesurer l'effet du contexte est le modèle contextuel [19]. Il s'agit de rajouter au modèle « individuel » une variable agrégée au niveau macro-unités ou au niveau supérieur. Il s'agit de rajouter au modèle « individuel » une variable agrégée au niveau macro-unités ou au niveau supérieur. Le coefficient de cette variable agrégée mesurera l'effet du contexte, tout en contrôlant les variables individuelles.

Cependant, ce type de modèle présente deux inconvénients. D'une part il mesure l'effet du contexte comme s'il s'agissait d'un effet fixe. Or, lorsque nous souhaitons étudier l'effet « région » sur la mortalité au niveau cantonal, les délimitations des régions ont une part « d'arbitraire ». Les régions françaises résultent de choix administratifs recoupant plus ou moins des réalités géographiques, sociales, économiques, historiques,... Pour exemple, le sud de la région Poitou-Charentes est à bien des égards davantage sous l'influence de Bordeaux (capitale régionale de l'Aquitaine) que sous celle de sa propre capitale régionale, Poitiers. Si, comme c'est le cas ici, nous souhaitons faire jouer à la région un rôle agrégeant certains effets contextuels (traditions alimentaires régionales, secteurs économiques employant de larges couches de la population, activités polluantes, etc), il y aurait donc des cantons susceptibles d'être classés autrement. Ceux du sud des Charentes en sont un exemple, mais l'on pourrait aussi évoquer ceux du sud de la Picardie qui subissent très fortement l'influence de l'aire urbaine de Paris, voire en font carrément partie. Autrement dit, la consistance de la définition des unités supérieures ou macros-unités (ici les régions) est sujette à discussion tout ayant une importante signification en matière de mortalité et surtout sans qu'il existe un système indiscutable de regroupement géographiquement continu des cantons. Nous ne pouvons donc parler d'effet fixe pour décrire la variable « Région », bien que d'un point de vue strictement nominal (administratif) les régions françaises aient une définition tout à fait stable. De façon générale, l'analyse spatiale, lorsqu'elle fait intervenir plus d'un niveau d'agrégation ne peut traiter l'effet de ces différents niveaux d'agrégation comme des effets fixes.

D'autre part, estimer les coefficients du modèle proposé par Bressoux [19] avec la méthode des moindres carrés ordinaires (MCO), ce qui est généralement réalisé, pose certains problèmes.

La même valeur de variable agrégée est attribuée à tous les individus du même groupe, elle est donc multipliée autant de fois qu'il y a d'individus. Son degré de liberté repose donc sur le nombre d'individus du niveau inférieur, alors qu'elle devrait reposer sur le nombre de groupes du niveau hiérarchique supérieur.

De plus, les cantons (micros-unités) de la même région ne sont pas indépendants les uns des autres ; par conséquent, les observations par groupe d'âges quinquennal ne le sont pas non plus. La problématique traitée est liée aux interactions entre les cantons et leur environnement. Les cantons qui appartiennent à la même région ont forcément des caractéristiques communes. Ils partagent la même histoire régionale, le même climat. Ils ont par ailleurs des caractéristiques culturelles communes, subissent ensemble certaines pollutions, présentent des spécificités productives semblables. Appartenir à la même région signifie donc partager des éléments contextuels. Les cantons d'une même région seront proches sur tout un ensemble de caractéristiques dont certaines demeurent non mesurables. Ces variables-là vont se disperser dans le résidu du modèle, provoquant une dépendance entre eux. Dans ce cas-là, l'hypothèse de l'indépendance des résidus n'est alors pas vérifiée, et les variances des coefficients seront donc sous-estimées par la MCO. De plus, cette dépendance est traitée par les MCO, comme une nuisance. Au contraire, elle doit être traitée comme un phénomène particulièrement intéressant à étudier, quand nous nous penchons sur les effets des unités supérieures sur le comportement des micros-unités).

