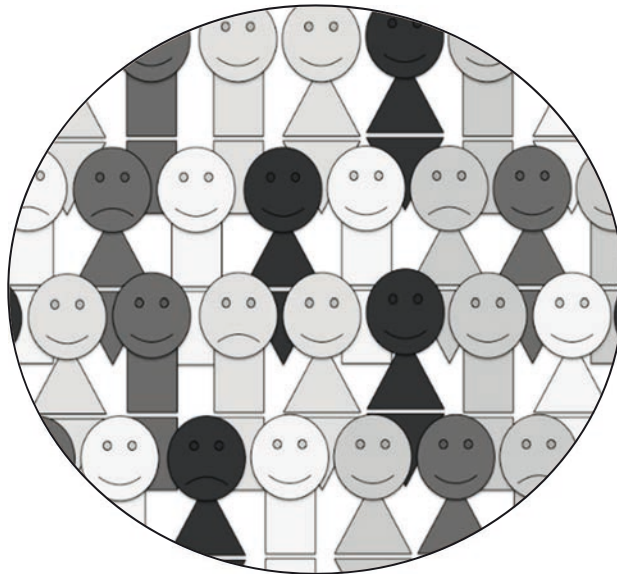


# COMPRENDRE LE CHANGEMENT SOCIAL

APPORT DES MÉTHODES DE DÉCOMPOSITION  
ET APPLICATION À L'ÉTUDE DU DIVIDENDE DÉMOGRAPHIQUE



PARFAIT M. ELOUNDOU  
SARAH C. GIROUX  
MICHEL TENIKUE



L'anecdote est bien connue, celle de ce brave homme qui, ayant perdu sa clef en pleine nuit, s'obstine à la chercher uniquement sous un lampadaire tout simplement parce que c'est le seul endroit illuminé. Les sciences sociales, et surtout celles appliquées au changement social, vivent aujourd'hui une situation similaire. De nombreuses études s'obstinent à expliquer le changement social en se basant uniquement sur les déterminants des comportements individuels parce qu'est à ce niveau que les méthodes d'analyse les plus rigoureuses existent. Ainsi pour estimer les effets de l'éducation sur la croissance économique nationale, ces études s'appuieraient sur les études liant le niveau d'éducation d'un travailleur à son revenu. En santé publique, pour estimer les effets du développement économique sur l'obésité nationale, elles examineraient comment le revenu d'un individu affecte sa masse pondérale. En démographie enfin, pour estimer les dividendes économiques de la transition démographique, elles examineraient comment certaines variables micro-démographiques telles la structure d'un ménage affectent son épargne.

Ces exemples illustrent le déphasage potentiel entre les problèmes de développement à étudier et les analyses proposées. Alors que le changement social est macro-dynamique, il est souvent analysé par une approche micro-statique et transversale. Un tel déphasage pose deux problèmes. Pour commencer, l'étude prend l'individu – et non la société entière – comme unité d'analyse; Ensuite, elle examine les différences entre individus – et non le changement historique. Une analyse micro-transversale peut certes révéler pourquoi certaines personnes ont une plus forte propension que d'autres à connaître un problème donné (ex. obésité), mais elle ne pourra pas établir de manière convaincante pourquoi la prévalence sociale de ce phénomène change au fil du temps.

Même si ce déphasage est souvent décrié, beaucoup d'analystes continuent d'interpréter leurs micro-résultats comme s'ils révélaient les macro-relations correspondantes. Une telle interprétation souffre de deux biais: un biais **écologique** qui consiste à utiliser des données micro pour tirer des conclusions macroscopiques; et un biais **historique**, c.-à-d. la lecture d'une comparaison transversale comme indiquant une évolution historique. Si, voulant éviter ces deux erreurs, notre analyste se tourne vers une analyse entièrement macroscopique, elle tombe de Charybde en Scylla. Son analyse sera critiquée pour son manque de rigueur statistique et sa propension à masquer la diversité des ménages (Rodrik 2005). La question qui se pose est « comment donc éviter les biais écologiques et historiques tout en conservant la rigueur statistique ? »

Idéalement, l'analyste doit intégrer la diversité entre individus et montrer comment diversité s'agrège pour produire un changement collectif. Dans cette optique, la décomposition est un outil de choix et un bon compromis entre les approches micro et macro. Même si elle n'identifie pas les causes ultimes du changement social, elle en localise les « sources » c.à.d. les groupes ou processus moteurs. Malgré ce potentiel, la décomposition est peu enseignée et sous-exploitée. Le présent ouvrage comble ce vide en offrant une introduction simple de la méthode. Il améliore une première version publiée en 2010 en incluant de nouvelles variantes et une application à l'étude du dividende démographique.

L'étude du dividende démographique – les bénéfices socioéconomiques attendus d'une transition démographique – offre un champ d'application idéal. Le dividende postule une relation entre deux processus macroscopiques (population et développement) mais les décisions y relatives se prennent au niveau des ménages ou individus. La méthode de décomposition permet d'examiner comment ces les comportements démographiques et économiques des diverses couches de la population s'agrègent pour générer un dividende.

Cet ouvrage a été réalisé dans le cadre d'un projet de renforcement des capacités de recherche démographique en Afrique, sous l'égide de l'Union Internationale pour l'Etude Scientifique de la Population (IUSSP) avec le soutien de la fondation William et Flora Hewlett. Il a bénéficié du soutien matériel de l'Université de Cornell (USA), l'Institut de Formation et de Recherche Démographiques (IFORD, Cameroun), et l'Institut Supérieur des Sciences de la Population (ISSP, Burkina Faso). Il a aussi bénéficié des apports intellectuels de nos collègues Françoise Vermeylen, Gervais Béninguissé, Jean-François Kobiane, Silvère Konan Yao, Crispin Mabika Mabika, Gilles Gohi, et Latif Dramane, ainsi que du feedback de dizaines de participants aux divers ateliers menés dans le cadre de ce projet IUSSP. L'assistance de Serge Bahoken, Firmin Zinvi, Charles Mouté, et Justin Dansou, Sandrine Mbele, Habibou Ouedraogo ainsi que l'initiative des premiers participants à appliquer ces méthodes – Yédodé Ahokpossi, Jacques Emina, Edzengte Pascal, Kouam Nadège, Bela Foe Chrystelle... – ont été déterminantes dans les révisions. En plus des apports de fond, les relectures de plusieurs collègues sus-cités et de Pie-Marie Belibi, Stephanie Hissoak et Vincent Onguéné, ont permis d'améliorer la forme du document et le rendre plus accessible aux membres du réseau. Malgré toutes ces contributions, les auteurs restent seuls responsables des insuffisances résiduelles.

**Ithaca, Août 2017**

<b>Préface</b>	<b>3</b>
<b>Chapitre I Introduction</b>	<b>7</b>
<b>I.1 Qu'est-ce que la décomposition ?</b>	<b>8</b>
I.1.1 Définitions	8
I.1.2 Domaines d'application	8
I.1.3 Mode d'explication	9
<b>I.2 Décomposition contre autres méthodes</b>	<b>10</b>
I.2.1 Avantages et limites de la décomposition	11
I.2.2 Pertinence politique et applications	12
<b>I.3 Principaux types de décomposition</b>	<b>13</b>
<b>I.4 Notations</b>	<b>15</b>
<b>I.5 Organisation du document</b>	<b>16</b>
<b>Chapitre II Décomposition démographique</b>	<b>17</b>
<b>II.1 Décomposition démographique simple</b>	<b>17</b>
II.1.1 Type de problème	17
II.1.2 Illustration graphique	18
II.1.3 Illustration chiffrée	18
II.1.4 Formulation mathématique	19
II.1.5 Application pratique	20
II.1.6 Application au dividende démographique	22
<b>II.2 Décompositions démographiques dérivées</b>	<b>23</b>
II.2.1 Décomposition d'une différence	23
II.2.2 Décomposition d'une inégalité	23
II.2.3 Décomposition ordinale	25
II.2.4 Décomposition emboîtée	25
<b>Chapitre III Décompositions de régression</b>	<b>27</b>
<b>III.1 décomposition de régression simple</b>	<b>27</b>
III.1.1 Type de problème	27

III.1.2 Formulation mathématique	27
III.1.3 Application	28
<b>III.2 Décompositions de régression dérivées</b>	<b>29</b>
III.2.1 Régression curvilinéaire	29
III.2.2. Régression multivariée	29
III.2.3 Régression multi-niveaux	29
<b>III.3 Application au dividende démographique</b>	<b>30</b>
<b>Chapitre IV Décompositions mathématiques</b>	<b>31</b>
<b>IV.1 Décomposition mathématiques simple</b>	<b>31</b>
IV.1.1 Type de problème	31
IV.1.2 Formulation mathématique	31
IV.1.3 Application	32
<b>IV.2 Décompositions mathématiques dérivées</b>	<b>32</b>
IV.2.1 Extension de chaines	32
<b>Chapitre V Combinaison démographique + régression</b>	<b>35</b>
<b>V.1 Extension de l'effet de comportement</b>	<b>35</b>
V.1.1 Présentation générale	35
V.1.2 Illustration	36
V.1.3 Lien avec la méthode des comptes et transferts nationaux	37
<b>V.2 Extension de l'effet de composition</b>	<b>37</b>
<b>V.3 Double extension</b>	<b>38</b>
<b>Chapitre VI Combinaison démographique + mathématique</b>	<b>41</b>
<b>Chapitre VII Combinaison régression + mathématique</b>	<b>43</b>
<b>Chapitre VIII Résumé et conclusions</b>	<b>45</b>
<b>Références</b>	<b>47</b>

---

# Chapitre I

## Introduction

---



La recherche peut-elle guider efficacement l'action sociale ? Qu'il s'agisse de réduire le chômage, étendre la scolarisation, contenir l'inégalité sociale, ou réduire la mortalité, le défi quotidien du planificateur est de promouvoir des changements sociaux désirables. Relever ce défi requiert une bonne connaissance des leviers du changement social et, dans ce cadre, la recherche scientifique a une mission-clé à remplir.

Cette mission est loin d'être accomplie. Même si la technologie de recherche en sciences sociales s'est fortement développée au cours des trente dernières années, ces progrès ont surtout bénéficié aux analyses microscopiques. L'accès aux ordinateurs, à l'internet, ainsi que le développement de logiciels d'analyse statistique ont décuplé les possibilités de collecte, de transfert, et de traitement de données statistiques sur les ménages et individus. Grâce à cet arsenal informatique, les chercheurs peuvent aujourd'hui étudier en détail les comportements individuels et fonctionnement des ménages, mais ce détail à lui seul ne satisfera pas un analyste préoccupé par l'étude du changement social.

Un tel analyste renverra dos-à-dos une approche entièrement microscopique et une approche entièrement macroscopique. Même s'il peut s'appuyer sur le détail de l'analyse microscopique, il devra effectuer un effort supplémentaire pour convertir ces informations micro-transversales en conclusions macro-longitudinales. Depuis plus d'un demi-siècle, Robinson (1950) a tiré la sonnette d'alarme sur ce biais écologique, soulignant que les relations observées au niveau macroscopique ne reflètent pas forcément celles qui existent au niveau individuel, et *vice versa*.<sup>1</sup> Plus récemment, Thornton (2001) a souligné une dérive toute aussi pernicieuse, qu'il a dénommée « lecture transversale de l'histoire » et qui consiste à se fonder sur des comparaisons transversales entre pays pour tirer des conclusions sur les trajectoires de développement. L'idéal serait de combiner le micro et macro pour concilier le détail et la robustesse statistique de l'un avec la pertinence de l'autre. C'est précisément cette conversion que certaines méthodes de décomposition présentées ici permettent de réaliser.

---

<sup>1</sup> Par exemple, le fait qu'au niveau macroscopique, les taux d'obésité augmentent avec le niveau de richesse des pays n'implique pas nécessairement qu'à l'intérieur des pays, les riches aient une plus grande propension à l'obésité que les pauvres. Ou encore, le fait qu'une augmentation du taux de divorce élève le taux de criminalité n'implique pas nécessairement que les personnes issues de familles divorcées soient plus prédisposées à commettre des crimes.

## I.1 Qu'est-ce que la décomposition ?

### I.1.1 Définitions

Décomposer, c'est expliquer un tout à travers ses composants élémentaires ; c'est séparer un ensemble en ses divers éléments constitutifs ; c'est, en somme, disséquer, démonter, découper, désassembler, ou désagréger.

L'analyse de décomposition prend des connotations différentes dans diverses disciplines, mais l'idée de fond reste la même à savoir, partitionner un ensemble fonctionnel en ses constituants élémentaires. Ainsi en biologie animale, la décomposition serait un processus de désintégration de matière organique. En chimie, ce serait l'éclatement d'une molécule complexe en molécules plus simples ou en atomes. En physique, notamment pour étudier une trajectoire de projectile, elle pourrait être la projection et formulation de cette trajectoire sous ses plans élémentaires, notamment un plan vertical, horizontal, et profondeur.

En sciences sociales enfin, la décomposition peut servir à comprendre un changement social à travers la contribution relative de deux ou plusieurs facteurs (ou groupes) complémentaires. Par exemple, comment toutes les régions d'un pays contribuent-elles au changement historique dans la dépense annuelle du pays pour les soins de santé ? ou encore comment diverses classes sociales contribuent-elles au changement de la fécondité nationale ?

Pour saisir les particularités de la décomposition par rapport à d'autres méthodes d'analyse, les deux sections suivantes précisent le type de situations dans lesquelles elle s'applique, et son mode d'explication. La décomposition reste sous-utilisée, peut être en raison d'une méconnaissance de ses possibilités : elle est présentée de manière fragmentée dans la littérature et peu de manuels en donnent une introduction systématique.<sup>2</sup> Peu d'ouvrages expliquent comment enrichir les décompositions élémentaires et les combiner à d'autres types d'analyse. Pourtant, comme nous le montrerons, des combinaisons avec la régression ou l'analyse qualitative donneraient des résultats plus robustes que l'application de l'une ou l'autre méthode prise individuellement.

### I.1.2 Domaines d'application

La décomposition couvre un vaste champ de processus allant des sciences physiques aux sciences sociales. En sciences sociales, elle est particulièrement utile dans l'analyse du changement social, entendu comme « toute transformation, induite ou spontanée, de la structure, du fonctionnement ou de la performance d'une collectivité sociale. » Même si cette définition couvre aussi des changements qualitatifs, l'accent est mis ici sur des changements quantifiables qui résultent d'une agrégation de comportements individuels.

La décomposition est donc peu recommandée pour des études portant soit l'individu comme centre d'intérêt, soit sur des phénomènes de nature intrinsèquement sociétale c.-à-d. ne résultant pas de l'agrégation de comportement individuels. En dehors de ces deux restrictions, la méthode trouve un grand champ d'application aussi bien en démographie, économie, sciences politiques, ou sociologie (Kitagawa 1955 ; Oaxaca 1973 ; DasGupta 1993 ; Vaupel and Romo 2003). La seule exigence est que le processus à étudier soit quantifiable, agrégé, et historiquement évolutif.

**Quantification** : Le phénomène à étudier devrait être mesurable par une moyenne, un pourcentage, une mesure d'inégalité, ou un ratio. Sont donc exclus les phénomènes qualitatifs tels le passage d'un pays d'un régime politique autocratique à un régime démocratique, ou la transition d'une économie de subsistance au marché, ou

---

<sup>2</sup> Voir Kitagawa (1955), Das Gupta (1993) et Vaupel et Romo (2003) pour quelques exceptions.



encore le passage d'un système familial du mode étendu au mode nucléaire. Même dans ces cas, les dichotomies autocratie/démocratie, subsistance/marché, ou étendu/ nucléaire peuvent être ré-exprimées comme un continuum. Par exemple, la dichotomie subsistance/marché serait remplacée par une variable continue telle le pourcentage de personnes travaillant pour un employeur autre que familial. Avec une telle reformulation, des phénomènes à priori qualitatifs peuvent se prêter à une analyse de décomposition.

