جامعة سيدي محمد بن عبد الله بفاس
ⵜⴰⵙⴷⴰⵡⵉⵜ ⵙⵉⴷⵉ ⵎⵓⵃⵎⵎⴰⴷ ⴱⵏ ⵄⴱⴷⵍⵍⴰⵀ ⵏ ⴼⴰⵙ
UNIVERSITÉ SIDI MOHAMED BEN ABDELLAH DE FES

كلية العلوم والتقنيات فاس
ⵜⴰⵖⵉⵡⴰⵏⵜ ⵏ ⵜⵎⴰⵙⵙⴰⵏⵉⵏ ⴷ ⵜⵙⵏⵉⴳⵉⵜⵉⵏ
FACULTÉ DES SCIENCES ET TECHNIQUES DE FÈS

Centre d'Etudes Doctorales : Sciences et Techniques de l'Ingénieur

N° d'ordre 09/2020

THESE DE DOCTORAT

Présentée par

Mr : Jamali Alaoui Amine

Spécialité : Probabilités et statistique

Sujet de la thèse : Modélisation et analyse statistique Et stochastique à Mouvement Brownien de Quelques problèmes réguliers: projections économétriques

Thèse présentée et soutenue le 21/01/2021 devant le jury composé de :

Nom Prénom	Titre	Etablissement	
Lahcen OUKHTITE	PES	Faculté des Sciences et Techniques de Fès	Président
Youssef BENGHABRIT	PES	Ecole Nationale Supérieure d'Arts et Métiers de Meknès	Rapporteur
Fatima EZZAKI	PES	Faculté des Sciences et Techniques de Fès	Rapporteur
Mohamed RHOUDAF	PH	Faculté des Sciences de Meknès	Rapporteur
Omar SIDKI	PES	Faculté des Sciences et Techniques de Fès	Examinateur
Abdelmajid HILALI	PES	Faculté des Sciences et Techniques de Fès	Examinateur
EL KHOMSSI Mohammed	PES	Faculté des sciences et techniques de Fès	Directeur de thèse

Laboratoire d'accueil : Modélisation et Structures Mathématiques

Etablissement : Faculté des Sciences et Techniques de Fès

Centre d'Etudes Doctorales : Sciences et Techniques de l'Ingénieur

Résumé de la thèse

Dans cette thèse nous avons développé les méthodes des moindres carrés et les méthodes de la régression régularisée dans le cas d'existence du mouvement brownien dans un modèle, puis nous avons appliqué ces méthodes pour certains problèmes économétriques. Aussi, nous avons proposé et résoudre deux modèles d'optimisation stochastique qui servent à maximiser l'espérance du PIB, en tenant compte des équations dynamiques budgétaires, du rattrapage technologique et de l'équilibre environnemental.

La thèse été composée en deux parties principales. La première partie est constituée de quatre chapitres.

Le premier chapitre est un état d'art, il présente l'évolution de la théorie de la croissance macroéconomique via des modèles mathématiques dynamiques, déterministes et stochastiques.

Dans le deuxième chapitre nous avons développé les méthodes des moindres carrés dans le cas brownien, et nous sommes arrivés à trouver des nouveaux estimateurs efficaces pour estimer un modèle comportant un mouvement brownien, en terminant le chapitre par une application économétrique.

Dans le troisième chapitre nous avons développé les méthodes de la régression régularisée dans le cas brownien, et nous sommes arrivés à trouver des nouveaux estimateurs efficaces pour estimer un modèle comportant un mouvement brownien dans le cas d'existence d'une multicolinéarité entre les variables explicatives, en terminant le chapitre par une application économétrique.

Dans le quatrième chapitre nous avons modifié le test de Fisher pour qu'il soit valable de tester les données en panel.

La deuxième partie est constitue de deux chapitres (chapitres cinq et six).

Dans le cinquième chapitre, nous avons proposé et résoudre un problème d'optimisation stochastique qui sert à maximiser l'espérance du PIB, en tenant compte des équations dynamiques budgétaires et du rattrapage technologique. Dans le sixième chapitre nous avons proposé et résoudre un modèle d'optimisation dynamique qui sert à maximiser le PIB, tout en minimisant la pollution générée par la production.

Mots clés : Estimation- Mouvement brownien- Régularisation- optimisation stochastique

Remerciements

Mes premiers mots de remerciements sont consacrés à mon directeur de thèse, Monsieur le professeur EL KHOMSSI Mohammed qui a orienté mon travail en me proposant des pistes de recherche tout en me laissant une grande liberté dans ma volonté d'explorer certains aspects afin que ce travail voie le jour. Une telle marque de confiance de sa part a stimulé ma motivation est à la base de notre excellente entente durant tout le fil des années de ma thèse. Je le remercie tout particulièrement pour son soutien moral et intellectuel lors des moments difficiles fréquentés par un chercheur pendant l'élaboration de son travail. Je lui suis reconnaissant pour toute l'aide et surtout pour les conseils précieux que vous m'avez apportés durant toutes ces années de thèse.

Mes plus sincères remerciements vont à tous les membres du jury d'avoir accepté d'évaluer ce travail, et de prendre part à cette thèse.

Tout d'abord, je tiens à remercier le Pr. Lahcen OUKHTITE qui m'a fait l'honneur de présider mon jury de thèse.

Je suis très reconnaissant et très honoré que messieurs Pr. Youssef BENGHABRIT, Pr. Fatima EZZAKI et Pr. Mohamed RHOUDAF d'avoir accepté de rapporter ma thèse, et de faire partie de mon jury de soutenance.

