

## **La qualité institutionnelle de la région MENA**

Les institutions sont définies comme un facteur fondamental de la croissance économique sur le long terme. Ces institutions sont considérées comme des règles juridiques et sociales qui régissent les systèmes économiques ; les institutions les plus importantes pour la croissance économique sont celles qui permettent à un pays de distribuer les ressources économiques et financières à ces usages les plus productives.

Le cadre institutionnel et réglementaire au sein duquel se déroulent les activités économiques qui affectent la manière dont les ressources sont réparties, les décisions d'investissement sont orientées et la créativité et l'innovation sont stimulées.

L'analyse et l'évaluation des institutions sont au premier rang des préoccupations en économie au cours de ces dernières années. Les institutions permettent d'expliquer les différences en termes des règles de droit et d'efficacité de leur application d'un pays à un autre, ainsi que leur rôle dans le développement du système financier, et par conséquent la croissance de l'économie tout entière.

Pour réussir une ouverture financière dans les économies de la région MENA, le cadre institutionnel doit s'améliorer dans la région. Ainsi, il convient d'apporter des améliorations, notamment par rapport à l'élimination de la paperasserie et des obstacles bureaucratiques, l'élimination de la corruption, le renforcement des systèmes juridiques et judiciaire, le fait de faire valoir le principe de la légalité.

### **5.1. La corruption dans la région MENA**

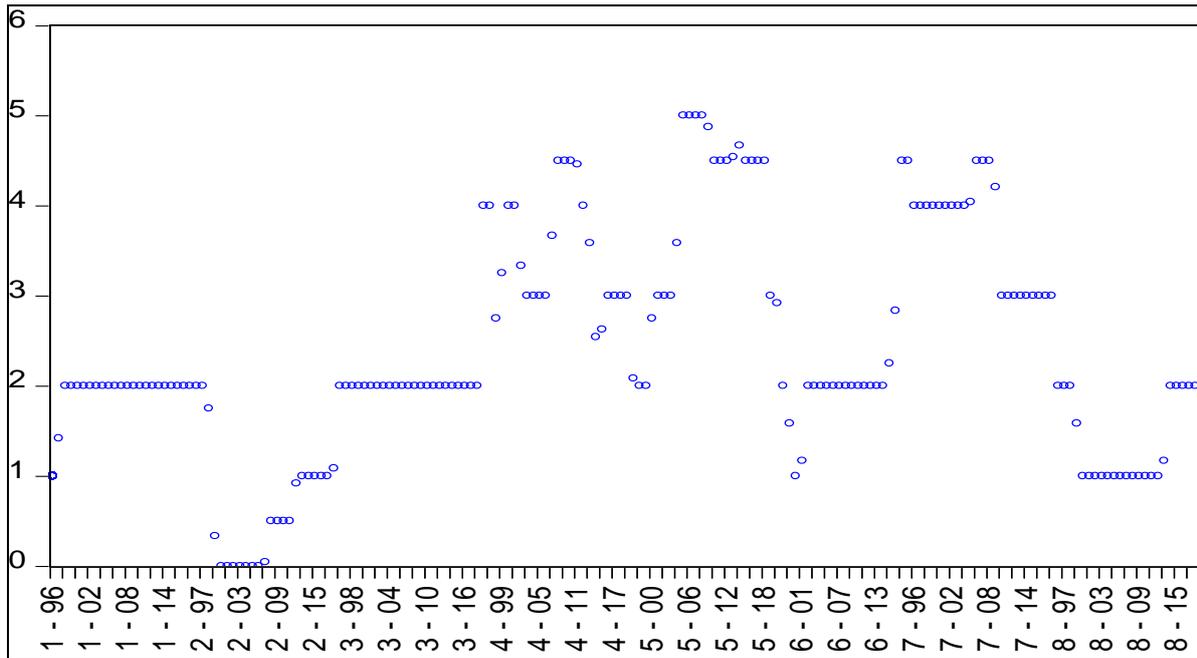
La corruption affaiblit un pays et porte atteinte à la stabilité et la sécurité pour les décisions des agents économiques.

le CPI (corruption perceptions index) qui est élaboré sur la base d'un certain nombre de sources disponibles qui rendent compte de la perception de la

corruption (sondages et évaluations de la corruption perçue collectés par diverses institutions réputées), varie entre 100 :le plus bas niveau de corruption perçue et 0 : le plus haut niveau).

D'après le graphique ci-dessus, les pays de la région MENA sont divisés en deux Catégories :

**GRAPHIQUE III.7 : LE CONTROLE DE LA CORRUPTION DANS LA REGION MENA**



**Source** : auteur, à partir de la base de données l'International Country Risk Guide (ICRG) 2020

- Une première catégorie composée de l'Arabie Saoudite, la Tunisie, Oman et la Jordanie ; ces économies sont caractérisées par un moyen de contrôle de la corruption. L'Arabie Saoudite est classée 57 rang mondial avec un score de 49, la Tunisie est classée 43ème rang mondial avec un score de 74, Oman est classé 52ème rang mondial avec un score de 36 et la Jordanie est classée 48ème avec un score de 60.
- Une deuxième catégorie composé de l'Algérie, l'Egypte, le Maroc et l'Iran ; ces pays sont caractérisés par un faible niveau de contrôle de la corruption ; dont l'Algérie est classée 8ème , l'Egypte est classée 9ème et le Maroc classée 10ème dans la région MENA ; ces pays sont parmi les dix pays les plus corrompus de la région MENA. Selon (transparency

international ,2017), l'Algérie est classée 106<sup>ème</sup> rang mondial et 19<sup>ème</sup> rang Africain avec un score de 35, l'Egypte est classée 107 rang mondial et 21<sup>ème</sup> rang Africain avec un score de 33, le Maroc est classé 86 rang mondial et 15<sup>ème</sup> rang Africain avec un score de 40, l'Iran est classée 130<sup>ème</sup> rang mondial avec un score de 27.