Lorsque notre objectif est de mesurer la variabilité territoriale de la mortalité, on s'intéresse tout particulièrement à l'effet « région ». Le modèle mixte ou le modèle à deux niveaux semble l'outil le plus adéquat. En effet, à l'aide des effets aléatoires qu'ils incluent aux différents niveaux, les modèles à deux niveaux tiennent compte de la structure hiérarchique des données lors de la procédure d'estimation des paramètres et de leurs écarts types. Mais au-delà, comme la variance des composants aléatoires du niveau supérieur apparaît significativement différente de zéro, ces modèles

APPLICATION DES MODELES HIERARCHIQUES EN DEMOGRAPHIE

permettent de traiter la variance du niveau supérieur comme une source d'information potentielle, au contraire des approches plus classiques qui la considèrent comme une nuisance [22]. Grâce à leurs composantes aléatoires, ces modèles sont en mesure de répartir les variations résiduelles du phénomène entre les différents niveaux, et leur intérêt spécifique est de permettre d'examiner comment évolue cette répartition lors des ajouts successifs de variables du niveau inférieur. Au total, ces modèles sont donc adaptés à l'analyse de la variabilité territoriale de la mortalité, pour deux raisons. Ils tiennent compte de la structure « hiérarchique » géographique des données et ils distinguent la variabilité du phénomène localisée à l'échelle cantonale, de la variabilité à l'échelle régionale.

Données et contexte

Les données utilisées pour cette seconde partie sont les données d'Etat civil gérées par l'Institut national de la statistique et des études économiques (INSEE), pour les effectifs de décès annuels répartis par commune du domicile, par sexe et par groupe d'âges quinquennal (en différence de millésimes). Les données obtenues concernent les années allant de 2005 à 2007.

Pour les effectifs de populations, le recensement rénové de population pour les vagues 2004 à 2008 dont l'année de référence est 2006, par interpolation et extrapolation, est utilisé. Puisque depuis 2004, le recensement en France est organisé par enquêtes annuelles suivant deux protocoles différents selon la taille des communes.

Le choix d'une échelle géographique la plus fine possible et qui assure des résultats statistiques néanmoins suffisamment robustes a pour but de rendre compte au mieux de l'état de santé de la population et d'étudier leur vulnérabilité au risque de décès. De ce fait le découpage communal n'a pas été retenu en raison des faibles effectifs que comportent certaines entités. A contrario, le découpage en « cantons-villes » (appelés anciennement « pseudo-cantons ») a été choisi car un canton-ville forme une zone plus peuplée que la commune notamment en milieu rural. Pour la ville de Paris, l'analyse a été menée au niveau de l'arrondissement municipal, considéré comme un canton. L'étude à l'échelle cantonale permet de révéler des particularités en matière de santé donc en matière de mortalité par l'analyse de 3 689 unités géographiques distinctes appartenant aux 22 régions de France métropolitaine. L'analyse porte sur la seule France métropolitaine en raison de la difficulté d'obtenir certaines données pour les départements/régions d'outre-mer (décès domiciliés par exemple, pour l'outre-mer les décès sont comptabilisés sur leur seul lieu d'enregistrement et non en fonction du domicile du décédé).

Population et indices utilisés

La population d'étude est la population âgée entre 30 et 60 ans. Ce choix est justifié par le fait que les différences régionales de mortalité les plus nettes s'observent entre 30 et 60 ans [21].

La vulnérabilité au risque de décès est mesurée par le TCM pour éliminer l'effet parasite de la structure par âge, pour comparer la mortalité des différents territoires et pour détecter ceux qui sont les plus vulnérables que d'autre. Le taux de mortalité est standardisé directement sur la structure par âges de la France métropolitaine, sexes réunis, 2006.

Ces TCM permettent de classer les territoires les uns par rapport aux autres, et aussi par rapport à la moyenne métropolitaine en termes de mortalité. Ils permettent aussi d'apprécier la fréquence de la mortalité, contrairement aux indices comparatifs de mortalité qui sont très fréquemment utilisés en épidémiologie (autre modalité de standardisation - standardisation indirecte) et qui sont exprimés par rapport à la moyenne métropolitaine (base de 100).

Toutefois, à l'échelle des petits territoires, comme les cantons-villes, la mesure de mortalité par taux sur une année peut s'avérer statistiquement problématique, en raison de l'instabilité engendrée par les petits effectifs. Plus le phénomène est rare et plus la population dans laquelle il survient est faible, plus le taux varie entre une année et une autre. Pour rendre les taux comparables entre ces petits territoires, il est nécessaire de tenir compte, à la fois de la structure par âges, et de la variabilité de ces taux. La méthode proposée pour contourner ce problème est de regrouper les décès

APPLICATION DES MODELES HIERARCHIQUES EN DEMOGRAPHIE

sur cinq années successives encadrant l'année du recensement, excepté la dernière période 2005-2007.