Je suis très reconnaissant et très honoré que les professeurs Omar SIDKI et Abdelmajid HILALI aient accepté de faire partie de mon jury, et aient examiné ma thèse.

Mes chaleureux remerciements vont aussi aux membres du Laboratoire de Modélisation et structures mathématiques (LMSM) de la faculté des sciences et techniques de Fès.

Évidemment, ces remerciements ne seraient pas complets sans mentionner ma famille. Un grand merci à mes parents pour m'avoir encouragé à poursuivre mes études doctorales, pour leur amour inestimable, leur confiance, leurs sacrifices et pour leur appui sans faille. Sans oublier mes chers frères et sœurs pour leur soutien et leur encouragement au cours de mes années des études doctorales.

Enfin, je n'oublie pas mes amis et tous mes collègues pour leurs encouragements sympathiques qui ont contribué à ma motivation pour ce travail.

dédicace

À ma très chère MÈRE.
Par les inestimables sacrifices que tu as consentis pour moi, tu as tant souhaité que je parvienne à ce but. Je te serai reconnaissant toute ma vie, qu'Allah t'accorde longue vie dans la santé !

À l'âme de mon très cher PÈRE
Grace à toi j'ai pu aller à l'école. En guise de reconnaissance, trouve ici mon amour filial. Ma réussite est la tienne. Même tu es mort, tu es toujours dans mon coeur.

À ma chère soeur et amie : Wafae
Ma belle soeur, mais aussi ma chère amie. Ta présence à mes cotés m'a toujours donné l'impression d'être proche de toute la famille. Je te remercie pour ton soutien et ton encouragement au cours de mes années des études doctorales. Je te dis merci et je te souhaite bonheur, réussite et prospérité. Je voudrais t'exprimer tout l'amour et toute l'affection que j'ai pour toi. Je t'aime ma soeur !

À ma chère soeur : Khadija
Ma chère soeur, je te remercie pour ton soutien et ton encouragement au cours de mes années des études doctorales. Je te dis merci et je te souhaite bonheur, réussite et prospérité. Je voudrais t'exprimer tout l'amour et toute l'affection que j'ai pour toi.

Au diament de notre famille, la petite : Aya
je te souhaite bonheur, réussite et prospérité, que Dieu te garde et te protège.

À mes chers frères et soeurs , que Dieu le Tout Puissant vous garde et vous procure santé et bonheur.

À ma chère fiancée : Zineb
Ma chère Zineb. Ta présence à mes cotés m'a toujours donné l'impression d'être motivé et courageux. Aucun mot ne pourrait exprimer mon amour et mon respect. Je remercie le dieu qui a croisé nos chemins. Puisse dieu nous procure santé et longue vie.

Table des matières

Introduction générale **10**

I Estimation des modèles comportant un mouvement brownien **16**

1 À propos des modèles déterministes et stochastiques en macroéconomie **17**
1.1 Introduction 17
1.2 Fonctions macroéconomiques et géométrie de la production 19
1.2.1 Généralité et modèles de la fonction de production 19
1.2.2 Propriétés géométriques de la production 21
1.2.3 Méthodes de calcul du PIB et RNB 22
1.3 Modèles dynamiques théoriques simples 24
1.3.1 Le modèle keynésien de Harrod (1939) 24
1.3.2 Le modèle néo-cambridgienne de Kaldor 27
1.3.3 Le modèle néoclassique de solow avec progrès technique 31
1.3.4 Le modèle néoclassique de Cass- Koopmans (1965) 32
1.3.5 Modèle de Romer 35
1.4 Estimation des frontières de production et modélisation de PIB 38
1.4.1 Frontières paramétriques déterministes et modélisation de PIB . . . 38
1.4.2 Frontières paramétriques stochastiques et modélisation de PIB . . . 40
1.4.3 Modélisation non paramétrique du PIB 42
1.4.4 Critiques des méthodes paramétriques et non paramétriques 44
1.5 Modèles dynamiques stochastiques en macroéconomie 45
1.5.1 Généralité sur L'approche Bayésienne et choix des croyances a priori 45
1.5.2 Modèle VAR 48
1.5.3 Modèles DSGE 49
1.5.4 Étude critique et comparative des modèles DSGE- VAR 52
1.6 Conclusion 52

2 Méthodes des moindre carré avec mouvement brownien et application économétrique **53**
2.1 Introduction 53
2.2 Linéarisation du modèle $\mathcal{M}_i$ 54
2.2.1 Méthode 1 54
2.2.2 Méthode 2 55
2.3 Hypothèses, notations et problématique 55
2.3.1 Hypothèses du modèle $(\mathcal{M}_i)$: 55
2.3.2 Notations 56

2.3.3 Problématique . . . 56
2.4 Élimination de l'hétéroscédasticité du mouvement brownien . . . 57
2.4.1 Rappel . . . 57
2.5 Test de la multicolinéarité et test d'autocorrélation . . . 63
2.5.1 Test de la multicolinéarité : Test de Farrar et Glauber modifié . . . 63
2.5.2 Étude de l'autocorrélation des erreurs . . . 64
2.6 Efficience des estimateurs des moindres carrés brownien . . . 66
2.7 Méthode de moindres carrées browniens à erreurs sur les variables . . . 71
2.7.1 Méthode des variables instrumentales avec mouvement browien . . . 71
2.7.2 Méthode de moment généralisé avec mouvement brownien . . . 77
2.8 Sélection du modèle optimal . . . 78
2.9 Simulation des méthodes : Estimation des élasticités des facteurs d'une fonction Cobb dauglass stochastique et modélisation du PIB . . . 80
2.9.1 Algorhithme de calcul de la matrice $\sqrt{\Sigma_T^{-1}}$ . . . 80
2.9.2 Ilustration graphique de dépendence des agrégats avec le PIB . . . 82
2.9.3 Estimation par moindres carrés et moindres carrés brownien des élasticités . . . 84
2.10 Conclusion . . . 87