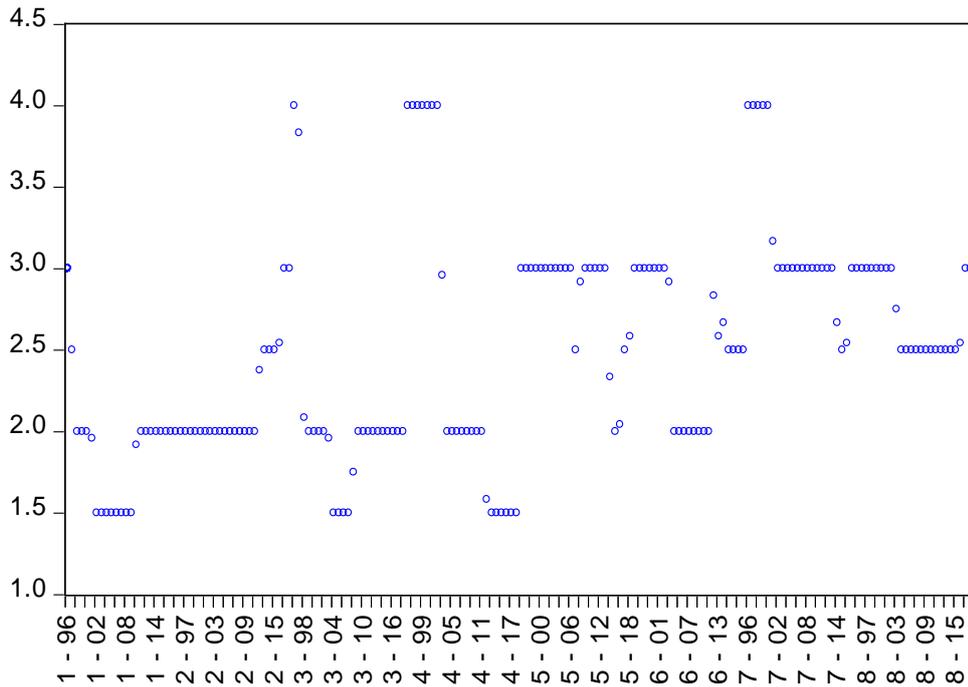
A partir de cette étude statistique les niveaux élevés de la corruption dans les pays de la région résultants du faible contrôle notamment après la révolution politique dans les pays de printemps arabes également la Tunisie et l'Egypte en 2014. Ces niveaux élevés sont traduits par l'augmentation de la bureaucratie ou encore certains fonctionnaires cherchent à créer des procédures bureaucratiques supplémentaires contre des pots de vin.

## **5.2. La qualité de la bureaucratie dans la région MENA**

La qualité de la bureaucratie qui sert à évaluer le degré d'autonomie du service bureaucratique et à apprécier le degré d'efficience de ce service bureaucratique quant à la fourniture des services publics. L'amélioration de la qualité de la bureaucratie et notamment la réduction des délais et des encombrements bureaucratiques vont éviter les lourdeurs des procédures administratives.

En revanche, la paperasserie bureaucratique constitue une entrave et un manque de vigilance à l'égard de certaines contraintes d'ordre politique et bureaucratique avait frustré l'implantation et limité l'impact de certains projets. (Banque mondiale ; 2000). Le graphique ci-après représente la qualité de la bureaucratie sur un échantillon de pays de la région MENA, nous avons utilisé la mesure de International Country Risk Guide.

### Graphique III.8 : la qualité de la bureaucratie dans la région MENA



**Source** : auteur, à partir de la base de données l'International Country Risk Guide (ICRG) 2020

La majorité des pays de la région ont une mesure faible de cet indicateur, les pays du Maghreb et l'Egypte sont pénalisées par de fortes lourdeurs administratives ; ce qui entrave les projets d'investissement et la diversification économique.

Les pays de la région doivent améliorer la performance bureaucratique car elle permet de réduire les incertitudes et les couts de la conduite des affaires ce qui permet l'accès aux projets rentables. Aussi, les investisseurs vont éviter les pays caractérisés par les retards bureaucratiques dans l'obtention des documents administratives.

## **Section 2 : évidence empirique**

Afin d'analyser l'effet de l'ouverture du compte de capital sur la croissance économique, nous avons utilisé l'économétrie des données de panel sur un ensemble de 8 pays de la région MENA (Algérie, Arabie Saoudite, Egypte, Iran, Maroc, Tunisie, Jordanie, Oman) couvrant la période 1996-2018. L'intérêt que nous portons pour la méthode de l'économétrie de panel réside dans le fait qu'elle permet d'étudier le phénomène de la libéralisation du compte de capital dans sa diversité comme dans sa dynamique, car cette méthode économétrique intègre deux dimensions, la dimension temporelle et la dimension individuelle. ; le choix de la période et de l'échantillon de pays sont contraints par la disponibilité des données pour éviter le panel incomplet ou non cylindré. Aussi, il existe plusieurs méthodes d'estimations en données de panel, mais cela dépend des résultats des tests préliminaires effectués.

### **1.Méthodologie de l'étude**

Beaucoup de modèles économétriques ont été développés dans le cadre des techniques d'économétrie de panel. Notre objectif est de déterminer la méthode la plus appropriée à retenir pour nos estimations. Comme on a cité précédemment, nos estimations seront effectuées dans le but d'illustrer la nature du lien entre l'ouverture du compte de capital et la croissance économique dans les pays de la région MENA, pour cela, on a effectué le test de spécification d'homogénéité ou d'hétérogénéité du processus générateur de données qui a pour but de tester l'égalité des coefficients du modèle étudié dans la dimension individuelle. Ensuite, on a testé la stationnarité des variables à travers les tests de racine unitaire sur panel, par la suite on a étudié la relation de cointégration de ces variables. En dernier lieu, on a passé à l'estimation du modèle avec la méthode ARDL panel.

## 1.1. Spécification du modèle à estimer

Cette étude optera deux modèles afin de mesurer l'effet de la libéralisation du compte de capital sur la croissance économique au niveau de notre échantillon. Notre étude empirique est conduite sur un panel de 8 pays de la région MENA sont : Algérie, Arabie Saoudite, Égypte, Iran, Maroc, Jordanie, Tunisie et Oman sur la période 1996-2018. Nous disposons des données relatives à  $N = 8$  pays de la région MENA sur  $T = 24$  périodes (1996 à 2018), donc le total panel observations ( $T \times N$ ) = 184 observations.

Le premier modèle utilise les variables de contrôles et la variable ouverture financière de jure « KAOPEN » dans un premier temps. Dans un second temps, en va essayer d'introduire les termes d'interaction entre l'indicateur de la libéralisation du compte de capital de jure et les variables de la qualité institutionnelle.

Le premier modèle à estimer est le suivant :

$$pib_{i,t} = \alpha_0 + \beta_i X_{i,t} + \alpha_1 KAOPEN + \varepsilon_{i,t} \dots (1-1)$$

$$pib_{i,t} = \alpha_0 + \beta_i X_{i,t} + \alpha_1 KAOPEN + \alpha_2 INST + \alpha_3 KAOPEN * INST + \varepsilon_{i,t} \dots (1-2)$$

Le deuxième modèle utilise les variables de contrôles et la variable ouverture financière « de facto ».

Le deuxième modèle à estimer est le suivant :

$$pib_{i,t} = \alpha_0 + \beta_i X_{i,t} + \alpha_1 LF + \varepsilon_{i,t} \dots (2)$$

**$pib_{i,t}$** : représente le PIB réel par habitant dans le pays  $i$  à la période  $t$ ;

**$X_{i,t}$**  : est la matrice des variables de contrôle couramment utilisées dans la littérature empirique traitant la problématique de la croissance économique ;

Parmi ces variables de contrôle on retient : le taux d'inflation, le degré

d'ouverture commerciale, l'indicateur de développement financier mesurer par les crédits orientés au secteur privé en pourcentage du PIB et enfin le stock du capital humain dans un pays  $i$  sur la période  $t$  ;

**KAOPEN** : mesure la libéralisation du compte de capital de jure dans le pays  $i$  à la période  $t$  ;

**INST** : représente le niveau de développement institutionnelle dans un pays,

**KAOPEN \* INST** : le terme d'interaction entre la qualité institutionnelle et ouverture du compte de capital

$\alpha_0$ : Le terme constant pour l'individu  $i$  ;

$\varepsilon_{i,t}$  : le terme d'erreur.

