

RÉPUBLIQUE ALGÉRIENNE DÉMOCRATIQUE ET POPULAIRE
MINISTÈRE DE L'ENSEIGNEMENT SUPÉRIEUR
ET DE LA RECHERCHE SCIENTIFIQUE
UNIVERSITÉ DR. TAHAR MOULAY-SAIDA-

**Faculté des Sciences Economiques, de Gestion et Sciences
Commerciales**

**Mémoire en vue d'obtention du diplôme de Master en sciences
économiques**

Option : Techniques Quantitatives de Gestion

THÈME :

**L'hétérogénéité des déterminants de dépenses
publiques de l'éducation dans la région MENA**

(de 1990 à 2012)

Étude économétrique à l'aide des données de Panel

Présenté par :

✍ **Falit Cheikh**

✍ **Morsli Aicha**

Sous la direction de :

✍ **Dr. Souar Youcef**

Membres du jury:

✍ **Dr. Diab Zeggay**

Université de Saida

Président.

✍ **Dr. Souar Youcef**

Université de Saida

Encadreur.

✍ **Mr. Boumediene Med Amine**

Université de Saida

Examineur.

Année Universitaire: 2013 / 2014

REMERCIEMENTS

Nous tenons tout d'abord à remercier ALLAH de nous avoir donné le courage et la patience pour accomplir ce travail.

Nous tenons tout particulièrement à remercier monsieur **Souar Youcef** pour l'encadrement de ce mémoire et pour la confiance qu'il nous a accordée, sa disponibilité et ses bons conseils.

Nous voudrions également remercier monsieur **Diab Zeggay** et monsieur **Boumediene Mohamed Amine** pour avoir accepté de faire partie du jury de ce mémoire.

Nos remerciements s'adressent également à tout le corps professoral et administratif de la faculté des Sciences Economiques, de gestion et sciences commerciales de l'Université de SAIDA.

Nos vifs remerciements vont également à :
À nos parents qui nous ont supportés et soutenues durant toutes nos études.

À nos amies : **Chafika, Habiba, Habib et Toufik.**

Nous voudrions remercier toutes les personnes qui, de près ou de loin, ont contribué à l'accomplissement de ce modeste travail.

Dédicaces

Je dédie ce travail à mes très chers parents qui m'ont apporté soutien et réconfort tout au long de ma vie, et qui m'ont orienté par leurs précieux conseils.

Un grand merci à mes frères qui m'ont aidé et encouragé pour l'élaboration de ce travail.

*Mon ami **Farid**, à qui je dois une fière chandelle en guise d'amitié de longues années.*

*À mon très cher ami monsieur **Nour ed-dine Abdallah** qui m'a supporté durant toute la période de mes études.*

CHEIKH

Dédicaces

Je dédie ce travail :

À mon Père qui m'a montré ce chemin...

À ma mère qui a tant veillé sur moi...

À tous mes frères et sœurs

À toute ma famille

À mon binôme et toute sa famille

À mes professeurs

À mes collègues de promotion 2014

À mes collègues de travail CNFPS

À mes proches

À mes amis

À mon frère Benmhamed

AICHA

Table des matières

Table des matières

Remerciements	
Dédicaces	
Table des matières	
Liste des tableaux et figures	
Liste des annexes	
Acronymes et abréviations	
Introduction Générale	01
 <u>Chapitre I</u> : Cadre théorique sur les dépenses de l'éducation et les données de panel	
Introduction	05
<u>Section I</u> : Présentation générale sur les dépenses de l'éducation et les données de panel.....	06
I - Théories et réalités sur les dépenses publiques de l'éducation.....	06
1-Les dépenses publiques de l'éducation	06
1-1-Dépenses de l'état	06
1-2-Dépenses privées	06
2-L'approche théorique des dépenses publiques en capital humain	07
2-1-Les rendements micro-économiques de l'éducation	07
2-2-L'approche Néo-classique	08
2-3-Le lien entre éducation et croissance	09
3-L'investissement dans le capital humain et l'impact sur la croissance dans la région MENA	10
3-1- L'investissement dans le capital humain	10
3-2- Éducation et croissance économique	13
3-3-Évolution des dépenses de l'éducation	14
II - Généralités sur l'économétrie des données de panel	15
1-Introduction à l'économétrie des données de panel	15
1-1-Histoire des données de panel selon Nerlove	15
1-2-Qu'est-ce qu'un panel ?	16

1-3-Avantages et Inconvénients	17
2-Spécification des données de panel	19
2-1-Modèle économétrique linéaire général	19
2-2-Variabilité des données de panel	20
3-Typologie des modèles et différents tests	24
3-1- Le test d'homogénéité vs hétérogénéité (test de Hsiao)	24
3-2- Modèle à effets individuels	33
3-2-1- Le modèle à effets fixes individuels	35
3-2- 2-Le modèle à erreurs composées	38
3-3 Tests de spécification des effets individuels (test de Hausman)	42
<u>Section II</u> : Revue de la littérature	43
Conclusion	55

Chapitre II : Cadre empirique sur les déterminants des dépenses Publiques de l'éducation en MENA

Introduction	56
<u>Section I</u> : présentation du modèle et des données	57
1-Échantillon et sources des données	57
2-Spécification du modèle	58
3-Identification des variables	59
4-Outils et Logiciels	60
<u>Section II</u> : Estimation et interprétation	61
1-L'analyse descriptive des variables	61
2-Le test de spécification d'homogénéité/hétérogénéité (Hsiao)	62
3-Les déterminants des dépenses de l'éducation en Algérie	69
Conclusion	78
Conclusion générale	79
Bibliographie	80
Annexes	82

LISTE DES TABLEAUX

Tableau	Titre	Page
01	Concepts de moyennes et d'écart	21
02	Composition des variabilités et des variances des observations	22
03	Estimation du modèle avec données empilées	64
04	Calcul du Total SCR_i	65
05	Estimation du modèle à effets fixes	67
06	Résultats de la régression pas-à-pas	69
07	Test d'hétéroscédasticité (test de White).	72
08	Test d'autocorrélation des erreurs	74
09	Test de stabilité du modèle (test de Chow)	76

LISTE DES FIGURES

Figure	Titre	Page
01	Taux bruts de scolarisation dans la région MENA et moyenne en nombre d'années de scolarisation dans cette région, en Asie de l'Est et en Amérique latine	12
02	<i>Procédure séquentielle des tests</i>	24
03	Absence d'hétérogénéité	30
04	Hétérogénéité de niveau	31
05	Hétérogénéité de pentes	32
06	Évolution des différents variables du modèle	61
07	Test de normalité	70

LISTE DES ANNEXES

Annexe	Titre	Page
01	Résultats de l'estimation du modèle pour l' Algérie	82
02	Résultats de l'estimation du modèle pour l'Arabie Saoudite	82
03	Résultats de l'estimation du modèle pour Bahreïn	83
04	Résultats de l'estimation du modèle pour l'Égypte	83
05	Résultats de l'estimation du modèle pour l'Émirats arabes unis	84
06	Résultats de l'estimation du modèle pour l'Iran	84
07	Résultats de l'estimation du modèle pour Jordanie	85
08	Résultats de l'estimation du modèle pour Koweït	85
09	Résultats de l'estimation du modèle pour Liban	86
10	Résultats de l'estimation du modèle pour Maroc	86
11	Résultats de l'estimation du modèle pour Oman	87
12	Résultats de l'estimation du modèle pour Qatar	87
13	Résultats de l'estimation du modèle pour Tunisie	88
14	Résultats de l'estimation du modèle pour Turquie	88
15	Résultats de l'estimation du modèle pour Yémen	89

Acronymes et Abréviations

Abréviation	Détails
IMRED	Introduction, Matériels et Méthodes, Résultats Et Discussion
LSDV	Squares Dummy Variable
MCG	Moindres Carrés Généralisés
MCO	Moindres Carrés Ordinaires.
MENA	Middle East and North Africa.
OCDE	Organisation de Coopération et de Développement Économique
PIB	Produit Intérieur Brut.
UNESCO	United Nations, Educational, Scientific and Cultural Organization

Introduction Générale

Introduction Générale

D'un point de vue théorique, les dépenses publiques dites « porteuses de croissance » sont soit les dépenses visant à développer le capital physique de l'état, ce qu'on appelle les dépenses d'investissements publics, soit les dépenses visant à développer le capital humain (dépenses de santé, d'éducation), soit les dépenses visant à développer le capital infra structurel (dépenses de transport et de communication) ou soit les dépenses visant à développer la technologie (dépenses de recherche et développement).

L'éducation est en MENA comme ailleurs l'un des principaux secteurs d'intervention de la puissance publique, qui y consacre une partie de ses ressources budgétaire, les autorités des pays de la région MENA font preuve d'un engagement sans faille vis-à-vis du financement public de l'éducation. Dans la région, la moyenne de l'investissement public dans ce domaine représente environ de 5% du PIB.

La région (MENA) a réalisé de grandes avancées dans l'éducation. Elle a multiplié par quatre le niveau moyen de scolarisation depuis 1960, divisé par deux le taux d'analphabétisme depuis 1980 et est parvenue à la parité presque parfaite entre les sexes à l'école primaire.

Les taux d'inscription dans le système scolaire ont significativement progressé au cours de la dernière décennie, au point que dans la plupart des pays de la région MENA, l'accès à l'enseignement primaire est désormais universel pour les garçons comme pour les filles. Le taux de scolarisation net moyen est passé de 86 à 94 % entre 2000 et 2010. Dans le secondaire également, on observe une amélioration, quoique moins marquée : le taux de scolarisation net y est passé de 62 à 70 % sur la même période.

Résultat direct des efforts déployés par la région pour renforcer l'accès à l'éducation, le taux d'alphabétisation de la population adulte (c'est-à-dire des plus de quinze ans) a fait un bond spectaculaire ces 20 dernières années, passant de 59 % en 1990 à 78 % en 2010.

Généralement il existe différents indicateurs pour déterminer les dépenses publiques consacrées à l'éducation selon plusieurs études empiriques. Ce qui nous permet de poser **la problématique** suivante :

⊕ « *Est-ce que les déterminants des dépenses publiques de l'éducation sont-ils identiques pour tout les pays de la région MENA ?* »

Hypothèses :

- ⊕ Les dépenses publiques de l'éducation sont déterminées avec des indicateurs précisées.
- ⊕ Ces indicateurs ont des impacts directs sur l'évolution de ces dépenses.
- ⊕ Ces déterminant sont : Le PIB par habitant ; le taux d'inscription à l'école secondaire ; la masse monétaire ; l'inscription à l'enseignement supérieur ; le taux d'alphabétisation
- ⊕ Les différents indicateurs sont les mêmes pour l'ensemble de pays de la région **MENA** et ont le même impact sur les dépenses publiques relatifs à l'éducation.

Choix et intérêt du thème :

- ❖ **Thème d'actualité pour les deux niveaux :** d'un part, l'importance des données de panel dans l'économétrie et d'autre part ; **La Déclaration universelle des droits de l'homme** a souligné l'importance de l'éducation comme un droit humain fondamental et un élément essentiel du développement, **L'éducation** englobe la portée des valeurs sociales, la morale, la tradition, la religion, la politique et l'histoire. Il est l'organe acquis des connaissances qui équipe la main-d'œuvre émergente avec les compétences nécessaires pour assurer sa participation active dans le développement économique.
- ❖ Le choix de notre thème n'est pas fortuit, son importance sur les différents plans, économétrique, de développement et économique,

son originalité ; ainsi que sa relation avec notre spécialité, sont des facteurs qui ont motivé notre choix.

- ❖ Thème porteur d'intérêt car il permet de voir en pratique ce que nous avons acquis dans notre cursus universitaire.

Les Objectifs du mémoire :

- Acquisition des concepts clés introduits dans notre mémoire.
- Maîtrise des différents tests de spécification et méthodes d'estimation pour quelques modèles des données de panel.
- Examen de la structure de panel pour les données des dépenses de l'éducation concernant les pays de la région MENA.
- Parmi les objectifs que nous sommes assignées dans le cadre de la réalisation de ce mémoire, c'est de montrer le lien existant entre les dépenses publiques de l'éducation et ses déterminants.
- Cette étude nous aide à situer le rôle et l'impact de ces indicateurs pour l'ensemble des pays de la région MENA.

Méthodologie de recherche :

Pour atteindre ces objectifs que nous nous sommes assignés et afin de répondre à la problématique précédemment citée, trois démarches méthodologiques appropriées sont adoptées :

- ✓ **Une approche descriptive** : consacrée aux définitions et aux aspects conceptuels relatifs à notre travail d'un part ; et revue de la littérature théorique et empirique d'autre part.
- ✓ **Une approche analytique** : utilisée pour les deux chapitres théorique et pratique.
- ✓ **Une approche économétrique** : adoptée par des méthodes quantitatives dans l'analyse empirique pour un panel de 15 pays de la région **MENA** durant la période 1990-2012

Plan de travail :

La démarche méthodologique de notre travail, est basée sur la méthode de « **IMRED** » ; Nous avons divisé notre étude en deux chapitres bien distincts :

Dans le premier chapitre on voudrait exposer deux sections ; la première est consacrée à l'approche théorique des dépenses publiques en capital humain, ainsi généralités sur l'économétrie des données de panel

Dans la seconde, nous avons insistés en particulier sur la revue de la littérature en fonction des études théoriques et empiriques relatives à notre sujet.

Nous avons consacré le deuxième chapitre à une étude empirique en utilisant les données de panel, nous avons intéressé d'une estimation d'un modèle proposé et de tester l'existence de la structure de panel dans notre échantillon.

L'étude empirique a proposé, également, d'évaluer l'impact des quelques déterminants sur l'évolution des dépenses publiques de l'éducation pour les pays de la région **MENA**.

CHAPITRE



Cadre théorique sur les dépenses de l'éducation et les données de panel

Introduction

Le présent chapitre représente un préambule sur les principaux concepts de base des théories de croissance et d'investissement en capital humain, notamment l'analyse des dépenses de l'éducation dans les pays de la région MENA, à l'aide d'économétrie des données de panel, il se divise en deux sections ; la première section débutera par une présentation générale sur les dépenses de l'éducation et les données de panel ; elle sera consacrée à l'approche théorique des dépenses publiques en capital humain, ainsi généralités sur l'économétrie des données de panel ; cependant la seconde section insistera en particulier sur la revue de la littérature en fonction des études théoriques et empiriques qui mettent en avant le rôle primordial concernant les dépenses publiques allouées à l'éducation.

Section I : Présentation générale sur les dépenses de l'éducation et les données de panel

Dans cette première section, qui constitue une simple introduction aux différentes théories sur les dépenses publiques de l'éducation et l'économétrie des données de panel, nous présentons à la sous section **I** l'approche théorique des dépenses publiques de l'éducation, le lien entre éducation et croissance ; et l'investissement dans le capital humain dans les pays de la région **MENA** ; puis en sous section **II**, généralités sur l'économétrie des données de panel.

I- Théories et réalités sur les dépenses publiques de l'éducation

1-Les dépenses publiques de l'éducation

Pour l'instant, nous nous sommes intéressés aux dépenses publiques pour l'éducation et dépenses totales du gouvernement.

Afin d'évaluer l'engagement des gouvernements en faveur de l'éducation, nous nous sommes concentrés sur la part des budgets de l'éducation dans le PIB, à savoir la capacité d'un pays à générer de la richesse. Mais il est également utile de comparer la part du budget de l'éducation dans les dépenses totales du gouvernement. Cette distribution permet de mettre en avant l'importance relative de l'éducation par rapport à d'autres priorités, comme les soins de santé ou la défense.

1-1-Dépenses de l'état

L'analyse des dépenses de l'éducation à la charge de l'État se fera du point de vue du volume et surtout dans son rapport d'évolution avec le budget national. Pour connaître l'importance des efforts consacrés à l'éducation par l'Etat, la mesure de la proportion de celle-ci par rapport au budget de l'Etat s'avère être un bon indicateur.

1-2-Dépenses privées

Les dépenses privées sont principalement les dépenses des familles pour leurs propres enfants. Elles comprennent les droits d'inscription, (généralement faibles, sauf lorsqu'il s'agit d'écoles privées non subventionnées), de transport, de cantines, d'hébergement, de

fournitures scolaires. Mais il existe aussi quelques dépenses privées non familiales. Il peut s'agir de dépenses financées par des entreprises, notamment dans le domaine de la formation professionnelle, de mécènes, qui peuvent faire des dons ou des legs à des institutions éducatives, ou de fondations (cas où l'on ne dépense que les revenus produits par une donation en capital). A l'échelle mondiale, l'ensemble des dépenses privées autres que celles des familles reste un mode de financement très marginal de l'éducation (moins de 1% du total), mais on observe quelques exceptions, comme les Etats-Unis pour le mécénat aux universités ou l'Allemagne pour les entreprises dans le domaine de la formation professionnelle. Enfin, dans les pays les moins développés, il existe une source de financement extérieure constituée par des aides en provenance des pays riches.

2-L'approche théorique des dépenses publiques en capital humain

2-1-Les rendements micro-économiques de l'éducation

Il existe toute une littérature, micro-économétrique, sur les rendements de l'éducation. Le point de référence de cette littérature est ce que l'on appelle « l'équation de Mincer ». Cette équation exprime le logarithme du salaire d'un individu comme fonction linéaire du nombre d'années d'études de cet individu. On mesure ainsi le rendement d'une année supplémentaire d'étude par son effet marginal sur le salaire, en partant du point de vue qu'à l'équilibre tout individu est indifférent entre l'option de poursuivre ses études et celle d'intégrer immédiatement le marché du travail : conformément à la théorie du capital humain, développée par **Becker (1964)**, les agents économiques, pour décider de poursuivre ou non leurs études, arbitrent en effet entre le surcroît de salaire qu'ils vont en retirer une fois entrés sur le marché du travail et la perte de revenu (salaire et frais d'études) qu'ils subissent en poursuivant leurs études pendant une année supplémentaire.

2-2-L'approche Néo-classique: seule compte l'accumulation de capital humain

Jusqu'à récemment, l'analyse de l'impact macro-économique du système éducatif sur la croissance a consisté pour l'essentiel à appliquer au niveau macro-économique l'approche micro-économique mentionnée plus haut. Ainsi l'hypothèse d'un salaire réel égal à la productivité marginale du travail sous-tend les premières études sur éducation et croissance basées sur le modèle de croissance néo- classique (**Solow, 1956**), en particulier l'article pionnier de **Mankiw, Romer et Weil (1992)**. L'idée de base dans cette approche est que le capital humain joue le même rôle dans la production que le capital physique, et qu'accumuler des années d'études revient à démultiplier la force de travail, autrement dit à en augmenter l'efficacité productive à technologie constante ; ce surcroît d'efficacité permet de compenser les rendements décroissants du capital et par conséquent de soutenir la croissance dans le long terme.

Le taux de croissance du **PIB** par tête pendant une période donnée est donc proportionnel au taux de croissance du niveau de l'éducation pendant la même période, le facteur de proportionnalité étant le rendement macroéconomique de l'éducation. Avec un progrès technique purement exogène, la contribution de l'éducation à la croissance se limite à ce seul impact sur l'efficacité de la force de travail. Pour maintenir une croissance positive à long terme il faut donc augmenter toujours le niveau de l'éducation de la population.

Cette approche néoclassique reçut un important soutien empirique dans l'article de **Mankiw, Romer et Weil**. Ceux-ci en effet mettent en évidence, sur données en coupe transversale, un effet significatif du taux de scolarisation des **12-17** ans (mesuré en moyenne sur la période **1960-1985**) sur le niveau du **PIB** par tête en **1985** (dans une régression par les moindres carrés ordinaires où figurent également le niveau initial du **PIB** par tête, le taux d'investissement et le taux de croissance du capital par tête).

2-3-Le lien entre éducation et croissance passe aussi par le progrès technologique

La critique la plus fondamentale de l'approche suivie par **Mankiw, Romer** et **Weil**, fut émise par **Benhabib** et **Spiegel (1994)** : contrairement à ce que suggère le modèle néo-classique, l'étude de ces auteurs montre un effet positif et significatif du niveau de capital humain (et non pas du taux de croissance de ce niveau), mesuré par le nombre d'années d'études moyen parmi la population active au début de la période considérée (**1965-1985**), sur le taux de croissance moyen du **PIB** par tête.

En remettant en cause l'approche néo-classique, purement basée sur d'accumulation du capital, l'article de **Benhabib** et **Spiegel** a remis à l'honneur une vision plus « technologique » du rôle de l'éducation dans la croissance économique, qui avait été développée de façon embryonnaire par **Nelson** et **Phelps (1966)**. Ces derniers considéraient en effet comme trop réductrice la vision néoclassique standard, selon laquelle les travailleurs très éduqués et peu éduqués sont des substituts parfaits, ne différant que par le nombre d'unités de travail « efficaces » dont ils sont dotés. Ils motivaient notamment leur critique en prenant l'exemple de la diffusion d'innovations dans le domaine agricole : des études ont en effet montré que ce sont les agriculteurs les plus éduqués qui adoptent les nouveaux produits et processus les premiers, et que les agriculteurs les moins éduqués ne s'adaptent au progrès technique que bien plus tard.