**Agrégation de comportements individuels :** Les sociologues distinguent deux types de caractéristiques sociales, selon qu'elles sont intrinsèques à la société entière (sans équivalent au niveau individuel) ou qu'elles résultent de l'agrégation de caractéristiques individuelles. Les premières incluent par exemple la loi d'un pays sur l'avortement. Les secondes incluent par exemple le niveau de fécondité d'un pays ou sa consommation moyenne de tabac. La société ne fume pas. Ce sont ses membres qui le font.

**Caractère graduel :** La décomposition s'applique difficilement aux processus rares ou soudains résultant d'une conjugaison accidentelle de multiples facteurs et se prêtant peu à un examen statistique. Par exemple, la décomposition serait peu indiquée pour analyser le nombre de victimes dans un tremblement de terre ou une épidémie soudaine de choléra.

### I.1.3 Mode d'explication

La décomposition tend à révéler les « sources » plutôt que les « causes » profondes du changement. Elle établit « *d'où provient le changement* » plus qu'elle n'élucide ce qui a causé le changement. Elle vise à « rendre compte, » à faire « une comptabilité des sources proches » d'un changement. Alors qu'une explication complète chercherait à élucider les causes ultimes (« pourquoi »), la décomposition identifie les sources proches (« par quoi » ou « par qui ») le changement intervient.

**Par quoi ?** Dans une décomposition démographique simple, les sources par lesquelles le changement survient incluront principalement deux effets complémentaires notamment a) un effet de composition qui reflète les changements dans la composition de la population ; b) et un effet de comportement qui, lui, reflète, le changement dans le comportement moyen des membres d'une ou plusieurs catégories sociales étudiés. Nous reviendrons amplement sur ces deux sources, mais nous pouvons déjà donner un premier exemple relatif au changement socio-culturel, notamment dans les attitudes au sujet du rôle politique des femmes.

Dans un pays africain francophone donné, l'opinion générale du public vis-à-vis des types de rapport à entretenir avec la France peut évoluer au fil du temps (par exemple, de 1980 à 2015) pour deux raisons. La première raison est liée à l'effet de composition : le pourcentage de personnes nées avant les indépendances (~1960) se réduit graduellement au fil du temps : L'on s'attend à avoir, en 2015, peu de survivants de l'ère pré-indépendance. Si l'on suppose que ces personnes, du fait de leur expérience directe de la colonisation, ont des avis nettement différents que leurs plus jeunes compatriotes, alors l'on s'attend à avoir une modification de l'opinion générale au fur et à mesure que cette génération pré-indépendance s'éteint. La seconde raison est un effet de comportement. Il pourrait y avoir, en 2015, des initiatives diplomatiques ou des événements nouveaux qui améliorent ou dégradent l'idée que ce pays (toutes générations confondues) se fait de la France.

**Par qui ?** En plus de rechercher par quoi le changement survient, l'analyste peut essayer de comprendre les acteurs ou groupes qui impulsent le changement. Elle recherchera les groupes sociaux, les régions, ou les catégories d'âge qui contribuent le plus à ce changement. Dans un cas comme dans l'autre, la décomposition établit moins les causes que les sources du changement.

## I.2 Décomposition contre autres méthodes

Imaginons qu'il soit aujourd'hui demandé à un jeune chercheur d'identifier les raisons d'un changement social, par exemple une baisse récente de mortalité. En toute probabilité, la décomposition ne serait ni le premier ni le seul outil envisagé. Notre chercheur pourrait recourir, par exemple, à une analyse *qualitative*, une analyse de tendances historiques ou une analyse de régression.

Si ce chercheur se tournait vers l'analyse qualitative, il/elle pourrait interviewer des personnes-ressource, organiser des discussions de groupe, ou consulter des archives pour identifier les événements historiques et acteurs susceptibles d'être à l'origine de ce changement. Y a-t-il eu de nouveaux projets et investissements, des découvertes scientifiques ou de nouveaux médicaments sur le marché, des événements inhabituels, des dirigeants exceptionnellement efficaces, une amélioration de l'approvisionnement en eau et en services médicaux ?

Si par contre notre chercheur se tournait vers l'analyse des tendances historiques, il/elle comparerait l'évolution historique de plusieurs processus sociaux pour discerner des co-évolutions éventuelles. Dans cette logique, les causes plausibles seraient à rechercher parmi les processus qui fluctuent dans le mêmes sens que – et précèdent – les changements de mortalité. Contrairement à l'analyse qualitative qui recherche souvent des explications locales et conjoncturelles, l'examen des tendances recherchera des facteurs structurels, récurrents, et généralisables dont l'apparition entraîne une réponse prévisible. Toutefois, cette analyse conserve un aspect qualitatif dans la mesure où les conclusions seront largement basées sur une évaluation visuelle – plutôt qu'un test statistique – de la co-évolution.

Plus formellement, notre chercheur pourrait recourir à une analyse de régression, et utiliser des tests statistiques pour identifier les corrélats du changement. Cette régression peut se baser sur des individus ou de larges agrégats. Si le problème posé est à l'échelle nationale, la macro-régression a l'avantage de la concordance d'échelle, mais elle masque malheureusement le détail des prises de décision individuelles. À l'opposé, la micro-régression donne ce détail mais elle n'a pas le regard macroscopique qui intéresse le décideur.

Les approches micro et macroscopique se complètent donc, l'une étant riche en détail et l'autre en pertinence. Cette complémentarité a encouragé deux formes d'intégration (Figure 1). Une première, allant du haut vers le bas, examine l'effet des politiques et conditions macroscopiques sur les individus, en utilisant par exemple la régression multi-niveaux (Luke 2004). Une deuxième, allant en sens inverse, explore comment le microscopique affecte le macroscopique, i.e., comment l'agrégation des comportements individuels produit le changement social. Une méthode emblématique de cette approche est la décomposition, et c'est elle qui retiendra notre attention dans cet ouvrage.

Plus généralement, les différences conceptuelles et pratiques entre les quatre approches d'analyse sont résumées au Tableau 1. Par exemple en termes de moteurs supposé du changement, ceux-ci – dans le cadre d'une décomposition démographique – résident fondamentalement dans le changement d'effectif ou de comportement des diverses sous-catégories de la population nationale. Par comparaison, l'analyse qualitative suppose souvent que le changement est précipité par une personne ou un événement-clé qui est particulier à la communauté étudiée et n'est donc pas forcément reproduit ailleurs.

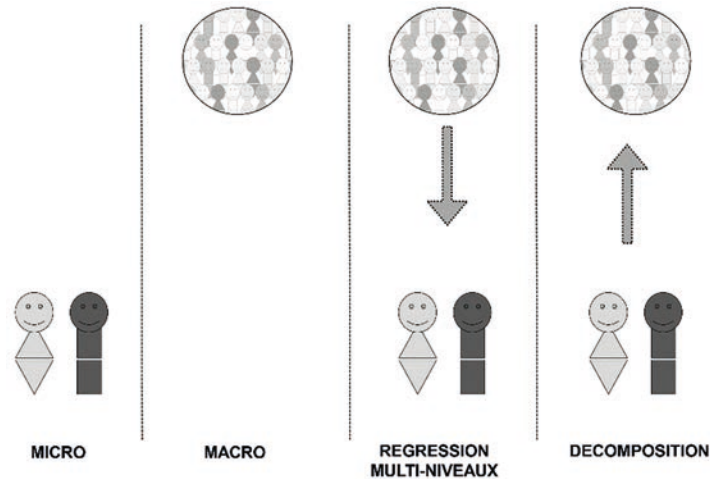


Figure 1. Quatre modalités d'intégration des analyses micro et macro.

Approche	Analyse qualitative	Analyse de Trend	Régression	Décomposition
Moteur Présumé du changement	Personne ou événements-clés	Changements sociaux concomitants	Facteurs statistiquement associés au phénomène	Changement dans la taille ou comportement de groupes sociaux-clés
Source de données	Focus groups, personnes ressource, archives	Statistiques nationales	Données d'enquêtes individuelles	Données d'enquêtes individuelles (groupes)
Question centrales	Pourquoi?	Quand?	Qui?	Comment?
Points Forts	Elucidation du processus	Séquence chronologique	Rigueur statistique, estimation de l'association statistique	Fiabilité, simplicité
Points faibles	1, 2, 4	4	3, 4	5

Points faibles: 1. Manque de représentation statistique; 2. Association statistique non mesurée; 3. Chronologie non établie; 4. Explications concurrentes non éliminées; 5. Processus causal non clarifié

Tableau 1. Comparaison de quatre approches d'explication du changement social.

### I.2.1 Avantages et limites de la décomposition

Par rapport à l'analyse causale, la décomposition présente l'inconvénient majeur de ne pas démontrer un lien de cause à effet. Dans notre exemple précédent sur la baisse de mortalité, la décomposition permet d'identifier les secteurs, les groupes ou les causes de décès au sein desquels la baisse a été particulièrement marquée, mais elle n'expliquera pas pourquoi ces baisses sont intervenues. Cette lacune est à relever dans la mesure où la mobilisation des ressources publiques doit idéalement se faire sur la base d'une bonne connaissance des liens de causalité. Ceci dit, en identifiant les sources changement, la décomposition offre des pistes utiles à l'action sociale. Bien plus, elle a l'avantage de la simplicité, flexibilité, facilité d'interprétation, et compatibilité avec d'autres méthodes.

**Simplicité.** La décomposition est simple à la fois dans son explication et son application. Dans son explication, elle identifie les sources proches du changement même si elle n'en révèle pas les causes profondes. Son application est toute aussi simple : elle ne requiert ni analyses statistiques sophistiquées, ni connaissances mathématiques approfondies, ni logiciels de pointe. L'essentiel des analyses pouvant se faire avec un tableur comme Excel. Les données nécessaires sont, dans beaucoup de cas, disponibles dans des rapports ou dans des systèmes de compilation et tabulation automatiques disponibles en ligne. Pour les données démographiques par exemple, les informations statistiques issues des Enquêtes Démographies Santé (EDS) ont été compilés et

sont disponibles sous une forme accessible sous StatCompiler ([www.statcompiler.com](http://www.statcompiler.com)). Pour toutes les autres statistiques socioéconomiques, les compilations actualisées de la Banque Mondiale (<http://data.worldbank.org/data-catalog/world-development-indicators>) sont un excellent point de départ. La disponibilité publique de telles données facilite les analyses tout en assurant la transparence des résultats.

**Flexibilité.** La décomposition est flexible dans la mesure où ses formes élémentaires peuvent être modifiées par le chercheur et donner naissance à des formes plus complexes et mieux adaptées aux besoins de l'utilisateur. Flexibilité et adaptation constituent à notre avis un des grands avantages de ces méthodes. Notre document illustrera quelques-unes de ces adaptations.

**Interprétation facile.** Les résultats de cette analyse sont d'une interprétation facile et intuitive. Contrairement à certaines statistiques issues de l'analyse de régression (intercept, logit, rapport de côtes), les résultats d'une décomposition sont exprimés dans un langage accessible. Ils indiquent simplement le pourcentage d'un changement social attribuable à un groupe donné ou un processus donné.

**Compatibilité.** La décomposition est compatible avec d'autres méthodes, telles les régressions classiques, les analyses multi-niveaux, la standardisation<sup>3</sup>, les projections démographiques ou même l'analyse qualitative. Elle peut être utilisée pour agréger les résultats des micro-régressions; elle peut se combiner aux analyses multi-niveaux; elle peut servir de prélude à une analyse qualitative : par exemple, après avoir identifié les sources « *par qui* ou *par quoi* » le changement survient, notre chercheur peut ensuite faire des analyses qualitatives complémentaires auprès de ces sources. Pour toutes ces raisons, la décomposition devrait intéresser un grand éventail de chercheurs aux orientations méthodologiques différentes. En clair, la décomposition ne remplace pas mais peut compléter d'autres méthodes de manière à améliorer la qualité des conclusions.

**Transparence.** Ceci est un atout majeur de la décomposition. Contrairement à ce qui se passe avec d'autres méthodes statistiques, le lecteur moyen peut aisément vérifier la cohérence interne des résultats d'une analyse de décomposition. Dans une régression multivariée par exemple, le lecteur est obligé de faire confiance au chercheur parce qu'il n'a aucun moyen de vérifier l'exactitude des coefficients de régression présentés, à moins qu'il n'ait accès à la même banque de données et au même programme d'analyse que le chercheur, et qu'il se donne la peine de retracer toutes les étapes parcourues depuis la codification jusqu'au modèle final. Ceci est rarement possible. Dans une analyse de décomposition en revanche, l'exactitude du résultat final peut être examinée à travers une lecture attentive des informations de départ, lesquelles peuvent être elles-mêmes vérifiées à leur source (rapports de synthèse, tableaux d'analyses descriptives, ou Compilateur statistique).

## I.2.2 Pertinence politique et applications

Il existe aujourd'hui en Afrique une forte demande pour des analyses fiables du changement social. Cette demande provient en partie des transitions démographiques et socioéconomiques en cours, mais reflète aussi des ambitions d'ingénierie sociale de plus en plus exprimées par les Etats. Les Objectifs du Millénaire de Développement et de Développement Durable illustrent ces ambitions. Sous l'égide des Nations Unies, les pays en développement se sont engagés à transformer leurs sociétés en concentrant l'action sur quelques objectifs quantifiés, notamment réduire la pauvreté de moitié, réduire la mortalité de 2/3, atteindre l'éducation universelle, etc., à l'horizon 2015 puis aujourd'hui à l'horizon 2030. Dans les pays à très faible revenu, il est impératif de faire un usage encore plus efficace des ressources disponibles, et ceci requiert une fine connaissance des moteurs du changement. Les méthodes de décomposition peuvent contribuer à améliorer cette connaissance.

---

<sup>3</sup> La décomposition est d'une certaine manière l'inverse des méthodes de standardisation qui, elles, visent plutôt à gommer les différences dans la composition, notamment par sexe et âge, des populations lorsqu'on compare des moyennes.

S'agissant des transitions sociales, l'Afrique est en mutation sur plusieurs fronts, notamment la fécondité, mais aussi d'autres processus démographiques (mortalité, nuptialité, migrations internationales, structures familiales et confiage d'enfants par exemple) et socioéconomiques (mobilité sociales, accès à l'éducation et l'emploi, mobilisation politique). Dans la mesure où ces diverses mutations peuvent fortement infléchir le cours du développement africain, toutes les méthodes de recherche susceptibles d'éclairer ces mutations sont importantes.

Compte tenu des intérêts de recherche des membres de notre réseau, nous ferons particulièrement référence au changement démographique et ses conséquences socioéconomiques. Nous consacrerons ainsi le chapitre final sur cette question de dividendes démographique, notamment dans les domaines de la croissance économique, éducation, et santé.

## I.3 Principaux types de décomposition

La décomposition n'est pas une méthode, mais en fait « un bouquet de méthodes » voisines. Parce que ces variantes sont traitées de manière disparate dans la littérature, leurs similarités et différences sont mal appréciées. Notre document clarifie les similarités entre les variantes majeures, celles-ci étant distinguées sur la base de quatre critères : le type de variable indépendante ; le type de variable dépendante, la fonction liant les variable dépendante et indépendante, et la complexité de l'analyse.