## 1.2. Présentation des variables de l'étude et leurs sources de données :

A partir de la revue de littérature théorique et empirique, nous avons sélectionné les variables de l'étude suivantes :

❖ **La variable dépendante (expliquée)** est la croissance économique du pays  $i$  à l'instant  $t$ , mesurée par le taux de croissance réel du PIB par habitant.

❖ **Les variables indépendantes (explicatives)** englobent les variables de contrôle, les indicateurs de la libéralisation du compte capital et les mesures institutionnelles.

1) **Les Variables de contrôles** : à partir de la revue de la littérature théorique et empirique, nous avons sélectionné quatre variables de contrôles agissant sur la croissance économique qui sont :

**Le ratio des crédits octroyés au secteur privé en pourcentage du PIB (CP)** : c'est un indicateur de développement financier. Théoriquement cette mesure finance les projets d'investissement à forte valeur ajoutée qui sont bénéfiques à la croissance économique. La littérature récente sur la croissance endogène s'accorde sur un impact positif du développement financier sur la croissance

économique King et Levine (1993) à travers une meilleure allocation des ressources vers les investissements les plus productifs qui sont généralement les investissements privés.

***Le degré d'ouverture commerciale (OUV)*** : correspond à la somme des exportations et des importations par rapport au PIB. L'ouverture commerciale doit précéder l'ouverture financière ; aussi l'ouverture commerciale impacte positivement la croissance économique.

***Le taux d'inflation (INF)*** est calculé par l'indice des prix à la consommation ; il représente l'efficacité des politiques macroéconomiques domestiques d'un pays ; En effet, la libéralisation du compte de capital ne peut avoir un effet positif sur la croissance économique que si le pays en question est doté d'un cadre macroéconomique solide. Le signe attendu est négatif.

***Le taux de scolarisation dans l'enseignement primaire (EDUC)*** : représente le stock du capital humain dans un pays  $i$  sur la période  $t$ , le capital humain a un impact positif sur la croissance économique.

## **2) Les indicateurs de libéralisation du compte de capital :**

Nous avons utilisé deux mesures de la libéralisation du compte de capital à savoir :

***La mesure de jure*** connus sous le nom « KAOPEN » mesuré par Chinn et Ito (2008) et actualisé en (2020). Il a pour objectif de mesurer le degré de restriction sur les opérations financières d'un pays  $i$  à l'instant  $t$  (pour savoir plus sur cet indicateur vous pouvez revenir au chapitre précédent).

***La mesure de facto est développé par*** Lane et Milesi-Ferretti (2007) et actualisé en (2015). Il a pour objectif de mesurer la libéralisation du compte de capital en volume ; elle comprend la somme des stocks d'avoir à l'étranger et d'engagements envers l'étranger rapportés au PIB comme indicateur à long terme. A partir de cette base, nous avons calculé différentes mesures de la libéralisation du compte de capital, en nous intéressant à la direction des flux (entrée ou sortie

de capitaux) et à la composition de ces derniers (IDE, investissements de portefeuilles et dettes). Les IDE mesurent des participations de contrôle acquises dans des entreprises étrangères, en plus des investissements Greenfield (ex : fusions-acquisitions). Le seuil est de 10 % du capital d'une entité. Quant à la mesure des investissements de portefeuille, elle comprend la part du capital des sociétés et fonds communs de placement qui sont en dessous du seuil de 10 % (Il s'agit de titre de participation). Enfin La variable de mesure de la dette comprend les portefeuilles de titres de créances, prêts et dépôts bancaires et autres instruments bancaires.

3) **Les Variables institutionnelles (INST)** représente le niveau de développement institutionnelle dans un pays, tiré de la base des données Internationales Country Risk Guide (ICRG); tel que:

***La corruption (corup)*** : mesure le degré de la corruption dans un pays ; elle indique que les agents de l'état exigent des versements illégaux sous forme de pot de vin afin de rendre un service exemple l'octroi des prêts ou l'octroi des licences d'importation ou d'exportation ....etc. Cet indice oscille entre 0 : gouvernement peu corrompu et 6: gouvernement hautement corrompu. On va inverser l'échelle avec 0 : gouvernement hautement corrompu, et 6 : gouvernement peu corrompu ; en s'inspirant de l'étude (Kunieda, Okada, & Shibata, 2014), afin de faciliter l'interprétation des résultats.

***La Bureaucratie (BURO)*** : mesure la compétence de l'administration et son autonomie par rapport aux changements politiques, ainsi que la qualité de la prestation du service public. Cet indice s'établi sur une échelle de 0 à 4, 0 : signifie que la qualité de la bureaucratie est mauvaise ; tandis que la valeur 4 : signifie que la qualité de la bureaucratie est efficace.

Les données utilisées dans cette étude proviennent d'une combinaison de plusieurs sources, la première source est la base de données world développement

indicateur publié par la Banque mondiale qui permet d'obtenir des statistiques concernant plusieurs variables tel que : le produit intérieur brut par habitant et les variables de contrôle (la ratio crédit accordés au secteur privé en pourcentage du PIB, le taux d'inflation et les statistiques importation et exportation afin de calculer l'indice de l'ouverture commerciale) . La deuxième source et la base de données de Chin et Ito (2020) qui permet d'obtenir la mesure de l'ouverture du compte de capital de jure (KAOPEN), la troisième source et la base de données Lane et Milesi-Ferretti (2015) qui permet d'obtenir la mesure de l'ouverture du compte de capital de facto qui mesure la libéralisation du compte de capital en volume. Enfin, la quatrième source est la base de données de l'International Country Risk Guide (ICRG), qui nous a permis d'extraire les mesures de développement institutionnelle de notre échantillon.

## **2. Techniques d'estimation économétrique**

L'économétrie des données de panel est une forme de régression multiple et les données de panel (ou données longitudinales) sont représentatives d'une double dimension : individuelle et temporelle.

Un panel équilibré (balanced panel) a le même nombre d'observation pour tous les individus, un panel déséquilibré (unbalanced panel) est un panel où il manque des observations pour certains individus.

La recherche de l'existence d'une relation entre plusieurs variables dans son cadre ne se fait qu'après l'identification de l'ensemble des effets que porte le modèle.

Il existe plusieurs méthodes d'estimation des données de panel, le choix est tributaire des hypothèses procédées sur les paramètres et les perturbations. Dans un échantillon de données de panel, la première étape est de distinguer entre la spécification homogène ou hétérogène.

Les avantages de l'analyse sur données de panel Selon Baltagi (2008) sont :

- Les données de panel permettent de contrôler l'hétérogénéité individuelle, car elles suggèrent que les individus, les entreprises et les pays sont tous

hétérogènes. Ainsi, l'analyse permet de prendre en compte la spécificité de chaque entité.

- Les données de panel fournissent plus d'information, plus de variabilité, moins de colinéarité entre les variables, plus de degrés de liberté à plus d'efficience.