3 Estimation régularisée et hétéroscédasticité des erreurs 88
3.1 Introduction . . . 88
3.2 Régression régularisée . . . 89
3.2.1 Problématique . . . 89
3.2.2 Norme brownienne et programme de régularisation . . . 90
3.3 Régréssion ridge brownien . . . 92
3.3.1 Régréssion ridge brownien . . . 92
3.3.2 Régréssion ridge brownienne généralisée . . . 97
3.4 Régularisations Elasticnets Browniennes . . . 99
3.4.1 Régularisation Elasticnet Brownienne . . . 99
3.4.2 Régularisation Elasticnet Brownienne Généralisée . . . 100
3.5 Tests de détection d'hétéroscédasticité . . . 101
3.5.1 Test de l'hétéroscédasticité de l'erreur : Test ARCH . . . 103
3.5.2 Test de l'hétéroscédasticité de l'erreur : Test B-ARCH . . . 104
3.5.3 Estimation paramètrique avec hétéroscédasticité supplémentaire . . . 105
3.6 Simulation et application économétrique . . . 106
3.7 conclusion . . . 109

4 Modélisation des données en panel 110
4.1 Introduction . . . 110
4.2 Test d'égalités des variances : VETEHBM . . . 110
4.3 Modèle des données en panel avec mouvement brownien multidimentionnel 111
4.4 Procédure générale de tests des donées en panel . . . 117
4.5 Conclusion . . . 119

II Modèles d'optimisation stochastiques de la croissance économique en tenant compte du rattrappage technologique et de

l'équilibre écologique 120

5 Modèle stochastique de la croissance économique avec technologie hétérogène et convergence du revenu 121
5.1 Introduction . . . 121
5.2 Notations et hypothèses macroéconomiques modèles . . . 122
5.2.1 Notations . . . 122
5.2.2 Hypothèses macroéconomiques du modèle . . . 122
5.3 Construction du modèle . . . 123
5.4 Analyse des équations dynamiques budgétaires . . . 125
5.4.1 Étude de trajectoire de la formation brute du capital en fonction du capital humain . . . 125
5.4.2 Analyse de l'équilibre budgétaire . . . 128
5.5 Principe du maximum stochastique et Résolution du modèle (P) . . . 129
5.6 Condition de rattrapage technologique et simulation du modèle (P) . . . 133
5.6.1 Condition de rattrapage . . . 133
5.6.2 simulation du modèle (P) . . . 134
5.7 Conclusion . . . 136

6 Contrôle optimal de la production économique sous condition de l'équilibre écologique 137
6.1 Introduction . . . 137
6.2 Modélisation macro-dynamique et équilibre écologique . . . 137
6.2.1 Notations et hyphothèses . . . 138
6.2.2 Modèle d'évolution avec équilibre écologique . . . 139
6.3 Modèle macro-dynamique optimale et équilibre écologique . . . 140
6.4 Conclusion . . . 144

Conclusion générale 145

Table des figures

1.1 Équilibre et répartition 29
1.2 Consensus social et équilibre 30
1.3 La “cruche de la veuve” 30
1.4 Diagramme de Solow 32

2.1 matrice de covariance pour 10 périodes 81
2.2 Racine et racine inverse de la matrice de covariance pour 10 périodes . . . 81
2.3 Nuage de points du PIB et de Formation brute du capital (Maroc) 82
2.4 Nuage de points du PIB et de population active occupée (Maroc) 82
2.5 Nuage de points de logarithme du PIB et de logharitme de Formation brute du capital (Maroc) 82
2.6 Nuage de points de logarithme du PIB et de logharitme de population active occupée (Maroc) 82
2.7 Nuage de points après élimination de l'effet de mouvement brownien de logarithme du PIB et de logharitme de Formation brute du capital (Maroc) 83
2.8 Nuage de points après élimination de l'effet de mouvement brownien de logarithme du PIB et de logharitme de population active occupée (Maroc) 83
2.9 Nuage de points du PIB et de Formation brute du capital (Royaume Uni) . 83
2.10 Nuage de points du PIB et de population active occupée (Royaume Uni) . 83
2.11 Nuage de points de logarithme du PIB et de logharitme de Formation brute du capital (Royaume Uni) 84
2.12 Nuage de points de logarithme du PIB et de logharitme de population active occupée (Royaume Uni) 84
2.13 Nuage de points après élimination de l'effet de mouvement brownien de logarithme du PIB et de logharitme de Formation brute du capital (Royaume Uni) 84
2.14 Nuage de points après élimination de l'effet de mouvement brownien de logarithme du PIB et de logharitme de population active occupée (Royaume Uni) 84

3.1 Exemple d'un modèle homoscédastique et hétéroscédastique 102

5.1 Trajectoire de la formation brute du capital en fonction du capital humain 127
5.2 Trajectoire de la formation brute du capital en fonction du capital humain 127
5.3 Trajectoire de la formation brute du capital en fonction du capital humain 128
5.4 Investissement optimal pour la production 135
5.5 Investissement optimal pour la recherche et développement 135
5.6 Investissement optimal pour le capital humain 136

Liste des tableaux

1.1 avantages et les inconvénients des méthodes paramétriques et non paramétriques 45