### *Critère 1 . Variable indépendante*

Selon que la variable indépendante est mesurée de manière nominale, quantitative, ou ordinale, nous distinguons les décompositions démographiques, de régression, et temporelle.

- Pour une décomposition démographique, la variable indépendante – ou variable de classification – est de type nominal. Elle regroupe les membres de la population en plusieurs catégories distinctes telles les provinces d'un pays, les groupes d'âge, l'ethnie, le statut matrimonial etc. ; notons néanmoins qu'il est possible de prendre une variable ordinale (ex. statut socioéconomique) et la traiter de manière nominale, c.-à-d., sans tenir compte de l'ordre implicite des catégories.
- Pour une décomposition temporelle, l'on tient compte de l'ordre explicite qui existe entre les catégories. En prenant par exemple le groupe d'âges comme variable, nous tiendrons compte du fait que les personnes âgées de 15-19 ans sont plus jeunes que les 20-24 ans lesquelles sont elles-mêmes plus jeunes que les 25-29 ans. Toutefois, l'on ne quantifiera pas la différence d'âge entre ces groupes. La seule information utilisée ici est l'ordre hiérarchique entre catégories.
- Dans une décomposition de régression enfin, la variable indépendante est quantitative (ex., le nombre d'années d'éducation, le nombre de frères et sœurs, le revenu en FCFA). L'objectif sera généralement d'examiner (avec une analyse de régression) la relation entre cette variable indépendante et la variable dépendante, ainsi que le changement historique dans cette relation. Par exemple, si l'on veut étudier les sources de changement dans la longévité des individus (variable dépendante), l'on peut poser que cette longévité dépend du niveau de revenus mais que cette relation elle-même peut changer dans le temps. Les changements dans la longévité dépendront donc à la fois des changements dans les revenus mais aussi des changements dans la relation longévité-revenus.

### *Critère 2 . Variable dépendante*

Pour les décompositions élémentaires les plus simples, la variable dépendante est une moyenne. Si cette variable est autre qu'une moyenne (par exemple une inégalité), nous aurons une forme de décomposition simple dérivée.

### *Critère 3 . Relation fonctionnelle*

Le critère central ici est la relation entre les variables indépendantes et dépendante. Le type et la complexité de cette relation fonctionnelle varient (voir figure 2) L'on peut ainsi distinguer trois types de relations, notamment.

- **relation démographique** : en gros, la valeur de Y de l'ensemble du pays est une moyenne obtenue par pondération démographique des  $y_j$  qui prévalent dans les diverses sous-populations du pays ;
- **relation statistique** : précisément une relation de régression entre Y et X ;
- **relation mathématique** : dans ce cas, les variables dépendante et indépendantes sont liées par une relation mathématique plus ou moins simple (qui implique un quotient, somme, produit, logarithme). Contrairement à la relation statistique, c'est une relation exacte qui ne varie pas selon les pays ou situations. Les seuls éléments qui varient sont les valeurs de ces variables.

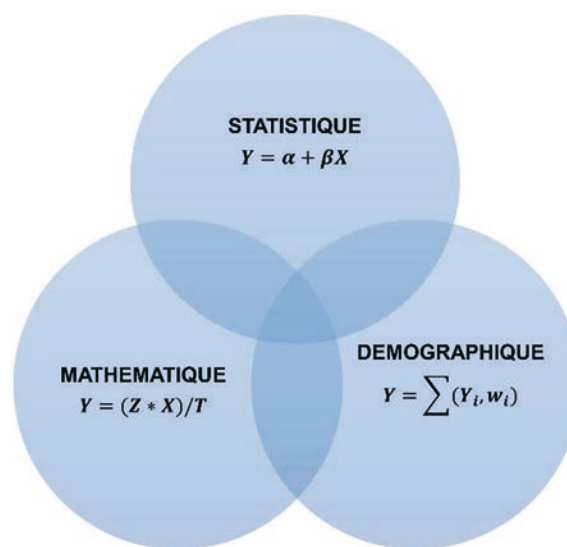


Figure 2. Types de décomposition élémentaires.

Chacune de ces relations fonctionnelles peut être plus ou moins complexe. Pour une régression par exemple, la relation peut simplement inclure une seule variable indépendante avec une spécification linéaire. Elle devient compliquée si l'on envisage un modèle à plusieurs variables indépendantes, des formes non linéaires, ou des analyses multi-niveaux.

### *Critère 4 . Degré de complexité*

Ce critère distingue les décompositions simples des décompositions avancées ou mixtes qui combinent deux ou plusieurs formes élémentaires. Les formes élémentaires (Figure 2) représentent donc les blocs élémentaires qui permettent de construire les formes plus complexes. Ainsi, comme indiqué en Figure 3, une décomposition élémentaire peut engendrer une décomposition dérivée (lorsque la variable dépendante ou la relation fonctionnelle est rendue plus complexe). La combinaison de deux ou trois formes élémentaires, elle, engendre une décomposition mixte. L'organisation de cet ouvrage est calquée sur cette typologie. Il commence par les formes élémentaires, puis les formes dérivées, et enfin les formes mixtes.

Notons enfin que des formules de décomposition spécialisées ont été développées pour étudier des problèmes particuliers, notamment l'espérance de vie (Vaupel 2003) ou encore la discrimination dans l'emploi (Oaxaca 1973). Ces formules ne soient pas abordées dans cet ouvrage, mais il est possible de les intégrer dans la taxonomie adoptée ici.

## I.4 Notations

Pour faciliter la présentation, les conventions de notation suivantes seront utilisées pour les variables, indices et changement historique.

**Variables.** Suivant une convention bien établie en sciences sociales, les variables dépendantes et indépendantes sont indiquées par les lettres Y et X respectivement. En plus, ces notations sont faites en lettres majuscules ou minuscules selon que les variables sont mesurées au niveau macroscopique ou microscopique, respectivement. Y désignera ainsi une variable dépendante au niveau macro; y désignera une variable dépendante au niveau micro ou méso; X désignera une variable indépendante au niveau macroscopique ; et x, enfin, désignera une variable indépendante au niveau micro ou méso. Ainsi, dans une étude sur les effets de l'éducation sur la mortalité par exemple, Y désignerait le taux de mortalité au niveau national tandis que y pourrait désigner le taux de mortalité dans une sous-population, par exemple les personnes âgées de 15 à 19 ans. Dans cette étude, X pourrait désigner le niveau d'éducation au niveau national, alors que x indiquerait le niveau d'éducation dans une sous-population, ex. parmi les personnes âgées de 15 à 19 ans.

**Pondération.** La lettre  $w$  mesure le facteur de pondération et est surtout utilisée dans les décompositions démographiques ou les décompositions avancées utilisant une élaboration démographique; Dans une telle décomposition,  $w$  représente le pourcentage de la population nationale appartenant à une catégorie donnée.

**Paramètres de régression.** Une décomposition de régression inclura les paramètres de régression usuels notamment

- $\alpha$ , l'intercept c.-à-d. la valeur de Y(ou y) lorsque X(ou x) = 0 ;
- $\beta$ , l'accroissement marginal de Y (ou y) lorsque X (ou x) augmente d'une unité ;
- $e$ , le terme d'erreur. Plusieurs valeurs de  $\beta$  ou de  $e$  seront intégrées dans le cadre de régressions plus compliquées, par notamment des régressions non linéaires ou multiniveaux.

**Indices.** La présentation utilisera aussi les indices suivants ,

- $j$  indexe les groupes; par exemple  $y_j$  dénote la valeur de la variable dépendante pour le groupe j, alors que  $x_j$  indiquera la valeur de la variable indépendante pour le même groupe ;
- $t$  indexe le temps; par exemple  $Y_t$  indiquera la valeur de la variable dépendante pour une année donnée et pour toute la population (ex. mortalité moyenne au Sénégal en 1990) ;
- $a$  indexe l'âge; par exemple  $Y_a$  indiquera la valeur de la variable dépendante pour un âge donné (ex. taux de fécondité spécifique au groupe d'âge concerné) ; dans ce contexte, le terme  $a+$  indiquera tous les âges supérieurs à l'âge a ; dans une analyse de fécondité,  $Y_{15-19+}$  indiquera la fécondité moyenne pour tous les âges au-dessus de 19 ans.

**Changement historique.**  $\Delta$  indique le changement dans le temps. Par exemple,  $\Delta Y$  représente le changement historique du phénomène à étudier, est égal à la différence entre les valeurs de Y pour deux années (par exemple  $Y_{t+1} - Y_t$ ).

**Moyennes.** Les analyses distingueront deux types de moyennes, selon qu'elles sont transversales ou historiques. Une moyenne transversale indique la valeur moyenne au sein de la population à un moment t donné (*calculée pour plusieurs groupes à un moment donné*). Puisque la variable dépendante dans la plupart des décompositions simples est une moyenne, ces moyennes transversales seront tout simplement dénotées Y(ou y). Les moyennes historiques en revanche représentent la moyenne de la variable entre deux périodes d'étude (*calculées pour un seul groupe mais deux périodes*). Elles seront marquées par une barre au-dessus des lettres. Ainsi,  $\bar{Y}$  indique la moyenne entre deux périodes pour la valeur de la variable dépendante au niveau national. De la même manière,  $\bar{y}_j$  indiquera la valeur moyenne entre deux années de la variable dépendante pour un groupe donné.  $\bar{Y} = (Y_{t1} + Y_t)/2$ .

## I.5 Organisation du document

Ce document étant introductif, sa présentation se veut simple et graduée. Il commence par une définition de la décomposition et sa comparaison rapide avec d'autres méthodes connues, notamment la régression. Il donne ensuite une typologie des méthodes de décomposition, selon le type de données et la complexité de l'analyse. Selon que la variable indépendante est nominale, quantitative or ordinale, nous distinguons les décompositions démographique, de régression ou temporelle.

Ces types élémentaires peuvent ensuite être légèrement modifiés pour produire des décompositions dérivées. Une fois de plus, la modification peut concerner soit le type de variable dépendante (par exemple, une mesure de l'inégalité au lieu d'une moyenne) soit la fonction reliant la variable indépendante à la variable dépendante (par exemple, une relation curvilinéaire ou multivariée au lieu d'une simple relation linéaire).

Enfin les types élémentaires peuvent être combinés de plusieurs manières pour produire des décompositions avancées (Figure 3). Nous présentons de telles décompositions avancées, avec des exemples tirés de travaux récents menés par des membres du réseau. En dehors de nos exemples, d'autres combinaisons peuvent être développées en fonction des besoins spécifiques de chaque étude et de l'imagination du chercheur.

Toujours dans un but pédagogique, l'ouvrage combine description verbale, formule mathématiques, exemples illustratifs, tableaux annotés, graphiques, voire dessins et caricatures. À travers cette pédagogie diversifiée, nous espérons atteindre un plus grand nombre de lecteurs voire de permettre à chaque lecteur d'appréhender chacun des concepts présentés sous plusieurs angles.

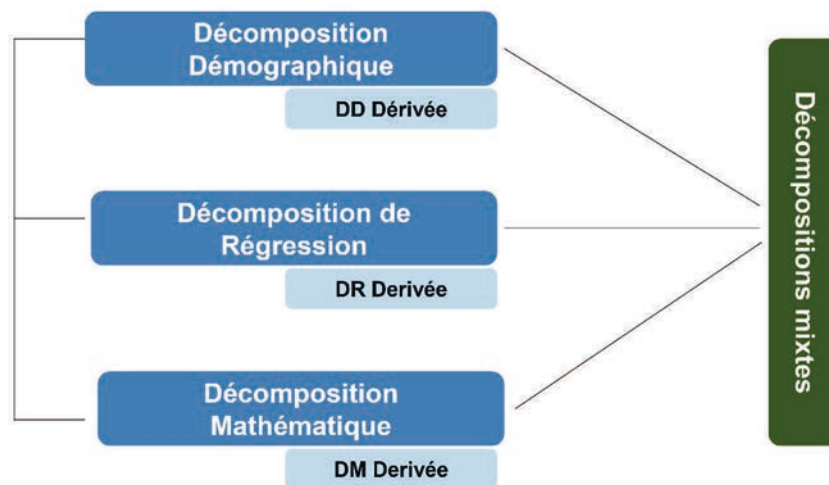


Figure 3. Décompositions dérivées.

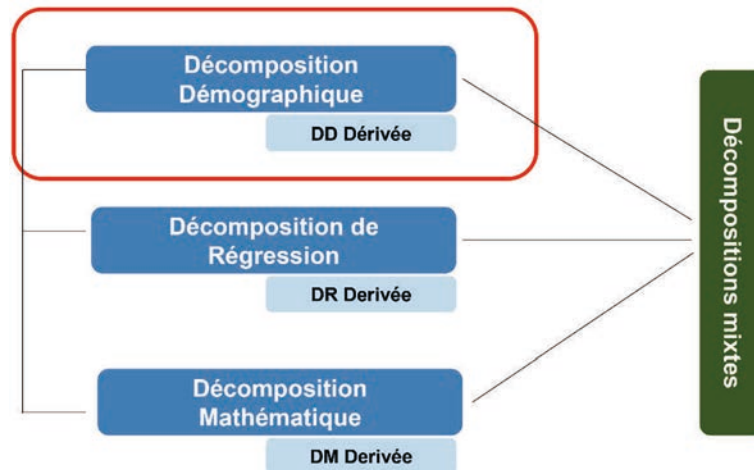


---

# Chapitre II

## Décomposition démographique

---



### II.1 Décomposition démographique simple

#### II.1.1 Type de problème

Dans cette première décomposition, tout processus à l'échelle nationale ( $Y$ ) est conçu comme une agrégation des comportements de plusieurs sous-populations ( $y_j$ ) avec une pondération par les effectifs relatifs de ces sous-populations ( $w_j$ ). En clair,

$$Y = f(y_j, w_j) \quad [\text{II.1}]$$

Ainsi par exemple, le taux de mortalité d'un pays est la moyenne pondérée des taux observés dans les différentes régions ou alors dans les différentes catégories socioéconomiques. Dans cette formulation,  $Y$ , la variable dépendante, doit satisfaire le critère habituel d'être quantifiable.  $X$ , la variable indépendante, doit remplir des critères de d'exhaustivité, distribution, variabilité, et pertinence.

**Exhaustivité** : les diverses catégories  $j$  doivent couvrir l'ensemble de la population tout en étant mutuellement exclusives. En d'autres termes, chaque membre de la population doit appartenir à une et une seule catégorie  $j$ .

**Distribution** : le nombre de catégories pour une bonne variable de classification ne doit être ni trop petit ( $> 2$ ) ni trop grand (une dizaine). Un nombre trop réduit de catégories produit une décomposition grossière. À l'inverse, un nombre trop grand émiette les effectifs. Ainsi, ni le sexe des individus (trop petit) ni leur âge (trop grand s'il est mesuré en années individuelles) ne sont généralement recommandés comme variables de classification.

**Variabilité** : les effectifs relatifs de ces catégories devraient fluctuer dans le temps. Une variable figée dans le temps ne permet pas, par définition, d'estimer la part du changement social liée au changement dans la composition de la population (effet de composition). Dans ce contexte, le sexe sera, une fois de plus, une mauvaise variable de classification à moins que le sex-ratio de la population étudiée soit susceptible de varier rapidement dans le temps.

**Pertinence** : une bonne variable de classification devrait avoir un lien théorique avec le phénomène étudié. À défaut de cela, elle peut être choisie selon sa pertinence politique. La région serait intéressante à ce titre dans un contexte d'administration décentralisée où les programmes se mettent en place au niveau régional.