- Les données de panel permettent de mieux étudier et de comprendre la dynamique d'ajustement.

- Elles permettent de détecter certains effets qui sont simplement indétectables sur des données en coupe transversale ou des séries temporelles.

- Elles permettent de mieux construire et de tester des comportements inter et intra individuels compliqués que les séries temporelles ne peuvent pas mettre en évidence.

Il existe plusieurs méthodes d'estimation des données de panel. Le choix est tributaire des hypothèses procédées sur les paramètres et sur les perturbations. Lorsque l'on considère un échantillon de données de panel, la toute première chose qu'il convient est de vérifier la spécification homogène ou hétérogène du processus générateur de données. Il s'agit, économétriquement parlant, de tester l'égalité des coefficients du modèle étudié dans la dimension individuelle. Sur le plan économique, les tests de spécification reviennent à déterminer si l'on est en droit de supposer que le modèle théorique étudié est parfaitement identique pour tous les pays, ou au contraire s'il existe des spécificités propres à chaque pays.

## 2.1. Test d'homogénéité des données

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

$$y_{i,t} = \alpha_{0,i} + \beta_i X_{i,t} + \varepsilon_{i,t}$$

$y_{i,t}$ : variable dépendante (endogène) observée pour l'individu  $i$  à l'instant  $t$ ,

$X_{i,t}$  : vecteur des K variables indépendantes (exogène),

( $X_{i,t} = X_{1i,t}, X_{2i,t}, \dots, X_{ki,t}$ ),  $X_{ki,t}$  est la valeur observée pour la  $K^e$  variable indépendante pour l'individu  $i$  à l'instant  $t$ ,

$\beta_i$  : vecteur des K coefficient des K variables exogène  $\beta_i = \beta_{1i}, \beta_{2i}, \dots, \beta_{ki}$ ,

$\alpha_{0,i}$ : terme constant pour l'individu  $i$ ,

$\varepsilon_{i,t}$  : terme d'erreur.

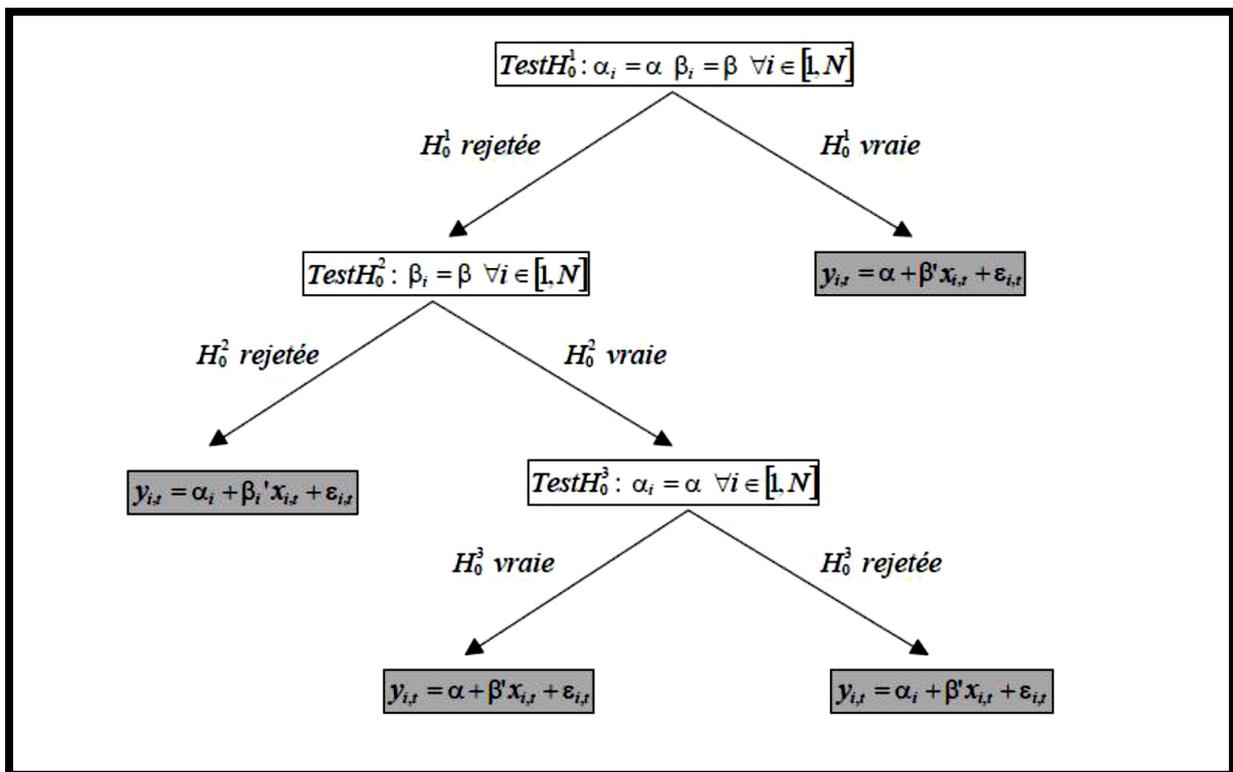
A partir de cette spécification générale, on peut envisager quatre possibilités

- 1- **Homogénéité totale** : Les constantes et les coefficients sont tous identiques pour tous les individus, nous avons pour toutes les valeurs de  $i$ . Le modèle ne comporte qu'une seule équation à estimer sur  $n= N.T$  observation empilées par MCO.
- 2- **Hétérogénéité totale** : Les constantes et les coefficients sont tous différents pour toutes les valeurs de  $i$ , la structure du panel est rejetée. Le modèle doit être estimé équation par équation pour N individus par la MCO.
- 3- **Hétérogénéité des coefficients des variables explicatives et homogénéité des termes constants** : Les constantes sont toutes identiques pour les individus, mais les coefficients des variables explicatives sont différents pour chaque, le modèle doit être estimé équation par équation pour N individus.

**4- Hétérogénéité des termes constants et homogénéité des coefficients des variables explicatives (modèle à effet individuels) :** Les constantes sont différentes pour les individus mais les coefficients des variables explicatives sont constants pour les individus donc on obtient un modèle à effet individuels.

Le choix de la spécification homogénéité ou hétérogénéité est très important, pour cela Hsiao (1986) propose une procédure séquentielle de test permettant de définir le cas dans lequel nous nous situons.

**Figure III.1 : Procédure générale de tests d'homogénéité**



Source : (Hsiao, 1986, p. 50)

**Le test d'homogénéité du processus générateur de données d'Hsiao:**

Avant de passer à l'estimation des coefficients du modèle, il convient d'abord de vérifier la présence des effets individuels dans nos données, et cela à travers les tests de spécification suivants :

- Test d'homogénéité globale :

Dans une première étape, on test l'hypothèse d'une structure parfaitement homogène c'est à dire constantes et coefficients identiques.

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

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

$$F_1 = \frac{SCR_1 - SCR / (N - 1)(K + 1)}{SCR / (N \times T - N(K + 1))}$$

$SCR_1$ : la somme des carrées des résidus du modèle contraint sous l'hypothèse  $H_0^1$ .  
Le degré de liberté est égal à  $(N \times T - (K + 1))$  nombre des coefficients à estimer.