Nous avons également calculé, pour les hommes puis pour les femmes, les proportions standardisées des bacheliers et proportions standardisées liées à la PCS.

Pour chaque période, pour chaque canton-ville, les proportions des différentes PCS sont standardisées sur l'âge ; comme pour les TCM ; pour obtenir ainsi dans la population totale de 30-59 ans :

- ✓ la proportion standardisée d'agriculteurs exploitants ;
- ✓ la proportion standardisée d'artisans, commerçants et chefs d'entreprises ;
- ✓ la proportion standardisée des cadres et des professions intellectuelles supérieures ;
- ✓ la proportion standardisée de professions intermédiaires ;
- ✓ la proportion standardisée d'employés ;
- ✓ la proportion standardisée d'ouvriers ;
- ✓ la proportion standardisée de bacheliers, de 20 ans et plus, dans la population totale des 20 ans et plus.

Ces proportions seront incluses dans le modèle à deux niveaux comme des variables explicatives au niveau cantonal.

Méthodologie d'analyse

Présentation du modèle à deux niveaux

Nous procédons par étape pour présenter le modèle linéaire mixte [18]. Pour simplification, nous retenons seulement deux niveaux (donc deux indices i et j).

On suppose que la variable expliquée TCM est gaussienne de vecteur d'espérance mathématique $X\gamma$ et matrice variances-covariances Γ .

Étape 1 : Modèle vide

Les variables explicatives n'ont pas incluses dans le modèle. Son but est de montrer comment la variance du phénomène étudié se répartit sur les différents niveaux qui structurent les données. On peut utiliser ce modèle pour voir comment la variabilité totale du phénomène étudié est répartie entre les deux niveaux (la variabilité intergroupe (niveau supérieur) et la variabilité intra-groupe (niveau inférieur)).

La forme structurelle du modèle vide s'écrit comme suit. Au niveau inférieur, le modèle contient une équation unique :

$$TCM_{ij} = \beta_{0j} + u_{ij} \quad \text{où } u_{ij} \sim N(0, \sigma_r^2)$$

Au niveau supérieur, il contient une équation qui explique la variabilité aléatoire de ce niveau :

$$\beta_{0j} = \gamma_{00} + u_{0j} \quad \text{où } u_{0j} \sim N(0, \sigma_{u0}^2)$$

En intégrant l'équation du niveau supérieur dans l'équation du niveau inférieur, nous obtenons l'écriture suivante : $TCM_{ij} = \gamma_{00} + u_{0j} + u_{ij}$

avec :

TCM_{ij} le taux comparatif de la mortalité du canton i de la région j ;

γ_{00} le taux comparatif moyen de la mortalité de toutes les régions ;

u_{0j} une erreur aléatoire associée à chaque région j supposée normalement distribuée, d'une moyenne nulle et d'une variance σ_{u0}^2 (variance inter-régions) ;

APPLICATION DES MODELES HIERARCHIQUES EN DEMOGRAPHIE

u_{ij} une erreur aléatoire associée à chaque canton i supposée normalement distribuée, d'une moyenne nulle et d'une variance σ_r^2 (variance intra-région).

Les variances estimées permettent de calculer un coefficient de corrélation ρ qui fournit la part de la variance inter-groupes (inter-régionale) de la mortalité dans la variabilité totale de la mortalité expliquée par le modèle.

$$\rho^2 = \frac{\sigma^2 u_0}{\sigma^2 u_0 + \sigma_r^2}$$

Étape 2 : Inclusion des variables explicatives.

Cette étape consiste à inclure les variables explicatives du niveau inférieur (cantonal) dans le modèle. Il s'agit de contrôler la dépendance des observations et de respecter leur structure hiérarchique spatiale.