L'analyse de **Nelson** et **Phelps** conduit donc à la conclusion suivante : dans une économie avec progrès technique, le niveau d'éducation affecte la croissance de long terme à travers ses effets sur la vitesse d'adaptation au changement technologique. Dans leur remise en cause de l'approche néo-classique, **Benhabib** et **Spiegel** ont sans doute été un peu loin en niant toute contribution de l'accumulation de capital humain à la croissance de long terme.

Dans un article influent dans lequel ils procèdent à une revue critique de la littérature sur le sujet, **Krueger** et **Lindhal (2003)** montrent que cette dernière conclusion n'est pas robuste, notamment parce qu'elle

repose sur une mesure erronée du capital humain – **Benhabib** et **Spiegel** utilisent le logarithme du nombre d'années d'éducation dans une version macro-économique de l'équation de Mincer, alors qu'il faut simplement utiliser le nombre d'années. Au terme d'un travail économétrique soigneux et utilisant de meilleures bases de données, **Krueger** et **Lindahl** mettent en évidence un rôle significatif sur la croissance à la fois de l'accumulation et du niveau initial de capital humain. (pour plus de détails référer à la section II).

3-L'investissement dans le capital humain et l'impact sur la croissance dans la région MENA

3-1- L'investissement dans le capital humain

Les pays de la région **MENA** ont commencé à investir plus tardivement dans le capital humain que ceux d'autres régions. Mais une fois commencé, ils ont de manière générale dépensé un important pourcentage de leur **PIB** à l'éducation et ont rapidement augmenté le niveau de scolarisation au sein de leurs populations. Au cours des dernières **40** années, les pays de la région **MENA** ont consacré en moyenne **5 %** de leur **PIB** et **20 %** des dépenses publiques à l'éducation, c'est-à-dire, plus que d'autres pays en développement à revenu similaire par tête. Ainsi, la région a été capable d'améliorer l'accès équitable à l'éducation à tous les niveaux d'instruction. Il s'agit de réalisations impressionnantes, tenant compte que les pays de la région **MENA**, au cours des années **60**, ont commencé avec des indicateurs éducatifs parmi les plus bas au monde. (Rapports sur le développement de la région MENA fait par la banque mondiale, 2007).

À quelques exceptions près, les pays de la région **MENA** ont réussi à atteindre une scolarisation universelle au niveau primaire, et à multiplier par trois l'inscription à l'enseignement secondaire entre **1970** et **2003** et par cinq l'inscription à l'enseignement supérieur. La plus grande réalisation des pays de la région **MENA** est d'avoir éliminé l'écart éducatif entre les sexes. La parité entre garçons et filles dans l'enseignement fondamental est pratiquement achevée. Bien que la région ait commencé à partir de niveaux relativement bas en matière d'égalité entre les sexes, les indices pour l'éducation secondaire et supérieure ne diffèrent pas de manière significative de ceux d'Amérique

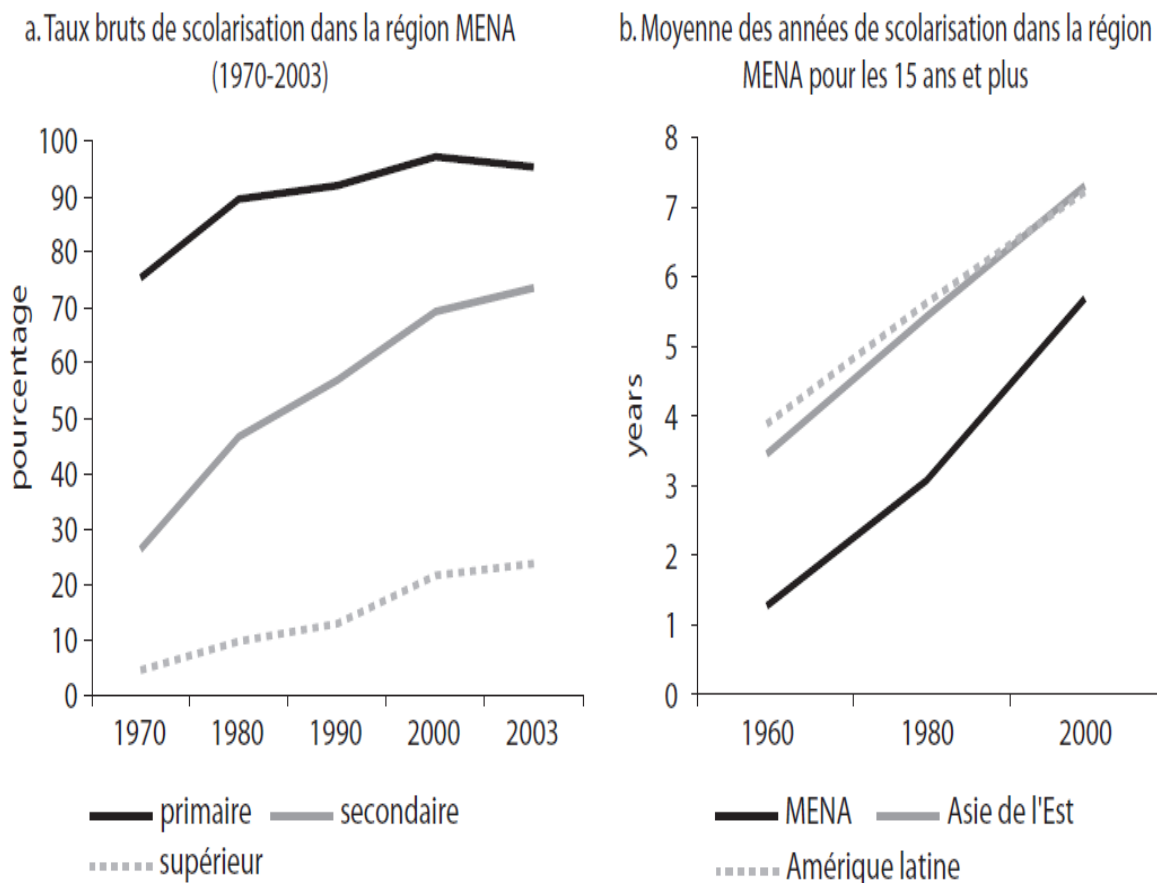
latine ou d'Asie de l'Est. L'investissement a également permis une amélioration des niveaux d'éducation ainsi que de certaines compétences de base. Les taux d'analphabétisme ont diminué de moitié au cours des dernières **20** années et l'écart absolu entre le taux d'alphabétisation des hommes et des femmes a diminué rapidement. En termes des résultats d'apprentissage des élèves, certains pays de la région **MENA** font également montre d'un bon résultat en matière de tests internationaux, notamment pour ce qui est des Tendances dans les études internationales sur les mathématiques et la science (**TIMSS**), lorsque l'on tient compte du niveau de revenus et des taux bruts de scolarisation.

En dépit de ces améliorations les progrès en matière d'éducation dans les pays de la région **MENA** restent inférieurs à ceux d'autres pays à niveaux similaires de développement économique. L'héritage de bas niveaux initiaux de développement de l'éducation n'a pas encore été surmonté. Bien que la région ait déjà rattrapé l'Asie de l'Est et l'Amérique latine en termes de scolarisation primaire universelle, elle est encore à la traîne en matière de scolarisation secondaire et supérieure (**figure 1 a**).

Le nombre d'années de scolarisation dans les pays de la région **MENA** est par conséquent inférieur de plus d'une année comparé aux deux autres régions (**figure 1b**). En outre, la distribution des résultats éducatifs (mesurée par l'écart type des années de scolarisation) est devenue plus inégale dans le temps en comparaison de l'Amérique latine ou de l'Asie de l'Est.

Par ailleurs, les taux d'analphabétisme restent inférieurs à ceux d'autres régions. Bien que les différences aient diminué depuis les années **50**, l'analphabétisme dans les pays de la région **MENA** est toujours deux fois plus élevé qu'en Asie de l'Est et en Amérique latine. En outre, bien que les résultats des tests internationaux indiquent que les niveaux atteints sont proches des prédictions tenant compte du **PIB** par tête et des taux d'inscription, ils restent inférieurs aux résultats enregistrés dans les pays à revenu moyen à développement économique rapide, tels que la Corée et la Malaisie.

Figure 1 : Taux bruts de scolarisation dans la région MENA et moyenne en nombre d'années de scolarisation dans cette région, en Asie de l'Est et en Amérique latine



Source : Appendice statistique

Source : Barro et Lee (2000)

Enfin, dans plus de la moitié des pays de la région **MENA**, environ deux tiers des étudiants se spécialisent dans les sciences sociales et humaines plutôt que dans les sciences et les mathématiques. Cette tendance est contraire à ce qui se passe en Asie de l'Est, et dans une moindre mesure en Amérique latine. Étant donné que l'innovation et l'adaptation technologiques jouent un rôle de plus en plus dominant dans le processus de développement. (Banque mondiale, Rapports sur le Développement de la région MENA).

3-2- Éducation et croissance économique

L'éducation est nécessaire mais seule ne suffit pas pour générer la croissance économique. Des recherches sur la région **MENA** viennent en appui à ce point de vue. La croissance économique par tête au cours des **20** dernières années a été relativement limitée malgré les améliorations survenues dans les résultats éducatifs. Paradoxalement, une plus forte croissance économique a été accompagnée de faibles niveaux d'éducation dans les années **60** et **70**.

De même, la productivité totale des facteurs mesurant l'impact des facteurs autres que le capital physique et humain, a été faible ou négative dans la région **MENA** au cours des années 80 et **90**, période au cours de laquelle les résultats éducatifs s'amélioraient. Cela donne à penser que le niveau éducatif n'a pas contribué de façon significative à la croissance économique ou à la productivité de la région.

Il y a des explications plausibles de la faible relation entre l'éducation et la croissance économique dans la région **MENA**. L'une d'entre elles est que la qualité de l'éducation est trop faible pour que la scolarisation puisse contribuer à la croissance et à la productivité. Une autre est que c'est le niveau *relatif* plutôt qu'*absolu* des résultats éducatifs qui explique la relation ténue entre l'éducation et la croissance économique dans les pays de la région **MENA**. L'investissement étranger direct, par exemple, a plutôt été orienté vers les pays ayant de meilleurs résultats éducatifs, toutes choses restant égales par ailleurs. Une troisième explication possible serait en relation avec la variance du niveau d'instruction, qui est, dans la région **MENA**, plus grande que dans les autres régions : les recherches internationales indiquent qu'une distribution plus égale du résultat éducatif est positivement corrélée avec une croissance économique supérieure. Enfin, la faible interaction entre les résultats pédagogiques et la croissance économique pourrait également être liée aux hauts niveaux d'emploi dans le secteur public et au petit nombre de secteurs économiques dynamiques et compétitifs sur le plan international.

3-3-Évolution des dépenses d'éducation

Les dépenses publiques d'éducation sont le plus souvent mesurées en pourcents du Produit Intérieur Brut (PIB) d'un pays donné. Jusqu'au vingtième siècle, ce pourcentage est resté modeste, soit en dessous d'un pourcent. La progression de cet indicateur est donc un phénomène relativement récent dans l'histoire. Les dépenses publiques d'éducation ont connu leur heure de gloire durant le troisième quart du vingtième siècle, entre 1950 et 1975. Durant cette période, elles sont passées à cinq pourcents du PIB (moyenne mondiale). Puis brutalement, à partir de 1975, cet indicateur a cessé de croître, et il s'est maintenu jusqu'à nos jours à cinq pourcents.

Pour la région **MENA**, la valeur de cet indicateur est souvent inférieure à la moyenne (un peu moins de 5%). (**Institut de statistique de l'UNESCO, Statistiques comparées sur l'éducation dans le monde, 2007**).

II- Généralités sur l'économétrie des données de panel

1-Introduction à l'économétrie des données de panel

Dans cette deuxième sous section, qui constitue une simple introduction à l'économétrie des données de panel, nous présentons l'histoire des données de panel, la définition d'un panel et les avantages versus les inconvénients des données de panel, le modèle linéaire générale, variabilité et modélisation de l'hétérogénéité.

1-1-Histoire des données de panel selon Nerlove

Gauss [1809], Legendre [1805]. Méthode des moindres carrés ordinaires appliqués aux mesures en astronomie.

Airy [1861]. Astronome anglais, ayant notamment réalisé des mesures précises pour le méridien de Greenwich. Première spécification d'un modèle à effets aléatoires: plusieurs observations du diamètre de Jupiter chaque nuit avec un télescope, avec des conditions atmosphérique entraînant une erreur de mesure spécifique à chaque nuit (constituant un groupe d'observation) en plus des erreurs des mesures spécifiques à chaque observations. Fait des transformations par rapport aux erreurs moyennes à chaque nuit avec d'évaluer l'erreur de mesure spécifique à chaque nuit.

Ronald A. Fisher [1918, 1925]. Moindres carrés ordinaires et analyse de la variance.

Hoch et Mundlak [1955, 1958]. Modèles à effets fixes pour estimer des fonctions de production sur données d'entreprise.

Balestra et Nerlove [1966]. Modèle à effets aléatoires (de plus dynamique) pour estimer la consommation de gaz dans différents États américains.

Conférence à Paris vers 1977, développement dans les années 90 avec plus grande disponibilité des données et applications disponibles dans les logiciels de statistiques et d'économétrie. Rapprochement récent avec les tests de racines unitaires et les séries temporelles.

Quelques autres noms en économétrie de panel théorique: Maddala (décédé), Hsiao, Baltagi, Arrelano.

Quelques autres noms en économétrie de panel appliquée: Griliches (décédé), Blundell, Bond, Meghir, Mairesse, Sevestre.

2-Qu'est-ce qu'un panel ?

Définition

A longitudinal, or panel, data set is one that follows a given sample of individuals over time, and thus provides multiple observations on each individual in the sample. (Hsiao, 1986, page 2).

On appelle panel (en anglais : cross sectional-time series data or longitudinal) une base de données dans laquelle deux dimensions sont présentes :

- une dimension individuelle –transversale- (au sens large) : il peut s'agir de personnes physiques, de ménages, d'entreprises, de pays, etc. On note ces individus $i = 1...N$.
- une dimension temporelle –longitudinale- : il existe plusieurs observations au cours du temps de ces individus. On note les périodes d'observations $t = 1...T$.

Les panels sont donc une combinaison de série temporelle et de coupe transversale ; ils permettent de combler en partie les limites de ces deux types d'échantillon :

- le fait qu'il y ait deux dimensions dans ces données permet d'augmenter le nombre d'observations et la variabilité de l'échantillon, ce qui permet d'obtenir des estimations plus précises.
- l'hétérogénéité individuelle peut être prise en compte de manière beaucoup plus satisfaisante dans le cadre de données de panel car on dispose de plusieurs observations des mêmes individus.

◇ Distinction entre panel cylindré ou non cylindré

➤ *Panel cylindré –équilibré- (balanced panel):*

Les observations de tous les individus du panel sont disponibles pour exactement les mêmes dates T .

➤ *Panel non cylindré –déséquilibré- (unbalanced panel):*

Au moins une observation temporelle pour un des individus n'est pas disponible par rapport aux autres individus. Ce problème peut dans certains cas compliquer les calculs des estimateurs. On peut bien sûr extraire un panel cylindré à partir d'un panel non cylindré en réduisant le nombre d'observations temporelles des individus.

3-Avantages et Inconvénients des données de panel

3-1- Rôle et avantages des données de Panel

Le modèle des données de Panel présente une série d'avantages entre autres :

◇ La double dimension des données (individuelle et temporelle) permet de rendre compte simultanément de la dynamique des comportements et de leur éventuelle hétérogénéité, ce qui constitue un avantage par rapport aux autres types de données que sont les séries temporelles et les coupes transversales. A ces avantages viennent s'ajouter d'autres liés au nombre très élevé de données et leur variabilité. (**Patrick Sevestre, 2002**).

L'hétérogénéité est subdivisée en hétérogénéité observée et inobservée. La première peut être contrôlée à travers les variables explicatives tandis que l'hétérogénéité inobservée peut être contrôlée à travers l'analyse en coupe instantanée et sur une série temporelle.

Ne pas tenir compte de l'hétérogénéité inobservée conduirait à une inefficacité des estimateurs des Moindres Carrées Ordinaires, soit un estimateur biaisé.

◇ L'autre avantage est l'accroissement de la taille de l'échantillon, ceci conduit à l'augmentation du nombre de degré de liberté et à la réduction de la colinéarité entre les variables explicatives améliorant par les estimations des résultats obtenus sur base des tests économétriques.

◇ Avec la double dimension des données de panel, il en résulte une interprétation facile des résultats.

En effet, la différence des individus, les uns par rapport aux autres crée une variabilité inter-individuelle. Cette dernière pouvant être décomposée en variabilité intra-individuelle temporelle proposée à chaque individu et en variabilité inter-individuelle.

3-2-Inconvénients des données de panel

◇ Le biais de sélection et problème des données de panel incomplet à cause de manque des données suffisantes sur l'ensemble des observations représentent les problèmes les plus récents des données de panels. Lorsque toutes les observations ne sont pas disponibles : le panel est dit non cylindré ou déséquilibré. Il y a tendance à cylindrer l'échantillon en éliminant certaines observations pour que chaque individu ait le même nombre d'observation.

◇ En effet, l'utilisation de panel non cylindré conduit à des problèmes d'hétéroscédasticité et ou auto-corrélation des erreurs. L'autre problème que présentent les données de panel réside dans la fréquence non négligeable des sources des données statistiques utilisées ou la présence d'observation aberrante. Huber(1981) a montré que seulement 3% des valeurs suffisantes dans l'ensemble d'observation a perturbé significativement la qualité d'observation.

(www.memoireonline.com Consulté le 15/01/2014).

⊕ En dépit de ces limites, les économistes s'accordent de dire que les avantages des données de panel sont plus importants que les inconvénients et préconisent que chaque fois qu'il y a moyen de faire une analyse en panel il faut la privilégier par rapport au modèle en coupe instantanée ou la série temporelle l'analyse.

2-Spécification des données de panel

2-1-Modèle économétrique linéaire général

Le modèle en données de panel peut s'écrire pour N individus ($i=1, \dots, N$) et T observations temporelles ($t=1, \dots, T$), soit $n=N \times T$ observations totales, de la manière suivante :

$$y_{it} = \alpha_{it} + \sum_{k=1}^K \beta_{kit} X_{kit} + \varepsilon_{it} \quad (i=1, \dots, N) \quad (t=1, \dots, T)$$

y_{it} : variable endogène observée pour l'individu (i) à la période (t).

X_{kit} : matrice des K variables exogènes ; X_{kit} est donc la valeur observée pour la $k^{\text{ème}}$ variable exogène pour l'individu (i) à l'instant (t).

α_{it} : terme constant pour l'individu (i) à l'instant (t).

ε_{it} : terme d'erreurs (les perturbations aléatoires).

☑ Les hypothèses du modèle :

On distingue les hypothèses stochastiques (liées à l'erreur ε) des hypothèses structurelles.

1. *Hypothèses stochastiques :*

♦ **H1** : les valeurs X_{it} sont observées sans erreur.

♦ **H2** : $E(\varepsilon_{it})=0$, l'espérance mathématique de l'erreur est nulle.

Où ε_{it} est une variable aléatoire distribuée selon une loi normale d'espérance nulle : $\varepsilon_{it} \sim N(0, \sigma^2_{\varepsilon}) \quad \forall i, \forall t$.

♦**H3** : $E(\varepsilon_{it}^2) = \sigma_\varepsilon^2$, la variance de l'erreur est constante ($\forall t$) (homoscédasticité).

♦**H4** : $E(\varepsilon_{it}, \varepsilon_{jt}) = 0$ si $i \neq j$, $E(\varepsilon_{it}, \varepsilon_{is}) = 0$ si $t \neq s$; les erreurs sont non corrélées (ou encore indépendantes).

♦**H5** : $\text{Cov}(X_{it}, \varepsilon_{it}) = 0$, l'erreur est indépendante des variables explicatives.

2. Hypothèses structurelles :

♦**H6** : absence de colinéarité entre les variables explicatives, cela implique que la matrice $(X'X)$ est régulière et que la matrice inverse $(X'X)^{-1}$ existe.

♦**H7** : $(X'X)/n$ tend vers une matrice finie non singulière.

♦**H8** : $n > k+1$, le nombre d'observations est supérieur au nombre des séries explicatives.

On notera que les erreurs du modèle linéaire sont supposées être i.i.d. (indépendantes et identiquement distribuées), de loi normale.

$$\varepsilon_{it} \text{ i.i.d. } \sim N(0, \sigma_\varepsilon^2) \forall i, \forall t.$$

2-2-Variabilité des données de panel

La grande variabilité des données de panel tient surtout à leur caractère individuel, et elle se manifeste dans leur dimension transversale (individuelle) et longitudinale (temporelle), c'est-à-dire en coupe comme en série. Elles permettent d'apprécier et de comparer leur variabilité dans ces deux dimensions.