$SCR$ : la somme des carrées des résidus du modèle non contraints, elle est égale à la somme des  $N$  somme des carrés des résidus des modèles estimés sur les  $T$  observations de chaque équation individuelles. Soit :  $SCR_1 = \sum_{i=1}^N SCR_i$ .

Le degré de liberté de chaque équation estimée soit

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

Le degré de liberté du numérateur est donc égale à la différence des degrés de liberté de  $SCR_i$  et  $SCR$ :

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

La statistique  $F_1$  est à comparer à la valeur lue dans la table de Fisher au degré de liberté du numérateur et dénominateur si  $F_1 > F_{ddl_n; dddl}$ , nous rejetons l'hypothèse  $H_0^2$  au seuil  $\alpha\%$ . Dans ce cas on passe à la deuxième étape qui consiste à déterminer si l'hétérogénéité provient des coefficients.

Si on accepte l'hypothèse  $H_0^1$ , on obtient un modèle totalement homogène

$$y_{i,t} = \alpha_{0,i} + \beta_i X_{i,t} + \varepsilon_{i,t}$$

- Test d'homogénéité des coefficients  $\beta_i$  :

La seconde étape, test l'égalité pour tous les individus des K composantes des vecteurs  $\beta_i$ .

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

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

$$F_2 = \frac{SCR_2 - SCR / (N - 1) \times K}{SCR / (N \times T - N(K + 1))}$$

$SCR_2$ : la somme des carrés des résidus du modèle contraint sous  $H_0^2$ . Le degré de liberté est égal à  $(N \times T - N(K + 1))$  nombre des coefficients à estimer, on estime K coefficients et N terme constants, le degré de liberté du numérateur est égal à :

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

La statistique  $F_2$  est à comparer à la valeur lue dans la table de Fisher au degré de liberté du numérateur et dénominateur si  $F_2 > F_{ddl_n, ddd}$ , nous rejetons l'hypothèse  $H_0^2$  au seuil  $\alpha$ ; la structure de panel est rejetée car au mieux seul les constantes peuvent être identiques entre les individus, dont on estime les paramètres vectoriels en utilisant les modèles différents pays par pays.

Si on accepte l'hypothèse  $H_0^2$  c'est-à-dire homogénéité des coefficients, on retient la structure de panel. Finalement, on passe à la troisième étape pour voir si les constantes ont une dimension individuelle.

- Test d'homogénéité des coefficients  $\alpha_i$  :

Cette étape consiste à tester l'égalité des N constante individuelles sous l'hypothèse de coefficients communs à tous les individus.

$$H_0^3: \alpha_i = \alpha \quad \forall i \in [1, N]$$

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

$$F_3 = \frac{SCR_{C1} - SCR_{C2} / (N - 1)}{SCR_{C2} / N \times ((T - 1) - K)}$$

$SCR_1$ : la somme des carrés des résidus du modèle contraint sous l'hypothèse  $H_0^1$ .

$SCR_2$ : la somme des carrés des résidus du modèle contraint sous  $H_0^2$ .

Le degré de liberté du numérateur est donc égale à :

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

La statistique  $F_3$  est à comparer à la valeur lue dans la table de Fisher au degré de liberté du numérateur et dénominateur si  $F_3 > F_{ddl_n, ddl_d}$ , nous rejetons l'hypothèse  $H_0^3$  au seuil  $\alpha$ , donc on obtiens un modèle de panel à effet individuels.

Si on accepte l'hypothèse  $H_0^3$ , on est devant une structure de panel totalement homogène (modèle pooled).

Le test  $H_0^3$  sert à confirmer ou infirmer les conclusions du test  $H_0^1$ .

## 2.2. Etude de stationnarité des variables sur les données de panel

Avant le traitement d'une étude empirique, il convient de tester la stationnarité des variables du modèle afin de déterminer l'ordre d'intégration de nos séries. Les tests de racine unitaire en panel reposent sur deux générations de tests ; la première repose sur l'hypothèse d'indépendance entre les individus, cette hypothèse apparait peu plausible surtout dans les séries macro-économiques ce qui a due à l'apparition d'une deuxième génération, cette dernière repose sur l'hypothèse de dépendance entre les individus.

En va se baser sur les tests les plus utilisés tels que Levin et Lin (LL) et d'IM, PESARAN et SHIN (IPS), ces derniers sont les plus fréquemment utilisé.

De façon générale les tests de racine unitaire en panel sont basés sur le modèle suivant :

$$\Delta y_{i,t} = \alpha_i + \beta_{i,t} + \varphi y_{i,t-1} + \varepsilon_{i,t}$$

$\varepsilon_{i,t}$  les termes d'erreur; ils sont distribués indépendamment entre les individus  $i$  dans les test de racine unitaire de la première génération.

**Figure III.2 : les tests de racine unitaire en panel**

<b>Tests de première génération : Indépendance entre individus</b>
<p><i>1- Spécification homogène de racine autorégressive (<math>H_1</math>) :</i></p> <p>Levin et Lin (1992, 1993) Lin, Levin et Chu (2002) Harris et Tzavalis (1999)</p> <p><i>2- Spécification hétérogène de racine unitaire</i></p> <p>Im, Pesaran et Shin (1997, 2002 et 2003) Maddala et Wu (1999) Choi (1999, 2001) Hadri (2000)</p> <p><i>3- Test séquentiel</i></p> <p>Hénin, Jolivaldt et Nguyen (2001)</p>
<b>Test de deuxième génération : Dépendances entre individus</b>
<p><i>1- Test fondé sur les modèles factoriels</i></p> <p>Bai et Ng (2001) Moon et Perron (2004) Phillips et Sul (2003a) Pesaran (2003) Choi (2002)</p> <p><i>2- Autres approches</i></p> <p>O'Connell (1998) Chang (2002, 2004)</p>

**Source :** (Hurlin & Mignon, 2005, p. 257)

## 1- Le test de Levin et Lin :

Levin et Lin 1992, Levin et Lin 1993 et Levin, Lin et Chu 2002 ont proposé le premier test de racine unitaire en panel ce dernier était inspiré de test de racine unitaire des séries temporelles de Dickey Fuller. (Hurlin & Mignon, 2005, p. 256)

A partir de ces trois modèles, Levin et Lin proposent de tester les hypothèses suivantes :

$$\text{Modèle1 : } \Delta y_{i,t} = \varphi y_{i,t-1} + \varepsilon_{i,t}$$

$$H_0 : \varphi = 0$$

$$H_0 : \varphi < 0$$

$$\text{Modèle2 : } \Delta y_{i,t} = \alpha_i + \varphi y_{i,t-1} + \varepsilon_{i,t}$$

$$H_0 : \varphi = 0 \text{ et } \alpha_i = 0 \quad \forall i = 1, \dots, N$$

$$H_1 : \varphi < 0 \text{ et } \alpha_i \in \mathbb{R} \quad \forall i = 1, \dots, N$$

$$\text{Modèle3 : } \Delta y_{i,t} = \alpha_i + \beta_{i,t} + \varphi y_{i,t-1} + \varepsilon_{i,t}$$

$$H_0 : \varphi = 0 \text{ et } \beta_i = 0 \quad \forall i = 1, \dots, N$$

$$H_1 : \varphi < 0 \text{ et } \beta_i \in \mathbb{R} \quad \forall i = 1, \dots, N$$

Il est important de noter que les hypothèses nulles des tests de Levin et Lin dans les modèles 2 et 3 sont des hypothèses jointes. Dans le modèle 2, l'hypothèse hypothèse nulle testée est l'hypothèse de racine unitaire pour tous les individus du panel.