On obtient, au niveau inférieur, une équation unique :

$$TCM_{ij} = \beta_{0j} + \beta_{1j} X_{ij} + u_{ij} \quad \text{où } u_{ij} \sim N(0, \sigma_r^2)$$

Et au niveau supérieur (régional), le modèle contient deux équations ; à savoir :

$$\beta_{0j} = \gamma_{00} + u_{0j} \quad \text{où } u_{0j} \sim N(0, \sigma_{u0}^2)$$

$$\beta_{1j} = \gamma_{10}$$

En intégrant les équations du niveau supérieur dans l'équation du niveau inférieur, nous avons :

$$TCM_{ij} = \gamma_{00} + \gamma_{10} X_{ij} + u_{0j} + u_{ij}$$

avec :

X_{ij} la variable explicative mesurée dans le canton i de la région j ;

γ_{00} la constante moyenne pour toutes les régions ;

γ_{10} la pente de la droite de régression des TCM sur X ;

u_{0j} une erreur aléatoire associée à chaque région j supposée normalement distribuée, d'une moyenne nulle et d'une variance σ_u^2 (variance inter-régions) ;

u_{ij} une erreur aléatoire associée à chaque canton i supposée normalement distribuée, d'une moyenne nulle et d'une variance σ_r^2 (variance intra-région).

Étape 3 - Modèle plus complexe

On peut compliquer le modèle en rajoutant une pente aléatoire au niveau supérieur pour introduire les variables explicatives au niveau supérieur. L'intérêt de cette étape est d'apprécier le gain en introduisant des facteurs explicatifs au niveau supérieur (ici région).

Au niveau inférieur, il contient une équation unique

$$TCM_{ij} = \beta_{0j} + \beta_{1j} X_{ij} + u_{ij} \quad \text{où } u_{ij} \sim N(0, \sigma_r^2)$$

Mais au niveau supérieur, le modèle contient deux équations :

$$\beta_{0j} = \gamma_{00} + u_{0j} \quad \text{où } u_{0j} \sim N(0, \sigma_{u0}^2)$$

$$\beta_{1j} = \gamma_{10} + u_{1j} \quad \text{où } u_{1j} \sim N(0, \sigma_{u1}^2)$$

En intégrant ces deux équations du niveau supérieur dans l'équation du niveau inférieur, nous obtenons $TCM_{ij} = \gamma_{00} + \gamma_{10}X_{ij} + u_{0j} + u_{1j}X_{ij} + u_{ij}$

Avec :

X_{ij} la variable explicative mesurée dans le canton i de la région j ;

γ_{00} la constante moyenne pour toutes les régions ;

γ_{10} la pente de la droite de régression des TCM sur X ;

u_{0j} une erreur aléatoire associée à chaque région j supposée normalement distribuée, d'une moyenne nulle et d'une variance σ_u^2 (variance inter-régions) ;

u_{ij} une erreur aléatoire associée à chaque canton i supposée normalement distribuée, d'une moyenne nulle et d'une variance σ_r^2 (variance intra-région).

Le coefficient associé à la variable X se décompose ici, comme l'ordonnée à l'origine ou constante, en une partie fixe et une partie aléatoire. Les parties fixes de l'ordonnée à l'origine (γ_{00}) et de la pente (γ_{10}) sont intéressantes en tant qu'effets « individuels », ici cantonaux, alors que les parties aléatoires (u_{0j} et u_{1j}) peuvent être interprétées en termes d'effets contextuels ou régionaux, puisqu'elles représentent des différences entre régions.

Composantes de la variance et hétéroscédasticité

On s'intéresse à la variabilité de la mortalité cantonale, mesurée par des taux comparatifs de mortalité TCM (caractère quantitatif) au niveau de chaque canton (pseudo-canton), au sein d'une région. On souhaite notamment savoir si la mortalité cantonale est fortement liée à la région ou non, c'est-à-dire si les cantons issus d'une même région sont plus semblables en termes de mortalité que ceux issus d'une autre région.

Pour cela on considère les 22 régions indexées $j=1,2,\dots,22$ chacune ayant n_j pseudo-cantons indexés $i = 1,2,\dots, n_j$.

On note TCM_{ij} le taux de mortalité comparatif de l'i-ème canton de la j-ème région.

Le modèle s'écrit : $TCM_{ij} = \gamma_{00} + \gamma_{10}X_{ij} + u_{0j} + u_{ij}$ (M2)

X_{ij} la variable explicative mesurée dans le canton i de la région j ;

γ_{00} la constante moyenne pour toutes les régions ;

γ_{10} la pente de la droite de régression des TCM sur X ;

u_{0j} une erreur aléatoire associée à chaque région j supposée normalement distribuée, d'une moyenne nulle et d'une variance σ_{u0}^2 (variance inter-régions) ;

u_{ij} une erreur aléatoire associée à chaque canton i supposée normalement distribuée, d'une moyenne nulle et d'une variance σ_r^2 (variance intra-région).