## II.1.2 Illustration graphique

Comme indiqué en introduction, l'objet de la décomposition est d'identifier les sources de changement, celles-ci pouvant être substantives (« par quoi » le changement arrive) ou sociologiques (« par qui » le changement arrive).

La figure ci-dessous offre une illustration tirée d'une étude sur la parité en éducation. Les cinq carrés représentés dans la partie A de ce diagramme montrent l'évolution de cette parité dans le pays tout entier, les couleurs de plus en plus sombres indiquant un rapprochement de la parité. Parti d'une forte inégalité dans la scolarisation entre garçons et filles (carré blanc à l'extrême gauche), le pays a graduellement atteint la parité (carré noir à l'extrême droite). La question maintenant est de savoir comment cette parité a été atteinte, notamment quelle a été la contribution relative des différentes classes socioéconomiques (SSE #1 à SSE #4). La partie B de ce diagramme propose deux scénarios opposés.

Dans le premier (B2), convergence verticale, le changement a été initié par le quartile supérieur de la population (SSE1) avant de s'étendre graduellement au reste de la population. En année 2 par exemple, la parité était déjà atteinte parmi ce quartile alors que le reste de la population montrait encore une forte inégalité de scolarisation entre garçons et filles. Dans le second scénario (B1) par contre, la convergence est dite horizontale : toutes les catégories sociales ont évolué à la même vitesse. À la question de savoir « par qui » le changement arrive, les deux scénarios de cette figure proposent des réponses différentes : dans le premier cas, le changement arrive par le haut, alors que dans le second cas, tous les groupes ont contribué de manière synchrone.

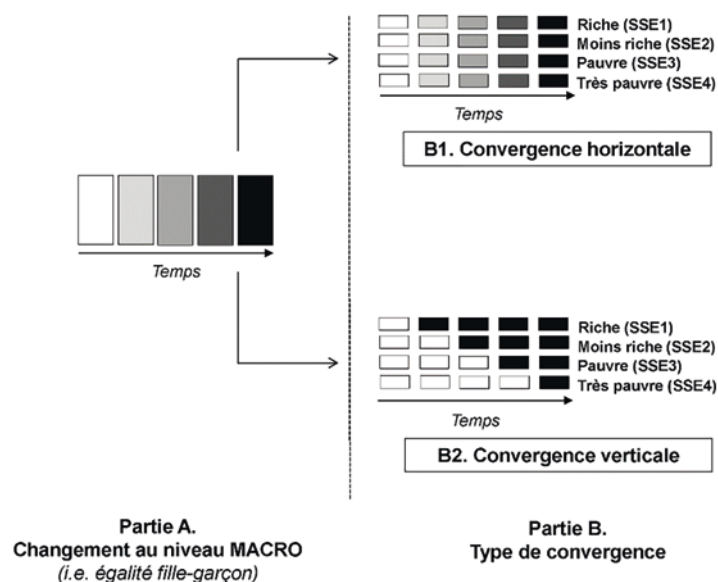


Figure 4. Convergence verticale versus horizontale dans l'éducation.

## II.1.3 Illustration chiffrée

Contrairement à la figure précédente qui montre les sources 'par qui' le changement arrive, le tableau 2, ci-après montre les sources 'par quoi' un changement arrive. Il illustre notamment la différence entre effets de composition et de comportement. Imaginons un pays avec un niveau initial de revenu mensuel (en milliers de FCFA) 142.5. Ce niveau national représente la moyenne pondérée des revenus au sein des différentes classes

sociales qui constituent la population nationale, des classes les plus hautes/riches aux classes les plus pauvres. Les deux colonnes sous l'année 1 montrent, pour chaque catégorie sociale, le revenu moyen et l'effectif relatif de cette classe. Par exemple, en année 1, la catégorie des plus pauvres constituait 20 % de l'effectif de la population et elle gagnait en moyenne 50,000 FCFA.

Pour l'année 2, le tableau montre deux scénarios de changement. Même s'ils conduisent tous les deux au même résultat final – une hausse du revenu moyen de 142.5 à 152.9 – ces deux scénarios sont qualitativement très différents. Sous le scénario 1, les revenus moyens des différentes classes n'ont pas varié ; ce qui a varié (et a ainsi induit la hausse des revenus) c'est l'effectif relatif des différentes classes, avec notamment la couche des pauvres qui est passée de 20 % à seulement 15 % de l'effectif national. Sous le scénario 2 par contre, l'on n'observe pas de changement dans la composition; les revenus moyens de certains groupes, par contre, ont évolué. Les plus riches (350 en moyenne) se sont encore plus enrichis (400) et c'est ce changement qui explique entièrement la hausse du revenu national moyen.

Les deux scénarios du Tableau 2 représentent des cas de figure extrêmes (100 % composition OU 100 % comportement). En pratique, les effets de composition et de comportement opèrent souvent simultanément. Par exemple, la composition peut expliquer 30 % du changement et le comportement expliquerait le reste. L'intérêt de la décomposition est précisément de quantifier les contributions relatives de ces deux processus.

Période 1		
Classe économiques	Revenu moyen	% de la population
La + haute	350	10%
Haute	200	15%
Moyenne	150	30%
Basse	90	25%
La +basse	50	20%
<b>Moyenne</b>	<b>142.50</b>	

Période 2 : SCENARIO 1		
Classe économiques	Revenu moyen	% de la population
La + haute	350	10%
Haute	200	20%
Moyenne	150	35%
Basse	90	20%
La +basse	50	15%
<b>Moyenne</b>	<b>153.93</b>	

100 % Change de composition

Période 2 : SCENARIO 2		
Classe économiques	Revenu moyen	% de la population
La + haute	400	10%
Haute	225	15%
Moyenne	155	30%
Basse	90.7	25%
La +basse	50	20%
<b>Moyenne</b>	<b>153.93</b>	

100 % Changement de comportement

Tableau 2. Effets de composition ou de comportement. Cas extrême.

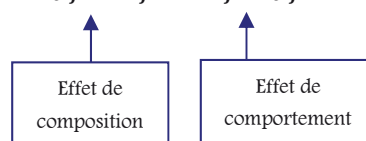
## II.1.4 Formulation mathématique

Dans ce premier cas, nous nous intéressons à une moyenne nationale ; Y sera donc formulée comme une moyenne pondérée (par  $w_j$ ) des valeurs des sous-populations individuelles ( $y_j$ ).

$$Y_t = \sum w_{jt} * y_{jt} \quad [II.2]$$

De cette formule, le changement national peut être décomposé :

$$\Delta Y = \sum \bar{y}_j * \Delta w_j + \sum \bar{w}_j * \Delta y_j$$



Cette décomposition de base identifie deux sources de changement. La première, **composition**, reflète un changement dans la représentation relative des diverses couches sociales de la population. Un accroissement

différencié des effectifs affecte la moyenne nationale par effet mécanique de pondération, dès lors que certains groupes sont plus vulnérables au processus en question.

La seconde source, **le comportement**, est moins mécanique. Elle indique un changement réel de la mortalité au sein d'un ou plusieurs groupes. Si la mortalité d'un groupe augmente, toutes choses étant égales par ailleurs, la mortalité nationale augmentera. Ce ne sont pas les effectifs relatifs qui changent ici, mais les niveaux de mortalité au sein de divers sous-groupes de la population.

## II.1.5 Application pratique

En pratique, une analyse de décomposition comprend quatre étapes majeures, notamment de définition du problème, calcul des informations de base et du changement, décomposition du changement, et présentation/discussion des résultats.

**Définition du problème.** Ici, le chercheur précise la nature de la variable substantive (dépendante), la variable de classification (indépendante), et la période considérée. Dans l'exemple proposé ci-dessous, la variable substantive est la mortalité infanto-juvénile ; la variable de classification est le statut socioéconomique ; et la période considérée est l'intervalle entre 1991 et 2001 en contexte camerounais.

**Calcul des informations de base.** Conformément à la formule (1), les seules informations nécessaires sont, pour chacune des **catégories de la variable de classification** (ici le SSE), le pourcentage de personnes ( $w_j$ ) et la valeur moyenne de la variable dépendante ( $y_j$ ). Ces données doivent être disponibles pour l'année de départ et l'année finale. Les procédures statistiques pour obtenir ces résultats sont simples : une analyse de fréquences (pour les  $w_j$ ) et une comparaison des moyennes (pour les  $y_j$ ).

En fait, ces informations de base peuvent parfois s'obtenir directement à partir de rapports d'études ou de sites qui génèrent des tableaux préfabriqués (e.g., [www.Statcompiler.com](http://www.Statcompiler.com)). En plus de faciliter l'obtention des données préliminaires, ces tableaux préfabriqués sont fiables s'ils sont générés directement par l'institution qui a collecté ces informations et maîtrise a priori la qualité des données. Par contre, les tableaux préfabriqués sont parfois inadaptés aux besoins précis du chercheur.

### Analyse proprement dite

Une fois l'information de base disponible, l'on applique tout simplement la formulation présentée en équation [3]. Vu le caractère répétitif des calculs, l'utilisation d'un tableur comme Excel est recommandée, mais en fonction des aptitudes du chercheur, une programmation sous un logiciel statistique est aussi envisageable. Le tableau 3 résume les données de base nécessaires, les calculs, et la présentation des résultats.

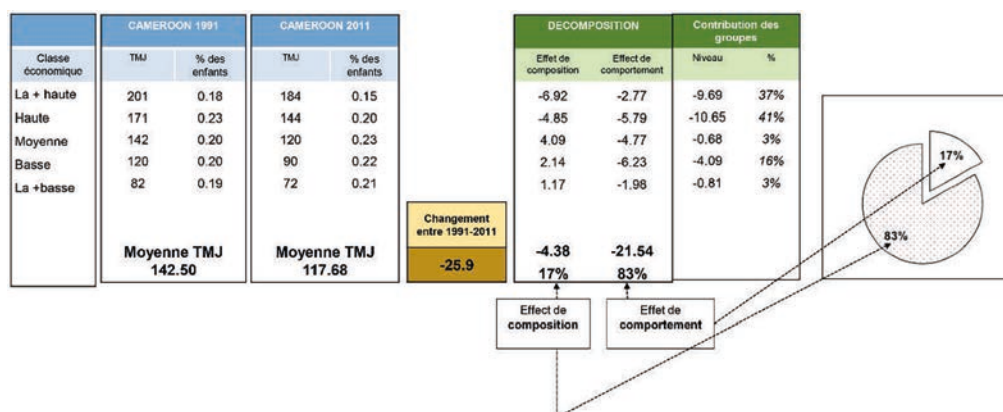


Tableau 3. Décomposition du taux de changement dans la mortalité infanto-juvénile, Cameroun 1991-2011.

**Information de base.** Les quatre premières colonnes contiennent l'information de base requise pour l'analyse, notamment le taux de mortalité infantile pour chacune des classes sociales et les pourcentages d'enfants inclus dans ces classes, et ceci pour l'année de départ (1991) et pour l'année finale (2011).

**Calcul du changement.** En partant de ces informations, l'on calcule la moyenne pour chacune des années et le changement dans les taux de mortalité infantile au niveau national : ceci est tout simplement la différence entre la moyenne pondérée nationale en année 1991 (143.6) et en année 2011 (117.7). Cette différence ici est de -25.9.

**Décomposition du changement.** L'étape finale consiste à expliquer cette baisse de mortalité, notamment les parts respectives dues à la composition et au comportement. Elles s'obtiennent en appliquant la formule [2]. Les résultats, inscrits au tableau suggèrent que dans ce cas particulier, 17 % de cette baisse était liée au changement de composition. Notez par exemple comment le pourcentage d'enfants dans les familles les plus pauvres a baissé de 18 à 15 %, alors que celui des familles les plus riches s'est élevé de 19 à 21 %. Le reste du changement (83 %) est lié à un effet de comportement. Notez, en particulier comment les taux de mortalité ont baissé dans toutes les classes sociales, y compris les plus pauvres.

### **Présentation des résultats**

Pour un auditoire scientifique, la présentation des résultats peut s'appuyer sur un tableau similaire au tableau II.2. La discussion portera sur deux groupes de pourcentages :

**Sources de changement I** (« Par quoi »). Les pourcentages en bas de tableau de tableau qui indiquent « par quoi » est venu le changement, notamment les contributions relatives des effets de composition (17 %) et de comportement (83 %).

**Sources de changement II** (« Par qui »). Tandis que les effets de composition et comportement indiquent « par quoi » le changement est survenu, cette décomposition peut aussi clarifier « par qui » le changement survient. Ici, la discussion s'attardera sur les contributions des diverses catégories sociales. Ces calculs, s'obtiennent par addition des effets de composition et comportement de chacune des catégories sociales. Par exemple, la contribution de la classe économique la « plus basse » au changement est égale à la somme de ses effets de composition (-6.92) et de comportement (2.77) pour un total de -9.69, soit 37 % du changement total. Les autres groupes, par ordre de revenus croissant, auront contribué à hauteur de 41 %, 3 %, 16 %, et 3 % respectivement.

Dans une analyse de décomposition, la somme de ces contributions est toujours égale à 100 %. Néanmoins, les contributions individuelles peuvent être négatives (moins de 0 %) ou supérieures à 100 %. Un pourcentage négatif indique une contribution allant dans le sens opposé du changement général. Par exemple, un effet de composition négatif dans un contexte de mortalité en baisse signifie que cet effet de composition concourt à accroître la mortalité, c.-à-d., il contrarie la tendance dominante. Par contre, un pourcentage supérieur à 100 % indique que la composante concernée explique entièrement le changement voire plus : le changement général aurait été encore plus fort s'il n'avait pas été contrarié par des effets allant dans le sens inverse.

Pour un auditoire non-scientifique, une présentation graphique basée sur des diagrammes 100 % ou en camembert serait recommandée. De tels diagrammes résument succinctement les résultats, et identifient clairement la composante dominante, c.à.d, celle ayant le plus fort pourcentage.

### **Implications politiques**

Les politiques à recommander varieront selon qu'un changement social reflète un effet de composition ou de comportement. Si la mortalité s'accroît par effet de composition, la réponse appropriée est de cibler les couches vulnérables de la population. Par contre, si ce changement est dû par un effet de comportement, le ciblage

devient moins indiqué et l'on envisagerait une politique de masse d'hygiène publique, d'éducation, ou de vaccination pour toutes les mères ou leurs enfants. Par la suite, nous verrons comment affiner les recommandations politiques avec des décompositions plus détaillées.

## II.1.6 Application au dividende démographique

Cette première décomposition, et notamment l'importance de l'effet de composition, peut s'appliquer dans l'analyse du dividende démographique selon la méthode des comptes de transferts nationaux (NTA) (Mason et Lee 2005). La méthode repose sur une reconnaissance de la variation systématique des revenus, consommation et épargne selon l'âge (Figure 5). En particulier, le solde revenus/consommation tend à être négatif parmi les jeunes et les personnes âgées, même si les âges-seuil varient d'un pays à un autre. Ce solde est par contre positif au sein de la population adulte. Il va de soi que si la proportion de la population nationale contenue entre ces deux âges augmente, le solde national sera plus positif qu'auparavant.

Le parallèle entre cette approche et la décomposition démographique est évident. Il suffit de prendre les groupes d'âges comme variable de classification et le solde revenu-consommation comme variable substantive pour entrevoir comme une décomposition démographique simple pourrait être appliquée à l'analyse du dividende démographique (voir tableau 4 ci-après). Dans ce tableau, l'effet de composition n'est autre que le dividende démographique (tout au moins la composante mécanique du dividende démographique). L'avantage de la méthode de décomposition ici se situe à trois niveaux : 1) elle ne suppose pas un profil de consommation et de revenu qui reste fixe dans le temps ; 2) elle quantifie le dividende non seulement de manière absolue, mais aussi de manière relative par rapport aux changements dans les revenus et la consommation ; 3) elle permet aussi de comparer les contributions des différents groupes d'âge.