2.1 Estimation de paramètres du modèle $\mathfrak{m}_1$ par l'estimateur β^{BOLS} 85
2.2 Estimation de paramètres du modèle $\mathfrak{m}_2$ par l'utilisation de l'estimateur β^{OLS} 86

3.1 Estimations par moindres carrés brownien : modèle $\mathfrak{M}_1$ (Maroc) 89
3.2 Estimations par moindres carrés : modèle $\mathfrak{M}_2$ (Maroc) 89
3.3 Estimation du modèle $\mathfrak{M}_1$ par β^{BRE} 107
3.4 Estimation du modèle $\mathfrak{M}_1$ par β^{B2E} 107

Résumé

Dans cette thèse nous avons développé les méthodes des moindres carrés et les méthodes de la régression régularisée dans le cas d'existence du mouvement brownien dans un modèle, puis nous avons appliqué ces méthodes pour certains problèmes économétriques. Aussi, nous avons proposé et résoudre deux modèles d'optimisation stochastique qui servent à maximiser l'espérance du PIB, en tenant compte des équations dynamiques budgétaires, du rattrapage technologique et de l'équilibre environnemental.
La thèse été composée en deux parties principales. La première partie est constituée de quatre chapitres. Le premier chapitre est un état d'art, il présente l'évolution de la théorie de la croissance macroéconomique via des modèles mathématiques dynamiques, déterministes et stochastiques.
Dans le deuxième chapitre nous avons développé les méthodes des moindres carrés dans le cas brownien, et nous sommes arrivés à trouver des nouveaux estimateurs efficaces pour estimer un modèle comportant un mouvement brownien, en terminant le chapitre par une application économétrique.
Dans le troisième chapitre nous avons développé les méthodes de la régression régularisée dans le cas brownien, et nous sommes arrivés à trouver des nouveaux estimateurs efficaces pour estimer un modèle comportant un mouvement brownien dans le cas d'existence d'une multicolinéarité entre les variables explicatives, en terminant le chapitre par une application économétrique.
Dans le quatrième chapitre nous avons modifié le test de Fisher pour qu'il soit valable de tester les données en panel.
La deuxième partie est constitue de deux chapitres (chapitres cinq et six).
Dans le cinquième chapitre, nous avons proposé et résoudre un problème d'optimisation stochastique qui sert à maximiser l'espérance du PIB, en tenant compte des équations dynamiques budgétaires et du rattrapage technologique. Dans le sixième chapitre nous avons proposé et résoudre un modèle d'optimisation dynamique qui sert à maximiser le PIB, tout en minimisant la pollution générée par la production.

Mots clés : Estimation- Mouvement brownien- Régularisation- optimisation stochastique

Abstract

In this thesis we developed least squares methods and regularized regression methods in the case of the existence of Brownian motion in a model, and then we applied these methods for some econometric problems. Also, we proposed and solved two stochastic optimization models that serve to maximize the GDP expectation, taking into account the dynamic fiscal equations, technological catch-up and environmental equilibrium.
This thesis was composed in two main parts. The first part consists of four chapters. The first chapter is a state of the art, it presents the evolution of the theory of macroeconomic growth via dynamic, deterministic and stochastic mathematical models.
In the second chapter we developed the least squares methods in the Brownian case, and we managed to find new efficient estimators for estimating a model with Brownian motion, ending the chapter with an econometric application.
In the third chapter we developed the methods of regularized regression in the Brownian case, and we were able to find new efficient estimators to estimate a model with Brownian motion in the case of the existence of multicolinearity between the explanatory variables, ending the chapter with an econometric application.
In the fourth chapter we have modified the Fisher test so that it is valid to test the data in a panel.
The second part consists of two chapters (chapters five and six).
In the fifth chapter, we proposed and solved a stochastic optimization problem that serves to maximize GDP expectations, taking into account dynamic fiscal equations and technological catch-up. In the sixth chapter we proposed and solved a dynamic optimization model that serves to maximize GDP, while minimizing the pollution generated by production.

Key words : Estimation- Brownian motion- Regularisation- Stochastic optimisation

Introduction générale

Introduction

Dès qu'un phénomène de divers natures : financier, économétrique, biologique ou autre, est très complexe ou trop bruité, l'utilisation d'une description analytique débouchant sur une modélisation déterministe sera inéfficace, pour décrire et prédire l'évolution de ce phénomène dans le future.

La modélisation statistique est l'art de formaliser et analyser mathématiquement un phénomène réel pour répondre à des problèmes liés à ce dernier. Ainsi elle nous permet à partir des observations, de prédire aproximativement les évolutions de certaines variables, de construire des conclusions à partir de l'analyse des résultats obtenus et finalement prendre des décisions.

Récemment, la modélisation statistique est indispensables pour l'aide à la décision, que ça soit ; la politique monétaire, la stabilité financière ou la surveillance bancaire, etc...

L'étude statistique d'un tel phénomène commence par l'estimation des paramètres liées à ce dernier. Puis on peut étudier un autre problème lié au même phénomène ; tel que le contrôle de la trajectoire de certaines variables d'états via d'autres variables de contrôles, afin d'optimiser l'espérance d'une fonction objectif, on parle donc de l'optimisation stochastique.

Parmi les modèles de l'optimisation stochastique, on peut trouver des modèles à mouvement brownien, qui modélisent des phénomènes contenant un mouvement interne aléatoire de ces petites particules (ou bien des variables explicatives), qui engendre des chocs entre les petites particules et la grosse particule (ou bien la variable endogène), et il en résulte un mouvement très irrégulier de la grosse particule (évolution irrégulière de la variable endogène).