Le tableau (1) fournit les différents concepts de moyennes et d'écarts, puis le tableau (2) fournit les différentes décompositions des variabilités et des variances des observations :

TABLEAU 1: Concepts de moyennes et d'écarts

Moyennes, écarts (terminologie)	Notations
Moyennes individuelles	$\bar{y}_i = \frac{1}{T} \sum_{t=1}^T y_{it}$
Moyennes temporelles	$\bar{y}_t = \frac{1}{N} \sum_{i=1}^N y_{it}$
Moyenne générale	$\bar{y} = \frac{1}{NT} \sum_{i=1}^N \sum_{t=1}^T y_{it}$
Écarts aux moyennes individuelles	$y_{it} - \bar{y}_i$
Écarts aux moyennes temporelles	$y_{it} - \bar{y}_t$
Moyennes individuelles en écarts à la moyenne générale	$\bar{y}_i - \bar{y}$
Moyennes temporelles en écarts à la moyenne générale	$\bar{y}_t - \bar{y}$
Écarts à la moyenne générale	$y_{it} - \bar{y}$

Source : Alain Pirotte, 2011

Variabilités	Notations	Variances
Variabilité totale	$Var_{tot} = \sum_{i=1}^N \sum_{t=1}^T (y_{it} - \bar{y})^2$	$\sigma_{tot}^2 = \frac{Var_{tot}}{NT - 1}$
Variabilité inter-individuelle	$Var_{inter-ind} = T \sum_{i=1}^N (\bar{y}_i - \bar{y})^2$	$\sigma_{inter-ind}^2 = \frac{Var_{inter-ind}}{T(N - 1)}$
Variabilité inter-temporelle	$Var_{inter-temp} = N \sum_{t=1}^T (\bar{y}_t - \bar{y})^2$	$\sigma_{inter-temp}^2 = \frac{Var_{inter-temp}}{N(T - 1)}$
Variabilité intra-individuelle	$Var_{intra-ind} = \sum_{i=1}^N \sum_{t=1}^T (y_{it} - \bar{y}_i)^2$	$\sigma_{intra-ind}^2 = \frac{Var_{intra-ind}}{N(T - 1)}$
Variabilité intra-temporelle	$Var_{intra-temp} = \sum_{i=1}^N \sum_{t=1}^T (y_{it} - \bar{y}_t)^2$	$\sigma_{intra-temp}^2 = \frac{Var_{intra-temp}}{T(N - 1)}$
Variabilité intra-individuelle-temporelle	$Var_{intra-ind-temp} = \sum_{i=1}^N \sum_{t=1}^T (y_{it} - \bar{y}_i - \bar{y}_t + \bar{y})^2$	$\sigma_{intra-ind-temp}^2 = \frac{Var_{intra-ind-temp}}{(N - 1)(T - 1)}$

TABEAU 2 : Décomposition des variabilités et des variances des observations.

Source : Alain Pirotte, 2011

On peut alors définir plusieurs décompositions de la variabilité totale des observations, ces décompositions sont données comme suit :

- Variabilité totale = variabilité inter-individuelle + variabilité intra-individuelle.
- Variabilité totale = variabilité inter-temporelle + variabilité intra-temporelle.
- Variabilité totale = variabilité intra-individuelle-temporelle
+ variabilité inter-individuelle
+ variabilité inter-temporelle.

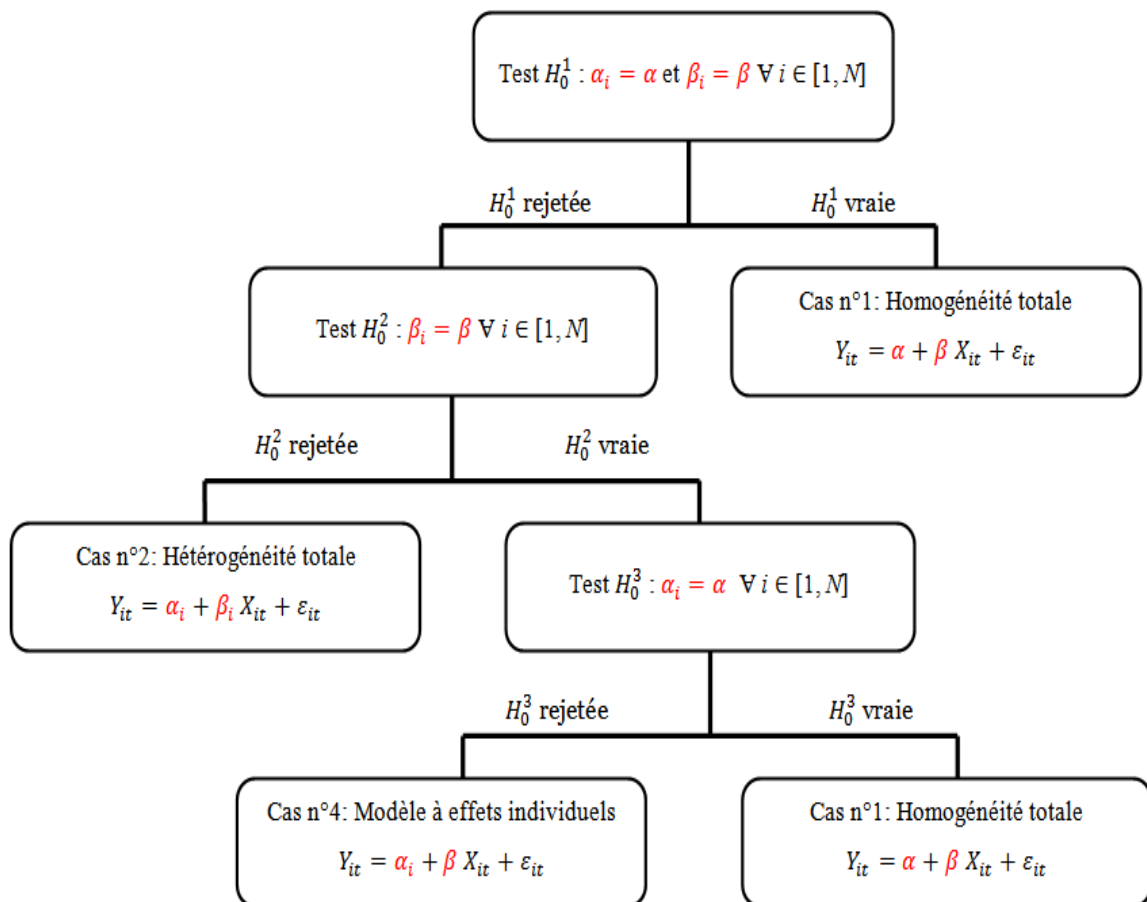
3-Typologie des modèles et différents tests

3-1-Le test d'homogénéité vs hétérogénéité

A. Procédure séquentielle de tests :

Le choix de la spécification (homogénéité, hétérogénéité) est donc très important (nous nous intéressons ici d'un modèle linéaire simple). Afin de déterminer la structure du panel, **Hsiao (1986)** propose une procédure séquentielle de tests (cf. **Figure 2**) permettant de définir le cas dans lequel nous nous situons. (**Régis Bourbonnais, 2011, P 347**).

Figure 2 : Procédure séquentielle des tests



◆ À partir de cette spécification de tests nous pouvons envisager quatre (4) possibilités.

Cas n° 1 : homogénéité totale.

Les constantes α_i et les coefficients β_i sont tous identiques pour tous les individus, nous avons $\alpha_i = \alpha$ et $\beta_i = \beta$ pour toutes les valeurs de (i). Le modèle ne comporte qu'une seule équation estimée sur $n=N \times T$ observations empilées par les MCO (ou les MCG selon la structure de la matrice des variances et covariances des erreurs).

Cas n° 2 : hétérogénéité totale.

Les constantes α_i et les coefficients β_i sont tous différents pour toutes les valeurs de (i), la structure en panel est rejetée. Le modèle doit être estimé équation par équation pour les N équations (une équation par individu) par les MCO (ou les MCG selon la structure de la matrice des variances et covariances des erreurs).

Cas n° 3 : hétérogénéité des coefficients des variables explicatives et homogénéité des termes constants.

Les constantes α_i sont toutes identiques $\alpha_i = \alpha$ pour les individus, mais les coefficients β_i des variables explicatives sont différents pour chaque individu. Comme au cas n°2, le modèle doit être estimé sur les N équations (une équation par individu) par les MCO (ou les MCG selon la structure de la matrice des variances et covariances des erreurs).

Cas n° 4 : hétérogénéité des termes constants et homogénéité des coefficients des variables explicatives (le modèle à effets individuels).

Les constantes α_i sont différentes pour les individus, mais les coefficients β_i des variables explicatives sont constants pour les individus $\beta_i = \beta$. Ce modèle est appelé « modèle à effets individuels ».

B. Construction des tests :

Les tests d'hypothèses sont construits à partir des statistiques de **Fisher** (test de Wald de restrictions sur les coefficients).

1) Test

$$H_0^1 : \alpha_i = \alpha \text{ et } \beta_i = \beta \quad \forall i \in [1, N]$$

$$H_a^1 : \exists (i, j) \in [1, N] / \alpha_i \neq \alpha_j \text{ ou } \beta_i \neq \beta_j$$

Ce test d'hypothèses jointes se ramène à un test de Fisher dont la statistique est donnée par :

$$F_1 = \frac{(SCR_{c1} - SCR)/(N - 1)(K + 1)}{SCR/[N * T - N(K + 1)]}$$

SCR_{c1} : somme des carrés des résidus du modèle contraint sous l'hypothèse **H₀¹**, soit à estimer par les **MCO** le modèle en empilant toutes les observations. Le degré de liberté est égal à : (**N * T**=nombre total d'observations)-(**K+1**=nombre de coefficients à estimer).

SCR : somme des carrés des résidus du modèle non contraint, elle est égale à la somme des **N** sommes des carrés des résidus des modèles estimés sur les **T** observations de chaque équation individuelle, soit **SCR = $\sum_{i=1}^N SCR_i$** . Le degré de liberté est donc la somme des **N** degrés de liberté de chaque équation estimée, soit

$$ddl = \sum_{i=1}^N [T - (K + 1)] = N * T - N(K + 1).$$

Le degré de liberté du numérateur est donc égal à la différence des degrés de liberté de **SCR_{c1}** et **SCR** :

$$ddln = [(N * T) - (K + 1)] - [N * T - N(K + 1)].$$

$$ddln = (N - 1)(K + 1).$$

- La statistique F_1 est à comparer à la valeur lue dans la table de Fisher aux degrés de liberté du numérateur et du dénominateur. Si $F_1 > F_{ddl_n, ddld}^\alpha$, nous rejetons l'hypothèse H_0^1 au seuil α .

2) Test

$$H_0^2 : \beta_i = \beta \quad \forall i \in [1, N]$$

$$H_a^2 : \exists (i, j) \in [1, N] / \beta_i \neq \beta_j$$

Ce test d'hypothèses jointes se ramène à un test de Fisher dont la statistique est donnée par :

$$F_2 = \frac{(SCR_{c2} - SCR) / [(N - 1) * K]}{SCR / [N * T - N(K + 1)]}$$

SCR_{c2} : somme des carrés des résidus du modèle contraint sous l'hypothèse H_0^2 , soit à estimer le modèle à effets fixes individuels. Le degré de liberté est égal à : $(N * T = \text{nombre total d'observations}) - (K + N = \text{nombre de coefficients à estimer})$, nous estimons K coefficients et N termes constants.

SCR : somme des carrés des résidus du modèle non contraint. Le degré de liberté du numérateur est donc égal à :

$$ddl_n = [(N * T) - (K + N)] - [(N * T) - N(K + 1)].$$

$$ddl_n = (N - 1) * K.$$

- La statistique F_2 est à comparer à la valeur lue dans la table de Fisher aux degrés de liberté du numérateur et du dénominateur. Si $F_2 > F_{ddl_n, ddld}^\alpha$, nous rejetons l'hypothèse H_0^2 au seuil α .

3) Test

$$H_0^1 : \alpha_i = \alpha \quad \forall i \in [1, N]$$

$$H_a^1 : \exists (i, j) \in [1, N] / \alpha_i \neq \alpha_j$$

Ce test d'hypothèses jointes se ramène à un test de Fisher dont la statistique est donnée par :

$$F_3 = \frac{(SCR_{c1} - SCR_{c2}) / (N - 1)}{SCR_{c2} / [N * (T - 1) - K]}$$

SCR_{c2} : somme des carrés des résidus du modèle contraint sous l'hypothèse H_0^2 .

SCR_{c1} : somme des carrés des résidus du modèle contraint sous l'hypothèse H_0^1 . Le degré de liberté du numérateur est donc égal à :

$$ddln = [(N * T) - (K + 1)] - [(N * T) - (K + N)].$$

$$ddln = (N - 1).$$

- La statistique F_3 est à comparer à la valeur lue dans la table de Fisher aux degrés de liberté du numérateur et du dénominateur. Si $F_3 > F_{ddln, ddld}^\alpha$, nous rejetons l'hypothèse H_0^3 au seuil α .

Pour repérer les différentes sources et formes d'hétérogénéité, nous devons définir la modélisation de cette hétérogénéité.

◆ Modélisation de l'hétérogénéité

On parle d'hétérogénéité lorsque les individus (et/ou les périodes) présentent des différences systématiques. Pour simplifier, nous ignorons pour l'instant l'hétérogénéité temporelle et nous raisonnons dans le cadre habituel du modèle linéaire simple. On a donc le modèle suivant :

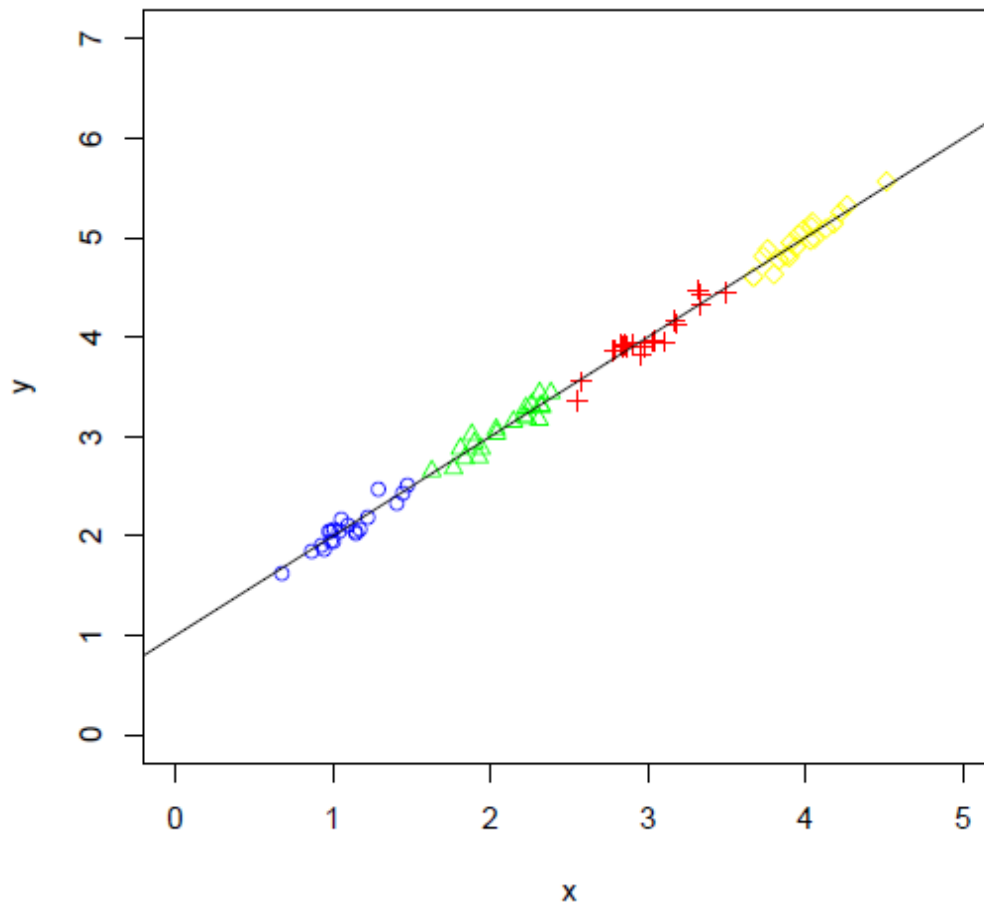
$$y_i = \alpha + \beta X_i + \varepsilon_i$$

L'hétérogénéité individuelle peut prendre différentes formes :

- ⊕ La moyenne de la variable explicative peut être différente d'un individu à un autre,
- ⊕ L'ordonnée à l'origine de la droite peut être différente d'un individu à un autre,
- ⊕ La pente de la droite peut être différente d'un individu à un autre.

Ces trois cas sont d'une nature très différente et nécessitent l'utilisation de modèles spécifiques. Pour l'instant, nous allons illustrer graphiquement ces trois formes d'hétérogénéité.

Figure 03 : Absence d'hétérogénéité



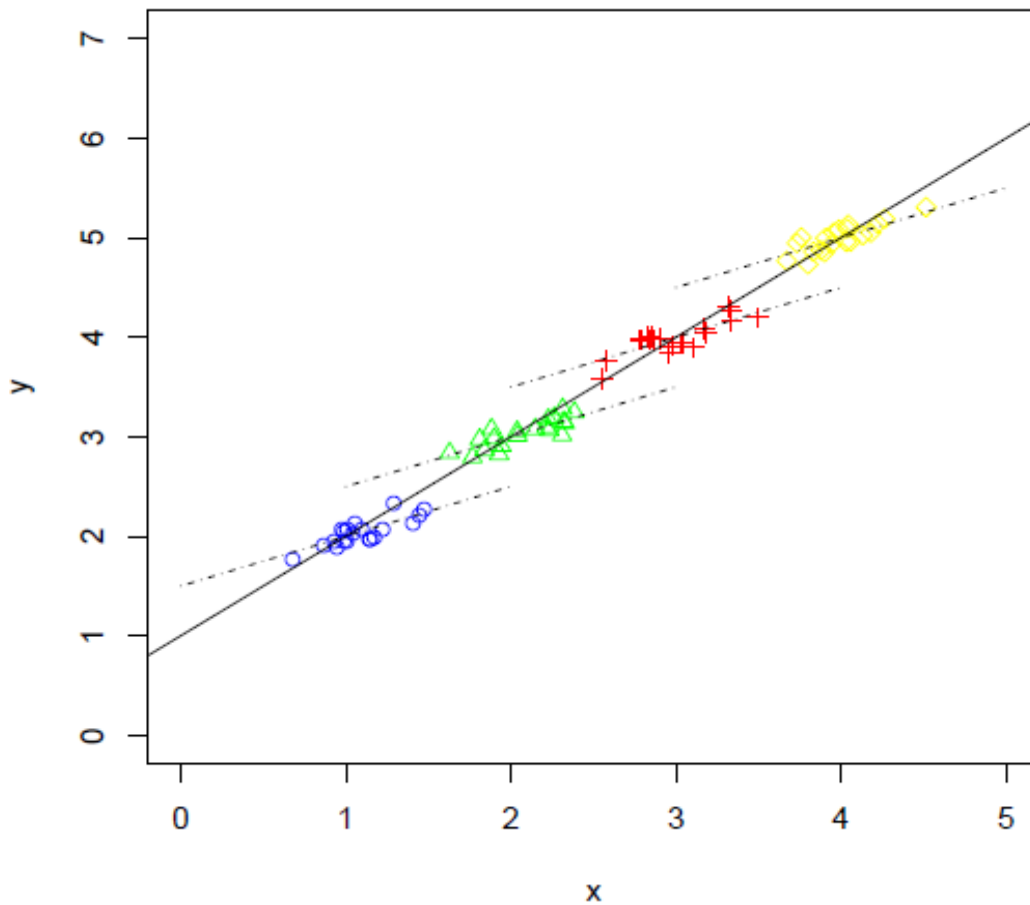
Source: The **plm** package. <http://www.jstatsoft.org>

Sur la **figure03**, on constate que la seule forme d'hétérogénéité entre les individus correspond à des niveaux moyens différents de la variable explicative, et donc de la variable expliquée. En revanche, l'ensemble des points semblent correctement alignés sur une même droite. Le modèle que l'on devra estimer peut donc s'écrire :

$$y_{it} = \alpha + \beta X_{it} + \varepsilon_{it}$$

Autrement dit, on se ramène à un modèle linéaire standard, et la méthode d'estimation des moindres carrés ordinaires est appropriée.