( $\varphi = \varphi_i = 0$ ) conjointement à l'hypothèse d'absence d'effets individuels, plus précisément à la nullité de toutes les constantes individuelles ( $\alpha_i = 0$ ). Dans le modèle 3, l'hypothèse hypothèse nulle consiste l'hypothèse de racine unitaire et d'absence de composante tendancielle déterministe pour tous les individus du panel ( $\beta_i = 0$ ). On retrouve ainsi exactement la structure des deux tests joints

proposaient dans le cas des séries temporelles par Dickey et Fuller (1981). (Hurlin & Mignon, 2005, p. 258)

Dans le cas général, en présence d'une éventuelle autocorrélation des résidus, le test de Levin et Lin est construit à partir de modèles de type Dickey-Fuller Augmentés (ADF) permettant de blanchir les résidus et de se ramener à des distributions connues pour les t-statistiques individuelles. (Levin A., 1993)

## 2- Le test d'Im, Pesaran et Shin

Une des principales limites du test de Levin et Lin réside dans le caractère homogène de la racine autorégressive  $\varphi_i$  sous l'hypothèse alternative. Il est peu probable en effet qu'en cas de rejet de l'hypothèse de racine unitaire on puisse accepter l'hypothèse d'une racine autorégressive  $\varphi$  commune à tous les individus si on applique des tests usuels de spécification.

Les tests proposés par Im, Pesaran et Shin dans une série de contributions (1997, 2002 et 2003) permettent de répondre à cette critique. En effet, ces auteurs furent les premiers à développer un test autorisant, sous l'hypothèse alternative, non seulement une hétérogénéité de la racine autorégressive  $\varphi_i \neq \varphi_j$ , mais aussi une hétérogénéité quant à la présence d'une racine unitaire dans le panel. (Hurlin & Mignon, 2005, pp. 265-266)

Im, Pesaran et Shin (IPS) considèrent un modèle avec effets individuels et sans tendance déterministe (équivalent du modèle 2 chez Levin et Lin). En absence d'autocorrélation des résidus, ce modèle s'écrit comme suit :

$$\Delta y_{i,t} = \alpha_i + \varphi y_{i,t-1} + \varepsilon_{i,t}$$

Sous:

$$H_0 : \varphi = 0 \quad \text{et} \quad \forall i = 1, \dots, N$$

$$H_1 : \varphi_i < 0 \quad \text{et} \quad \forall i = 1, 2, \dots, N_1$$

$$\varphi_i < 0 \quad \text{et} \quad \forall i = N_1 + 1, \dots, N$$

L'hypothèse alternative coexiste deux types d'individus, les individus indicés  $i=1, \dots, N_1$ , dont la variable  $y_{i,t}$  est stationnaire et les individus indicés  $i= N_1+1, \dots, N$  pour lesquels la dynamique de la variable  $y_{i,t}$  accepte une racine unitaire

Le premier avantage de l'approche d'IPS par rapport à Levin et Lin tient à la prise en compte de l'hétérogénéité de la racine autorégressive sous l'alternative.

Le deuxième avantage, les auteurs proposent une statistique de test très simple fondée sur la moyenne des statistiques de Dickey-Fuller ou de Dickey-Fuller Augmentées individuelles.

Plus précisément, Im, Pesaran et Shin (2003) ont introduit un test basé sur la moyenne des statistiques de Dickey-Fuller calculées pour chaque individu du panel.

Sous l'hypothèse d'absence d'autocorrélation des résidus, IPS dérivent la loi asymptotique de leur statistique moyenne (lorsque T et N convergent vers l'infini) mais aussi la loi semi- asymptotique lorsque T est fixe et N converge vers l'infini. Dans ce cas, il est en effet possible de dériver la loi exacte de la statistique de test de racine unitaire pour une taille T quelconque, contrairement à la statistique de Levin et Lin. IPS proposent même dans ce cas des approximations des seuils de rejet à distance finie pour T et N fixes<sup>1</sup>.

En revanche, sous l'hypothèse d'autocorrélation des résidus, on ne peut plus caractériser la loi exacte de la statistique moyenne pour une taille T donnée : IPS dérivent dans ce cas les lois asymptotiques pour T et N tendant vers l'infini soit de façon séquentielle, soit le long d'une diagonale et proposent deux statistiques moyennes standardisées. On retrouve encore une fois une distribution normale. (Im K.S., 2003, pp. 70-74)

### **3- Le test de Maddala et Wu**

C'est un test non paramétrique de Fisher (1993) son principe repose sur une combinaison des travaux de significativité (P value) de N test individuels de

racine unitaire indépendants. Ce test peut être appliqué même dans le cas de panel non cylindrés. (Christophe Hurlin, 2006, p. 266)

Maddala et Wu (1999) retiennent la statistique de test est définie comme suit :

$$P_{MW} = -2 \sum_{i=1}^N \ln(\rho_i)$$

Si les statistiques individuelles de test sont continués, les p-values sont distribuées selon des lois uniformes sur [0,1] et  $\ln(\rho_i)$  est distribuée selon un  $\chi^2(1)$ ,  $\forall i = 1, \dots, N$ . Maddala et Wu (1999) considèrent le cas où les statistiques individuelles sont indépendantes. Sous cette hypothèse, la statistique  $P_{MW}$  suit un  $\chi^2(2N)$ , sous l'hypothèse hypothèse nulle de racine unitaire, quelle que soit la taille N de l'échantillon. si la réalisation de est supérieure au seuil d'un  $\chi^2(2N)$ , on rejette l'hypothèse hypothèse nulle N de racine unitaire pour les individus du panel. (Maddala & Wu, 1999, pp. 639-645.)

#### 4- Test de Hadri

Le test de (Hadri, 2000) est basé sur l'hypothèse hypothèse nulle de la stationnarité, ce dernier est une extension de test de stationnarité dans le cadre d'économétrie des séries temporelle, il s'agit d'un test multiplicateur de Lagrange visant à tester l'hypothèse hypothèse nulle de la stationnarité des séries. (Hadri, 2000)

Hadri se base sur une équation obtenue à partir de l'équation  $y_{i,t}$  avec constante ou une constante et un trend

$$y_{i,t} = r_{i,t} + \beta_{i,t} + \varepsilon_{i,t}$$

Ou  $r_{i,t}$  est une marche aléatoire :  $r_{i,t} = r_{i,t-1} + \mu_{i,t}$

$\mu_{i,t}$  est un bruit blanc de moyenne nulle de variance  $\sigma_{\Sigma}^2$  sous l'hypothèse

$$\text{hypothèse nulle } H_0 : \sigma_{\Sigma}^2 = 0$$

Si la variance  $\mu_{i,t} = 0$ , alors  $r_{i,t}$  devient constante et par conséquent  $y_{i,t}$  est stationnaire. le test de Hadri (2000) permet la correction de l'hétéroscédasticité. (Hurlin & Mignon, 2005, p. 271)

### **2.3. Etude de cointégration**

Le concept de cointégration peut être définie comme un Co-mouvement systématique à long terme entre deux ou plusieurs variables économiques (Yoo, 2006). Plusieurs tests sont élaborés dans le cadre des panels dont les tests d'absence de cointégration sur données de panel proposés par Pedroni (1995, 1997, 1999, 2004), Kao (1999) sont des tests résiduels analogues aux tests proposés par Engle et Granger (1987) dans le cadre des séries temporelles.