Dans cette thèse, nous nous intéressons à l'étude des modèles qui contiennent un mouvement brownien, dans les deux dimensions ; estimations des paramètres sous l'existence du mouvement brownien et optimisation stochastique de l'espérance de la variable endogène :

- Estimation des paramètres sous l'existence du mouvement brownien : on considère que les perturbations des évolutions des variables explicatives dans le temps, peut provoquer des chocs sur l'évolution de la variable endogène, et on estime paramétriquement les coefficients du modèle, en tenant compte de l'existence de la composante browniene dans le modèle.
- Optimisation stochastique de l'espérance de la variable endogène : en prenant en considération les estimations paramétriques de la première dimension, ainsi on contrôle les trajectoires des variables explicatives d'états via des variables explicatives de contrôle, tout en optimisant l'espérance de la variable endogène sur une gamme future.

La thèse est composée de deux parties :

La première partie consiste après une étude bibliographique et critique de certains mo-

dèles économétriques existants dans la litérature, à étudier le comportement du mouvement brownien, et de modifier certains tests statistiques, pour qu'ils soient compatibles avec l'existence du mouvement brownien dans le modèle, tout en appliquant les résultats théoriques obtenus à certains phénomènes économétriques.
Dans la deuxième partie, on propose deux modèles de l'optimisation stochastique, qui servent à maximiser l'évolution de la variable endogène en tenant compte de l'évolution technologique et de l'équilibre environemental, tout en se basant sur les estimations de la première partie.
Première partie **:**
Elle comporte quatre chapitres ; le premier est un état d'art et une étude critique de certains modèles liés à l'économie, et les trois chapitres qui suivent, portant sur les estimations des paramètres d'un modèle comportant un mouvement brownien.
Motivation **:**
Depuis le début du 20ème siècle, et sous l'effet des développements successifs de nouvelles techniques tant matérielles que formelles, ainsi que l'insuffisance des modèles detérministes d'expliquer plusieurs phénomènes, surtout dans le cas d'existence des chocs dans les évolutions de certaines variables dans le modèle, le développement des modèles statistiques permet de plus en plus d'expliquer et de prédire aproximativement et avec un niveau de confiance fixé, divers phénomènes économétriques, financiers, biologiques etc... , en tenant compte des aléas stochastiques imprévisibles.
Dans la littérature, on trouve plusieurs méthodes et modèles statistiques, avec divers versions. Citons par exemples, les méthodes des moindres carré élaborées par Legendre et Gauss [1-3] au début du 20ème siècle, elles permettent de comparer des observations entachées d'erreurs de mesure, à un modèle mathématique qui sert à décrire ces données. Les modèles à la base des séries chronologiques [4-8], les modèles basés sur les régression régularisée, comme les modèles LASSO[9], etc... .
Le problème qui existe dans ces modèles et méthodes et qu'ils ne tiennent pas en compte l'existence d'un mouvement brownien. Or, en réalité, on peut rencontrer plusieurs phénomènes qui peuvent contenir des chocs des variables explicatives sur la variable endogène, et qui provoquent des évolutions irrégulières de la variable endogène, par exemple les chocs conjoncturels qui désignent des événement ponctuels modifiant sensiblement la situation économique (comme les chocs monétaires et les chocs d'offre).
Nous nous sommes donc intéressés dans cette première partie, après un états d'art, au modèles qui contiennent des processus de Wiener (brownien), en présentant, nos méthodes d'estimations que nous proposons selon des tests statistiques spécifiques.
Problématique **:** Les estimateurs des moindres carré ne convergent pas vers la vraie valeur des coefficients, et ils sont instables.
Même les régressions ridges, ne seront pas optimales et de variances non minimale.
Les tests de l'hétéroscédasticité et de l'autocorrélation des erreurs ne seront pas utiles dans le cas d'existence de mouvement brownien.
Nous visons donc à répondre à deux objectifs :

- Modifier les tests pour qu'ils soient compatibles avec l'existence du mouvement brownien dans le modèle.
- Trouver des estimateurs optimaux convergents, qui servent à estimer les paramètres dans le modèle et ceci dans le cas d'existence du mouvement brownien.

La première partie est constituée de quatre chapitres.
Dans le premier chapitre, nous présentons une revue de la littérature des modèles déterministes et stochastiques liées à l'économie, ces modèles servent à l'étude de la croissance