Figure 04 : Hétérogénéité de niveau



Source: The **plm** package. <http://www.jstatsoft.org>

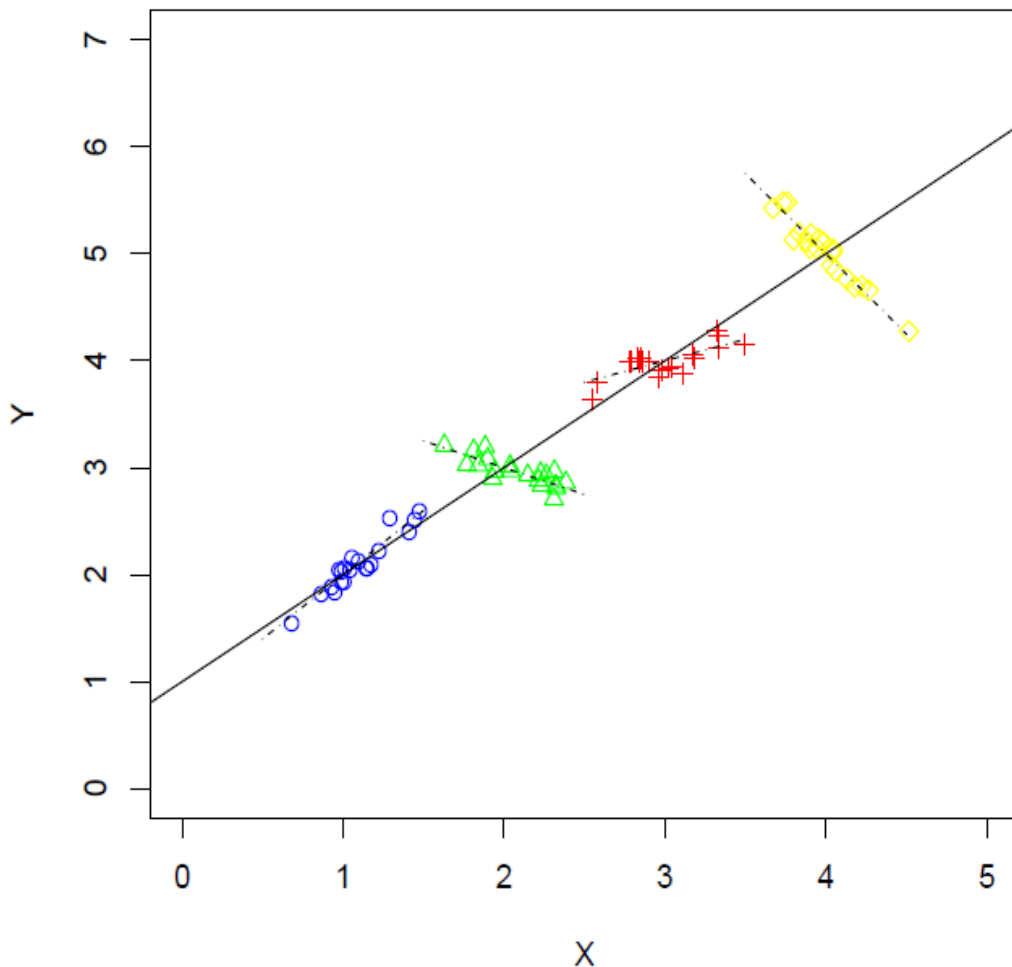
Sur la **Figure 04**, on constate que les points correspondant à chaque individu semblent alignés sur des droites parallèles, c'est-à-dire des droites dont les pentes β sont identiques, mais dont les ordonnées à l'origine sont propres à l'individu.

Dans ce cas, le modèle à estimer peut s'écrire sous la forme :

$$y_{it} = \alpha + \beta X_{it} + \varepsilon_{it} \text{ Avec } \varepsilon_{it} = \mu_i + w_{it}$$

et l'ordonnée à l'origine propre à l'individu i s'écrit : $\alpha + \mu_i$. On est alors dans le cadre du modèle à erreurs composées.

Figure 05 : Hétérogénéité de pentes.



Source: The **plm** package. <http://www.jstatsoft.org>

Sur la **Figure 05**, on constate que les points correspondants à chaque individu sont alignés sur des droites différentes et non parallèles. Dans ce cas, le modèle à estimer s'écrit :

$$y_{it} = \alpha + \beta_i X_{it} + \varepsilon_{it} \text{ Avec } \varepsilon_{it} = \mu_i + w_{it}$$

Les points correspondant à l'individu i sont alors alignés sur une droite d'ordonnée à l'origine $\alpha + \mu_i$ et de pentes β_i . Le modèle approprié à ce cas est le modèle à coefficients variables et une estimation correcte de ce modèle nécessite une dimension temporelle suffisamment longue.

3-2-Modèle à effets individuels

Nous allons à présent intéresser à des modèles de panel hétérogènes, où la seule source d'hétérogénéité provient des constantes individuelles.

Hypothèse (H1) :

- ✚ On suppose que les coefficients des différentes variables explicatives sont identiques pour tous les individus du panel

$$\beta_{kit} = \beta_k$$

- ✚ On suppose en outre que ces coefficients sont des constantes déterministes, et les constantes individuelles α_i ; quant à elles, diffèrent selon les Individus

$$\alpha_i = \alpha + b_i$$

Nous allons distinguer deux cas : ***les modèles à effets fixes*** (l'effet individuel est constant au cours du temps) et ***les modèles à effets aléatoires*** ou encore ***modèles à erreurs composés*** (le terme constant est une variable aléatoire).

Toutefois, avant de présenter ces deux modèles, nous commencerons tout d'abord par introduire les différentes méthodes d'empilement de données de panel qui autorisent une écriture vectorielle du modèle à effet individuel.

En effet, il existe deux possibilités d'écriture vectorielle de ce modèle. Autrement dit, il existe deux façons d'empiler les données :

1. Empilement par individus :

Pour une variable donnée, les T réalisations historiques de chaque individu sont stockées dans un vecteur colonne, et les N vecteurs colonnes ainsi obtenus sont ensuite empilés à la suite des uns des autres dans l'ordre des individus.

$$y_i = \begin{pmatrix} y_{i,1} \\ y_{i,2} \\ \dots \\ y_{i,T} \end{pmatrix}_{(T,1)} \quad X_i = \begin{pmatrix} x_{1,i,1} & x_{2,i,1} & \dots & x_{K,i,1} \\ x_{1,i,2} & x_{2,i,2} & \dots & x_{K,i,2} \\ \dots & \dots & \dots & \dots \\ x_{1,i,T} & x_{2,i,T} & \dots & x_{K,i,T} \end{pmatrix}_{(T,K)} \quad \varepsilon_i = \begin{pmatrix} \varepsilon_{i,1} \\ \varepsilon_{i,2} \\ \dots \\ \varepsilon_{i,T} \end{pmatrix}_{(T,1)}$$

On peut toutefois écrire le modèle de façon totalement vectorielle en empilant les vecteurs y_i et les matrices X_i :

$$Y = \begin{pmatrix} y_1 \\ y_2 \\ \dots \\ y_N \end{pmatrix}_{(TN,1)} \quad X = \begin{pmatrix} X_1 \\ X_2 \\ \dots \\ X_N \end{pmatrix}_{(TN,K)} \quad \varepsilon = \begin{pmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \dots \\ \varepsilon_N \end{pmatrix}_{(TN,1)}$$

2. Empilement par dates :

Pour une variable donnée, les N réalisations individuelles pour une date donnée sont stockées dans un vecteur colonne, et les T vecteurs colonnes ainsi obtenus pour toutes les dates sont ensuite empilés à la suite des uns des autres.

$$y_t = \begin{pmatrix} y_{1,t} \\ y_{2,t} \\ \dots \\ y_{N,t} \end{pmatrix}_{(N,1)} \quad X_t = \begin{pmatrix} x_{1,1,t} & x_{2,1,t} & \dots & x_{K,1,t} \\ x_{1,2,t} & x_{2,2,t} & \dots & x_{K,2,t} \\ \dots & \dots & \dots & \dots \\ x_{1,N,t} & x_{2,N,t} & \dots & x_{K,N,t} \end{pmatrix}_{(N,K)} \quad \varepsilon_t = \begin{pmatrix} \varepsilon_{1,t} \\ \varepsilon_{2,t} \\ \dots \\ \varepsilon_{N,t} \end{pmatrix}_{(N,1)}$$

De la même façon, il est possible d'exprimer le modèle sous une forme vectorielle complète.

$$Y_{(TN,1)} = \begin{pmatrix} y_1 \\ y_2 \\ \dots \\ y_T \end{pmatrix} \quad X_{(TN,K)} = \begin{pmatrix} X_1 \\ X_2 \\ \dots \\ X_T \end{pmatrix} \quad \varepsilon_{(TN,1)} = \begin{pmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \dots \\ \varepsilon_T \end{pmatrix}$$

Remarque :

1- Les modèles à effets individuels peuvent être transposés au modèles à effets temporels, la constante peut s'écrire alors : $\alpha_i = \alpha + b_i$ (**Sevestre**)

2- Il est très important de noter que les principaux logiciels d'économétrie optent généralement pour une méthode d'empilement par individus. Si le logiciel n'empile pas lui-même les données, les séries utilisées dans le cadre des applications doivent, de façon impérative, être ordonnées sous la forme préconisée par les concepteurs du logiciel.

3-2-1-Le modèle à effets fixes individuels

A. Spécification du modèle

On fait maintenant l'hypothèse que les effets individuels α_i sont représentés par des constantes (d'où l'appellation modèle à effets fixes). Nous allons déterminer la forme générale des estimateurs des paramètres α_i et β dans ce modèle à effets fixes. On considère donc sous l'hypothèse (H1) le modèle suivant :

$$y_{it} = \alpha_i + \sum_{k=1}^K \beta_k x_{kit} + \varepsilon_{it}$$

Avec: $\alpha_i = \alpha + b_i$

Finalement le modèle à effets fixes individuels s'écrit :

$$y_{it} = \alpha + b_i + \sum_{k=1}^K \beta_k x_{kit} + \varepsilon_{it}$$

Tous les paramètres du modèle sont des constantes et l'on suppose pour simplifier qu'il n'existe pas d'effet temporel.

- Le modèle à effets fixes individuels présente une structure des résidus qui vérifient les hypothèses standards des MCO. Il s'agit en fait d'un modèle classique avec variables indicatrices individuelles.

B. Méthodes d'estimation

La méthode d'estimation des paramètres va dépendre de la structure des termes d'erreurs :

- Si les erreurs sont homoscédastiques, non autocorrélées dans la dimension temporelle $\text{Cov}(\varepsilon_{it}, \varepsilon_{is}) = 0$ pour : $t \neq s$ et dans la dimension individuelle $\text{Cov}(\varepsilon_{it}, \varepsilon_{jt}) = 0$ pour : $i \neq j$: méthode des **Moindre Carrés Ordinaires (MCO)** sur les variables muettes indicatrices **LSDV (Least Square Dummy Variable)** ou sur les estimateurs **Within**.
- Si les erreurs sont hétéroscédastiques et/ou autocorrélées dans la dimension temporelle mais indépendantes dans la dimension individuelle : méthode des **Moindres Carrés Généralisés (MCG)** sur les variables indicatrices **LSDV** ou sur les estimateurs **Within**.

1- Estimateur LSDV :

L'estimateur **LSDV** consiste à appliquer la méthode des **MCO** sur le modèle avec variables indicatrices spécifiques pour chacun des N individus. Nous construisons donc N variables indicatrices tel que :

$D_i = 1$ pour l'individu i et 0 pour les autres. Le modèle s'écrit :

$$y_{it} = \alpha + \alpha_1 D_1 + \alpha_2 D_2 + \dots + \alpha_N D_N + \sum_{k=1}^k \beta_k x_{kit} + \varepsilon_{it}$$

Pratiquement, nous estimons le modèle sans la constante α (sinon, il y aurait alors colinéarité entre le vecteur du terme constant α et la somme des N variables indicatrices, la matrice $(X'X)$ serait alors singulière (non inversible))

$$y_{it} = \alpha_1 D_1 + \alpha_2 D_2 + \dots + \alpha_N D_N + \sum_{k=1}^k \beta_k x_{kit} + \varepsilon_{it}$$

Ce modèle à estimer par les **MCO** ou les **MCG** si les erreurs sont hétéroscédastiques et/ou autocorrélées.

Nous pouvons ensuite calculer les coefficients $\alpha_i = \alpha + b_i$ du modèle initial, la valeur de la constante α étant égale à la moyenne des coefficients b_i estimés.

2- **Estimateur *Within*** : (le théorème de **Frisch-Waugh**)

L'estimateur ***Within*** (estimateur intra-individuel) consiste à centrer préalablement toutes les variables à expliquer et explicatives sur leurs moyennes individuelles et à appliquer la méthode des **MCO** ou les **MCG** sur le modèle ainsi transformé :

$$(y_{it} - \overline{y_i}) = \sum_{k=1}^k \beta_k (x_{kit} - \overline{x_{ki}}) + \varepsilon_{it} \text{ Pour } (i=1, \dots, N)(t=1, \dots, T)$$

Les réalisations des estimateurs des constantes α_i sont alors déduites des relations :

$$\widehat{\alpha}_1 = \overline{y_1} - \sum_{k=1}^K \widehat{\beta}_k \overline{x_{k1}} ; \widehat{\alpha}_2 = \overline{y_2} - \sum_{k=1}^K \widehat{\beta}_k \overline{x_{k2}} ; \dots ; \widehat{\alpha}_N = \overline{y_N} - \sum_{k=1}^K \widehat{\beta}_k \overline{x_{kN}}$$

$$\widehat{\alpha}_i = \overline{y_i} - \sum_{k=1}^K \widehat{\beta}_k \overline{x_{ki}} \text{ avec : } (i=1, \dots, N)$$

Nous calculons ensuite les coefficients $\alpha_i = \alpha + b_i$ du modèle initial comme précédemment.

Cette méthode d'estimation conduit exactement aux mêmes résultats que la méthode **LSDV**.

3-2-2-Le modèle à erreurs composées

A. Spécification du modèle

Le modèle à effets aléatoires individuels a été introduit en **1966** par **Balestra** et **Nerlove**. Il consiste à introduire un effet individuel dans la perturbation du modèle. Ceci implique que l'estimation par les moindres carrés ordinaires n'est pas optimale. De plus, les statistiques de test déduites des formules habituelles ne sont plus valables de sorte qu'il faut modifier le calcul de la matrice de covariance de l'estimateur.

La solution réside dans l'application des moindres carrés généralisés (**MCG**). Toutefois, pour pouvoir mettre en œuvre cet estimateur, il faudra d'abord estimer la matrice de covariance des perturbations du modèle à erreurs composées, ce qui implique d'introduire deux autres estimateurs : l'estimateur inter-individuel (**between**) et l'estimateur intra-individuel (**within**).

Le modèle à erreurs composées suppose que la spécificité individuelle est sous forme aléatoire. Le terme constant spécifique à l'individu i est aléatoire. Il se décompose en un terme fixe et un terme aléatoire spécifique à l'individu permettant de contrôler l'hétérogénéité individuelle. En regroupant les termes aléatoires du modèle, on obtient une structure à erreurs composées.

Comme nous décomposons la constante dans le modèle à effets fixes, il s'agit dans le modèle à effets aléatoires de décomposer les résidus. C'est en effet dans ces derniers qu'interagissent les variables explicatives omises. Le modèle s'écrit toujours :

$$y_{it} = \alpha + \sum_{k=1}^k \beta_k x_{kit} + \varepsilon_{it}$$

Le terme d'erreur aléatoire ε_{it} est alors décomposé de la manière suivante : $\varepsilon_{it} = \mu_i + \lambda_t + w_{it}$

α désigne la composante fixe et μ_i la composante stochastique individuelle, non observable telle que la qualité du management dans le cadre d'un panel d'entreprises et dans le cas simple ou l'effet temporel n'existe pas $\lambda_t = 0$, Le modèle à erreurs composées s'exprime de la manière suivante:

$$y_{it} = \alpha + \mu_i + \sum_{k=1}^k \beta_k x_{kit} + w_{it}$$

Note : Attention, contrairement à ce qui se passe dans le cadre du modèle à effet fixe pour lequel les individus se démarquent les uns des autres par un élément constant, la composante μ_i qui apparaît ici n'est pas une constante mais bien la réalisation d'une variable aléatoire.

Généralement, on est conduit à faire un certain nombre d'hypothèses sur cette structure de résidus.

Les hypothèses sur les résidus :

On suppose que les résidus sont i.i.d. et satisfont les conditions suivantes :

- $E(\mu_i) = E(w_{it}) = 0$
- $E(\mu_i, w_{it}) = 0$
- $E(\mu_i, \mu_j) = \sigma_\mu^2, i=j, 0 \text{ sinon}$
- $E(w_{it}, w_{js}) = \sigma_w^2, i=j \text{ et } t=s, 0 \text{ sinon}$
- $E(\mu_i, X_{it}) = E(w_{it}, X_{it}) = 0$

Sous ces hypothèses, la variance de la variable endogène y_{it} conditionnellement aux variables explicatives x_{it} est alors égale à :

$$\sigma_y^2 = \sigma_\mu^2 + \sigma_w^2$$

B. Méthode d'estimation

Contrairement au modèle à effets fixes, le modèle à erreurs composées utilise les deux dimensions du panel. Par conséquent, on peut élargir le champ de recherche empirique, en mesurant l'effet de facteurs invariants dans le temps, (secteur d'activité de l'entreprise, à titre d'exemple), en plus des facteurs explicatifs variables dans le temps ; puisque les erreurs composées, sont corrélées par individu dans le temps, on utilise la méthode des moindres carrés généralisés.

L'estimateur des **MCG** combine donc les variabilités inter-individuelle (**Between**) et intra-individuelle (**Within**) des observations.

Il est aisé de vérifier que l'application des **MCG**, revient à appliquer les **MCO** sur le modèle transformé suivant:

$$y_{it} + (\sqrt{\theta} - 1)\bar{y}_i = \alpha\sqrt{\theta} + \mu_i\sqrt{\theta} + \sum_{k=1}^k \beta_k [x_{kit} + (\sqrt{\theta} - 1)\bar{x}_{ki}] + w_{it} + (\sqrt{\theta} - 1)\bar{w}_i$$

$$\theta = \frac{\sigma_w^2}{\sigma_w^2 + T\sigma_u^2}$$

♦ La méthode SUR :

La méthode SUR (Seemingly Unrelated Regressions) de **Zellner (1962)** est utilisée lorsque les erreurs des équations individuelles sont corrélées : la covariance individuelle $\text{Cov}(\varepsilon_{it}, \varepsilon_{jt}) = \sigma^2_{ij} \neq 0$ pour $i \neq j$. Les individus sont alors interdépendants.

Cette méthode consiste à appliquer les **MCG** :

$$\hat{\beta}_{\text{sur}} = (X' \Omega_w^{-1} X)^{-1} (X' \Omega_w^{-1} Y)$$

Avec :

$$\Omega_w = \begin{pmatrix} \sigma_1^2 I & \sigma_{12} I & \dots & \sigma_{1N} I \\ \sigma_{21} I & \sigma_2^2 I & \dots & \sigma_{2N} I \\ \dots & \dots & \dots & \dots \\ \sigma_{N1} I & \sigma_{N2} I & \dots & \sigma_N^2 I \end{pmatrix}$$

(N^*T, N^*T)

Chaque élément $\sigma_{ij} I$ étant de dimension (T, T) .

Pratiquement, la procédure est la suivante :

- ☞ Estimation par les **MCO** des N équations individuelles.
- ☞ Calcul des N résidus e_{it} .
- ☞ Calcul des $\widehat{\sigma}_i^2 = \frac{\sum_{t=1}^T w_{it}^2}{T}$ et $\widehat{\sigma}_{ij} = \frac{\sum_{t=1}^T w_{it} w_{jt}}{T}$.
- ☞ Application des **MCG** : $\hat{\beta}_{\text{sur}} = (X' \Omega_w^{-1} X)^{-1} (X' \Omega_w^{-1} Y)$.

3-3-Le test des effets fixes vs effets aléatoires

Le test d'**Hausman (1978)** est un test de spécification qui permet de déterminer si les coefficients des deux estimations (fixes et aléatoires) sont statistiquement différents.

Sous l'hypothèse nulle H_0 d'orthogonalité entre les variables explicatives et le terme d'erreur du modèle à effets aléatoires, les deux estimateurs (**Within** et **MCG**) sont des estimateurs non biaisés et dans ce cas il ne doit pas y avoir de différence significative entre les estimations **Within** et **MCG** des différents coefficients. La méthode des **MCG** est alors retenue : le modèle est à effets aléatoires.

➤ Soit le test d'hypothèses :

$H_0 : \hat{\beta}_{LSDV} - \hat{\beta}_{MCG} = 0 \longrightarrow$ le modèle est à effets aléatoires.

$H_a : \hat{\beta}_{LSDV} - \hat{\beta}_{MCG} \neq 0 \longrightarrow$ le modèle est à effets fixes.