#### **2.3.1. LES TESTS DE PEDROUNI :**

Pedrouni (1995-1997) repose sur l'hypothèse nulle d'absence de cointégration sur les données de panel homogène et hétérogène. Sou l'hypothèse alternative, il existe une relation de cointégration pour chaque individu et que les paramètres de cette relation de cointégration ne sont pas nécessairement les mêmes pour chacun des individuels du panel la mise en considération d'une hétérogénéité constitue un avantage indésirable, car en réalité, il est rare que les vecteurs de cointégrations soient identiques d'un pays à l'autre.

Dans ces conditions, imposer de manière erronée une homogénéité des vecteurs de cointégration aurait pour conséquence un non-rejet de l'hypothèse nulle d'absence de cointégration, alors même que les variables sont cointégrées. (Hurlin & Mignon, 2007, p. 252)

(Pedrouni, 1995) a proposé sept tests visant l'hypothèse nulle, quatre sont basés sur la dimension intra individuels et trois sont basés sur la dimension

interindividuel; les deux catégories du test de pedroni repose sur une hypothèse nulle identique

$H_0: \rho_i = 1$  et  $\forall_i$ ; sachant que  $\rho_i$  désigne le terme autorégressive des résidus estimés tel que :

$$\varepsilon_{i,t} = \rho_i \varepsilon_{i,t-1} + \mu_{i,t}$$

La distinction se situe au niveau de la spécification de l'hypothèse alternative :

$$\rho_i = \rho < 1 \text{ et } \forall_i$$

- Pour les tests basés sur la dimension intra, l'hypothèse alternative s'écrit :

$$\rho_i = \rho < 1 \text{ et } \forall_i$$

- Ainsi, Pour les tests basés sur la dimension inter, l'hypothèse alternative s'écrit :

$$\rho_i < 1 \text{ et } \forall_i$$

### 2.3.2. LE TEST DE KAO

Kao (1999) a également proposé des tests de l'hypothèse nulle d'absence de cointégration : test de type Dickey-Fuller et test de type Dickey-Fuller Augmenté. Contrairement aux tests de Pedroni, Kao considère le cas particulier où les vecteurs de cointégration sont supposés homogènes entre les individus. En d'autres termes, ces tests ne permettent pas de tenir compte de l'hétérogénéité sous l'hypothèse alternative et ne sont par ailleurs valables que pour un système bivarié c'est à dire lors qu'un seul régresseur est présent dans la relation de cointégration. (Hurlin & Mignon, 2007, p. 254)

(Kao, 1999) considère le modèle suivant avec effet individuel :

$$y_{it} = \alpha_i + \beta x_{it} + u_{it} \quad (1)$$

(Kao, 1999) propose deux types de tests, en utilisant les estimateurs Least Squares Dummy Variable (LSDV).

Le premier test est de type Dickey-Fuller appliqué aux résidus  $\hat{u}_{it}$  de l'équation (1)

$$\hat{u}_{it} = \rho \hat{u}_{it-1} + e_{it} / H_0 : \rho=1$$

le second test est de type ADF basé sur la régression suivante :

$$\hat{u}_{it} = \rho \hat{u}_{it-1} + \sum_{j=1}^p \delta_j \Delta \hat{u}_{it-j} + e_{itp} / H_0 : \rho=1$$

P est choisie de sorte que les résidus  $e_{itp}$  ne soient pas autocorrélés. Comme le test de Pedroni, si on ne peut rejeter l'hypothèse nulle alors on conclut que les séries sont non cointégrées. A l'opposée, le rejet de l'hypothèse nulle permet de conclure que les séries sont cointégrées.

En résumé, le test se présente de la façon suivante :

$$H_0: \rho = 1 \text{ (Absence de cointégration)}$$

$$H_1: \rho < 1 \text{ (Présence de cointégration)}$$

## 2.4. Modélisation ARDL sur les données de panel

Pesran et autre (2001) ont développé une nouvelle technique d'estimation pour tester l'existence d'une relation de long terme entre les variables caractérisées par un ordre d'intégration différent; cette technique s'appelle modèle autorégressive a retard échelonné ou modèle ARDL c'est un modèles dynamique qui prend en compte la dynamique temporelle exemple anticipation ou délai d'ajustement comme il nous permet d'analyser aussi bien la relation à court terme qu'à long terme entre la variable expliquée et les variables explicatives.

La méthode repose sur la spécification dynamique de panel du modèle Auto Regressive Distributive Lags (ARDL) de la forme suivante :

$$y_{i,t} = \sum_{j=1}^p \gamma_{ij} y_{i,t-j} + \sum_{j=0}^q \delta_{ij} x_{i,t-j} + U_i + \varepsilon_{i,t}$$

I : le nombre d'individus avec  $i = 1, \dots, N$

t : le nombre de période avec  $T=1, \dots, T$

P : le nombre de retard de la variable dépendante

q : le nombre de retard des variables indépendantes

$X_{it}$  : un vecteur à  $K^* - 1$  variables explicatives

$\delta_{ij}$  : Sont k - 1 vecteur de coefficient

$\gamma_{ij}$  : sont des scalaires

$U_i$  : l'effet spécifique pays

Si les variables de l'équation sont cointégrées, donc ces variables ont une réaction à toute déviation d'équilibre à long terme ce qui nécessite un modèle à correction d'erreur. Dans lequel les dynamiques de court terme des variables de système sont influencées par la déviation d'équilibre ; donc on va ré-paramétriser l'équation précédente dans l'équation à correction d'erreur comme suit :

$$\Delta y_{it} = \phi_i (y_{t-1} - \sigma x_{it}) + \sum_{j=1}^p \gamma_{ij} \Delta y_{i,t-j} + \sum_{j=0}^q \delta_{ij} \Delta x_{i,t-j} + U_i + \varepsilon_{i,t}$$

- Le paramètre  $\phi_i$  et la vitesse de correction des erreurs du terme d'ajustement.
- Si  $\phi_i = 0$ , il n'y a pas de preuve de présence d'une relation de long terme ;
- Si  $\phi_i < 0$ , il n'y a pas une correction d'erreur qui implique que les variables sont cointégrées.