Les termes u_{0j} et u_{ij} sont supposés centrés, gaussiens et indépendants :

APPLICATION DES MODELES HIERARCHIQUES EN DEMOGRAPHIE

$$u_{0j} \sim N(0, \sigma_{u0}^2)$$

$$u_{ij} \sim N(0, \sigma_r^2)$$

La variabilité entre les régions est donc mesurée par la variance due à l'effet « région » σ_{u0}^2 . L'effet aléatoire associé à chaque région introduit la corrélation entre les pseudo-cantons qui appartiennent à cette région. En effet, si dans ce modèle l'espérance a une forme plus simple :

$$E(TCM_{ij}) = \gamma_{00} + \gamma_{10} X_{ij} + E(u_{0j}) + E(u_{ij}) = \gamma_{00} + \gamma_{10} X_{ij}$$

La variance de chaque observation est la somme de la variance due à la région et de la variance résiduelle (cantonale) :

$$V(TCM_{ij}) = V(u_{0j}) + V(u_{ij}) = \sigma_{u0}^2 + \sigma_r^2$$

Les pseudo-cantons de la même région ne sont pas indépendants. Ainsi ce modèle suppose que les pseudo-cantons de deux régions différentes sont indépendants et que la corrélation au sein d'une même région est uniforme. La covariance s'écrit :

$$COV(TCM_{ij}, TCM_{i'j'}) = \begin{cases} \sigma_{u0}^2 & \text{si } j = j' \\ 0 & \text{sinon} \end{cases}$$

Au total, les composantes de la variance sont la variance de l'effet région σ_{u0}^2 et la variance résiduelle σ_r^2 .

Les résidus sont non indépendants les uns des autres ce qui complique la vérification de l'adéquation du modèle multi-niveaux. Néanmoins, on peut valider le modèle visuellement par plusieurs graphes et le plus utilisé est le graphe des résidus standardisés (résidus divisés par leurs écarts-type) versus une variable quantitative permet d'évaluer si l'effet de cette variable est correctement ajusté par le modèle ou non. Cette présentation graphique permet aussi de repérer les observations extrêmes.

On peut également, vérifier l'absence de l'homoscédasticité par le graphe des résidus standardisés versus les prédictions de la variable étudiée (endogène). Si toutes les observations se répartissent plus ou moins dans un intervalle de]-2,+2[, on conclut l'absence de l'hétéroscédasticité. Souvent lorsque le diagnostic est négatif, des transformations mathématiques (exemple log, racine...) sont nécessaires pour contourner le problème de l'hétéroscédasticité et améliorer le modèle [20].

Résultats

Entre 30 et 60 ans, les personnes sont plus stables et moins mobiles qu'aux âges précédents (sauf probablement pour certaines catégories socioprofessionnelles, notamment les cadres) ; ce qui pourrait induire une part interrégionale de la variabilité de la mortalité des cantons plus importante qu'aux âges précédents.

On constate que les disparités intra-région sont plus importante que les disparités de la mortalité entre les régions. C'est du moins ce que l'on observe pour les hommes de 30-59 ans, avec une part interrégionale de la variabilité de la mortalité cantonale de 13% ($\rho^2 = 13\%$) en 2005-2007 (Modèle A du tableau 1 en annexe). La variabilité de la mortalité féminine (pour les mêmes âges) entre les cantons est plus homogène que pour les hommes, en 2005-2007. En effet, la part interrégionale de la variabilité de la mortalité cantonale des femmes de 30-60 ans est de seulement 6% ($\rho^2 = 6\%$) en 2005-2007 (Modèle A du tableau 2 en annexe).

De plus, en 2005-2007, l'influence de la proportion cantonale standardisée de bacheliers parmi les 30-59 ans et les proportions cantonales standardisées liées à la structure par PCS est assez importante sur la variabilité territoriale de la mortalité cantonale masculine.