➤ Nous calculons la statistique :

$$H = (\hat{\beta}_{LSDV} - \hat{\beta}_{MCG})' [\text{Var}(\hat{\beta}_{LSDV}) - \text{Var}(\hat{\beta}_{MCG})]^{-1} (\hat{\beta}_{LSDV} - \hat{\beta}_{MCG})$$

La statistique H est distribuée selon un khi-deux à K degré de liberté. Si $H > \chi^2(k)$ pour un seuil à α % fixé, nous rejetons l'hypothèse H_0 , l'estimateur **LSDV (Within)** est non biaisé, nous rejetons alors la spécification à effets aléatoires et nous choisissons un modèle à effets fixes individuels. (**Régis Bourbonnais, 2011, Page 356**).

Section II : Revue de la littérature

Étant donné que nous ne pouvons pas prétendre être les premiers chercheurs à aborder un sujet appartenant les données de panel, nous avons consulté les travaux suivants :

1. Mankiw, Romer & Weil (1992) :

Leurs estimations principales portent sur **98** pays, pour lesquels ils disposent de données sur la période **1960-1985**.

Ceux-ci en effet mettent en évidence, sur données en coupe transversale, un effet significatif du taux de scolarisation des **12-17** ans (mesuré en moyenne sur même période) sur le niveau du **PIB** par tête en **1985** (dans une régression par les moindres carrés ordinaires où figurent également le niveau initial du **PIB** par tête, le taux d'investissement et le taux de croissance du capital par tête)

2. Article de Scherer et Huh (1992) :

À partir d'un échantillon de données concernant 221 entreprises américaines sur la période **1970-1985**, cette étude met en évidence l'existence d'une corrélation positive significative entre le niveau d'éducation technique des cadres supérieurs et les dépenses en R&D (Recherche et développement) des entreprises concernées.

3. Barro et Sala-i-Martin (1995):

Ils présentent d'autres estimations de la contribution du capital humain à la croissance, utilisant les mêmes données sur la même période. La croissance du PIB par tête est fonction entre autres du nombre moyen d'années d'études primaires, secondaires et supérieures dans la population, et des dépenses d'éducation en proportion du PIB au début de la période. Alors que le coefficient de l'éducation primaire n'est pas significatif, ceux de l'éducation secondaire et supérieure des hommes sont significatifs et positifs.

4. Islam (1995) :

Il a mis en application une formulation de données de panel de la fonction de production de Solow étendue pour tenir compte du capital humain. Au lieu d'une seule estimation transnationale, Islam a pris les données couvrant la même période de **1960 à 1985** et les a réparties en cinq sous-périodes pour tirer parti de cinq points de données par pays.

En outre, Islam a tenu compte des effets (fixes) propres à chaque pays pour corriger le biais dû à la variable omise.

En guise d'approximation de l'équilibre stable du capital humain, Islam **(1995)** a choisi la variable humain de Barro et Lee **(1993)** qui fournit des renseignements sur le nombre moyen d'années de scolarité aux niveaux primaire, secondaire et supérieurs dans l'ensemble de la population âgée de plus de **25** ans. Il a réparti les pays en trois échantillons, soit les mêmes : NONOIL, INTER et OCDE.

Les résultats des estimations d'Islam tenant compte des effets sur les pays supposaient des valeurs de la vitesse annuelle de la convergence conditionnelle **(OCDE : 0,0913)** supérieures à celles obtenues par **MRW (OCDE : 0,0203)**.

La littérature empirique traditionnelle avec les données de Panel standard ou de coupe transversale enrichit la littérature de la croissance économique surtout avec les tentatives de **Mankiw, Romer et Weil en (1992)**, **Benhabib Spiegel (1994)** et d'autres économistes.

Les résultats et les estimations faites par ces économistes restent limités selon la nature de l'échantillon, leurs caractéristiques aussi la longueur de la durée d'étude et les variables utilisés et la disponibilité de données, pour finir par la recourt de même modèle de Solow augmenté.

5. Gérard Tchouassi, Université de Yaoundé II, Cameroun **« Capital humain et croissance économique au Cameroun et au Sénégal : une modélisation VAR »**

L'objectif de cette étude est de mettre en évidence empiriquement la relation entre le capital humain (**éducation, santé**) et la croissance économique. Il a effectué une analyse comparative entre le Cameroun et le Sénégal à l'aide d'une modélisation **VAR** avec cointégration des variables utilisées. Ce qui lui a permis d'établir la comparaison des effets du capital humain sur la croissance économique entre les deux pays africains. Les résultats de la décomposition de la variance pour le Cameroun indiquent que la variance de l'erreur de prévision du **PIB** par tête est due à **57%** à ses propres innovations, à **30%** à celle de l'inflation et à **5%** aux dépenses publiques d'éducation, à **4%** à celle de l'**IDE (investissement direct étranger)** et également du même montant à celle de la population. Les résultats de la décomposition de la variance pour le Sénégal indiquent que la variance de l'erreur de prévision du **PIB** par tête est due à **90%** à ses propres innovations, à **4%** à celle de l'inflation et à **1%** aux dépenses publiques d'éducation, à **1%** à celle de l'**IDE**. Ce qui traduit le fait que dans les deux pays les innovations constituent un canal important permettant de conduire la croissance économique. Il a conclu que l'innovation est ainsi la variable sur laquelle les autorités du Cameroun et du Sénégal doivent agir pour atteindre le niveau de croissance économique souhaité.

6. Pritchett (2001) :

Dans un article très remarqué ; **Pritchett** présente des résultats d'estimation utilisant les données sur l'éducation de **Barro et Lee (1993)** et **Nehru, Swanson et Dubey (1995)**, dans lesquels l'effet de la croissance du nombre moyen d'années d'éducation par travailleur sur la croissance du PIB par tête est négatif et significatif ; Au contraire des auteurs précédemment cités, il considère que ce résultat n'est pas dû à la mauvaise qualité des données disponibles ou à des méthodes empiriques défaillantes. Il existerait bel et bien un « paradoxe micro/macro » c'est-à-dire que les effets microéconomiques avérés de l'éducation sur les revenus individuels n'auraient pas d'équivalent macroéconomique en termes de croissance. **Pritchett** insiste sur l'hétérogénéité des pays :

l'effet négatif obtenu dans une régression utilisant des données internationales n'est qu'une moyenne d'effets positifs dans certains pays, négatifs dans d'autres ; l'explication des résultats négatifs trouveront leur réponse dans les trois hypothèses avancées par **Pritchett** et concernant : l'environnement institutionnel des pays, les transformations structurelles de l'économie, et la mauvaise qualité des écoles dans plusieurs pays en voie de développement.

7. Arcand et D'Hombres (2002) :

Ils ont remarqué que la corrélation entre éducation et croissance semble être positive lorsque l'on utilise des données « en coupe », où l'on ne dispose que d'une seule observation par pays (cf. Barro, 1991 ; Barro ; 1997, McMahon, 1998 ; Temple, 1999), et non significative ou négative lorsque l'on utilise des données « de panel », où chaque pays est observé sur plusieurs années ou intervalles de quelques années (cf. Benhabib et Spiegel, 1994 ; Bond, Hoeffler et Temple, 2001 ; Bräuninger et Pannenberg, 2002 ; Caselli, Esquivel et Lefort, 1996; Hamilton et Monteagudo, 1998 ; Islam, 1995 ; McDonald et Roberts, 2002)

8. Krueger et Lindahl : (2003)

Dans un panel de **110** pays observés entre **1960** et **1990**, Krueger et **Lindahl (2003)** soulignent le fait que les estimations de l'influence du capital humain sur les taux de croissance de la productivité et de l'ensemble de l'économie sont de beaucoup inférieures aux estimations de cette influence sur les salaires et sur d'autres résultats sur le marché du travail observés à l'échelle individuelle. Cette disparité est due à la qualité des mesures du capital humain employées dans les modèles macroéconomiques de la croissance.

**9. Philippe Aghion et Élie Cohen ; (Paris, 2004)
le rapport «Éducation et croissance »**

Dans ce rapport ; les deux chercheurs indiquent et analysent des estimations économétriques effectuées sur un panel de 20 pays de l'OCDE ; l'étude a procédé à des analyses économétriques de la relation entre distance à la frontière technologique, répartition de l'effort d'éducation entre primaire ; secondaire et enseignement supérieur, et croissance de la productivité.

Les résultats empiriques confirment bien le rôle que joue l'éducation dans l'assimilation et la formation du progrès technique. Ils confirment également que, plus un pays se rapproche de la frontière technologique, plus l'importance de l'enseignement supérieur croît par rapport à celle de l'enseignement secondaire.

10. Ben et Hassad (2006) :

Dans leur analyse en coupe transversale sur l'efficacité du financement des services publics et croissance dans 45 pays en développement, sur la période **(1990- 2002)**, ils ont montré que les dépenses publiques d'éducation et de santé sont porteuses de croissance économique si ces dépenses sont utilisées d'une manière efficace dans les Pays en Voie de Développement

**11. BEKIOUA Fateh : INPS Alger-Ingénieur en statistique appliquée (2006)
« Education et croissance économique en Algérie: Une analyse en terme de causalité à l'aide des modèles VAR »**

Le but de cette étude économétrique est d'étudier les liaisons entre la croissance économique et les variables éducatives .et pour tester la présence de cette relation ; l'étude a proposé de retenir deux types de variables : les variables éducatives et économiques .Ces différentes catégories de variables seront analysées dans le cas de l'**Algérie** sur la période **1963-2004**.

Le chercheur a utilisé l'approche des vecteurs autorégressifs (VAR)

Dans un premier temps, un modèle VAR à quatre variables est estimé pour étudier la relation entre les variables éducatives (jusqu'à l'obtention du baccalauréat) et le PIB

- **SCO** : Effectifs scolarisés à tous les niveaux confondus
- **BAC** : Nombre de bacheliers
- **DEP** : Dépenses d'éducation
- **PIB** : Produit intérieur brut

Dans un second temps, un modèle VAR à deux variables est estimé pour étudier la relation entre le nombre de diplômés et le PIB

- **DIP** : nombre de diplômés du supérieur
- **PIB** : produit intérieur brut

L'étude des deux circuits de causalité il a indiqué d'une part, que la croissance économique influence directement le nombre de bacheliers et le nombre de diplômés, et d'autre part que ces relations de causalité sont positives. C'est à dire que la croissance économique influence positivement le nombre de bacheliers et de diplômés. Cependant, on mentionne qu'il n'existe pas de relations de causalités entre les différentes variables de la sphère éducative.

Les décompositions de la variance du premier et du deuxième modèle VAR il a indiqué qu'un choc sur le PIB a plus d'impact sur les variables éducatives (en particulier sur le nombre de bacheliers et le nombre de diplômés), que celles-ci sur la croissance économique.

12. Thomas Jobert : (Septembre 2008)

« Dépenses publiques d'éducation, dépenses militaires et croissance en Turquie : Une analyse empirique sur séries longues »

L'étude a mesuré le lien positif entre les dépenses publiques et l'activité économique est une condition nécessaire à des mécanismes de croissance endogène par un test de causalité au sens de Granger des dépenses publiques vers la croissance. Il a utilisé une base de données couvrant une période allant de 1924 à 2003. Elle est composée des Dépenses publiques d'éducation, des dépenses militaires et du PIB à partir d'une application de la méthode de **Johansen (1988)** et **Les tests de causalité au sens de Granger et le modèle Vectoriel Autorégressif (VAR)**; Il a trouvé une causalité positive des dépenses

d'éducation vers le PIB et une causalité négative des dépenses militaires. Ce travail a complété sur les effets de l'éducation avec une analyse basée sur les effectifs d'enseignants. La décomposition de ces effectifs par niveau d'éducation montre les effectifs du primaire et du supérieur sont responsables du lien positif entre éducation et croissance.

13. Romuald sostaine Foueka Tagne : Université de Yaoundé Master/ DEA- NPTCI -Cameroun (2009)
« Croissance des dépenses publiques et incidence sur le développement au Cameroun: le cas du secteur éducatif »

Cette étude identifie les facteurs qui fondent l'évolution des dépenses publiques au **Cameroun** et leurs effets sur le bien-être des individus. Plus spécifiquement il s'agit d'analyser les principaux déterminants de la croissance des dépenses publiques au **Cameroun** d'une part ; et d'apprécier l'incidence de ces dépenses dans la promotion du développement au **Cameroun** à travers le secteur de l'éducation d'autre part.

Les résultats empiriques indiquent, à partir d'un modèle de déséquilibre où les aspects d'offre et de demande de dépenses publiques sont pris en compte simultanément, que le niveau de vie appréhendé à partir du revenu réel, est le principal facteur qui explique la demande. Ce qui valide la loi de **Wagner** dans le cas du **Cameroun**. Du côté de l'offre, les facteurs tels que l'effet de **Baumol** et le montant des taxes et impôts ont une influence significative. Mais une observation faite dans le secteur éducatif, à partir de l'analyse bénéfice incidence permet de constater que ces dépenses sont inégalement réparties selon la zone géographique.

14. Damas HOUNSOUNON : 2009
« *Efficiencia des dépenses publiques d'éducation, de santé et croissance économique dans l'espace UEMOA* »

Dans cette étude, le chercheur a tenté d'analyser sur une période de 35 ans (**1970-2004**), les scores d'efficiencia des dépenses publiques d'éducation et de santé à l'échelle de l'**UEMOA (Union Economique et Monétaire Ouest Africaine)** et rechercher si l'efficiencia de ces

dépenses permet un accroissement de la production plus vite que le volume des dépenses engagées.

Pour traité cette base de données ; il a d'abord procédé à l'estimation des scores d'efficience par la méthode **DEA-Malmquist** avant d'étudier l'impact de ces scores estimés sur la croissance à travers un modèle de croissance endogène. les différents modèles estimés sont des modèles de panel (8 pays de **UEMOA** sur une période de 35 ans)

Modèle 1: $lpib_{it} = \alpha_1 licp_{it} + \alpha_2 ltsse_{it} + \alpha_3 tpam_{it} + \alpha_4 tc_{it} + \alpha_5 rp_{it}$

Modèle 2 : $lpib_{it} = \alpha_1 licp_{it} + \alpha_2 ltsse_{it} + \alpha_3 tpam_{it} + \alpha_4 tc_{it} + \alpha_5 rp_{it} + \alpha_6 ds_{it}$

Modèle 3 : $lpib_{it} = \alpha_1 licp_{it} + \alpha_2 ltsse_{it} + \alpha_3 tpam_{it} + \alpha_4 tc_{it} + \alpha_5 rp_{it} + \alpha_6 ses_{it}$

Modèle 4 : $lpib_{it} = \alpha_1 licp_{it} + \alpha_2 ltsse_{it} + \alpha_3 tpam_{it} + \alpha_4 tc_{it} + \alpha_5 rp_{it} + \alpha_6 de_{it}$

Modèle 5 : $lpib_{it} = \alpha_1 licp_{it} + \alpha_2 ltsse_{it} + \alpha_3 tpam_{it} + \alpha_4 tc_{it} + \alpha_5 rp_{it} + \alpha_6 see_{it}$

Les résultats montrent d'une part que les dépenses socio-publiques d'éducation et de santé sont peu efficaces dans les pays de l'UEMOA durant la période considérée et d'autre part que c'est une utilisation efficace des ressources consacrées à l'éducation et à la santé qui est plus importante que le volume de ces dépenses en tant que facteur contribuant à la croissance.

15. Guillaume ALLÈGRE ; Thomas MÉLONIO ; Xavier TIMBEAU Document de travail de OFCE: Dépenses publiques d'éducation et inégalités (Une perspective de cycle de vie) (Paris-Mars 2010)

Dans cet article, le chercheur a mesuré l'impact redistributif des dépenses éducatives et de leur financement pour une promotion de jeunes actifs ayant achevé leur formation initiale en 2001 et en 2002 (échantillon de **1872** étudiants) Sur données françaises.

Il a sélectionné une perspective de cycle de vie qui consiste à estimer pour chaque individu d'une part, l'ensemble des dépenses publiques d'éducation dont il a bénéficié, et d'autre part, l'imposition nécessaire pour financer cette dépense.

Afin d'isoler les questions d'inégalités et de redistribution intra-générationnelles et l'hypothèse que la promotion d'étudiants sous revue finance ses propres dépenses d'éducation ; Il a conclu que les dépenses et le financement du système d'éducation au-delà de l'âge de l'enseignement obligatoire (**16 ans**) aggravent légèrement les inégalités entre ménages classés par origine sociale. Toutefois, ces transferts sont un enjeu marginal par rapport à celui des différences de parcours scolaires et professionnels en fonction du milieu social d'origine. Il a découvert également que dans une perspective de cycle de vie, ce système est progressif dans la mesure où les individus à hauts revenus sont des contributeurs nets alors que ceux à bas revenus sont des bénéficiaires nets.

16. David TSHILEO, Université de Goma RDC (2010)

« Déterminants de long terme de dépenses publiques de l'éducation en RDC de 1980 à 2009 »

Selon **David TSHILEO**, les facteurs déterminants à long terme les dépenses publiques de l'éducation en R.D.C sont : le taux d'alphabétisation ; la masse monétaire ; le produit intérieur brut par habitant et l'inflation. En vu de vérifier cette hypothèse, il a fait l'analyse de l'évolution statistique des variables du modèle ; en plus le recours à la technique économétrique lui a permis de déterminer les variables pertinentes du modèle et enfin l'estimation par la méthode des moindres carrés ordinaires (**MCO**) lui a facilité de dégager la relation entre les variables.

Après l'analyse, il a abouti aux résultats suivants: le taux d'alphabétisation a un impact positif sur le financement des dépenses publiques de l'éducation. Les résultats montrent que le coefficient de la corrélation est de **4,52** ce qui implique que lorsque le taux d'alphabétisation augmente de **1%**, cela fait accroître le financement des dépenses publiques de l'éducation de **4,52%** ce qui conforme au signe attendu.

La masse monétaire influe positivement sur le financement des dépenses publiques de l'éducation, nous constatons que lorsqu'il y a augmentation de **1%** de la masse monétaire, cela entraîne une hausse des financements des dépenses publiques de l'éducation de l'ordre de **1,05%** cela confirme le signe prédit de la relation entre ces deux variables. L'inflation a une implication négative sur le financement des dépenses publiques de l'éducation en R.D.C, cela a comme conséquence qu'une hausse de l'inflation de **1%** entraîne une diminution de **0,135%** du financement des dépenses publiques de l'éducation. Ce qui n'est pas conforme à son attente car dans son prédictions il a estimé qu'une hausse du taux d'inflation entraînerait une hausse de financement des dépenses publiques d'éducation. Il estime que cette situation serait due à une distorsion dans la transmission de la politique monétaire.

17. Régis Bourbonnais(2011)

Il a intéressé à l'évolution des dépenses d'éducation sur 25 ans pour neuf (9) pays. Le modèle à étudier est donc :

$$Y_{it} = a_{oi} + a_{1i} X_{1it} + a_{2i} X_{2it} + \varepsilon_{it}$$

Avec:

Y_{it} : évolution des dépenses d'éducation en milliers de dollars pour le pays (i) à l'année (t).

X_{1it} : évolution des dépenses militaires en milliers de dollars pour le pays (i) à l'année (t).

X_{2it} : évolution du **PIB** (Produit intérieur brut) en milliers de dollars pour le pays (i) à l'année (t).

ε_{it} : erreur de spécification pour le pays (i) à l'année (t), ces erreurs répondent aux hypothèses classiques.

Cette étude a utilisé les différents estimateurs de spécification des modèles à données de panel cylindré et les différents tests (Hsiao, Hausman) ; le résultat c'est que le modèle est à effets individuels fixes.

Avec:
$$Y_{it} = -125.65 + 62.93 X_{1it} + 29.41 X_{2it} + \varepsilon_{it}$$

**18. Borel Foko, Beifith Kouak Tiyab, Guillaume Husson
(Mars 2012) « Les Dépenses des Ménages en Éducation »**

Ce travail donne une vision élargie des dépenses nationales d'éducation dans une quinzaine de pays d'Afrique, en soulignant que la contribution des ménages est substantielle, avec un investissement qui représente, en termes de dollars équivalents, près de la moitié des dépenses publiques d'éducation. L'analyse confirme le caractère nécessaire de la bien éducation, dont le poids dans le budget des ménages varie sans lien étroit avec le niveau de développement économique des pays. Aussi, l'effort budgétaire consenti par les ménages pauvres pour la scolarisation de leurs enfants est loin d'être négligeable, en comparaison avec l'effort consenti par les ménages les plus riches.

**19. Article de Samir Makdisi, Zeki Fattah et Imed
Limam
«Determinants of growth in the MENA countries»**

Dans leurs article ; ils ont essayé de mesurer la contribution relative des sources principales de croissance économique des pays de la région de **MENA**. Ils ont trouvé que la performance globale de la croissance dans la région du **MENA** au cours de la période **1960-1998** a été caractérisée par un degré plus élevé de volatilité comparé à d'autres régions du monde. En comparant le modèle de croissance de la région de **MENA** dans une perspective international, ils ont trouvé que le capital est moins efficace; la malédiction de ressource naturelle davantage prononcée, la croissance de la productivité totale des facteurs dans la région de **MENA** n'était pas une source importante de croissance comparée a d'autres régions, les pays non pétroliers et les économies diversifiées se sont bien mieux comportées que les pays pétroliers en termes de croissance de leur **PIB/tête**, et de celle des **PTF**.