Le paramètre d'ajustement  $\phi$  est censé être significativement négative pour qu'il puisse expliquer que les variables présentent un retour à l'équilibre de long terme. Afin de faire face aux problèmes liés à l'analyse des séries qui ne sont pas intégrées de même ordre. Cette procédure présente plusieurs avantages. Premièrement, la méthodologie de test d'ARDL est applicable indépendamment si les variables explicatives sont stationnaires ou intégrées d'ordre 1. Ainsi, le problème de l'ordre d'intégration associée au test de Johansen (1995) est évité. Deuxièmement, il a de bonnes propriétés par rapport aux petits échantillons. Troisièmement, la méthode ARDL corrige le problème de la corrélation et

d'endogénéité, par une augmentation appropriée de l'ordre des variables explicatives.

Le modèle ARDL en panel est basé sur trois estimateurs : l'estimateur fondé sur la moyenne de groupe (MG : Mean Group), l'estimateur à effet fixe dynamique (DFE : dynamique Fixed Effect) et l'estimateur moyenne de groupe agrégée (PMG : pooled Mean Groupe).

#### **2.4.1. L'estimateur de la moyenne de groupe (MG) :**

La méthode des MG appelle à estimer des régressions séparées pour chaque pays, en calculant les coefficients comme des moyennes non pondérées des coefficients estimés pour chaque pays. Ceci n'impose aucune restriction. Il permet à tous les coefficients de varier et d'être hétérogènes dans le long-terme et le court-terme. Toutefois, la condition nécessaire à la validité et à la cohérence de cette approche est d'avoir une dimension de séries temporelles suffisamment grandes des données.

#### **2.4.2. L'estimateur à effet fixe dynamique (DEF) :**

Blackburne et Frank (2007), l'estimateur des effets fixes dynamiques (DFE) impose l'identité de tous les coefficients de pente et des variances de l'erreur ne tolérant que des différences des effets individuels entre pays. C'est-à-dire les DEF restreint les coefficients du vecteur cointégré à être égaux entre tous les Panels. Également, elle restreint le coefficient de la vitesse d'ajustement et les coefficients de court terme à être égaux. Cependant, elle présente des intercepts pays-spécifiques. Mais, les modèles en effet fixe sont soumis à un biais d'équations simultanées venant de l'endogénéité entre le terme d'erreur et la variable dépendante retardée.

#### **2.4.3. L'estimateur par la moyenne de groupe agrégée :**

Peseran et Smith 1995 proposent la méthode d'estimation moyenne de groupe agrégée (pooled mean group (PMG)) avec le modèle ARDL sur les données de

panel. L'estimateur fondé sur la moyenne de groupe agrégée (PMG) impose les mêmes restrictions à long terme que l'estimateur DFE, il considère que les coefficients de long terme sont uniformes, mais les coefficients de court terme ne sont pas nécessairement, cette méthode nous permet d'établir la distinction entre dynamique à court terme et dynamique à long terme en tenant compte de l'hétérogénéité des pays. Ainsi, elle permet de résoudre les problèmes d'endogénéité des variables dans la spécification dynamique.

La caractéristique principale de la méthode des PMG est qu'elle permet aux coefficients de courts termes, y compris les intercepts, à la vitesse d'ajustement aux valeurs d'équilibre de long terme et aux erreurs de variances à être hétérogènes pays par pays pendant que les coefficients de la pente de long-terme sont restreints à être homogènes entre les pays. L'ajustement de court terme a permis d'être pays-spécifique en raison de l'impact très différent à la vulnérabilité aux crises financières et aux chocs extérieurs, aux politiques de stabilisation, à la politique monétaire, etc (Samargandi et *al.*, 2013).

Certaines conditions sont nécessaires pour que la méthode PMG soit efficace :

- L'existence de relation de long terme entre les variables requiert que les coefficients du terme d'erreur soit négatif.
- Le résidu du modèle à correction d'erreur soit non corrélé et les variables explicatives peuvent être traités comme exogènes.

Pour que ces conditions peuvent être rempli en concluant les retards de la variable dépendante et le retard des variables indépendantes dans le modèle a correction d'erreur.

$$\Delta y_{it} = \phi_i (y_{i,t-1} - \theta_i' X_{it}) + \sum \gamma_{ij} \Delta y_{i,t-j} - \beta_j = 1 + \sum \delta_{ij} \Delta X_{i,t-j} + \mu_i + \epsilon_i$$

Les raisons pour l'utilisation du PMG est une question importante, pour l'estimation et de déterminer s'il faut agréger ou non les données, et donc s'il faut restreindre les paramètres pour les rendre uniformes ( $\lambda = \lambda_i$ ). En effet, l'agrégation

peut conduire à des estimations biaisées et incohérentes lorsque la restriction n'est pas valable c'est-à-dire lorsque les coefficients de pente varient entre les pays. (Pesaran et Smith, 1995).

S'agissant du problème, une solution consiste à estimer une équation pour chaque pays pendant la période considérée puis à calculer la moyenne des coefficients estimés (estimateur fondé sur la méthode dite « Mean Group »). Cependant, cette dernière méthode ne donne de résultats cohérents que si la dimension du panel tend vers l'infini (Pesaran et Smith, 1995). Pour les échantillons de petite taille, une autre option consiste à utiliser un estimateur fondé sur la méthode dite « Pooled Mean Group » (PMG). Dans ce cas, les coefficients de long terme sont uniformes, mais les coefficients de court terme ne le sont pas nécessairement. Cet estimateur permet d'établir une distinction entre dynamique de court terme et dynamique de long terme et tient compte de l'hétérogénéité des pays. Pour résoudre les problèmes d'endogénéité et d'hétérogénéité dans une spécification dynamique, nous avons choisi d'utiliser les estimateurs fondés sur la PMG.

Par conséquent, l'estimateur PMG peut être interprété comme intermédiaire entre les estimateurs DFE et MG car il combine agrégation et moyennisation.

Le choix parmi ces estimateurs consiste en un compromis entre consistance et efficacité. En effet, les estimateurs qui imposent des restrictions sont plus efficaces avec les modèles hétérogènes, pour peu que les restrictions soient valables. Par exemple, si les coefficients de long terme sont identiques d'un pays à l'autre, les estimations PMG seront consistantes et efficaces, tandis que les estimations MG seront consistantes, mais non efficaces. En revanche, si les restrictions de long terme sont imposées à mauvais escient, les estimations PMG ne seront pas consistantes tandis que les estimations MG fourniront des estimations consistantes de la moyenne des coefficients de long terme parmi les pays. En outre, imposer une homogénéité non valable des paramètres dans des modèles dynamiques conduit généralement à sous-estimer la vitesse d'ajustement (Robertson et Symons, 1992; Pesaran et Smith, 1995). Toutefois, la

méthode PMG est suffisamment souple pour permettre d'introduire un sous-groupe de restrictions d'homogénéité à long terme. De plus, par rapport à l'estimateur MG, l'estimateur PMG est notablement moins sensible à l'inclusion de cas extrêmes (outliers) en raison du mécanisme de moyennes pondérées. Pour ces différents motifs, il constitue donc un bon compromis entre consistance et efficacité dans le groupe d'estimateurs considérés.