APPLICATION DES MODELES HIERARCHIQUES EN DEMOGRAPHIE

Neutraliser ces différentes proportions a entraîné une baisse de la variabilité résiduelle intra-régionale de la mortalité cantonale des hommes à 30-59 ans de 13% et une baisse de la variabilité interrégionale de la mortalité cantonale des hommes à 30-59 ans de 20% (passage du modèle A au modèle C) (cf. tableau 1 en annexe).

La partie fixe du modèle C indique que les cantons où la proportion cantonale standardisée de bacheliers parmi les 30-59 ans combinée linéairement à la proportion cantonale standardisée des cadres et des professions intellectuelles supérieures parmi les 30-59 ans et la proportion cantonale standardisée des professions intermédiaires parmi les 30-59 ans sont élevées, sont en moyenne aussi ceux où la mortalité des 30-59 ans est la plus faible. A l'inverse, les cantons où la proportion cantonale standardisée d'employés parmi les 30-59 ans et la proportion standardisée d'ouvriers parmi les 30-59 sont élevées, sont en moyenne aussi ceux où la mortalité des 30-59 ans est la plus forte.

La variabilité interrégionale de la mortalité cantonale masculine à 30-59 ans dépend significativement de la proportion cantonale standardisée des ouvriers parmi les 30-59 ans (modèle D).

Pour les femmes les coefficients estimés (tableau 2 en annexe), dans la partie fixe du modèle C, pour les différentes proportions standardisées incluses dans le modèle ont aussi des signes attendus : les cantons où la proportion cantonale standardisée de bachelières parmi les 30-59 ans combinée linéairement à la proportion cantonale standardisée des cadres et des professions intellectuelles supérieures parmi les 30-59 ans et la proportion cantonale standardisée des professions intermédiaires parmi les 30-59 ans sont élevées, sont en moyenne aussi ceux où la mortalité féminine des 30-59 ans est la plus faible. A l'inverse, les cantons où la proportion cantonale standardisée d'employées parmi les 30-59 ans et la proportion standardisée d'ouvrières parmi les 30-59 sont élevées, sont en moyenne aussi ceux où la mortalité féminine des 30-59 ans est la plus forte.

Nous notons qu'en 2005-2007, la proportion cantonale standardisée de bachelières influence seule et d'une façon assez importante la variabilité interrégionale de la mortalité cantonale des femmes de 30-59 ans (20%). La partie fixe du modèle B indique que les cantons où la proportion cantonale standardisée de bachelières parmi les 30-59 ans est élevée, sont en moyenne aussi ceux où la mortalité féminine des 30-59 ans est la plus faible.

La répartition spatiale inégale des populations masculines, mesurée par les différentes proportions standardisées de la catégories socio-professionnelles CSP, a une influence assez importante sur les différences résiduelles intra-régionales de la mortalité cantonale masculine à 30-59 ans. Pour les femmes de 30-59 ans, c'est la proportion standardisée de bachelières parmi les 30-59 ans qui exerce une influence sur les différences résiduelles intra-régionales de la mortalité cantonale féminine à 30-59 ans.

4 Conclusion

Il est intéressant d'appliquer les modèles hiérarchiques pour la recherche « causale » aux données de panel qui permettraient une meilleure estimation de l'effet causal.

L'intérêt majeur de mesurer la variabilité territoriale de la mortalité par le modèle à deux niveaux est de quantifier la part des différences de la mortalité de chaque niveau et détecter ainsi l'effet du contexte (ici région)) et la part des différences de la mortalité liée à l'individu statistique (ici canton). L'analyse multi-niveau permet de synthétiser l'ensemble de constats réalisés et mesurés isolément (étude cartographique de mortalité, corrélation entre la TCM et les proportions standardisées des PCS,...) puisqu'elle intègre dans un même modèle d'étude des mortalités cantonales, la mesure des variabilités interrégionales, intra-régionales et les facteurs sociaux.

Le recours aux modèles multi-niveaux est primordial puisqu'ils sont les seuls capables de tester rigoureusement la signification statistique des corrélations observées dans le contexte de données à structure géographique « hiérarchique ».

Néanmoins, si, pour expliquer les différences territoriales de mortalité, les spécificités régionales

APPLICATION DES MODELES HIERARCHIQUES EN DEMOGRAPHIE

(comportements, type d'activités, ...) persistent tout en constatant l'impact de l'inégale répartition de la population selon les PCS se réduire au profit des concentrations socio-spatiales, il reste néanmoins de nombreux autres facteurs explicatifs à introduire dans le modèle (par exemple la répartition territoriale des professionnels de santé et d'autres spécificités cantonales).