Dans leur analyse, les auteurs ont utilisé comme mesure de capital humain (pour expliquer à la fois le taux de la croissance du PIB/tête et le taux de croissance des **PTF**), **les taux de scolarisation primaire en 1960** pour chaque pays.

la croissance de la PTF : c'est le résidu de l'estimation dans le modèle de **Solow** (la part de la croissance économique restée non

expliquée par les différentes variables) qui symbolise la mobilisation du savoir acquis par la population à travers l'éducation

20. Coulibaly Mamadou Université Felix Houphouet-Boigny, Abidjan-Cocody (Septembre 2013)

« IMPACT DES DÉPENSES PUBLIQUES D'ÉDUCATION SUR LA CROISSANCE ÉCONOMIQUE EN CÔTE D'IVOIRE »

Cette étude a recours à l'analyse des séries temporelles pour étudier la relation fonctionnelle véritable entre les dépenses publiques d'éducation et la croissance économique en Côte d'Ivoire durant la période **1970-2005**. Les résultats des estimations montrent que ces dépenses ont de façon significative un impact positif sur la croissance économique. A court terme, l'estimation du modèle à correction d'erreur montre que des augmentations de **1%** des taux de croissance de la **FBCF** (gross fixed capital formation and education) et des dépenses d'éducation entraînent des hausses respectives de **0.1659%** et de **5.795%** du taux de croissance. Les dépenses d'éducation apparaissent donc comme la seconde force génératrice du processus de croissance économique.

Conclusion

Dans ce chapitre théorique nous avons rappelé, dans un premier lieu, les différentes approches théoriques des dépenses publiques en capital humain et l'impact de ces dépenses sur la croissance dans la région MENA ; ainsi que des généralités sur l'économétrie des données de panel. Dans un second lieu nous avons indiqué la revue de la littérature afin de justifier le choix du modèle dans le deuxième chapitre concernant l'étude empirique.

CHAPITRE



**Cadre empirique sur les déterminants
des dépenses publiques de l'éducation en MENA**

Introduction

Le présent chapitre se propose de présenter le cadre pratique de notre étude, il se subdivise en deux section, la première section est intitulé «Présentation du modèle et de données» où nous allons bien viser l'échantillon et sources des données, la spécification du modèle et l'identification des variables ; ainsi les outils économétriques et logiciels utilisés pour le traitement des données. Quant à la deuxième section, elle s'intéresse à l'analyse descriptive de l'évolution des variables et surtout elle focalise sur le test de spécification d'homogénéité versus hétérogénéité (**Test de Hsiao**).

Section I : présentation du modèle et des données

1- Échantillon et sources des données

MENA est l'acronyme de « Middle East and North Africa » (littéralement, « Moyen-Orient et Afrique du Nord ») régulièrement utilisé dans les écrits académiques et d'affaires. Il désigne une grande région, depuis le Maroc au nord-ouest de l'Afrique jusqu'à l'Iran à sud-ouest de l'Asie, qui comprend généralement tous les pays du Moyen-Orient et de l'Afrique du Nord.

Pour la dimension «individuelle» ; notre échantillon est composé de **(15)** pays de la région de **MENA** lesquels :

- Algérie
- Arabie saoudite
- Bahreïn
- Égypte
- Émirats arabes unis
- Iran
- Jordanie
- Koweït
- Liban
- Maroc
- Oman
- Qatar
- Tunisie
- Turquie
- Yémen

Pour chaque pays on dispose d'observations dans le temps (de 1990 à 2012) : c'est la dimension chronologique (temporelle).

⊕ Sources des données :

Les données qui seront utilisées dans toutes les régressions de la section suivante, sont collectées à l'aide des séries macroéconomiques internationales disponibles dans «Indicateurs de développement mondiale» (WDI) de la banque mondiale et les statistiques de l'UNESCO.

2- Spécification du modèle

À la lumière des théories sur les dépenses publiques de l'éducation évoquées au premier chapitre, nous allons vérifier la relation qui existe entre les dépenses publiques de l'éducation et un certain nombre de ses déterminants pour le cas de la région MENA.

En basant sur l'étude de **David TCSHILEO** (voir la revue de la littérature page 51) ; on peut spécifier le modèle suivant :

$$y_{it} = \alpha_{it} + \beta_{1it} x_{1it} + \beta_{2it} x_{2it} + \beta_{3it} x_{3it} + \beta_{4it} x_{4it} + \varepsilon_{it}$$

Avec :

i : indique la dimension individuelle (pays) ; ($i=1, \dots, 15$)

t : indique la dimension temporelle (années) ; ($t=1990, \dots, 2012$)

y_{it} : évolution des dépenses publiques de l'éducation comme variable dépendante (endogène) mesurée en pourcentage (%) du **PIB**.

⊕ L'ensemble des variables explicatives sont :

x_{1it} : Croissance du PIB par habitant (% annuel).

x_{2it} : Croissance de la masse monétaire (% annuel).

x_{3it} : évolution vers l'école secondaire (%).

x_{4it} : inscription à l'école, enseignement supérieure (% brut).

$\beta_{1it}, \beta_{2it}, \beta_{3it}, \beta_{4it}$ sont les coefficients des variables explicatives.

α_{it} : terme constant pour l'individu (i) à l'instant (t).

ε_{it} : terme d'erreurs (les perturbations aléatoires).

3- Identification des variables

3-1- Les dépenses publiques d'éducation (% du PIB)

Les dépenses publiques de l'éducation sont les dépenses publiques courantes et en immobilisations au titre de l'éducation et comprennent les dépenses publiques relatives aux établissements d'enseignement (publics et privés) et à l'administration de l'éducation ainsi que les subventions à des entités privées (étudiants/ménages et autres entités privées). Les données sont en % du PIB. (Institut des statistiques de l'UNESCO)

3-2- Croissance du PIB par habitant (% annuel)

Le PIB par habitant est le produit intérieur brut divisé par la population en milieu d'année. Le PIB est la somme de la valeur ajoutée brute de tous les producteurs résidents d'une économie plus toutes taxes sur les produits et moins les subventions non incluses dans la valeur des produits. Elle est calculée sans effectuer de déductions pour la dépréciation des biens fabriqués ou la perte de valeur ou la dégradation des ressources naturelles. Les données sont en % annuel.

(Indicateurs du développement dans le monde ; la Banque mondiale et l'OCDE)

3-3- Croissance de la masse monétaire (% annuel)

La masse monétaire correspond à la somme des circulations fiduciaires hors banque, des dépôts à vue autres que ceux du gouvernement central, des dépôts d'épargne à terme et des dépôts en devise étrangère des secteurs résidents autres que le gouvernement central, les chèques de banque et de voyage, ainsi que d'autres titres comme les certificats de dépôt et les billets de trésorerie.

(Indicateurs du développement dans le monde ; la Banque mondiale et l'OCDE)

3-4- Évolution vers l'école secondaire (%)

Le passage au secondaire est le nombre des élèves entreprenant leur première année d'études secondaires au cours d'une année en tant que pourcentage des écoliers inscrits à la dernière année du primaire l'année précédente. (Institut des statistiques de l'UNESCO)

3-5- Inscription à l'école, enseignement supérieure (% brut)

Taux bruts de scolarisation. Total correspond au total des inscriptions dans l'éducation tertiaire, indépendamment de l'âge, exprimé en pourcentage de la population totale du groupe de cinq ans après la sortie de l'école secondaire. (Institut des statistiques de l'UNESCO)

4- Outils économétriques et logiciels

Pour le traitement des données de notre étude, nous allons utiliser l'économétrie des données de panel notamment un panel équilibré composé de **(15)** pays observés durant la période **(1990-2012)**.

D'abord, nous avons effectué l'analyse descriptive de l'évolution de l'ensemble des variables dans la région **MENA** en générale à l'aide d'une graph ; ainsi pour tester l'existence de la structure de panel nous avons procédé des tests de spécification **(Hsiao, Hausman)**, et la méthode des **MCO** avec tout les étapes d'estimation.

Concernant les logiciels, nous avons utilisé **Excel 2010** pour classer et empiler les observations par pays. Cependant **EViews 8** est utilisé pour l'estimation économétrique.

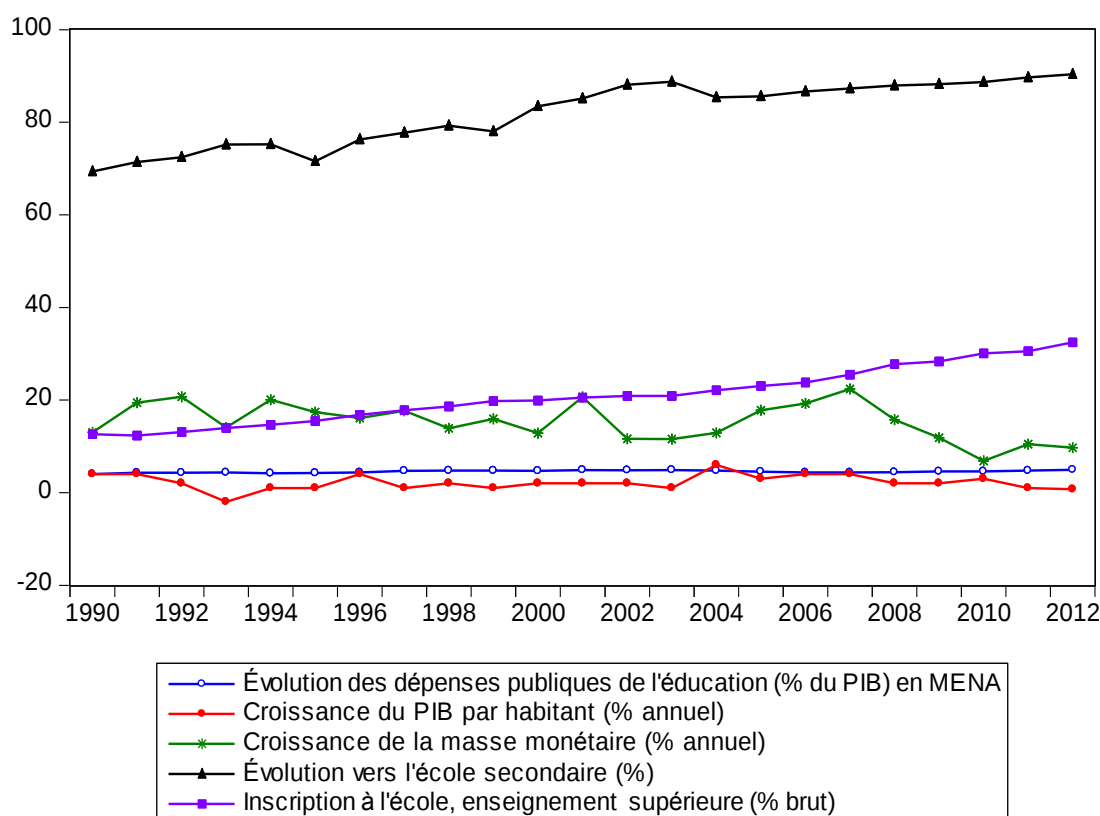
⊕ **Excel 2010** 

⊕ **EViews 8** 

Section II : Estimation et interprétation

1- L'analyse descriptive des variables

Figure 06 : Évolution des différents variables du modèle.



Source : Outputs d'E-views-8

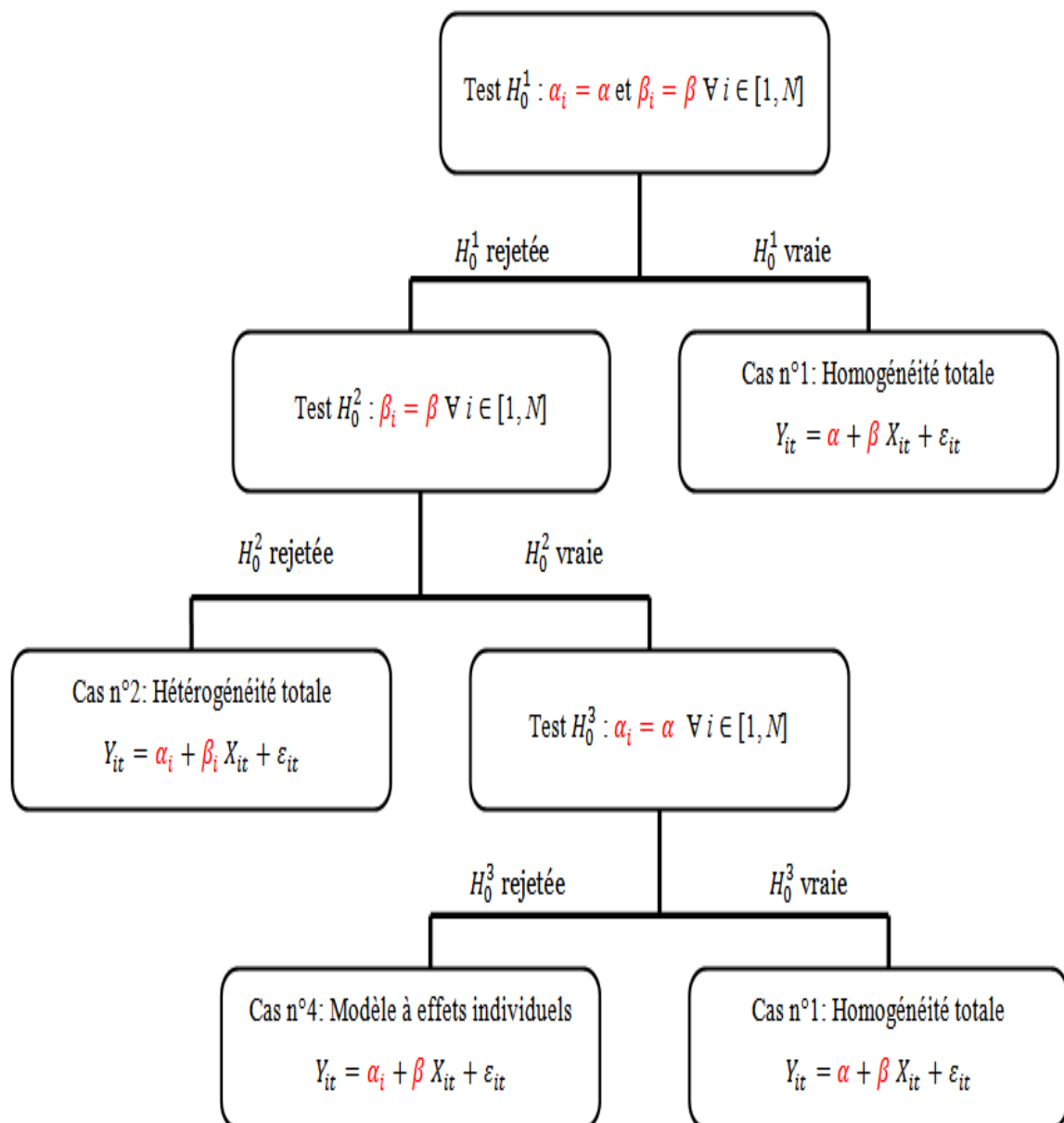
On remarque une sorte de stabilité des dépenses publiques de l'éducation durant la période 1990-2012 d'une moyenne de 4.5% qui est inférieure à 5% du PIB (moyenne mondiale) pour l'ensemble des pays de la région étudiée.

Le pourcentage le plus bas de croissance de PIB par habitant a été négative l'année 1993 d'une valeur (-2%), au contraire le pourcentage le plus élevé a été (6%) pour l'année 2004 et entre ces deux années ; les pays de MENA affichent de faible niveau de croissance de PIB par habitant d'une moyenne de (1.8%) mais après l'année de 2004 ; la croissance de PIB a abaissé jusqu'à (0.8%).

La croissance de la masse monétaire varie selon un aspect aléatoire, on constate que le taux de croissance atteint sa valeur minimale 6.87% l'année 2010 et sa valeur maximale 22.37% l'année 2007.

L'évolution vers l'école secondaire (de 69.35% à 90.36% durant la période indiquée) et l'inscription à l'école de l'enseignement supérieure (de 12.6% à 32.44%) suivent une tendance croissante, ce qui montre bien l'énorme effort entrepris par les différentes politiques adoptées par les gouvernements de la région MENA pour promouvoir une éducation accessible pour tous les habitants.

2- Le test de spécification d'homogénéité/hétérogénéité (Hsiao 1986)



Source : Régis Bourbonnais (2011) p347

Nous allons appliquer la stratégie des tests d'homogénéité ci-dessus afin de déterminer la structure du panel en utilisant **E-views 8**.

Construction des tests :

1) Test 1 :

$$H_0^1 : \alpha_i = \alpha \text{ et } \beta_i = \beta \quad \forall i \in [1, N]$$

$$H_a^1 : \exists (i, j) \in [1, N] / \alpha_i \neq \alpha_j \text{ ou } \beta_i \neq \beta_j$$

Ce test d'hypothèses jointes se ramène à un test de Fisher dont la statistique est donnée par :

$$F_1 = \frac{(SCR_{c1} - SCR)/(N - 1)(K + 1)}{SCR/[N * T - N(K + 1)]}$$

D'abord, nous calculons SCR_{c1} : somme des carrés des résidus du modèle contraint sous l'hypothèse H_0^1 , nous estimons par les **MCO** le modèle en empilant toutes les observations.

Le degré de liberté est égal à :

($N * T$ =nombre total d'observations)-($K+1$ =nombre de coefficients à estimer).

($15*23$ =nombre total d'observations)-($4+1$ =nombre de coefficients à estimer).

⊕ **Les résultats sont les suivants :**

Tableau 03 : Estimation du modèle avec données empilés

Dependent Variable: Y?

Method: Pooled Least Squares

Date: 05/31/14 Time: 13:01

Sample: 1990 2012

Included observations: 23

Cross-sections included: 15

Total pool (balanced) observations: 345

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	4.688534	0.636890	7.361606	0.0000
X1?	-0.016359	0.016937	-0.965897	0.3348
X2?	-0.023035	0.004443	-5.184605	0.0000
X3?	0.012777	0.007470	1.710348	0.0881
X4?	-0.024886	0.007599	-3.274725	0.0012
R-squared	0.134643	Mean dependent var	4.845549	
Adjusted R-squared	0.124463	S.D. dependent var	1.511859	
S.E. of regression	1.414649	Akaike info criterion	3.546026	
Sum squared resid	680.4185	Schwarz criterion	3.601730	
Log likelihood	-606.6895	Hannan-Quinn criter.	3.568210	
F-statistic	13.22538	Durbin-Watson stat	0.306095	
Prob(F-statistic)	0.000000			

Source : outputs d'E-views 8

$$SCR_{c1} = 680.4185$$

Puis, nous calculons **SCR** : somme des carrés des résidus du modèle non contraint, elle est égale à la somme des **15** sommes des carrés des résidus des modèles estimés sur les **23** observations de chaque équation individuelle, soit $SCR = \sum_{i=1}^{15} SCR_i$

Tableau 04 : Calcul du Total SCR_i

Pays	SCR_i
Algérie	2,46184
Arabie saoudite	10,34808
Bahreïn	16,58148
Égypte	7,64493
Émirats arabes unis	20,61225
Iran	3,27185
Jordanie	11,86622
Koweït	9,91344
Liban	7,01943
Maroc	3,17590
Oman	2,17983
Qatar	6,24513
Tunisie	3,16537
Turquie	2,38811
Yémen	35,35063
Total SCR_i	142,22449

Source : les annexes (1-15)

$$SCR = 142,22449 \quad ddld = N * T - N(K + 1) = 270$$

$$ddln = (N - 1)(K + 1) = (15 - 1)(4 + 1) = 70$$

$$F_1 = \frac{(680.4185 - 142,22449)/70}{142,22449/270}$$

$$F_1 = \frac{7.68849}{0.52676} = 14.596$$

$$F_{Cal} = 14.596$$

$$F_{Tab} = F_{ddl n, ddld}^{\alpha} = F_{70, 270}^{5\%}$$

$$F_{Tab} = 1.32$$

La décision:

$F_{Cal} > F_{Tab}$. Alors nous rejetons l'hypothèse H_0^1 au seuil de 5%.

1- ♦ Afin de déterminer la source de cette hétérogénéité, nous nous dirigeons vers la branche de gauche du schéma de **(Hsiao 1986)**

2) Test 2 :

$$H_0^2 : \beta_i = \beta \quad \forall i \in [1, N]$$

$$H_a^2 : \exists (i, j) \in [1, N] / \beta_i \neq \beta_j$$

Ce test d'hypothèses jointes se ramène à un test de Fisher dont la statistique est donnée par :

$$F_2 = \frac{(SCR_{c2} - SCR) / [(N - 1) * K]}{SCR / [N * T - N(K + 1)]}$$

Nous calculons SCR_{c2} : somme des carrés des résidus du modèle contraint sous l'hypothèse H_0^2 , soit à estimer le modèle à effets fixes individuels., nous estimons $K=4$ coefficients et $N=15$ termes constants.