économique, et expliquent l'évolution économique dans les pays qui aspirent au bien-être, et que ce dernier n'est pas envisageable sans la croissance. Les pays présentant les meilleurs indicateurs de bien-être et de développement sont ceux qui présentent les meilleures performances en termes de croissance.
François Perroux (voir[10]) définit la croissance économique comme un processus continu et soutenu d'élévation du PIB réel d'un pays dans le temps. Ceci revient à dire que la croissance économique n'est pas un fait du hasard mais plutôt une oeuvre soutenue et qu'elle se traduit par un accroissement continu du produit réel de toute l'économie. Il est important de noter que si d'une année à une autre, le produit de l'économie s'accroît à un taux inférieur à l'accroissement de la population active, alors la croissance réalisée n'entraînera pas un accroissement du revenu par habitant. Il serait donc préférable de la définir comme un processus d'élévation continu et soutenu du PIB par habitant du pays. L'étude de la croissance économique est justifiée par la nécessité de comprendre comment par un ensemble d'actions harmonieuses et concertées, une nation peut améliorer de manière soutenue, le niveau de vie de sa population ou se libérer de l'ornière de la pauvreté. Une identification des déterminants de la croissance et une bonne compréhension de leurs interactions devraient permettre de bien concevoir une politique de développement.
La macroéconomie est l'approche théorique qui étudie l'économie à travers les relations existantes entre les grands agrégats économiques, le revenu, l'investissement, la consommation, le taux de chômage, l'inflation, etc.
Les modèles présentés dans le premier chapitres, sont classés généralement en trois catégories ; modèles théoriques simples, modèles de prévision empiriques et modèles macroéconomiques stochastiques.
Les Modèles théoriques simples, sont basés sur quelques équations comportant quelques variables, qui peuvent être expliquées avec de simples schémas. Les variables qui apparaissent dans ces modèles représentent souvent des agrégats macro-économiques (tels que le PIB, le total de l'emploi,..) plutôt que des variables de choix individuel et bien que les équations reliant ces variables décrivent des décisions économiques, elles ne sont généralement pas dérivées directement de modèles individuels. Ils sont assez simples pour être utilisés comme illustrations de points théoriques dans les explications introductives des idées macroéconomiques ; Par conséquent, l'application quantitative à la prévision, aux essais ou à l'évaluation des politiques est habituellement impossible sans augmenter substantiellement la structure du modèle. Nous citons dans cette catégorie, les modèles de ; Harod [14] Solow [15], Kaldor [16], Cass- Koopmans [17], Ramsy [18], Romer [19], Aghion- Howitt [20], etc ...
Les modèles de prévision empiriques estiment les relations entre différentes variables macroéconomiques à l'aide d'une analyse de séries temporelles (principalement linéaire).Comme les modèles théoriques plus simples, ces modèles empiriques décrivaient les relations entre les quantités agrégées, mais trop traitaient d'un niveau de détail plus fin (par exemple, étudier les relations entre la production, l'emploi, l'investissement et d'autres variables). Ainsi, ces modèles ont augmenté pour inclure des centaines ou des milliers d'équations décrivant l'évolution de plusieurs prix et de quantités au fil du temps, rendant les ordinateurs essentiels pour leur solution. Alors que le choix des variables à inclure dans chaque équation était en partie guidé par la théorie économique (par exemple, le revenu passé comme déterminant de la consommation, comme suggéré par la théorie des anticipations adaptatives), l'inclusion variable était principalement déterminée empiriquement. Nous citons dans cette catégorie, les modèles de, Jan Tinbergen [21], Kein [21], Milton Friedman [22], Edmund Phelps [23] etc...

Les modèles macroéconomiques stochastiques, ce sont des modèles plus proches de la réalité économique, que les modèles théoriques simples malgré leurs complexités. En effet, ils prennent en considérations le fait aléatoire de plusieurs agrégats économiques (par exemple : le progès technologique et l'emploi), nous citons dans cette catégorie, les modèles VAR [24],EGC [25] DSGE [26], DSGE pour le B du VAR [27], etc...
On termine chaque section du chapitre 1, par une étude critique des modèles présentés.
Ce chapitre est organisé comme suit :
Il est constitué de quatre sections, après l'introduction. Dans la première section, nous donnons des notions basiques sur les agrégats macroéconomiques et la géométrie de la production. Dans la seconde section, nous abordons les résultats de quatre modèles théoriques simple importants, dans la troisième section, nous présentons les méthodes paramétriques et non paramétriques économétrique, et nous terminons ce chapitre par un regard bayésien sur les modèles dynamiques stochastiques de la macroéconomie.
Dans le deuxième chapitre nous nous intéréssons à l'étude du modèle suivant :

$$(\mathcal{M}_i) \quad Y_{i,t} = F_i(X_{i,t}, \beta_i, u_{i,t}) + e_{i,t}$$

Où :
$Y_{i,t}$ est la variable endogène que l'on veut expliquer par les variables explicatives.
$X_{i,t} = (x_{i,1,t}, ..., x_{i,m,t})$ à l'instant t
$\beta_i \in \mathbb{R}^{m+1}$ est le vecteur de $m+1$ paramètres à estimer.
$u_i = \left(u_{i,t}\right)_{t\geq 0}$ est un mouvement brownien unidimensionnel
$F_i : \mathbb{R}^m \times \mathbb{R}^{m+1} \times \mathcal{E} \to \mathbb{R}$ est une application différentiable par rapport à β_i et $x_{i,k,t}$, $k \in [1, m] \cap \mathbb{N}$, avec $\mathcal{E}$ est l'ensemble des applications de l'espace des fonctions continues $\mathcal{C}(\mathbf{IR}^+, \mathbf{IR})$ dans l'espace probabilisé $(\Omega, \mathfrak{F}, P_W)$ et P_W est la mesure de Wiener.
$e_{i,t}$: l'erreur de modélisation.
Ce chapitre est constitué de huit sections. Notre idée ici, est d'éliminer l'hétéroscédasticité du processus brownien sur une gamme du temps fini, puis nous supposons que l'élimination de l'hétéroscédasticité du processus brownien, entrainera une élimination de l'hétéroscédasticité de l'erreur dans l'erreur du modèle.
Nous modifions la méthode de moindres carré existante dans la litérature, ainsi que le test de l'autocorrélation des erreurs, "Test de Breusch-Godfrey avec multiplicateur de lagrange", nous obtenons des nouveaux estimateurs efficaces du modèle $(\mathcal{M}_i)$.
Le troisième chapitre, est un prolongement du chapitre deux. Il est constitué de cinq sections. Dans ce chapitre on rejette l'hypothèse d'homoscédasticité des erreurs après l'élimination de l'hétéroscédasticité du processus brownien, ainsi que l'hypothèse de l'absence d'une multicolinéarité dans les variables explicatives. Nous commençons le chapitre par l'étude de la régression ridge dans le cas d'existence de multicolinéarité et du processus brownien dans le modèle. Puis nous présentons une version du modèle GARCH, qui prend en compte le fait d'existence du mouvement brownien dans le modèle.
Nous terminons la première partie, par le quatrième chapitre, on s'intérèsse à l'étude des modèles contenants plusieurs individus (variables endogènes) et plusieurs variables explicatives. avec un mouvement brownien multidimentionnel.
Ce chapitre est composé en trois section ; dans la première section on étudie l'hypothèse d'égalité de la variance entre les individus. Dans la deuxième section nous nous intéréssons à trouver des estimateurs optimaux des modèles en panel, en généralisant les estimateurs qu'on a abordé dans les chapitres 2 et 3. Dans la troisième section nous abordant le test des données en panel, pour trouver la catégorie des modèles modélisant l'ensembles des