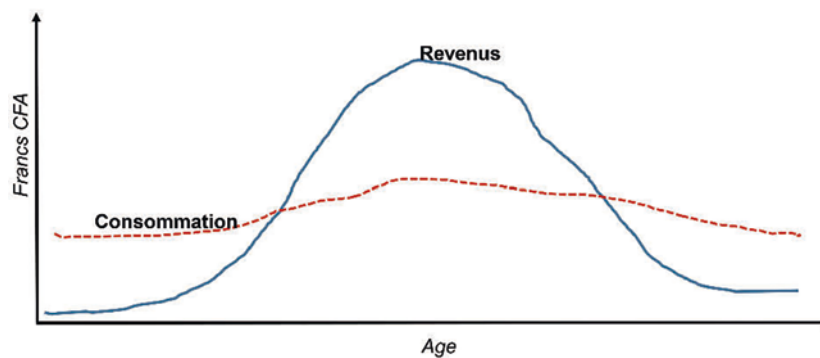


Figure 5. Profils de revenus et de consommation suivant l'âge. Données hypothétiques.

Groupe d'âge	Période 1		Période 2		DECOMPOSITION		Contribution des groupes d'âges	
	Revenu net (\$)	% des individus	Revenu net (\$)	% des individus	Effet de composition	Effet de comportement	Niveau	%
0-14	-15	.30	-20	.20	1.75	-1.25	.50	4%
15-29	-3	.25	-5	.20	.20	-.45	-.25	-2%
20-49	35	.25	50	.30	2.13	4.13	6.25	48%
50-64	50	.15	75	.20	3.1	4.38	7.5	58%
65+	-10	.05	-15	.10	-.63	-.38	-1.00	-8%
	<b>AVERAGE REVENUE \$10.50</b>		<b>AVERAGE REVENUE \$23.50</b>		<b>6.58</b>	<b>6.43</b>		
					<b>51%</b>	<b>49%</b>		

Changement total de revenu

**\$13.00**

Dividende démographique = Gain de croissance dû au changement de la structure par âge de la population

Tableau 4. Estimation du dividende démographique selon la méthode de décomposition.

## II.2 Décompositions démographiques dérivées

### II.2.1 Décomposition d'une différence

Au lieu d'examiner le changement historique, notre chercheur pourrait s'intéresser à la différence entre deux pays ou deux provinces au sein d'un pays. Fort heureusement, la même démarche s'applique. Il suffira d'avoir, pour chaque province et pour chacune des deux années-butoir, la composition de la population selon la classe sociale et le taux de mortalité pour chacun de ces groupes. Les calculs d'effectuent de la même manière que dans le cas précédent.

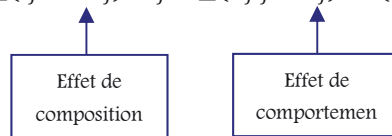
### II.2.2 Décomposition d'une inégalité

Au lieu de la valeur moyenne de la variable substantive (ex, le niveau de mortalité infantile), l'on s'intéresse ici à l'inégalité nationale dans la mortalité, qu'elle soit mesurée par rapport à la région, au SSE, ou toute autre dimension. Nombre d'études s'attèlent ainsi à étudier l'inégalité économique mondiale, comme une fonction des différences dans les populations nationales ( $p$ ) et les revenus par habitant ( $j$ ). Si l'on prend la déviation logarithmique moyenne (MLD) comme mesure d'inégalité, alors le niveau d'inégalité au niveau mondial peut s'exprimer comme suit :

$$I_t = \sum w_{jt} * \ln(1/r_{jt}) \quad \text{[II.4]}$$

Dans ce contexte, le changement historique dans l'inégalité mondiale est :

$$\Delta I = \sum (\bar{r}_j - \overline{\ln r_j}) \Delta w_j + \sum (\overline{w_j r_j} - \bar{w}_j) \Delta \ln(r_j) \quad \text{[II.5]}$$



L'exemple choisi ici montre le changement historique de l'inégalité économique entre les pays d'Afrique de 1985 à 2005 (Tableau 5). Entre ces deux années, l'inégalité économique, telle que mesurée par le Mean Log Deviation (une mesure courante de l'inégalité) dans le PIB par habitant s'est accru de 0.391 à 0.425, soit d'environ 9 %. Pour comprendre les sources de cette aggravation de l'inégalité, notre chercheur peut recourir à l'analyse de décomposition. Dans ce cas précis, il pourra examiner les contributions des effets de composition et de comportement, respectivement. Il pourra également observer les contributions individuelles des pays, notamment ceux qui auront le plus contribué à accroître l'inégalité, et ceux au contraire qui auront contribué à la diminuer. Les résultats de cette analyse sont consignés au tableau ci-contre.

Les quatre premières colonnes après la liste des pays montrent les données sur le PIB par habitant ainsi que la population totale de chaque pays pour les années 1985 et 2005, respectivement. Ce sont les données brutes de départ. Pour calculer l'inégalité, ces données absolues doivent être changées en valeurs relatives : la population totale par exemple est transformée en part de la population africaine vivant dans le pays et elle s'obtient en divisant la population de chaque pays par la population totale de l'Afrique pour cette année. De même, le PIB est transformé en PIB relatif et celui-ci c'obtient en divisant le PIB du pays par la moyenne pondérée du PIB africain pour cette année. Les pays ayant des valeurs de PIB relatif supérieures à 1 sont plus riches que la moyenne et ceux en dessous de 1 sont plus pauvres que la moyenne. Sur la base des valeurs relatives, l'inégalité est ensuite calculée et les valeurs consignées au bas des deux colonnes suivantes, donnant des valeurs de 0.39 et 0.43 respectivement, pour un changement d'environ 0.04. Ce changement total est ensuite décomposé en termes de son effet de composition (- 18 %) et de comportement (118 %).

En clair, la croissance démographique différentielle dans les pays africains aura contribué à réduire l'inégalité économique, même si cette contribution est relative modeste en termes absolus par rapport à la contribution des différences de croissance. Comme nous le verrons par la suite, il est possible de réaliser une décomposition beaucoup plus complète et complexe qui renseigne davantage sur les sources de changement, mais ce premier découpage représente un bon départ.

Il est aussi possible d'examiner les contributions individuelles des pays, présentées dans la dernière colonne du tableau 5.

PAYS	PIB par habitant		Populati on totale		PIB relatifs		Parts de population		Niveau d'inégalité (MLD)	DECOMPOSITION				
	1985	2005	1985	2005	1985	2005	1985	2005	1985	Effet de composition	Effet de comportement	Contribution Nationales Nominal%		
Congo, Dem. Rep.	715	273	32443229	58740547	0.34	0.12	0.064	0.069	0.07	0.15	0.010	0.0561	0.0657	183%
Guinee Eq	2227	24770	314190	608807	1.06	10.47	0.001	0.001	0.00	0.00	0.000	0.0079	0.0083	23%
Niger	696	584	6708883	13264190	0.33	0.25	0.013	0.016	0.01	0.02	0.004	0.0030	0.0067	19%
Madagascar	1044	882	9778464	17614261	0.50	0.37	0.019	0.021	0.01	0.02	0.002	0.0033	0.0051	14%
Egypt, Arab Rep.	2954	4319	50654901	77154409	1.40	1.83	0.100	0.091	-0.03	-0.05	-0.010	0.0152	0.0048	13%
Kenya	1291	1349	19673682	35598952	0.61	0.57	0.039	0.042	0.02	0.02	0.003	0.0012	0.0047	13%
Botswana	4620	12088	1173803	1835938	2.19	5.11	0.002	0.002	0.00	0.00	0.000	0.0049	0.0046	13%
Malawi	717	648	7264558	13226091	0.34	0.27	0.014	0.016	0.02	0.02	0.002	0.0023	0.0041	11%
Ethiopia	500	633	41049476	74660901	0.24	0.27	0.081	0.088	0.12	0.12	0.011	-0.0076	0.0038	10%
Cote d'Ivoire	2155	1560	10475579	19244866	1.02	0.66	0.021	0.023	0.00	0.01	0.002	0.0016	0.0037	10%
Liberia	1384	323	2214623	3334222	0.66	0.14	0.004	0.004	0.00	0.01	-0.001	0.0039	0.0031	9%
Tunisia	3905	6445	7260360.969	10029000	1.85	2.72	0.014	0.012	-0.01	-0.01	-0.004	0.0063	0.0026	7%
Zambia	1333	1127	6784944	11738432	0.63	0.48	0.013	0.014	0.01	0.01	0.001	0.0017	0.0022	6%
Togo	886	772	3344926	5992080	0.42	0.33	0.007	0.007	0.01	0.01	0.001	0.0011	0.0017	5%
Angola	3109	3611	9331250	16617589	1.48	1.53	0.018	0.020	-0.01	-0.01	0.001	0.0003	0.0016	4%
Benin	1240	1309	4122108	7867626	0.59	0.55	0.008	0.009	0.00	0.01	0.001	0.0002	0.0015	4%
Burundi	469	340	4884506	7378129	0.22	0.14	0.010	0.009	0.01	0.02	-0.002	0.0033	0.0015	4%
Mauritius	4184	9975	1016000	1243253	1.99	4.22	0.002	0.001	0.00	0.00	-0.001	0.0025	0.0014	4%
RCA	899	644	2678023	4191429	0.43	0.27	0.005	0.005	0.00	0.01	0.000	0.0015	0.0010	3%
Chad	1044	1468	5227487	10145609	0.50	0.62	0.010	0.012	0.01	0.01	0.002	-0.0011	0.0008	2%
Senegal	1463	1614	6513728	11281296	0.69	0.68	0.013	0.013	0.00	0.01	0.000	0.0001	0.0005	1%
Gambie	1141	1142	735039	1526138	0.54	0.48	0.001	0.002	0.00	0.00	0.000	0.0001	0.0005	1%
Namibia	4371	5361	1130855	2019677	2.08	2.27	0.002	0.002	0.00	0.00	0.000	0.0002	0.0004	1%
Guinea-Bissau	653	497	919005	1472626	0.31	0.21	0.002	0.002	0.00	0.00	0.000	0.0005	0.0004	1%
Swaziland	2519	4335	705657	1124410	1.20	1.83	0.001	0.001	0.00	0.00	0.000	0.0003	0.0002	1%
Mauritania	1599	1684	1715028	2963105	0.76	0.71	0.003	0.004	0.00	0.00	0.000	0.0001	0.0002	0%
Guinea	860	1056	5267169	9220768	0.41	0.45	0.010	0.011	0.01	0.01	0.001	-0.0005	0.0001	0%
Cape Verde	1589	2695	318417	477438	0.75	1.14	0.001	0.001	0.00	0.00	0.000	0.0000	-0.0001	0%
Cameroon	2716	1959	10514988	17795149	1.29	0.83	0.021	0.021	-0.01	0.00	0.000	-0.0005	-0.0003	-1%
Sierra Leone	720	640	3631130	5106977	0.34	0.27	0.007	0.006	0.01	0.01	-0.002	0.0011	-0.0007	-2%
Burkina Faso	711	1026	7709074	13933363	0.34	0.43	0.015	0.016	0.02	0.01	0.002	-0.0024	-0.0008	-2%
Congo, Rep.	4033	3497	2116659	3416654	1.91	1.48	0.004	0.004	0.00	0.00	0.000	-0.0007	-0.0009	-3%
Lesotho	835	1266	1472306	1980831	0.40	0.54	0.003	0.002	0.00	0.00	-0.001	-0.0004	-0.0011	-3%
Rwanda	768	793	6111361	8992140	0.36	0.34	0.012	0.011	0.01	0.01	-0.002	0.0006	-0.0014	-4%
Mali	707	1004	6793924	11611090	0.34	0.42	0.013	0.014	0.01	0.01	0.000	-0.0020	-0.0016	-4%
Uganda	525	901	14795432	28699255	0.25	0.38	0.029	0.034	0.04	0.03	0.007	-0.0091	-0.0022	-6%
Nigeria	1306	1731	81598130	141356083	0.62	0.73	0.162	0.167	0.08	0.05	0.006	-0.0088	-0.0030	-8%
Gabon	16557	13029	791848	1368229	7.86	5.51	0.002	0.002	0.00	0.00	0.000	-0.0032	-0.0030	-8%
Ghana	816	1193	13005766	21915168	0.39	0.50	0.026	0.026	0.02	0.02	0.000	-0.0038	-0.0036	-10%
Morocco	2485	3589	21779134	30142708.8	1.18	1.52	0.043	0.036	-0.01	-0.01	-0.008	0.0033	-0.0046	-13%
Sudan	889	1601	24051873	38698472	0.42	0.68	0.048	0.046	0.04	0.02	-0.002	-0.0100	-0.0122	-34%
Algeria	6847	7176	22097343	32854159	3.25	3.03	0.044	0.039	-0.05	-0.04	-0.010	-0.0061	-0.0160	-44%
Mozambique	312	677	13324105	20532675	0.15	0.29	0.026	0.024	0.05	0.03	-0.004	-0.0131	-0.0169	-47%
South Africa	8100	8504	31307880	46892428	3.84	3.59	0.062	0.055	-0.08	-0.07	-0.016	-0.0108	-0.0266	-74%
MOYENNE OU														
TOTAL	2106.6	2365.8	504,806,844	845,868,171					0.39	0.43	-0.006	0.0423		
												0.04	-18%	118%
												Effet de composition	Effet de comportement	

Tableau 5. Décomposition du changement dans l'inégalité économique entre pays africains (1985-2005).



### II.2.3 Décomposition ordinale

La décomposition ordinale se rapproche de la décomposition démographique à la différence que la variable de classification est ordinale. L'illustration choisie ici est le calcul de l'indice synthétique de fécondité (ISF) peut servir d'exemple classique. La valeur de cet indice au moment  $t$  est la somme des valeurs spécifiques pour tous les groupes d'âge de 15-19 à 45-49.

$$Y_t = \sum y_{at} \quad (7)$$

Supposons que cet indice vienne à baisser dans le temps, un analyste pourrait vouloir estimer dans quelle mesure un changement historique de cet ISF reflète un effet de « *quantum* » ou de « *tempo*. » Le premier effet quantifie le changement généralisé à tous les groupes d'âge, et le second reflète un report des naissances vers les âges les plus avancés.

Pour faire cette décomposition temporelle, l'analyste peut exprimer la fécondité spécifique à chaque groupe d'âge ( $y_a$ ) en fonction de la fécondité observée au niveau des groupes d'âge plus avancés que  $a$  ( $y_{a+}$ ). Ainsi par exemple, la fécondité des jeunes de 15-19 ans sera exprimée en fonction de la fécondité moyenne pour toutes les femmes de 20 à 49 ans ( $y_{15-19} = r_{15-19} * y_{20-49}$ );

$$y_a/y_{a+} = r_a \rightarrow y_a = r_a * y_{a+}$$

L'ISF peut donc être ré-exprimée comme ci-dessous :

$$Y_t = \sum r_{at} * y_{a+t} \quad (II.6)$$

Cette nouvelle formulation permet de décomposer le changement de l'ISF en deux termes qui reflètent le quantum et le tempo, respectivement.

$$\Delta Y = \sum \bar{y}_{a+} * \Delta r_a + \sum \bar{r}_a * \Delta y_{a+} \quad (II.7)$$

#### *Illustration*

L'illustration ici vient de Ouedraogo (2012) (données non présentées ici) qui propose un suivi détaillé des transitions démographiques en Afrique. La procédure habituelle pour avoir ce détail consiste à comparer les courbes de fécondité par groupe d'âge pour les périodes à examiner, comme indiqué dans son étude sur le Cameroun des années 1991, 1998, et 2004. Un examen de ces courbes montre un parallélisme imparfait : la baisse de fécondité était généralisée mais pas linéaire. Bien que cette figure offre des informations détaillées, cette information n'est pas facile à résumer. Pour un chercheur qui voudrait par exemple étudier dans quelle mesure la baisse des naissances a été disproportionnellement enregistrée parmi les groupes les plus jeunes, la décomposition temporelle est un outil de choix.