Le degré de liberté est égal à :

$(N * T = \text{nombre total d'observations}) - (K + N = \text{nombre de coefficients à estimer})$

$(15 * 23 = \text{nombre total d'observations}) - (4 + 15 = \text{nombre de coefficients à estimer})$

⊕ **Les résultats sont les suivants :**

Tableau 05 : Estimation du modèle à effets fixes

Dependent Variable: Y?

Method: Pooled Least Squares

Date: 06/04/14 Time: 12:16

Sample: 1990 2012

Included observations: 23

Cross-sections included: 15

Total pool (balanced) observations: 345

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	3.988035	0.449542	8.871337	0.0000
X1?	0.000747	0.010014	0.074616	0.9406
X2?	-0.006633	0.003125	-2.122519	0.0345
X3?	0.007265	0.005786	1.255629	0.2101
X4?	0.015072	0.006827	2.207735	0.0280
Fixed Effects (Cross)				
_ALG--C	-1.546477			
_SAU--C	1.075104			
_BHR--C	-0.514315			
_EGY--C	-0.467877			
_ARE--C	1.080629			
_IRN--C	-0.397105			
_JOR--C	0.856647			
_KWT--C	0.989109			
_LBN--C	-2.356535			
_MAR--C	0.970855			
_OMN--C	-1.145071			
_QAT--C	-0.968935			
_TUN--C	1.425242			
_TUR--C	-1.323185			
_YEM--C	2.321915			
Effects Specification				
Cross-section fixed (dummy variables)				
R-squared	0.734595	Mean dependent var	4.845549	
Adjusted R-squared	0.719941	S.D. dependent var	1.511859	
S.E. of regression	0.800085	Akaike info criterion	2.445300	
Sum squared resid	208.6842	Schwarz criterion	2.656973	
Log likelihood	-402.8142	Hannan-Quinn criter.	2.529598	
F-statistic	50.12848	Durbin-Watson stat	0.821260	
Prob(F-statistic)	0.000000			

Source : outputs d'E-views 8

$$SCR_{c2} = 208.6842$$

$$ddld = N * T - N(K + 1) = 270$$

$$ddln = (N - 1) * K = (15-1)*4 = 56$$

$$F_2 = \frac{(208,6842 - 142,22449)/56}{142,22449/270}$$

$$F_1 = \frac{66.45971}{0.52676} = 2.253$$

$$F_{Cal} = 2.253$$

$$F_{Tab} = F_{ddln, ddld}^{\alpha} = F_{56, 270}^{5\%}$$

$$F_{Tab} = 1.36$$

La décision:

$F_{Cal} > F_{Tab}$. Alors nous rejetons l'hypothèse H_0^2 au seuil de 5%.

Ce résultat confirme l'hétérogénéité totale, c'est-à-dire les constantes α_i et les coefficients β_i sont tous différents pour toutes les valeurs de (i) , la structure en panel est rejetée. Le modèle doit être estimé équation par équation pour les **15** pays.

Dans la perspective économique, cette hétérogénéité peut être justifiée par les différences principales entre les pays de notre échantillon étudié, notamment de tailles (population, infrastructure,..) et de systèmes éducatifs et politiques.

3- Les déterminants des dépenses de l'éducation en Algérie :

A-Résultats de la régression pas-à-pas est bien détaillé ci-dessous

Tableau 06 : Résultats de la régression pas-à-pas

Dependent Variable: Y_ALG

Method: Stepwise Regression

Date: 06/16/14 Time: 14:37

Sample: 1990 2012

Included observations: 23

Number of always included regressors: 5

No search regressors

Selection method: Stepwise forwards

Stopping criterion: p-value forwards/backwards = 0.5/0.5

Variable	Coefficient	Std. Error	t-Statistic	Prob.
X1_ALG	0.056895	0.039262	1.449089	0.1645
X2_ALG	-0.008210	0.009257	-0.886930	0.3868
X3_ALG	0.013361	0.029427	0.454058	0.6552
X4_ALG	0.114366	0.029937	3.820178	0.0013
C	0.182963	1.913529	0.095616	0.9249
R-squared	0.892889	Mean dependent var	3.183800	
Adjusted R-squared	0.869086	S.D. dependent var	1.022118	
S.E. of regression	0.369823	Akaike info criterion	1.038076	
Sum squared resid	2.461843	Schwarz criterion	1.284922	
Log likelihood	-6.937869	Hannan-Quinn criter.	1.100157	
F-statistic	37.51235	Durbin-Watson stat	1.720104	
Prob(F-statistic)	0.000000			
Selection Summary				
No regressors were chosen by the stepwise routine				

Source : Outputs d'E-views-8

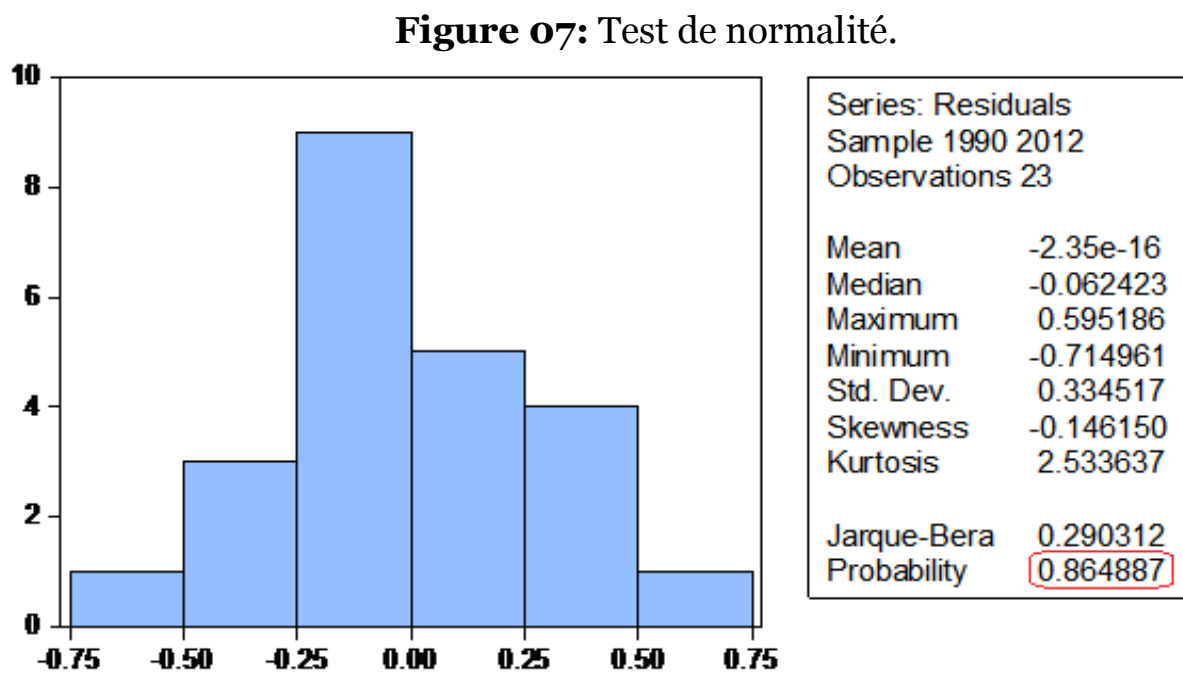
B- Tests sur les résidus :

1) Test de normalité :

On procède au test de normalité afin de vérifier que les erreurs (**résidus**) suivent la loi normale ou non. **(Philippe Casin, 2009)**

Les hypothèses du test sont les suivantes :

$$\begin{cases} H_0 : \text{Les erreurs suivent une loi normale. } (\varepsilon_t \sim N(0, \sigma^2_\varepsilon) \forall t). \\ H_a : \text{Les erreurs ne suivent pas une loi normale.} \end{cases}$$



Source : Outputs d'E-views-8.

- ⊕ Il apparaît que les résidus peuvent s'ajuster à une distribution normale, car la probabilité fournie, est égale à **0,864887**. Donc nous acceptons l'hypothèse nulle de normalité des résidus pour un seuil de signification $\alpha = 5\%$.

2) Test d'hétéroscédasticité des erreurs (Test de White) :

Il existe plusieurs tests pour détecter l'hétéroscédasticité des erreurs :

- ⊕ Test de Goldfeld-Quandt.
- ⊕ Test de White.
- ⊕ Test de Breusch-Pagan-Godfrey.
- ⊕ Test de Glejser.

Nous n'étudierons ici que le test de White.

Les hypothèses du test sont les suivantes :

$$\begin{cases} H_0 : E(\varepsilon_t^2) = \sigma_\varepsilon^2 I. & \longrightarrow \text{Homoscédasticité des erreurs.} \\ H_a : E(\varepsilon_t^2) = \sigma_{\varepsilon_t}^2 \neq \sigma_\varepsilon^2 I & \longrightarrow \text{Hétéroscédasticité des erreurs.} \end{cases}$$

La statistique de test est donnée par :

$$W = T \cdot R^2$$

Avec : T : le nombre d'observations.

R^2 : le coefficient de détermination.

Tableau 07 : Test d'hétéroscédasticité (test de White).

Heteroskedasticity Test: White

F-statistic	3.027221	Prob. F(14,8)	0.0599
Obs*R-squared	19.34784	Prob. Chi-Square(14)	0.1521
Scaled explained SS	9.086872	Prob. Chi-Square(14)	0.8254

Test Equation:

Dependent Variable: RESID^2

Method: Least Squares

Date: 06/16/14 Time: 15:21

Sample: 1990 2012

Included observations: 23

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	26.91463	12.70137	2.119033	0.0669
X1_ALG^2	-0.003093	0.005945	-0.520279	0.6170
X1_ALG*X2_ALG	-0.001142	0.002113	-0.540508	0.6036
X1_ALG*X3_ALG	0.012568	0.006914	1.817914	0.1066
X1_ALG*X4_ALG	-0.008360	0.007593	-1.101031	0.3029
X1_ALG	-0.828367	0.475688	-1.741407	0.1198
X2_ALG^2	-0.000244	0.000280	-0.870381	0.4094
X2_ALG*X3_ALG	-0.007187	0.002348	-3.060636	0.0156
X2_ALG*X4_ALG	0.005810	0.002009	2.892184	0.0201
X2_ALG	0.494241	0.156749	3.153067	0.0135
X3_ALG^2	0.008977	0.003365	2.667528	0.0285
X3_ALG*X4_ALG	-0.020105	0.006557	-3.066022	0.0154
X3_ALG	-1.001642	0.415291	-2.411905	0.0424
X4_ALG^2	0.010521	0.003280	3.207794	0.0125
X4_ALG	1.172694	0.416333	2.816722	0.0226
R-squared	0.841210	Mean dependent var		0.107037
Adjusted R-squared	0.563328	S.D. dependent var		0.135533
S.E. of regression	0.089562	Akaike info criterion		-1.739476
Sum squared resid	0.064171	Schwarz criterion		-0.998936
Log likelihood	35.00397	Hannan-Quinn criter.		-1.553232
F-statistic	3.027221	Durbin-Watson stat		2.236946
Prob(F-statistic)	0.059905			

Source : Outputs d'E-views 8.

☞ La valeur de **W** est donnée par **Obs*R-squared**, est égale ici à **19,34784** et on conclut à l'homoscédasticité des erreurs (résidus) c'est-à-dire on accepte **H₀** , puisque la probabilité critique est égale à **0,1521** > 5%.

3) Test d'autocorrélation des erreurs :

1. Test d'autocorrélation d'ordre 1 (Test de Durbin-Watson) :

La statistique de test est donnée par :

$$DW = \frac{\sum_{t=2}^T (e_t - e_{t-1})^2}{\sum_{t=1}^T e_t^2}$$

Les hypothèses sont les suivantes :

$$\begin{cases} H_0 : \rho = 0. \longrightarrow \text{Les erreurs ne sont pas autocorrélées.} \\ H_a : \varepsilon_t = \rho * \varepsilon_{t-1} + v_t. \longrightarrow \text{Les erreurs sont autocorrélées d'ordre 1.} \end{cases}$$

La valeur de **DW** est égale à **1,720104** $\approx$ **2** (**Tableau 06**) donc nous acceptons l'hypothèse nulle, c'est-à-dire les erreurs ne sont pas autocorrélées d'ordre 1.

2. Test d'autocorrélation d'ordre supérieur à 1 : (Test de Ljung-Box)

La statistique de test est donnée par :

$$Q_{LB} = T(T+2) \sum_{j=1}^k \frac{\rho_j^2}{T-j}$$

Où **T** est le nombre d'observations et ρ_j est le coefficient d'autocorrélation d'ordre j.

Les hypothèses sont les suivantes :

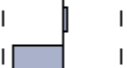
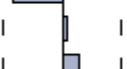
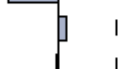
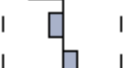
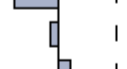
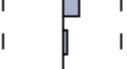
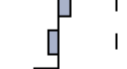
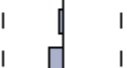
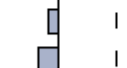
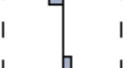
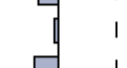
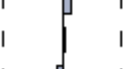
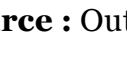
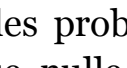
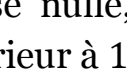
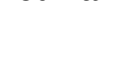
$$\begin{cases} H_0 : \rho_1 = \rho_2 = \dots = \rho_k = 0. \\ H_a : \text{il existe au moins un } \rho_i \text{ significativement différent de 0.} \end{cases}$$

Tableau 08 : Test d'autocorrélation des erreurs.

Date: 06/19/14 Time: 13:06

Sample: 1990 2012

Included observations: 23

Autocorrelation	Partial Correlation	AC	PAC	Q-Stat	Prob
		1 0.034	0.034	0.0294	0.864
		2 -0.337	-0.338	3.1315	0.209
		3 0.034	0.069	3.1651	0.367
		4 0.116	-0.004	3.5699	0.467
		5 -0.276	-0.283	6.0054	0.306
		6 -0.307	-0.291	9.1936	0.163
		7 0.161	-0.005	10.128	0.181
		8 0.106	-0.106	10.563	0.228
		9 -0.237	-0.245	12.879	0.168
		10 -0.097	-0.232	13.297	0.208
		11 0.113	-0.309	13.912	0.238
		12 0.186	-0.048	15.724	0.204
		13 0.115	0.095	16.486	0.224
		14 0.029	-0.058	16.539	0.282
		15 0.008	-0.171	16.543	0.347
		16 -0.102	-0.286	17.397	0.360
		17 -0.024	-0.066	17.453	0.424
		18 -0.089	-0.127	18.364	0.432
		19 0.009	-0.023	18.375	0.498
		20 0.069	-0.161	19.284	0.503
		21 0.022	-0.141	19.427	0.558
		22 -0.033	-0.051	20.045	0.580

Source : Outputs d'E-views 8.

✍ Il apparaît que toutes les probabilités sont supérieures à 5%, donc nous acceptons l'hypothèse nulle, c'est-à-dire les erreurs ne sont pas autocorrélées d'ordre supérieur à 1.

4) Test de stabilité (Test de Chow):

❖ Le test de Chow permet de savoir si un modèle est stable ou non sur une période donnée. Un modèle instable a ses coefficients qui varient durant la période considérée.

◆ Les hypothèses du test sont les suivantes le cas ci-dessous :

$$y_t = a_1 x_{1t} + a_2 x_{2t} + a_3 x_{3t} \Rightarrow SCR_0$$

$$y_t = a'_1 x_{1t} + a'_2 x_{2t} + a'_3 x_{3t} \Rightarrow SCR_1$$

$$y_t = a''_1 x_{1t} + a''_2 x_{2t} + a''_3 x_{3t} \Rightarrow SCR_2$$

$$\left\{ \begin{array}{l} H_0 : a_1 = a'_1 = a''_1 \\ \quad a_2 = a'_2 = a''_2 \rightarrow \text{le modèle est stable} \\ \quad a_3 = a'_3 = a''_3 \\ \Rightarrow \text{modèle contraint avec } SCR_c = SCR_0 \\ \\ H_1 : a'_1 \neq a''_1 \\ \quad a'_2 \neq a''_2 \rightarrow \text{le modèle est instable} \\ \quad a'_3 \neq a''_3 \\ \Rightarrow \text{modèle non contraint avec } SCR_{nc} = SCR_1 + SCR_2 \end{array} \right.$$

SCR_c : Somme des carrées des résidus du modèle contraint.

SCR_{nc} : Somme des carrées des résidus du modèle non contraint.

◆ Le test de Chow se construit comme un test de Fisher.

$$F = \frac{(SCR_c - SCR_{nc}) / (dl_c - dl_{nc})}{SCR_{nc} / dl_{nc}} \sim F(dl_c - dl_{nc}, dl_{nc}) = F(p, q)$$

$$F = \frac{(SCR_0 - (SCR_1 + SCR_2)) / (dl_c - dl_{nc})}{(SCR_1 + SCR_2) / dl_{nc}} \sim F(dl_c - dl_{nc}, dl_{nc}) = F(p, q).$$

Tableau 09 : Test de stabilité du modèle (Test de Chow).

Chow Breakpoint Test: 2000

Null Hypothesis: No breaks at specified breakpoints

Varying regressors: All equation variables

Equation Sample: 1990 2012

F-statistic	0.381444	Prob. F(5,13)	0.8527
Log likelihood ratio	3.148618	Prob. Chi-Square(5)	0.6771
Wald Statistic	1.907222	Prob. Chi-Square(5)	0.8618

Source : Outputs d'E-views 8.

La probabilité de Fisher est égale à **0,8527**, elle est supérieure à 5% ; donc nous acceptons l'hypothèse nulle de stabilité du modèle.

C-Diagnostic du modèle : (Tableau 06)

La signification partielle des coefficients :

D'après le tableau 06, on remarque que :

- 1-** Pour la constante : La **Prob=0.9249** est supérieure à la valeur critique au niveau de **5%** cela nous conduit d'accepter l'hypothèse nulle, c'est-à-dire la constante n'est pas différente significativement de zéro « 0 ».
- 2-** Pour l'ensemble des variables exogènes (x_1, x_2, x_3) ; la Prob est supérieure de **5%**, donc on accepte H_0 dont les coefficients sont non-significatifs.
- 3-** Par contre la variable (x_4) qui représente l'inscription d'enseignement supérieure est significativement différente de zéro, car la Prob est inférieure de la valeur critique
(**Prob(t-Statistic)= 0.0013 < 0.05**).

La spécification globale du modèle :

- 1- Le modèle estimé est accepté à partir de
 $\text{Prob}(\text{F-statistic}) = 0.000000 < 5\%$.
- 2- $R^2 = 0.892889$ c'est-à-dire l'ensemble des variables exogènes (X_k) explique **89.29%** de la variable endogène (Y).

Pour le cas de l'Algérie ; et d'après les différent tests, on conclue que :

-  Les erreurs suivent la loi normale ; sont non autocorrélées et ses variances sont homoscédastiques.
-  le modèle estimé est généralement accepté et stable.
-  Il existe une relation positive et significative entre l'évolution des dépenses de l'éducation et l'inscription à l'école, enseignement supérieure.

Conclusion

Dans ce chapitre empirique ; on a conclu que la spécification des indicateurs des dépenses publiques de l'éducation est différente globalement pour tous les pays de la région MENA ; et pour bien indiquer ces déterminants ; il faut estimer un propre modèle pour chaque pays.

Pour le cas de l'Algérie ; les déterminants des dépenses publiques de l'éducation sont déterminés selon notre étude par une seule variable qui représente l'inscription à l'école, enseignement supérieur selon la régression **(Stepwise forwards)** ; on peut justifier cette relation positive que le système d'éducation en Algérie est un système étatique et les dépenses publiques sur l'éducation se déterminent par le taux d'inscription dans l'enseignement supérieur et non pas la qualité de l'éducation.

Conclusion Générale

Conclusion Générale :

Les nouvelles théories de développement économique précisent une nécessité de l'investissement en capital humain dans l'économie de l'éducation. Les dépenses socio-publiques d'éducation sont de nature à générer des gains de productivité pour une croissance rapide comme le supposent les modèles de croissance endogène.

L'analyse des déterminants de dépenses publiques de l'éducation et la recherche de ses facteurs dans les pays de la région MENA , s'inscrivent dans le cadre global des économies qui ont connu au cours de ces trois dernières décennies de grandes mutations et de profonds changements, liés à la mise en œuvre de réformes sur le plan politique, économique et en particulier sur le plan éducatif.

La présente étude s'est proposé de déterminer empiriquement les déterminants de dépenses publiques de l'éducation dans les pays de la région MENA, pour cela, nous avons adopté l'analyse des données de panel, nous avons référé au test de **Hsiao** pour tester l'existence de la structure de panel.