Annexes

Tableau 1 : Modèles à deux niveaux pour la période 2005-2007 - TCM masculine à 30-59 ans (pour 100 000)

	Modèle A : Modèle sans var. explicatives		Modèle B : Modèle avec var. explicatives PCS et constante aléatoire		Modèle C : Modèle avec var. explicatives PCS, diplôme et constante aléatoire		Modèle D : Modèle avec var. explicatives PCS, diplôme, constante aléatoire et pente aléatoire	
Composantes de la variance	Estimateur	p(Z)	Estimateur	P(z)	Estimateur	p(Z)	Estimateur	p(Z)
<u>Variance inter-régions</u>								
Constante aléatoire	2 541,3 (823,0)	**	2045,3 (677,4)	**	2042,0 (676,8)	**	164,87 (424,2)	NS
pentes aléatoires : PS ouvriers							2,52 (0,95)	**
<u>variance inter-cantons</u>								
résidu : variance inter- cantons	17510 (409,2)	** *	15341,0 (358,7)	***	15313,0 (358,1)	** *	15213 (356,4)	** *
Effets fixes	Estimateur	p(t)						
Constante	391,6 (11,01)	** *	421,9 (31,4)	***	402,16 (29,18)	** *	395,8 (27,4)	** *
PS artisans, comm., chef ETP			-2,15 (1,09)	*	-2,2 (1,08)	*	-2,01 (1,07)	NS
PS cadres			-3,8 (0,6)	***	—		—	
PS prof. intermédiaire			-6,98 (0,7)	**	-6,63 (0,67)	** *	-6,39 (0,67)	** *
PS employés			+9,6 (1,00)	***	+9,6 (0,99)	** *	+9,47 (0,99)	** *
PS ouvriers			+1,79 (0,6)	**	+1,03 (0,64)	NS	+1,05 (0,7)	NS
PS cadres_diplômés			—		-20,1 (3,0)	** *	-19,9 (2,9)	** *
<u>critère d'ajustement :</u>								
AIC	46 527,6		46 028,5		46 018,7		45 998,6	
- 2 log L	46 523,6		46 024,5		46 014,7		45 992,6	

Ps : proportion standardisée sur l'âge.

APPLICATION DES MODELES HIERARCHIQUES EN DEMOGRAPHIE

*** signification au seuil inférieur à 1 pour mille. ** signification au seuil entre 1 pour mille et 1 %. * signification au seuil entre 1 % et 5 %.

Tableau 2 : Modèles à deux niveaux pour la période 2005-2007 - TCM féminine à 30-59 ans (pour 100 000)

	Modèle A : Modèle sans var. explicatives		Modèle B : Modèle avec var. explicatives diplôme et constante aléatoire		Modèle C : Modèle avec var. explicatives PCS et constante aléatoire	
Composantes de la variance	Estimateur	p(Z)	Estimateur	p(Z)	Estimateur	p(Z)
<u>Variance inter-régions</u>						
Constante aléatoire	339.0 (117)	**	271.14 (95.6)	**	393.7 (133.6)	**
<u>variance inter-cantons</u>						
résidu : variance inter-cantons	5 738.9 (134)	***	5 554.3 (129.8)	***	5 672.6 (132.6)	***
Effets fixes	Estimateur	p(t)	Estimateur	p(t)	Estimateur	p(t)
Constante	172.5 (4.15)	***	241.1 (7.1)	***	114.9 (13.5)	***
PS bachelières			-2.18 (0,19)	***	—	—
PS agricultrices exp.					-0.54 (0.5)	NS
PS art. com. Chef ETP					+4.8 (0.9)	***
PS employées					+1 .10 (0.38)	**
PS ouvrières					+1.73 (0,39)	***
critère d'ajustement : AIC	42 435.2		41 883.8		42 389.4	
-2 log L	42 431.1		41 879.8		42 385.4	

Ps : proportion standardisée sur l'âge.

*** signification au seuil inférieur à 1 pour mille. ** signification au seuil entre 1 pour mille et 1 %. * signification au seuil entre 1 % et 5 %.