individus en fonction des variables explicatives.
Deuxième partie **:**
Elle comporte deux chapitres. Dans le cinquième chapitre, nous présentons un modèle d'optimisation stochastique appliqué à l'évolution de la croissance économique avec technologie hétèrogène. et dans le sixième chapitre nous présentons un modèle d'optimisation stochastique de la maximisation de la production industrielle tout en minimisant la polution en tenant compte de l'équilibre écologique.
Motivation **:**
L'optimisation stochastique est une branche des mathématiques cherchant à modéliser, à analyser et à résoudre analytiquement ou numériquement les problèmes qui consistent à minimiser ou maximiser une fonction contenant un ou des processus stochastiques sur un ensemble donné.
L'optimisation stochastique joue un rôle important dans divers domaines, tels que ; la finance l'économie incertaine, la physique, etc.... dans les mathématiques appliquées (fondamentales pour l'industrie et l'ingénierie), en analyse et en analyse numérique, en statistique pour l'estimation du maximum de vraisemblance d'une distribution, pour la recherche de stratégies dans le cadre de la théorie du contrôle et de la commande stochastique.

Plusieurs systèmes susceptibles d'être décrits par un modèle mathématique sont optimisés. La qualité des résultats et des prédictions dépend de la pertinence du modèle, du bon choix des variables que l'on cherche à optimiser, de l'efficacité de l'algorithme et des moyens pour le traitement numérique.
Dans les chapitres de la première partie, nous nous intéréssons à l'estimation statistique, et nous arrivons à trouver le modèle qui caractérise bien l'évolution du PIB dans le temps. Pour cette raison nous nous sommes donc intéressés dans la deuxième partie, à maximiser le PIB, sur une gamme du temps en tenant compte de l'impact de la technologie puis de l'équilibre environnementale sur la production
Problématique **:**
-La maximisation du PIB, consiste à améliorer le niveau technologique du pays, pour produire plus.
-Tant qu'on produit plus, plus que la polution augemente, ce qui implique un déséquilibre environnementale qui touche notre santé.
Nous nous visons donc à répondre à deux objectifs :

- Proposer un modèle d'optimisation stochastique qui sert à maximiser le PIB, au cours du temps, en essayant d'améliorer la frontière technologique du pays.
- Proposer un modèle d'optimisation stochastique qui sert à maximiser le PIB, mais cette fois en gardant un niveau minimal de la polution. Nous proposons de réserver une partie de l'investissement à la remédiation de la polution générée par la production

Contenu des chapitres de la deuxième partie **:**

Dans cinquième chapitre, on propose un modèle stochastique de la croissance économique en tenant en compte la différence des niveaux technologiques entre les pays. Nos objectifs dans ce chapitre sont :
Optimiser le PIB par emploi, en tenant compte de l'équilibre budgetaire et des chocs stochastiques externes sur l'emploi et la frontière technologique de chaque pays.
Établir l'importance de la qualification de la population active occupée et l'insertion du secteur de recherche et développement dans la production pour améliorer le revenu par

habitant des citoyens et par suite leurs niveau de vie.
pour attiendre ces deux objectifs, nous organisons ce chapitre comme suit :
Après la section d'introduction, le chapitre contient trois sections et se termine par une conclusion.
Dans la première section ; nous présentons les notations et les hypothèses macroéconomiques de la modélisation de notre problème.
La deuxième section est consacrée à l'élaboration et la proposition du modèle, qui sert à répondre aux deux objectifs fixés au dessus. La troisième section est consacrée à l'analyse et à la résolution du modèle.
Dans le sixième chapitre, nous nous intéréssons à l'étude de la croissance économique sous condition de l'équilibre écologique, nous construisons deux modèles dynamiques qui servent à maximiser le PIB par tête, en tenant compte de l'équilibre écologique. Nous considérerons que la production est divisée en deux ; production de base et production auxiliaire détruisant la pollution générée par la production de base. Puis on contrôle la trajectoire de la formation brute du capital par l'investissement.
Les travaux de cette thèse ont fait l'objet de trois publications [11-13], et deux articles soumis :

♦ **Publications :**

- Jamali Alaoui, A., El khomssi, M., El goumi, B., (2018). ''Stochastic model of economic growth with heterogeneous technology and technological upgrading". Lecture Notes in Networks and Systems- Springer. pp. 1–9, 2018. https ://doi.org/10.1007/978-3-319-74500-8_90.
- Jamali Alaoui, A., El khomssi, M., El goumi, B., (2019). ''Optimal control model of economic production under condition of ecological balance". AIP Conference Proceedings 2074, 020006 (2019) ; https ://doi.org/10.1063/1.5090623.
- Jamali Alaoui, A., El khomssi, M., El goumi, B., (2019). ''Dynamic Optimization Model of the Economic Behavior of Pension Fund". IEEE, 10.1109/ICOA.2019.8727677.