Les calculs pour cette décomposition montrent une baisse dominée par le quantum (90 %) : en général, la baisse a touché toutes les couches d'âge mais, ainsi que le montre le graphique, elle était plus accentuée parmi les femmes les plus jeunes. Il y avait donc un certain report des naissances (10 %). Notez que cette analyse pourrait aussi être reprise en utilisant le statut socioéconomique – par exemple – comme variable ordinale de classification pour étudier la concentration de la reproduction vers les classes les plus pauvres.

### II.2.4 Décomposition emboîtée

Cette extension, plus détaillée, exprime le changement dans l'effectif des groupes en fonction des moteurs classiques du changement démographique, à savoir la fécondité, mortalité, nuptialité et/ou migration différentielles des groupes. Bien évidemment, l'application de cette extension requiert beaucoup plus

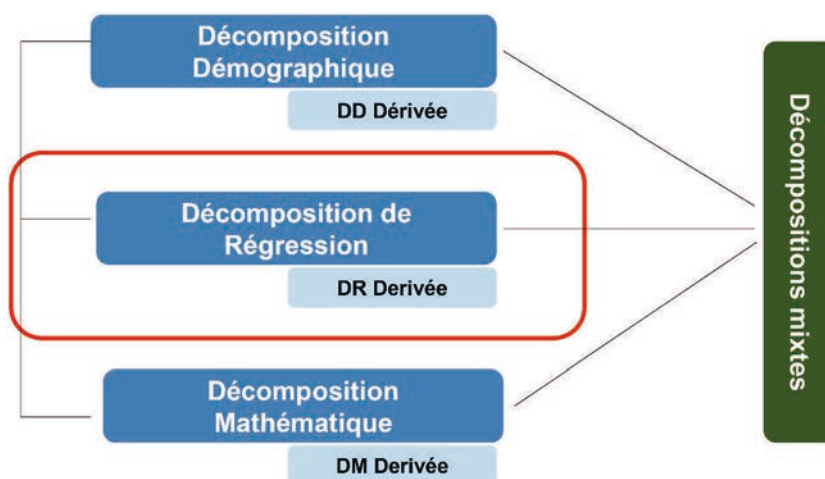


---

# Chapitre III

## Décompositions de régression

---



### III.1 décomposition de régression simple

#### III.1.1 Type de problème

La principale différence d'avec la méthode de décomposition démographique tient à la nature de la variable indépendante qui, ici, est quantitative. Par ailleurs, le lien avec la variable dépendante est décrit non par une sommation, mais par une équation de régression. Cette équation comporte un coefficient de régression (effet brut de la variable indépendante sur Y) et un intercept, mais le terme d'erreur est omis pour cette décomposition simple qui analyse une moyenne nationale estimée.

#### III.1.2 Formulation mathématique

L'expression générique est de la forme  $Y = f(\alpha, \beta, X)$  mais une formulation plus explicite serait par exemple

$$Y_t = \alpha_t + \beta_t X_t \quad \text{[III.1]}$$

Dans ce cas, la décomposition consiste à expliquer le changement dans la variable dépendante en fonction du changement dans les divers paramètres de l'équation de régression. Ce changement s'exprime comme suit :

$$\Delta Y = \Delta \alpha + \bar{\beta} \Delta X + \bar{X} \Delta \beta \quad \text{[III.2]}$$

Une fois de plus, la même procédure s'applique aussi bien pour une analyse transversale (différence entre deux groupes au cours d'une année donnée) que pour une analyse longitudinale (changement vécu par un groupe donné entre deux années). Dans un cas comme dans l'autre, la démarche est la même, et seule l'interprétation diffère.

### Première Illustration

Une question classique dans laquelle la décomposition de régression s'applique est l'analyse des différences de revenus entre hommes et femmes et dans quelle mesure cette différence reflète une discrimination sexuelle sur le marché de l'emploi. L'idée générale ici est que la rémunération de l'éducation sur le marché de l'emploi n'est pas la même pour les hommes et les femmes, mais il est aussi possible que hommes et femmes partent initialement avec des niveaux d'éducation différents. Il faut donc estimer dans quelle mesure les différences de revenus reflètent une discrimination plutôt qu'une différence de capital humain. Ces deux fonctions de production de revenus (et la formule de décomposition correspondante) sont décrites ci-dessous.

$$Y_h = \alpha_h + \beta_h * X_h$$

$$Y_f = \alpha_f + \beta_f * X_f$$

$$\Delta Y = \Delta\alpha + \bar{\beta}\Delta X + \bar{X}\Delta\beta \quad [III.3]$$

### III.1.3 Application

Pour prendre un exemple numérique, le tableau 6 présente des données sur les revenus en fonction du niveau d'instruction. Cet exemple hypothétique montre des fonctions de revenus dans lesquelles la rémunération moyenne de l'éducation est de 4,000 FCFA par année d'éducation pour les hommes et de 5,000 FCFA pour les femmes. Par ailleurs, le salaire de base est de 20,000 et 25,000 pour les hommes et femmes, respectivement. Compte tenu du fait que les niveaux d'éducation moyens sont de 10 et 12 années, respectivement, pour les hommes et les femmes, le salaire moyen d'un homme serait de 60,000 tandis que celui d'une femme serait de 85,000 FCFA, soit une différence de 25,000. Le calcul de décomposition montre que cette différence de salaire entre hommes et femmes est imputable pour 20 % à l'écart entre les salaires de base ( $\Delta\alpha$ ) ; et pour 36 % à une différence entre les niveaux d'instruction selon le sexe, et enfin 44 % à une différence dans la rémunération différentielle de l'éducation selon le sexe. En d'autres termes, même si l'on atteignait la parité dans l'éducation, l'inégalité salariale entre femmes et hommes se réduirait uniquement de 36 %. L'inégalité résiduelle serait due aux inégalités dans les salaires de base et la rémunération de l'éducation.

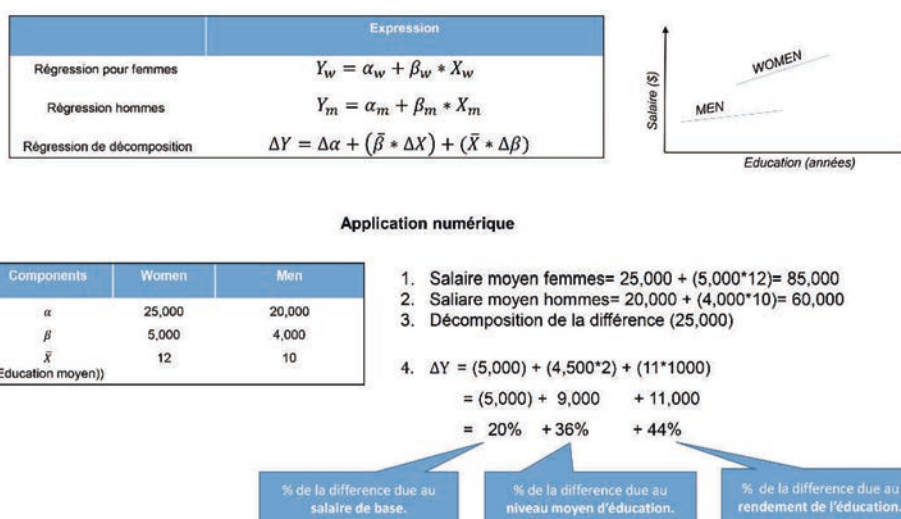


Tableau 6. Décomposition de régression pour l'analyse des écarts de salaire entre femmes et hommes.

## III.2 Décompositions de régression dérivées

### III.2.1 Régression curvilinéaire

Pour la plupart des phénomènes étudiés en sciences sociales, un modèle linéaire et bivarié, tel que présenté dans la section précédente, est simpliste. Dans cet exemple qui modélise le revenu selon le niveau d'instruction, un effet curvilinéaire semblerait plus plausible. Fort heureusement, la décomposition s'étend aisément à des modèles non-linéaires. Pour un effet quadratique par exemple,

$$Y = \alpha + \beta_1 X + \beta_2 X^2 \quad (\text{III.4})$$

L'analyse de décomposition donnera les termes suivants :

$$\begin{aligned} \Delta Y &= \Delta\alpha + (\bar{X} * \Delta\beta_1) + (\bar{\beta}_1 * \Delta X) + (\bar{X}^2 * \Delta\beta_2) + (\bar{\beta}_2 * 2\Delta X) \\ \Delta Y &= \Delta\alpha + (\bar{X} * \Delta\beta_1) + (\bar{X}^2 * \Delta\beta_2) + ((\bar{\beta}_1 + 2\bar{\beta}_2) * \Delta X) \end{aligned} \quad (\text{III.5})$$

Comme exemple, l'on peut retourner sur l'exemple hypothétique des écarts salariaux entre hommes et femmes, mais en spécifiant cette fois-ci une relation curvilinéaire pour l'effet de l'instruction sur le revenu.

### III.2.2 Régression multivariée

De manière similaire, la décomposition d'une régression multivariée inclura un ou des termes supplémentaires qui, cette fois-ci seront des variables indépendantes supplémentaires, plutôt que le carré de la première variable. Comme dans l'exemple précédent, le nombre de termes sera simplement plus long. Pour prendre le cas simple de deux facteurs, l'on pourrait étendre l'analyse des niveaux de salaires en considérant cette fois-ci le niveau d'instruction et le nombre d'années d'expérience.

$$Y = \alpha + \beta_1 X_1 + \beta_2 X_2 \quad (\text{III.6})$$

Dans ce cas particulier, les sources du changement dans le niveau des salaires incluraient les composantes suivantes :

$$\Delta Y = \Delta\alpha + (\bar{X}_1 * \Delta\beta_1) + (\bar{\beta}_1 * \Delta X_1) + (\bar{X}_2 * \Delta\beta_2) + (\bar{\beta}_2 * \Delta X_2) \quad (\text{III.7})$$

### III.2.3 Régression multi-niveaux

Comme son nom l'indique, la régression intègre des analyses à deux ou plusieurs niveaux qui peuvent inclure l'individu, son groupe et sa société toute entière. Dans l'analyse de la scolarisation par exemple, l'on peut examiner comment la performance des élèves dépend des caractéristiques individuelles des élèves (niveau 1) ainsi que celles de la société (niveau 2). Les équations pour estimer un tel modèle s'expriment comme suit :

$$\text{Niveau 1 : } Y_{jk} = \beta_{0k} + \beta_{1k} * x_{jk} + r_{jk}$$

$$\text{Niveau 2 : } \beta_{0k} = \gamma_{00} + \gamma_{01} Z_k + \mu_{0k}$$

$$\beta_{1k} = \gamma_{10} + \gamma_{11} Z_k + \mu_{1k}$$

En intégrant les valeurs du niveau 2 dans l'équation de niveau 1, l'on obtient l'équation mixte qui exprime la performance du groupe à la fois en fonction des caractéristiques du groupe, du contexte national et de leur interaction :

$$Y_{jk} = \gamma_{00} + \gamma_{10} x_{jk} + \gamma_{01} Z_k + \gamma_{11} Z_k x_{jk} + \mu_{0k} + \mu_{1k} x_{jk} + r_{jk} \quad (\text{III.8})$$

Pour étudier le changement de  $Y_{jk}$  dans le temps, il suffit de différencier la formule ci-dessus et l'incorporer dans le deuxième terme de l'équation [1].

### III.3 Application au dividende démographique

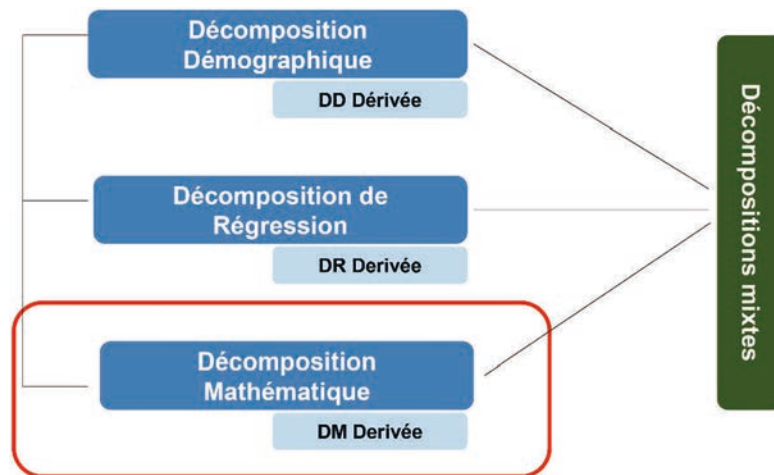
Une question classique en rapport avec le dividende démographique est la relation entre taille de la famille et niveau d'éducation atteint. De nombreuses études montrent un lien inverse entre ces deux variables. Supposons qu'entre 1990 et 2010, un pays X connaisse une baisse dans la taille moyenne de la famille, de 6 à 4 enfants, en même temps qu'une amélioration du niveau moyen d'éducation de 8 à 12 années. Dans quelle mesure cette amélioration de l'éducation est imputable à la réduction de la taille moyenne des familles ? Il est possible que cette réduction ait joué un rôle mais que d'autres facteurs aient aussi contribué. La décomposition permet d'estimer les contributions relatives de divers facteurs, mais une bonne réponse requiert des informations fiables sur la relation statistique entre éducation et taille de la famille mais aussi comment cette relation a évolué dans le temps. Une étude de Eloundou et Giroux (2012) indique bien comment intégrer la relation entre l'éducation et la fécondité au niveau micro et les informations sur les changements d'effectif pour estimer les implications du changement démographique sur l'éducation au niveau national.

---

# Chapitre IV

## Décompositions mathématiques

---



### IV.1 Décomposition mathématiques simple

En plus des formulations démographiques et statistiques, l'on peut envisager une décomposition qui étudie un phénomène décrit par une relation mathématique exacte entre deux variables. Par exemple, le revenu moyen par habitant dans un pays est égal au revenu total divisé par la population totale. Le changement dans ce revenu moyen peut donc être a priori être décomposé comme résultant du changement dans le revenu total (numérateur) et du changement dans la taille de la population (dénominateur).

#### IV.1.1 Type de problème

Ce type d'analyse s'applique à des processus qui peuvent s'exprimer comme une relation mathématique entre deux/plusieurs variables sociales qui sont elles-mêmes importantes sur le plan conceptuel ou plan politique. Le PNB par habitant est un exemple. En ce qu'il combine une composante économique (PNB) et une composante démographique (nombre d'habitants), il peut être décomposé en termes simples qui montrent comment ce PNB par habitant évolue en fonction des changements de l'une ou l'autre de ces deux composantes. D'autres variables similaires qui mesurent le bien être individuel en rapportant une ressource donnée à la population correspondante.