Bibliographie

- [1]Goldstein H. (1995), Multilevel statistical models (2ème éd.), Edwards Arnold, Londres; Halstead Press, New-York.
- [2] Sjniders T. A.B. et Bosker R. (1999), Multilevel analysis: An introduction to basic and advanced multilevel modeling, Sage, Londres.
- [3]Mahmood S., Zainab B. et Mahbub Latif A.H.M (2013), Frailty modeling for clustered survival data : an application to birth interval in Bangladesch, Journal of Applied Statistics, 40(12), 2670-2680.
- [4] Rahaman Khan H. et Shaw E. (2011), Multilevel logistic regression analysis applied to binary contraceptive prevalence data, Journal of Data Science, 9, 93-110.
- [5] Case A., Paxson C. et Ableidinger J. (2004), Orphans in Africa: Parental Death Poverty and School Enrollment, Demography, 41(3), 483-508.
- [6]Rita Testa M. (2014), On the positive correlation between education and fertility intentions in Europe: Individual- and country-level evidence, Advances in Life Course Research, 21, 28-42.
- [7]Pebley A.R., Goldman N. et Rodriguez G. (1996). Prenatal and delivery care and childhood immunization in Guatemala: Do family and community matter?, Demography, 33(2), 231-247
- [8]Courgeau D. (2004), Du groupe à l'individu. Synthèse multiniveau, INED, Paris.
- [9]Oleke C., Blystad A., Fylkesnes K. et Tumwine J.K. (2007), Constraints to educational opportunities of

APPLICATION DES MODELES HIERARCHIQUES EN DEMOGRAPHIE

orphans: A community-based study from northern Uganda, *AIDS Care*, 19(3), 361-368.

[10]Case A. et Ardington C. (2006), The impact of parental death on school outcomes: longitudinal evidence from South Africa, *Demography*, 43(3), 401–420.

[11]Timaues Ian M. et Boler T. (2007), Father figures: the progress at school of orphans in South Africa, *AIDS*, 21 (Sup. 7), S83-S93.

[12]Blayo C., Bergouignan C., Etchelecou A., Cordazzo V., Llopert M. et Okobo N. (2004), Situation démographique et sociale du Burundi Résultats de l'enquête ESDSR de 2002, Rapport IEDUB-SET.

[13]Ardington C. (2010), Orphanhood and schooling in South Africa: Trends in the vulnerability of orphans between 1993 and 2005, *Economic Development and Cultural Change*, 58(3), 507-536.

[14]Evans D. et Miguel E. (2007), Orphans and Schooling in Africa: A Longitudinal Analysis, *Demography*, 44(1), 35-57.

[15]Moulton, B. R. (1990), An Illustration of a Pitfall in Estimating the Effects of Aggregate Variables on Micro Unit, *The Review of Economics and Statistics*, 72(2), 334–338.

[16]Bryk A.S. et Raudenbush S.W. (1992), *Hierarchical linear models*, Sage, Newbury Park, CA.

[17] Allison P.D. (2006), Fixed effects regression method in SAS, *Statistics and Data Analysis*, SUGI 31.

[18] Arrègle Jean-Luc (2003), Les modèles linéaires hiérarchiques : 1. principes et illustration », *M@n@gement* 1/2003 (Vol. 6) , p. 1-28 URL : www.cairn.info/revue-management-2003-1-page-1.htm. DOI : 10.3917/mana.061.0001

[19] Bressoux P (2008), *Modélisation statistique appliquée aux sciences sociales. Méthodes en sciences humaines*. De boeck ;.

[20] Commenges D. et Jacqmin-Gadda H (2015), *Modèles biostatistiques pour l'épidémiologie*, Edition de boeck supérieur, .

[21] Daudin J.J (2015), *Statistique- Le modèle linéaire et ses extensions : Modèle linéaire général, modèle linéaire généralisé, modèle mixte, plans d'expériences*. Edition ELLIPSES,

[21] Laoudj F.(2010), *Les disparités territoriales de mortalité et leur évolution depuis le début des années quatre-vingt : Apports et limites de la prise en compte de l'inégale répartition spatiale des groupes sociaux en France métropolitaine* ». Thèse dirigée par M. Christophe BERGOUIGNAN Professeur des universités, Université Montesquieu Bordeaux,

[22] Rice N, Jones A, Multilevel models and health economics. *Health Econ* 1997, 6, 561-575.