En effet, les résultats du test montrent qu'il existe une hétérogénéité totale des déterminants de dépenses publiques de l'éducation ; cette hétérogénéité peut être justifiée par les différences principales entre les pays de notre étude, notamment de tailles (population, infrastructure,..) et de systèmes éducatifs et politiques ; donc le modèle doit être estimé équation par équation de chaque pays.

Références bibliographiques

Ouvrages :

En français :

- ✚ Alain Pirotte, Économétrie des données de panel « Théorie et application », Ed ECONOMICA, Paris, 2011.
- ✚ Patrick Sevestre, Économétrie des données de panel, Dunod, Paris, 2002.
- ✚ Philippe Casin, Économétrie « Méthodes et applications avec EViews », Editions Technip, Paris, 2009.
- ✚ Régis Bourbonnais, Économétrie « Manuel et exercices corrigés », 8^e édition, Dunod, Paris, 2011.
- ✚ Valérie Mignon, Économétrie « Théorie et applications », Ed ECONOMICA, Paris, 2008.

En anglais :

- ✚ Badi H. Baltagi, Econometrics, 5th edition, Springer, New York, USA, 2011.
- ✚ Becker (G.S.) “Human Capital: A Theoretical and Empirical Analysis, With Special Reference to Education”, 3rd edition, University of Chicago Press, USA, 1993.
- ✚ Benhabib J. and M. Spiegel, “The Role of Human Capital in Economic”, 1994.
- ✚ Hsiao C. “Analysis of Panel Data”, Econometric Society Monographs, N°11, Cambridge University Press, 1996.
- ✚ Schultz (T.W.) “Investment in Human Capital”. *American Economic Review*, LI, March 1961.

Rapports et revues :

- + BANQUE MONDIALE, Rapports sur le développement de la région MENA, « Un parcours non encore achevé : La réforme de l'éducation au Moyen-Orient et en Afrique du Nord », Washington, 2007.
- + UNESCO, Institut de statistique, Recueil de données mondiales sur l'éducation « Statistiques comparées sur l'éducation dans le monde », 2007.

Webographie :

- + www.memoireonline.com
- + www.worldbank.org
- + www.uis.unesco.org



ANNEXES

ANNEXE N°01 : Résultats de l'estimation du modèle pour l'Algérie

Dependent Variable: Y_ALG
Method: Least Squares
Date: 06/15/14 Time: 15:36
Sample: 1990 2012
Included observations: 23

Variable	Coefficient	Std. Error	t-Statistic	Prob.
X1_ALG	0.056895	0.039262	1.449089	0.1645
X2_ALG	-0.008210	0.009257	-0.886930	0.3868
X3_ALG	0.013361	0.029427	0.454058	0.6552
X4_ALG	0.114366	0.029937	3.820178	0.0013
C	0.182963	1.913529	0.095616	0.9249
R-squared	0.892889	Mean dependent var		3.183800
Adjusted R-squared	0.869086	S.D. dependent var		1.022118
S.E. of regression	0.369823	Akaike info criterion		1.038076
Sum squared resid	2.461843	Schwarz criterion		1.284922
Log likelihood	-6.937869	Hannan-Quinn criter.		1.100157
F-statistic	37.51235	Durbin-Watson stat		1.720104
Prob(F-statistic)	0.000000			

ANNEXE N°02 : Résultats de l'estimation du modèle pour l'Arabie Saoudite

Dependent Variable: Y_SAU
Method: Least Squares
Date: 05/31/14 Time: 15:28
Sample: 1990 2012
Included observations: 23

Variable	Coefficient	Std. Error	t-Statistic	Prob.
X1_SAU	-0.155389	0.054093	-2.872655	0.0101
X2_SAU	-0.006269	0.033673	-0.186175	0.8544
X3_SAU	0.116952	0.054819	2.133413	0.0469
X4_SAU	-0.009621	0.024719	-0.389222	0.7017
C	-4.150151	4.650508	-0.892408	0.3839
R-squared	0.434605	Mean dependent var		6.042294
Adjusted R-squared	0.308962	S.D. dependent var		0.912100
S.E. of regression	0.758217	Akaike info criterion		2.473966
Sum squared resid	10.34808	Schwarz criterion		2.720813
Log likelihood	-23.45061	Hannan-Quinn criter.		2.536047
F-statistic	3.459040	Durbin-Watson stat		2.259005
Prob(F-statistic)	0.028999			

ANNEXE N°03 : Résultats de l'estimation du modèle pour Bahreïn

Dependent Variable: Y_BHR

Method: Least Squares

Date: 05/31/14 Time: 15:09

Sample: 1990 2012

Included observations: 23

Variable	Coefficient	Std. Error	t-Statistic	Prob.
X1_BHR	0.024094	0.081875	0.294282	0.7719
X2_BHR	-0.022972	0.026284	-0.873975	0.3936
X3_BHR	-0.134457	0.214860	-0.625788	0.5393
X4_BHR	0.001558	0.064271	0.024237	0.9809
C	17.54337	19.31488	0.908282	0.3757
R-squared	0.205689	Mean dependent var		4.459043
Adjusted R-squared	0.029175	S.D. dependent var		0.974103
S.E. of regression	0.959788	Akaike info criterion		2.945452
Sum squared resid	16.58148	Schwarz criterion		3.192298
Log likelihood	-28.87270	Hannan-Quinn criter.		3.007533
F-statistic	1.165286	Durbin-Watson stat		0.564857
Prob(F-statistic)	0.358875			

ANNEXE N°04 : Résultats de l'estimation du modèle pour l'Égypte

Dependent Variable: Y_EGY

Method: Least Squares

Date: 05/31/14 Time: 15:11

Sample: 1990 2012

Included observations: 23

Variable	Coefficient	Std. Error	t-Statistic	Prob.
X1_EGY	-0.096929	0.106302	-0.911821	0.3739
X2_EGY	-0.053918	0.031103	-1.733490	0.1001
X3_EGY	-0.034422	0.034187	-1.006860	0.3273
X4_EGY	0.047784	0.044530	1.073073	0.2974
C	6.993202	2.125663	3.289893	0.0041
R-squared	0.251136	Mean dependent var		4.430435
Adjusted R-squared	0.084722	S.D. dependent var		0.681199
S.E. of regression	0.651704	Akaike info criterion		2.171208
Sum squared resid	7.644926	Schwarz criterion		2.418054
Log likelihood	-19.96889	Hannan-Quinn criter.		2.233289
F-statistic	1.509101	Durbin-Watson stat		1.191751
Prob(F-statistic)	0.241476			

ANNEXE N°05 : Résultats de l'estimation du modèle pour l'Émirats arabes unis

Dependent Variable: Y_ARE

Method: Least Squares

Date: 05/31/14 Time: 15:03

Sample: 1990 2012

Included observations: 23

Variable	Coefficient	Std. Error	t-Statistic	Prob.
X1_ARE	-0.024047	0.045258	-0.531333	0.6017
X2_ARE	-0.032088	0.027028	-1.187199	0.2506
X3_ARE	-0.086476	0.102576	-0.843039	0.4103
X4_ARE	0.066791	0.062518	1.068347	0.2995
C	13.18722	8.571804	1.538442	0.1413
R-squared	0.173872	Mean dependent var		5.945928
Adjusted R-squared	-0.009713	S.D. dependent var		1.064946
S.E. of regression	1.070105	Akaike info criterion		3.163051
Sum squared resid	20.61225	Schwarz criterion		3.409897
Log likelihood	-31.37508	Hannan-Quinn criter.		3.225132
F-statistic	0.947095	Durbin-Watson stat		0.901553
Prob(F-statistic)	0.459715			

ANNEXE N°06 : Résultats de l'estimation du modèle pour l'Iran

Dependent Variable: Y_IRN

Method: Least Squares

Date: 05/31/14 Time: 15:12

Sample: 1990 2012

Included observations: 23

Variable	Coefficient	Std. Error	t-Statistic	Prob.
X1_IRN	-0.014948	0.027629	-0.541027	0.5951
X2_IRN	0.000437	0.006674	0.065437	0.9485
X3_IRN	0.040114	0.013549	2.960549	0.0084
X4_IRN	-0.028612	0.011315	-2.528646	0.0210
C	1.643855	1.091679	1.505804	0.1495
R-squared	0.363023	Mean dependent var		4.456522
Adjusted R-squared	0.221472	S.D. dependent var		0.483196
S.E. of regression	0.426344	Akaike info criterion		1.322521
Sum squared resid	3.271848	Schwarz criterion		1.569367
Log likelihood	-10.20899	Hannan-Quinn criter.		1.384602
F-statistic	2.564615	Durbin-Watson stat		1.260589
Prob(F-statistic)	0.073665			

ANNEXE N°07 : Résultats de l'estimation du modèle pour Jordanie

Dependent Variable: Y_JOR
Method: Least Squares
Date: 05/31/14 Time: 15:13
Sample: 1990 2012
Included observations: 23

Variable	Coefficient	Std. Error	t-Statistic	Prob.
X1_JOR	-0.097005	0.049869	-1.945207	0.0675
X2_JOR	-0.062735	0.030002	-2.090984	0.0510
X3_JOR	-0.469044	0.134823	-3.478964	0.0027
X4_JOR	0.152722	0.042874	3.562133	0.0022
C	47.22419	11.98246	3.941111	0.0010
R-squared	0.447502	Mean dependent var		5.947826
Adjusted R-squared	0.324725	S.D. dependent var		0.988051
S.E. of regression	0.811932	Akaike info criterion		2.610861
Sum squared resid	11.86622	Schwarz criterion		2.857707
Log likelihood	-25.02490	Hannan-Quinn criter.		2.672942
F-statistic	3.644825	Durbin-Watson stat		1.906329
Prob(F-statistic)	0.024117			

ANNEXE N°08 : Résultats de l'estimation du modèle pour Koweït

Dependent Variable: Y_KWT
Method: Least Squares
Date: 05/31/14 Time: 15:15
Sample: 1990 2012
Included observations: 23

Variable	Coefficient	Std. Error	t-Statistic	Prob.
X1_KWT	0.001579	0.030908	0.051098	0.9598
X2_KWT	-0.045481	0.027152	-1.675064	0.1112
X3_KWT	-0.004086	0.010599	-0.385535	0.7044
X4_KWT	0.072934	0.031562	2.310798	0.0329
C	5.095441	1.111610	4.583839	0.0002
R-squared	0.279182	Mean dependent var		5.917391
Adjusted R-squared	0.119000	S.D. dependent var		0.790657
S.E. of regression	0.742123	Akaike info criterion		2.431057
Sum squared resid	9.913443	Schwarz criterion		2.677904
Log likelihood	-22.95716	Hannan-Quinn criter.		2.493138
F-statistic	1.742906	Durbin-Watson stat		1.015642
Prob(F-statistic)	0.184529			

ANNEXE N°09 : Résultats de l'estimation du modèle pour Liban

Dependent Variable: Y_LBN
Method: Least Squares
Date: 05/31/14 Time: 15:17
Sample: 1990 2012
Included observations: 23

Variable	Coefficient	Std. Error	t-Statistic	Prob.
X1_LBN	0.014966	0.019826	0.754852	0.4601
X2_LBN	0.008613	0.007589	1.134926	0.2713
X3_LBN	-1.64E-06	0.054810	-3.00E-05	1.0000
X4_LBN	0.020275	0.029619	0.684538	0.5024
C	1.643667	3.930215	0.418213	0.6807
R-squared	0.160966	Mean dependent var		2.686957
Adjusted R-squared	-0.025486	S.D. dependent var		0.616666
S.E. of regression	0.624474	Akaike info criterion		2.085847
Sum squared resid	7.019428	Schwarz criterion		2.332694
Log likelihood	-18.98724	Hannan-Quinn criter.		2.147928
F-statistic	0.863313	Durbin-Watson stat		1.133738
Prob(F-statistic)	0.504602			

ANNEXE N°10 : Résultats de l'estimation du modèle pour Maroc

Dependent Variable: Y_MAR
Method: Least Squares
Date: 05/31/14 Time: 15:18
Sample: 1990 2012
Included observations: 23

Variable	Coefficient	Std. Error	t-Statistic	Prob.
X1_MAR	-0.030145	0.020112	-1.498860	0.1512
X2_MAR	-0.018598	0.021887	-0.849707	0.4066
X3_MAR	0.032264	0.040783	0.791122	0.4392
X4_MAR	0.032388	0.049991	0.647876	0.5252
C	2.990157	3.111469	0.961011	0.3493
R-squared	0.270709	Mean dependent var		5.639130
Adjusted R-squared	0.108644	S.D. dependent var		0.444910
S.E. of regression	0.420047	Akaike info criterion		1.292758
Sum squared resid	3.175904	Schwarz criterion		1.539604
Log likelihood	-9.866714	Hannan-Quinn criter.		1.354839
F-statistic	1.670376	Durbin-Watson stat		1.677333
Prob(F-statistic)	0.200544			

ANNEXE N°11 : Résultats de l'estimation du modèle pour Oman

Dependent Variable: Y_OMN

Method: Least Squares

Date: 05/31/14 Time: 15:25

Sample: 1990 2012

Included observations: 23

Variable	Coefficient	Std. Error	t-Statistic	Prob.
X1_OMN	-0.008139	0.025243	-0.322438	0.7508
X2_OMN	-0.015606	0.011265	-1.385308	0.1829
X3_OMN	0.022878	0.018267	1.252407	0.2264
X4_OMN	0.036392	0.017076	2.131267	0.0471
C	1.221222	1.512015	0.807678	0.4298
R-squared	0.638945	Mean dependent var		3.652174
Adjusted R-squared	0.558711	S.D. dependent var		0.523858
S.E. of regression	0.347996	Akaike info criterion		0.916411
Sum squared resid	2.179828	Schwarz criterion		1.163258
Log likelihood	-5.538731	Hannan-Quinn criter.		0.978493
F-statistic	7.963489	Durbin-Watson stat		2.232140
Prob(F-statistic)	0.000704			

ANNEXE N°12 : Résultats de l'estimation du modèle pour Qatar

Dependent Variable: Y_QAT

Method: Least Squares

Date: 05/31/14 Time: 15:27

Sample: 1990 2012

Included observations: 23

Variable	Coefficient	Std. Error	t-Statistic	Prob.
X1_QAT	-0.051511	0.037326	-1.380043	0.1845
X2_QAT	-0.018771	0.012030	-1.560342	0.1361
X3_QAT	0.043309	0.032423	1.335775	0.1983
X4_QAT	-0.034038	0.027518	-1.236960	0.2320
C	0.705859	3.166640	0.222905	0.8261
R-squared	0.252897	Mean dependent var		3.878261
Adjusted R-squared	0.086875	S.D. dependent var		0.616409
S.E. of regression	0.589026	Akaike info criterion		1.968967
Sum squared resid	6.245128	Schwarz criterion		2.215814
Log likelihood	-17.64312	Hannan-Quinn criter.		2.031048
F-statistic	1.523270	Durbin-Watson stat		1.028190
Prob(F-statistic)	0.237561			

ANNEXE N°13 : Résultats de l'estimation du modèle pour Tunisie

Dependent Variable: Y_TUN

Method: Least Squares

Date: 05/31/14 Time: 15:29

Sample: 1990 2012

Included observations: 23

Variable	Coefficient	Std. Error	t-Statistic	Prob.
X1_TUN	0.007262	0.049389	0.147044	0.8847
X2_TUN	0.002038	0.026352	0.077345	0.9392
X3_TUN	0.009426	0.013437	0.701505	0.4920
X4_TUN	-0.006991	0.024284	-0.287879	0.7767
C	5.627350	0.609645	9.230536	0.0000
R-squared	0.072568	Mean dependent var		6.217391
Adjusted R-squared	-0.133528	S.D. dependent var		0.393876
S.E. of regression	0.419349	Akaike info criterion		1.289434
Sum squared resid	3.165366	Schwarz criterion		1.536281
Log likelihood	-9.828492	Hannan-Quinn criter.		1.351515
F-statistic	0.352108	Durbin-Watson stat		1.892566
Prob(F-statistic)	0.839153			

ANNEXE N°14 : Résultats de l'estimation du modèle pour Turquie

Dependent Variable: Y_TUR

Method: Least Squares

Date: 05/31/14 Time: 15:31

Sample: 1990 2012

Included observations: 23

Variable	Coefficient	Std. Error	t-Statistic	Prob.
X1_TUR	-0.029088	0.017527	-1.659649	0.1143
X2_TUR	-0.006348	0.003138	-2.023085	0.0582
X3_TUR	0.002072	0.008818	0.235011	0.8169
X4_TUR	0.039340	0.010666	3.688381	0.0017
C	2.393809	0.659004	3.632465	0.0019
R-squared	0.851421	Mean dependent var		3.317391
Adjusted R-squared	0.818404	S.D. dependent var		0.854747
S.E. of regression	0.364243	Akaike info criterion		1.007668
Sum squared resid	2.388111	Schwarz criterion		1.254515
Log likelihood	-6.588185	Hannan-Quinn criter.		1.069749
F-statistic	25.78699	Durbin-Watson stat		1.445698
Prob(F-statistic)	0.000000			

ANNEXE N°15 : Résultats de l'estimation du modèle pour **Yémen**

Dependent Variable: Y_YEM

Method: Least Squares

Date: 05/31/14 Time: 15:32

Sample: 1990 2012

Included observations: 23

Variable	Coefficient	Std. Error	t-Statistic	Prob.
X1_YEM	0.078326	0.093079	0.841503	0.4111
X2_YEM	0.010108	0.032347	0.312485	0.7583
X3_YEM	-0.020333	0.032285	-0.629807	0.5367
X4_YEM	0.248194	0.164058	1.512837	0.1477
C	6.119009	2.151925	2.843504	0.0108
R-squared	0.150598	Mean dependent var	6.908696	
Adjusted R-squared	-0.038158	S.D. dependent var	1.375405	
S.E. of regression	1.401401	Akaike info criterion	3.702482	
Sum squared resid	35.35063	Schwarz criterion	3.949328	
Log likelihood	-37.57854	Hannan-Quinn criter.	3.764563	
F-statistic	0.797846	Durbin-Watson stat	0.563010	
Prob(F-statistic)	0.542101			

Résumé

L'objet de ce mémoire est de chercher la structure de panel, c'est pourquoi cette étude a recours à l'analyse des données de panel pour étudier la relation fonctionnelle entre les dépenses publiques de l'éducation et ses déterminants pour un échantillon composé de (15) pays de la région MENA au cours de la période (1990-2012). Dans ce travail une attention particulière est accordée au test de spécification d'homogénéité ou hétérogénéité (test de Hsiao) des paramètres estimés relatifs aux déterminants des dépenses publiques de l'éducation. Les principaux résultats retenus à travers le test précédemment noté, montrent qu'il existe une hétérogénéité totale bien évidente, qui peut être justifié par les différences principales entre les pays de la région, au niveau de la population, les infrastructures et surtout les systèmes éducatifs et politiques qui ont également un impact sur les dépenses publiques de l'éducation en pourcentage du PIB.

Mots-clés : Dépenses publiques de l'éducation, La région MENA, Données de panel, Homogénéité ou hétérogénéité, Test de Hsiao.

Abstract

The purpose of this thesis is to find the structure of the panel, which is why this study uses the analysis of panel data to study the functional relationship between public education spending and its determinants for a sample of (15) countries in the MENA region over the period (1990-2012). In this work special attention is given to the test of homogeneity or heterogeneity (Hsiao test) for estimated parameters related to the determinants of public spending on education. The main results given through the test noted above, show that there is a well obvious total heterogeneity, which can be justified by the major differences between countries in the region, in terms of population, infrastructure and especially education systems and policies, which also have an impact on public spending on education as a percentage of GDP.

Keywords: Public spending on education, MENA region, Panel data, Homogeneity or heterogeneity, Hsiao test.

المخلص:

الهدف من هذه المذكرة هو البحث عن بنية معطيات " panel " ، وسبب اللجوء إلى هذا النوع من المعطيات هو دراسة العلاقة الوظيفية بين النفقات العمومية على التعليم ومحدداتها، لعينة مكونة من (15) بلد من منطقة الشرق الأوسط وشمال إفريقيا « MENA » خلال الفترة الزمنية الممتدة ما بين (1990 - 2012) ، هذا العمل أولى الاهتمام باختبار وجود تجانس أو عدمه للمعالم المقدرة الخاصة بمحددات الإنفاق العمومي على التعليم (اختبار Hsiao) ، حيث أسفرت نتائج هذا الاختبار عن وجود عدم تجانس كلي ، والذي يمكن تبريره بالفروقات الأساسية بين دول المنطقة المدروسة ، وذلك على مستوى النمو السكاني ، البنية التحتية وخاصة الأنظمة السياسية والتعليمية والتي لها تأثير على النفقات العمومية في التعليم كنسبة من الناتج الداخلي الخام.

الكلمات المفتاحية:

الإنفاق العام على التعليم، منطقة الشرق الأوسط وشمال إفريقيا MENA، بيانات Panel ، التجانس أو عدم التجانس،

اختبار Hsiao.