#### IV.1.2 Formulation mathématique

Prenons un exemple légèrement plus complexe que le PNB par habitant, notamment le niveau des dépenses publiques d'éducation par enfant ( $r$ ). Cette dépense est proportionnelle au niveau total des ressources nationales ( $g$ ) et au pourcentage de ces ressources alloué à l'éducation ( $k$ ). Par contre, elle est inversement proportionnelle au nombre d'enfants scolarisables dans ce pays ( $p$ ).

$$r = g^k/p \quad (\text{IV.1})$$

Dans ce cas, le changement historique dans cette dépense peut être décomposé comme indiqué ci-après.

$$\Delta r \cong -\left[\overline{(kp/p^2)} * \Delta p\right] + \left[\overline{(k/p)} * \Delta g\right] + \left[\overline{(g/p)} * \Delta k\right] \quad (IV.2)$$

↑  
 Effet du changt  
dans la populat.  
D'age scolaire

↑  
 Effet du changt  
dans le revenu  
national

↑  
 Effet du changt dans  
la part du budget  
allant a l'éducation

### IV.1.3 Application

Dans cette étude, et en s'appuyant sur les statistiques nationales publiées par la Banque Mondiale (2014), Eloundou, Tenikue et Ryu (2014) décomposent les changements dans les dépenses publiques par enfant parmi les pays de l'Afrique Australe entre 1990 et 2010 et comparent la situation de ces pays à celle de la Corée du Sud, quinze années auparavant. Les données montrent qu'au courant des périodes concernées, la Corée du Sud a connu une croissance économique (250 %) de loin supérieure à celle enregistrée dans ces pays de l'avant-garde africaine (entre 3 et 94 %). Par ailleurs, le changement démographique y a aussi été plus marqué, notamment avec un passage du nombre d'enfants par adulte de 0.65 à 0.23 en Corée contre 0.67 à 0.46 par exemple en Afrique du Sud. Malgré ces différences nominales dans la démographie et l'économie, il y a une certaine similarité dans les contributions relatives du changement démographique dans l'accroissement des dépenses publiques en éducation par enfant (60 % pour la Corée au courant de cette période, contre 70 % et 61 % en Afrique du Sud et au Swaziland respectivement. Notons aussi un fait souvent occulté, à savoir que ces pays africains allouent souvent une plus grande partie de leur budget national à l'éducation (ex. 10 % au Lesotho et près de 6 % au Botswana contre 2 % en Corée). Il n'est donc pas surprenant que les politiques budgétaires aient été d'un apport relatif plus important dans l'amélioration des valeurs de r au Lesotho (32 %) ou au Botswana (24 %) par rapport à la Corée (12 %).

	PIB par adulte		Part du budget pour l'éducation		Nombre d'enfant par adulte		Dépense publique en éducation par enfant		$\Delta r$	% du changement en dépense par enfant (r) associé au changement en		
	g		k		p*		r			Age Dependency	Income	Budget Commitment
	1990	2010	1990	2010	1990	2010	1990	2010				
	[1]	[2]	[3]	[4]	[5]	[6]	[7]	[8]	[9]	[10]	[11]	[12]
Botswana	6257	9890	0.057	0.078	0.85	0.55	421.2	1409.4	988.20	41.7%	34.4%	24.0%
Ile Maurice	4581	8889	0.032	0.037	0.44	0.30	330.8	1093.5	762.70	37.1%	50.6%	12.2%
Afrique du Sud	8390	8906	0.053	0.060	0.67	0.46	659.7	1166.8	507.10	69.9%	9.9%	20.3%
Swaziland	4099	4229	0.055	0.069	0.98	0.67	230.0	436.2	206.20	61.1%	4.7%	34.2%
Lesotho	1074	1514	0.101	0.130	0.79	0.65	137.3	303.2	165.90	26.0%	42.3%	31.8%
	1975	1995	1975	1995	1975	1995	1975	1995	1975-1995			
Corée du Sud	4,885	17,254	0.021	0.032	0.654	0.230	\$158.5	\$2,415.9	\$2,257.4	60%	28%	12%

Tableau 7. Décomposition mathématique de l'évolution des dépenses publiques par enfant (comparaison entre Corée du Sud et pays « avant-garde » dans la transition africaine).

## IV.2 Décompositions mathématiques dérivées

### IV.2.1 Extension de chaines

Retournons maintenant vers la décomposition du PIB par habitant telle qu'introduite en section IV.1.1. Cette première décomposition donne une description succincte mais qui n'est pas optimale sur le plan conceptuel : la raison est que les deux composantes ne nous renvoient pas à des variables de décision ou des phénomènes en transition. Il est possible de transformer ces expressions mathématiques en expressions légèrement plus longues



mais conceptuellement plus riches. Par exemple, la formulation du PIB par habitant décrite plus haut était  $Y = G/P$  et différenciée comme

$$\Delta Y = (\bar{G} * \Delta(1/P)) + ((\overline{1/P}) * \Delta G) \quad (IV.3)$$

Elle peut se réécrire de manière plus intéressante comme

$$Y = \frac{G}{P} = \frac{G}{A} * \frac{A}{P} = \pi * \alpha \quad (IV.4)$$

où G et P désignent toujours le revenu national et la population nationale totale, respectivement. Le nouveau terme introduit est A, la population active (ou adulte) du pays. Dans cette réécriture, G/A (ou  $\pi$ ), la productivité des adultes, est une variable intéressante sur le plan conceptuel et de planification. C'est une variable souvent invoquée dans les théories économiques de la croissance et elle peut faire l'objet de politiques spécifiques d'accroissement de productivité, soit à travers l'éducation, la recherche ou le développement technologique. De même le nouveau terme A/P (ou encore  $\alpha$ ) renvoie à la structure par âge de la population, plus précisément le ratio du nombre d'adultes sur la population totale, et il est souvent invoqué dans les théories sur le dividende démographique. L'analyste se retrouve donc avec deux variables conceptuellement intéressantes ( $\pi$  et  $\alpha$ ) et il/elle peut décomposer les changements du PIB par habitant en fonction de ces deux variables.

$$\Delta Y = (\bar{\pi} * \Delta\alpha) + (\bar{\alpha} * \Delta\pi) \quad (IV.5)$$

Bien évidemment cette nouvelle expression peut elle-même être développée davantage et donner des décompositions de plus en plus détaillées et potentiellement optimales sur le plan de l'intervention publique. Par exemple, la productivité des adultes peut être partitionnée en deux composantes : le taux de chômage des adultes et la productivité des travailleurs adultes.

$$Y = \frac{G}{P} = \frac{G}{E} * \frac{E}{A} * \frac{A}{P} = \rho * \varepsilon * \alpha \quad (IV.6)$$

Et donc son changement historique deviendrait :

$$\Delta Y = (\bar{\varepsilon}\bar{\rho} * \Delta\alpha) + (\bar{\rho}\bar{\alpha} * \Delta\varepsilon) + (\bar{\varepsilon}\bar{\alpha} * \Delta\rho) \quad (IV.7)$$

De même, le ratio des adultes sur la population totale peut être exprimé comme une fonction des populations de jeunes et de vieux, respectivement. Jusqu'où le chercheur ira dans ce développement dépendra de son propre arbitrage entre le souci de détail et de parcimonie, mais aussi de la disponibilité des données et de la simplicité des nouvelles expressions.

Toujours dans le but de décomposer la croissance économique en des termes pertinents aussi bien sur le plan théorique que celui de l'intervention publique, il est possible de recourir aux formulations classiques de la croissance économique, notamment la fonction Cobb-Douglas, qui exprime la croissance en fonction de du capital physique (K), humain (h), de l'emploi (L) et de la productivité globale des facteurs (A) :

$$Y = AK^\alpha(hL)^\beta \quad (IV.8)$$

Partant de cette formulation (voir Tenikué 2014), il est possible de décomposer la croissance économique entre deux périodes en fonction de ces divers paramètres. Spécifiquement,

$$\Delta y_t \cong \Delta A_t + \alpha \Delta k_t + (1 - \alpha) \Delta h_t + \Delta l_t + \Delta w_t \quad (IV.9)$$

Où  $y$  est le GDP par tête (Y/P)

$k$  est le stock de capital physique par personne employée (K/L)

$h$  est le capital humain

$l$  est le taux d'emploi (L/W)

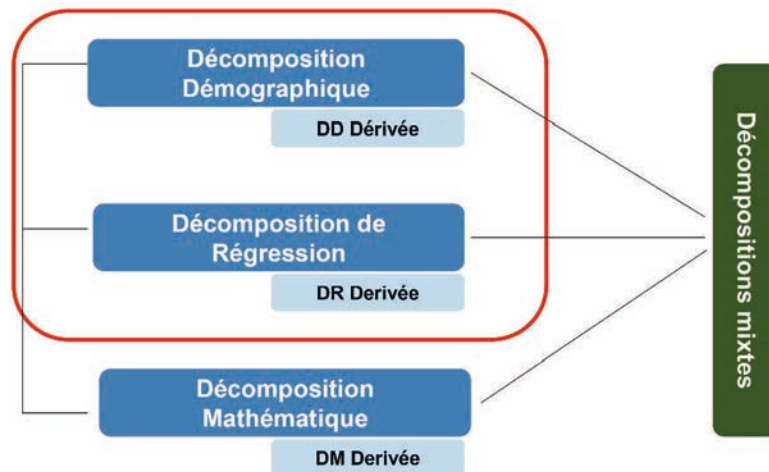
$w$  est la structure d'âge de la population (W/P)

A est la productivité globale des facteurs (PGF)



# Chapitre V

## Combinaison démographique + régression



Les décompositions élémentaires ont le mérite de la simplicité, mais elles manquent de détail. Ce problème peut être résolu avec des décompositions mixtes qui combinent deux ou plusieurs formes de décomposition élémentaires. Il est ainsi possible de combiner une décomposition démographique et une décomposition de régression pour obtenir une forme plus détaillée et intéressante. Ce type de combinaison fera l'objet de ce chapitre. Partant de la décomposition démographique simple, nous verrons comment développer chacune des composantes principales, en commençant par l'effet de comportement ( $\Delta y_j$ ) puis ensuite l'effet de composition ( $\Delta w_j$ ).

### V.1 Extension de l'effet de comportement

#### V.1.1 Présentation générale

Reprenons la formule de base (3) qui exprime le changement de la moyenne nationale en fonction des effets de composition et de comportement des divers groupes,

$$\Delta Y = [\sum \bar{y}_j * \Delta w_j] + [\sum \bar{w}_j * \Delta y_j]$$

Effet de  
composition

Effet de  
comportement

Cette formule peut être étendue en exprimant la performance d'un groupe donné ( $y_j$ ) comme une fonction d'une ou plusieurs autres variables. Pour une relation de régression simple,

$$y_j = \alpha + \beta x_j + \mu_j \text{ avec,}$$

- $\alpha$  (intercept) représente la performance de base, lorsque  $x=0$ ,
- $\beta$  est l'accroissement de mortalité lié à un accroissement unitaire de la variable  $x$ , et
- $\mu_j$  l'erreur, (surperformance/ sous performance relative du groupe, ou comme l'effet résiduel des facteurs autres que  $x$ , non considérés dans l'analyse.

Le changement dans la valeur de  $y_j$  entre deux périodes s'obtient ainsi

$$\Delta y_j = \Delta\alpha + \bar{\beta}\Delta x_j + \bar{x}_j\Delta\beta + \Delta\mu_j \quad [V.1]$$

En insérant [V.1] dans l'équation de base (3), l'on obtient [V.2]

$$\Delta Y = \underbrace{\sum \bar{y}_j \Delta w_j}_A + \underbrace{\sum \bar{w}_j \Delta \alpha}_{B1} + \underbrace{\sum \bar{w}_j \bar{\beta} \Delta x_j}_{B2} + \underbrace{\sum \bar{w}_j \bar{x}_j \Delta \beta}_{B3} + \underbrace{\sum \bar{w}_j \Delta \mu_j}_{B4}$$

- A, l'effet de composition qui reste inchangé par rapport à la situation précédente,
- B l'effet de performance qui se subdivise désormais en quatre sous-composantes qui reflètent respectivement les changements dans
  - la performance de base (A1),
  - le niveau de la variable indépendante (B2),
  - l'effet de la variable indépendante (B3),
  - l'effet résiduel (B4).

### V.1.2 Illustration

Pour illustrer ce cas, nous pouvons retourner au premier exemple de décomposition démographique de la mortalité infantile de 1990 à 2011 (Tableau 3). À titre de rappel, une décomposition simple indiquait une baisse dans la mortalité infantile de près de 26 points, laquelle provenait à 17 % d'un effet de composition et à 83 % d'un effet de performance. Pour affiner l'étude de cette relation, envisageons que la mortalité dépende des niveaux d'éducation. Il devient possible d'exprimer la mortalité (performance) au sein de chaque groupe en fonction de l'éducation. Si l'on peut estimer cette relation, notamment les valeurs de  $\alpha$  (mortalité de base),  $\beta$  (effet de l'éducation sur la mortalité), et  $\mu$  (terme d'erreur), il est ensuite possible de décomposer plus finement le changement dans la mortalité nationale en fonction des changements dans ces diverses composantes élémentaires, chacune desquelles renseigne utilement sur la nature des forces qui ont occasionné le changement au niveau national. Cette décomposition permet d'identifier de manière plus fine les secteurs de politique qui ont été décisifs dans le changement. Ainsi, dans l'équation V.2,

- l'effet de performance de base (B1) reflète l'amélioration des conditions de santé de base, par exemple les innovations qui améliorent le niveau minimum de santé au sein de la population, indépendamment de leur niveau d'éducation ;
- B2 mesure l'amélioration de la santé associée à l'amélioration du niveau d'éducation des diverses catégories ;
- B3 mesure l'amélioration de la survie lié au changement de l'effet d'éducation ;
- B4 mesure l'effet résiduel des autres variables non considérées.

Bien que cet exemple se limite à une seule variable indépendante et un effet linéaire, l'on peut imaginer comment cette extension de l'effet de performance pourrait considérer deux ou plusieurs variables indépendantes ou alors des effets curvilinéaires. La seule différence est que l'analyse impliquera une équation de plus en plus longue !

### V.1.3 Lien avec la méthode des comptes et transferts nationaux

La méthode des comptes de transferts nationaux est une méthode courante dans l'étude du dividende démographique (Mason et Lee 2005). Elle repose sur un principe simple qui consiste à documenter le comportement économique (revenus, épargne, transferts) au sein d'une population et voir comment il varie systématiquement avec l'âge. Compte tenu de ce profil de consommation par âge (Figure 5), l'on peut calculer un surplus par âge et, en cumulant ces surplus, obtenir le surplus national pour une année donnée. En supposant que le profil de comportement économique par âge reste constant, tout changement dans la structure par âge de la population entraîne ipso facto une modification du surplus national.

Toutefois, il n'est pas réaliste de supposer que le profil de consommation et revenus reste fixe pendant la transition démographique et donc, il est nécessaire de tenir compte de la possibilité d'un changement à la fois dans la structure par âge et dans le profil de consommation. Nous montrons que ceci est facile à envisager dans l'approche de décomposition et que le NTA peut en partie être considéré comme un exemple de décomposition mixte (Eloundou, Tenikue, Giroux 2014). En effet, le surplus national pour une année donnée est la moyenne pondérée des surplus spécifiques à chaque groupe d'âge.

$$S_t = \sum w_{jt} s_{jt} \quad [V.3]$$

Son changement s'exprime, comme dans une décomposition simple, ainsi :

$$\Delta S = \sum \bar{s}_j \Delta w_j + \sum \bar{w}_j \Delta s_j \quad [V.4]$$

Cette décomposition simple permet d'estimer l'influence mécanique du changement dans la structure par âge de la population (effet de composition), mais elle ignore le fait qu'un changement dans la structure par âge pourrait aussi affecter le comportement économique des gens. Par exemple, dans une population où le nombre de dépendants est élevé, la possibilité d'épargner pourrait se trouver réduite. Il est donc utile d'explorer la possibilité additionnelle qu'un changement dans la structure d'âge puisse aussi avoir un effet sur le comportement. Si notre analyste peut obtenir, à travers l'analyse de régression, une estimation fiable de l'effet de la structure d'âge sur ( $J$ ) sur le comportement économique de chaque groupe d'âge [Equations x et x], alors, le changement dans le surplus national pourrait être estimé à travers une décomposition mixte [Equation x].

$$s_{jt} = a_{jt} + b_{jt} J_t + e_{jt} \quad [V.5]$$

$$\Delta s_j = \Delta a_j + \bar{J} \Delta b_j + \bar{b} \Delta J + \Delta e_j \quad [V.6]$$

$$\Delta S = \sum \bar{s}_j \Delta w_j + \sum \bar{w}_j \Delta a_j + \sum \bar{w}_j \bar{J} \Delta b_j + \sum \bar{w}_j \bar{b} \Delta J + \sum \bar{w}_j \Delta e_j \quad [V.7]$$



Dans cette équation mixte, le changement dans la structure d'âge a à la fois un effet mécanique (premier terme) et un effet substantif sur le comportement (avant dernier terme). La question bien évidemment est de voir dans quelle mesure l'on peut obtenir une estimation fiable du coefficient  $b$ , qui exprime l'effet de la structure par âge de la population sur le comportement économique d'une catégorie d'âge donnée.

## V.2 Extension de l'effet de composition

Tout comme l'effet de comportement, l'effet de composition peut être désagrégé. Ceci peut se faire à travers plusieurs stratégies qui considèrent soit les groupes démographiques primaires, soit la structure par âge des

diverses sous-populations, soit les dynamiques de reproduction démographique (fécondité, mortalité, migration).

### (a) Extension selon les groupes primaires

Si l'on adopte une extension selon les groupes primaires, la taille du groupe étudié est exprimée comme une fonction de la taille d'un groupe primaire auquel appartiennent les sujets. Pour notre exemple, le pourcentage d'enfants dans les familles pauvres ( $w_j$ ) est fonction du nombre de familles pauvres ( $n_j$ ) et la fécondité relative des familles pauvres ( $f_j$ ) qui est la fécondité des pauvres par rapport à la moyenne nationale.

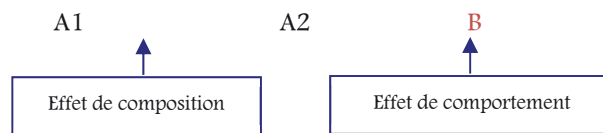
$$w_j = n_j * f_j \quad [V.8]$$

Le changement dans les pourcentages des enfants issus de chacun des groupes dépendra du changement dans (1) la proportion de familles appartenant à différentes catégories sociales et (2) la fécondité relative de ces familles.

$$\Delta w_j = \bar{f} \Delta n_j + \bar{n} \Delta f_j \quad [V.9]$$

On peut ainsi insérer [V.9] dans [3] pour obtenir

$$\Delta Y = [\sum \bar{y}_j * \bar{f} \Delta n_j] + [\sum \bar{y}_j * \bar{n} \Delta f_j] + [\sum \bar{w}_j * \Delta y_j] \quad [V.10]$$



Dans cette équation, A1 représente le changement dans le pourcentage de familles pauvres au sein de la société ; A2 représente le changement dans la fécondité relative des diverses classes sociales. Ces deux variables sont conceptuellement plus riches et renseignent plus l'analyste sur plusieurs changements sociaux d'intérêt. La première (A1) renvoie au changement économique et aux politiques sociales de réduction de pauvreté. La deuxième aura davantage trait à la capacité des programmes de planification familiales à niveler les écarts de fécondité surtout s'ils sont dûs à une différence d'accès à la planification familiale.

## V.3 Double extension

Pour finir, l'on peut jumeler [V10] et [3] et obtenir une formulation détaillée qui affine l'analyse à la fois du côté de  $w$  et du côté de  $y$ . Le résultat de cette combinaison (donné ci-dessous) produit une expression encore plus longue et détaillée. Dans celle-ci,

- A1 = effet de changement dans la distribution des mères par classe sociale ;
- A2 = changement dans la fécondité relative des mères ;
- B1 = amélioration des conditions de santé de base ;
- B2 = changement de revenu ;
- B3 = différenciation de mortalité selon le revenu (vulnérabilisation des pauvres) ;
- B4 = effet résiduel des facteurs omis dans l'équation de régression.

$$\Delta Y = \underbrace{[\sum \bar{w}_j \Delta \alpha] + [\sum \bar{w}_j \bar{\beta}_j \Delta x_j] + [\sum \bar{w}_j \bar{x}_j \Delta \beta] + [\sum \bar{w}_j \Delta \mu_j]}_{\text{B1, B2, B3, B4}} + \underbrace{[\sum \bar{y}_j \bar{f}_j \Delta n_j] + [\sum \bar{y}_j \bar{n}_j \Delta f_j]}_{\text{A1, A2}}$$

Cette décomposition avancée décortique encore plus les premiers résultats issus d'une décomposition démographique simple. Elle montre que chacun des effets primaires (démographiques ou de comportement) peut lui-même être découpé en éléments plus simples encore.

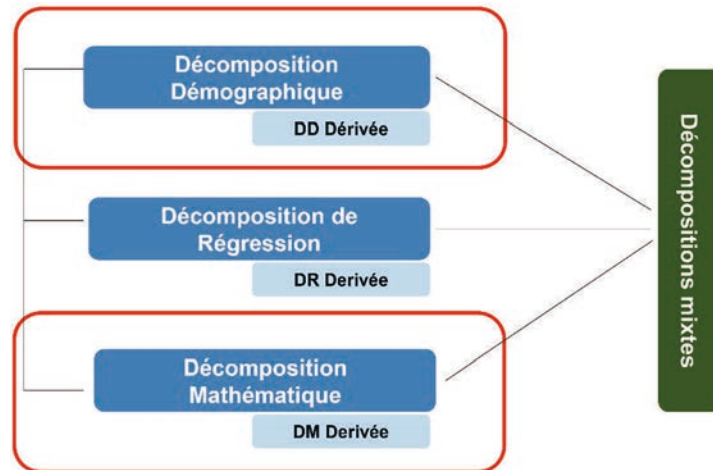
Sur la base de cette comptabilité plus détaillée, le planificateur peut se faire une meilleure idée des programmes à envisager de manière prioritaire, dans ce cas la promotion de la santé publique à travers des programmes d'ensemble (amélioration de la qualité d'eau, vaccination, couverture hospitalière...).





# Chapitre VI

## Combinaison démographique + mathématique



Dans cette nouvelle situation, l'on part toujours d'une décomposition démographique, mais la nouvelle expression développée de  $y_j$  sera de type mathématique. Comme illustration, nous envisagerons une étude sur le changement du PIB moyen de l'Afrique. Ce PIB peut s'exprimer comme une moyenne pondérée des différents pays concernés, et son changement comme une la somme d'un effet de composition et un effet de changement, selon la formulation habituelle.

$$\Delta Y = \left[ \sum \bar{y}_j * \Delta w_j \right] + \left[ \sum \bar{w}_j * \Delta y_j \right]$$

↑

Effet de  
composition

↑

Effet de  
comportement

Dans la deuxième phase, l'on peut maintenant exprimer les PIB individuels des pays comme une fonction de la productivité de leurs adultes ( $\pi$ ) et de leur structure d'âge ( $\alpha$ ). Le changement de ces PIB individuels (qui est aussi inclus dans l'équation ci-dessus) s'écrit comme ci-dessous.

$$\Delta y_j = (\bar{\pi}_j * \Delta \alpha_j) + (\bar{\alpha}_j * \Delta \pi_j) \tag{VI.1}$$

En insérant VI.1 dans l'équation de base de la décomposition démographique (II.3), le changement dans le PIB moyen de l'Afrique serait donc

$$\Delta Y = \left[ \sum \bar{y}_j \Delta w_j \right] + \left[ \sum \bar{w}_j \bar{\pi}_j \Delta \alpha_j \right] + \left[ \sum \bar{w}_j \bar{\alpha}_j \Delta \pi_j \right] \tag{VI.2}$$

↑

Effet de la structure  
d'âge des populations

↑

Effet de la productivité  
dependency

↑

Effet de la taille  
des populations

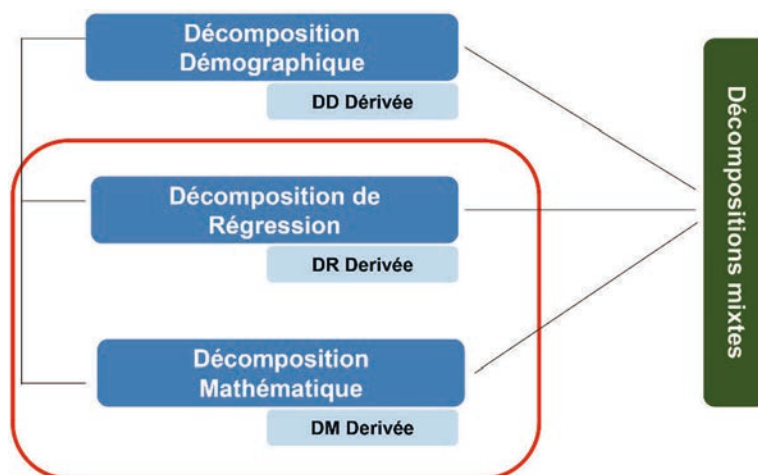


---

## Chapitre VII

# Combinaison régression + mathématique

---



Contrairement aux combinaisons précédentes, les cas couverts dans cette seront moins courants. Considérons la relation supposée entre l'argent et le bonheur. Plusieurs travaux de recherche concluent que cette relation est curvilinéaire, l'argent devenant de moins en moins suffisant pour augmenter son bonheur à mesure que l'on s'enrichit. Supposons néanmoins pour simplifier que cette relation soit linéaire.

$$Y = \alpha + \beta\$ + e \quad \text{[VII.1]}$$

Envisageons ensuite un individu ayant plusieurs sources de revenus (ex., salaire et dons) ou alors que cet individu fasse la différence entre son salaire brut et son salaire net après impôts. Il est possible d'imaginer que ces deux sources de revenus ne soient pas équivalentes sur le plan de la satisfaction qu'elles engendrent. Même si la formule VII.1 est mathématiquement exacte, elle n'est pas intuitivement satisfaisante. L'on peut néanmoins incorporer cette satisfaction différentielle dans l'équation ci-dessous et exprimer les changements historiques dans la satisfaction générée par les salaires brut et net, respectivement.

À l'inverse, l'on peut partir d'une relation mathématique et incorporer une relation de régression. Pour tenter d'illustrer cette situation, reprenons l'exemple des dépenses publiques par enfant. Supposons maintenant, comme cela est probable, que la part du budget alloué aux enfants change en fonction du revenu, ou alors du niveau moyen d'éducation des parents. Dans cette situation, l'analyste peut affiner son travail de décomposition des sources de changements dans les revenus par enfant. Une fois de plus, même si ces possibilités d'extension existent, le chercheur devra faire un arbitrage entre détail et parcimonie, possibilité théorique et application pratique, ou comptabilité rigoureuse et essai d'explication. Les méthodes présentées ici ne sont donc pas des recettes à appliquer mécaniquement, mais plutôt des outils à impliquer sélectivement par le chercheur dans sa quête de compréhension du changement social.



---

## Chapitre VIII

### Résumé et conclusions

---

L'objectif de ce document était d'offrir une brève introduction aux méthodes de décomposition. Ces méthodes peuvent enrichir l'arsenal méthodologique des chercheurs en sciences sociales, et leur apport devrait combler un vide pour les analystes soucieux d'étudier le changement social, mais aussi pour ceux qui sont conscients des limites des méthodes d'analyse actuelles. Même si la décomposition ne permet pas une analyse causale, elle permet de localiser l'origine (sociale, géographique, ou générationnelle) du changement social. Ce faisant, elle réduit la marge d'erreur dans la compréhension du changement social. Dans les limites de ce qu'elle promet, la décomposition est un outil fiable qui, seul ou combiné à d'autres méthodes, peut aider les planificateurs à faire des choix mieux informés dans l'allocation des ressources. Lorsqu'ils sont confrontés par une multitude d'options, ils peuvent s'appuyer sur la décomposition comme l'un de guides pour identifier les options susceptibles d'avoir le plus grand impact.

La flexibilité des méthodes permet d'envisager des adaptations et combinaisons détaillées et créatives, pour répondre à des problèmes précis. Bien que nous n'en ayons présenté ici qu'un échantillon limité, les lecteurs pourront, en fonction de leurs questions de recherche, imaginer d'autres applications. De même, les exemples pris comme illustration dans ce document se réduisent au champ des questions démographiques, mais ces méthodes sont applicables à un large champ de phénomènes sociaux.

Un avantage de la décomposition est sa compatibilité avec d'autres méthodes. En combinant judicieusement la décomposition avec d'autres méthodes, le chercheur peut enrichir l'explication au-delà des possibilités offertes par l'une ou l'autre de ces méthodes prises séparément. Notre objectif n'était donc pas de proposer la décomposition comme substitut, mais comme complément. Dans sa complémentarité avec d'autres méthodes quantitatives et qualitatives, elle permet d'identifier les processus et groupes-clé, tandis que les autres méthodes pourraient élucider les causes de prise de décision individuelle. Cette triangulation de méthodes permet aux chercheurs de mieux remplir leur mission d'explication du changement social, et au planificateur de « transformer le monde » sur la base d'une meilleure compréhension des moteurs du changement social.



---

## Références

---

- Das Gupta, P. 1993. Standardization and Decomposition of Rates: A User's Manual. Current Population Reports Series P23-186. Washington, DC: U.S. Bureau of the Census.
- Eloundou-Enyegue, PM, et Sarah Giroux. 2012. Fertility Transitions and Schooling in SSA: From Micro to Macro-level Effects. *Demography* 49(4):1407-1432.
- Firebaugh, G. and Goelsing, B. 2004. Accounting for the Recent Decline in Global Income Inequality. *American Journal of Sociology* (110): 283-312.
- Kitagawa, EM. 1955. Components of a Difference Between Two Rates, *Journal of the American Statistical Association*, 50(272):1168-1194.
- Luke, D. 2004. *Multilevel Modeling*. Quantitative Applications in the Social Sciences. Sage Publications Inc., Thousand Oaks, Ca.
- Robinson, W.S. (1950). "Ecological correlations and the behavior of individuals." *American Sociological Review* 15:351-57.
- Rodrik, D. 2005. "Why We Learn Nothing from Regressing Economic Growth on Policies" Mimeo, Kennedy School of Government.
- Thornton, A. 2001. "The Developmental Paradigm, Reading History Sideways and Family Change." *Demography*, 38(4): 449-465.
- UNICEF. 2008. The State of Africa's Children 2008. United Nations, New York.
- Vaupel James W. et Vladimir C. Romo. 2003. "Decomposing Change in Life Expectancy: A Bouquet of Formulas in Honor of Nathan Keyfitz's 90th Birthday." *Demography*, 40(2):201-216.
- World Bank. 2005. *World Development Indicators*. Available online at: <http://devdata.worldbank.org/dataonline>.