♦ **Articles soumis :**

- Jamali Alaoui, A., El khomssi, M., El goumi, B., "Deterministic and stochastic mathematical models in macroeconomics".
- Jamali Alaoui, A., El khomssi, M., El goumi, B., "Brownian regulation model and econometric applications".

Première partie

Estimation des modèles comportant un mouvement brownien

Chapitre 1

À propos des modèles déterministes et stochastiques en macroéconomie

1.1 Introduction

Il est essentiel de comprendre les principaux mécanismes qui sont sources de la croissance économique, car des faibles écarts de taux de croissance peuvent correspondre à une différence qualitative forte sur la vie des générations futures.
La croissance économique est considérée par certains comme l'un des phénomènes les plus passionnants de la macroéconomie parce que tous les pays aspirent au bien-être, et que ce dernier n'est pas envisageable sans la croissance. Les pays présentant les meilleurs indicateurs de bien-être et de développement sont ceux qui présentent les meilleures performances en termes de croissance.
François Perroux (voir[10]) définit la croissance économique comme un processus continu et soutenu d'élévation du PIB [1] réel d'un pays dans le temps. Ceci revient à dire que la croissance économique n'est pas un fait du hasard mais plutôt une oeuvre soutenue et qu'elle se traduit par un accroissement continu du produit réel de toute l'économie. Il importe de noter que si d'une année à une autre, le produit de l'économie s'accroît à un taux inférieur à l'accroissement de la population, la croissance ainsi réalisée n'entraînera pas un accroissement du revenu par habitant. Il serait donc préférable de la définir comme un processus d'élévation continu et soutenu du PIB par habitant du pays.
L'étude de la croissance économique est justifiée par la nécessité de comprendre comment par un ensemble d'actions harmonieuses et concertées, une nation peut améliorer de manière soutenue, le niveau de vie de sa population ou se libérer de l'ornière de la pauvreté. Une identification des déterminants de la croissance et une bonne compréhension de leurs interactions devraient permettre de bien concevoir une politique de développement.
La macroéconomie est l'approche théorique qui étudie l'économie à travers les relations existantes entre les grands agrégats économiques, le revenu, l'investissement, la consommation, le taux de chômage, l'inflation, etc.
La macroéconomie constitue l'outil essentiel d'analyse des politiques économiques des

1. Le PIB est l'abréviation du "Produit intérieure Brute", il représente aussi le revenu total d'un tel états. Généralement en distingue entre deux types du PIB ; PIB réel (ou PIB en volume) et le PIB nominale (ou PIB valeur). Le PIB réel est mesuré à prix constants, autrement dit, le PIB réel d'une année $t_0 + t$ est calculé à partir des prix en vigueur au cours d'une année de référence t_0, et des volumes de l'année $t_0 + t$. Le PIB nominale est mesuré à prix courant (c-à-d, en tenant compte de l'inflation). Généralement, le PIB réel est utilisé pour mesurer la croissance du PIB d'une année à l'autre, par contre, pour les données de long terme, le PIB nominal restant la mesure de référence.

États ou des organisations internationales. Il s'agit d'expliquer les mécanismes par lesquels sont produites les richesses à travers le cycle de la production, de la consommation, et de la répartition des revenus au niveau national ou international. Pour cette raison, les économistes chercheurs créent des modèles, pour analyser et expliquer les principaux mécanismes résponsables de la croissance macroéconomique, puis de prévoir cette croissance dans le futur afin d'en tirer des conséquences pour des prises de décision.
Un modèle macroéconomique est un outil d'analyse destiné à décrire le fonctionnement de l'économie d'un pays ou d'une région. un tel modèle macroéconomique est généralement conçu pour étudier la dynamique de quantités globales telles que la quantité totale de biens et services produits, le total des revenus, le niveau d'emploi des ressources productives, et le niveau des prix.
Dans la littérature, il y a plusieurs modèles macroéconomiques, qui sont classés généralement en trois catégorie ; modèles théoriques simples, modèles de prévision empiriques et les modèles macroéconomiques stochastiques.
Les Modèles théoriques simples, sont basés sur quelques équations comportant quelques variables, qui peuvent être expliqué avec de simples schémas. Les variables qui apparaissent dans ces modèles représentent souvent des agrégats macro-économiques (tels que le PIB, le total de l'emploi,..) plutôt que des variables de choix individuelles et bien que les équations reliant ces variables décrivent des décisions économiques, elles ne sont généralement pas dérivées directement de modèles individuels. Ils sont assez simples pour être utilisés comme illustrations de points théoriques dans les explications introductives des idées macroéconomiques ; Par conséquent, l'application quantitative à la prévision, aux essais ou à l'évaluation des politiques est habituellement impossible sans augmenter substantiellement la structure du modèle. Nous citons dans cette catégorie, les modèles de ; Harod, Solow , Kaldor, Cass- Koopmans, Ramsy, Romer, Aghion- Howitt, etc ...
Les modèles de prévision empiriques estiment les relations entre différentes variables macroéconomiques à l'aide d'une analyse de séries temporelles (principalement linéaire). Comme les modèles théoriques plus simples, ces modèles empiriques décrivaient les relations entre les quantités agrégées, mais beaucoup traitaient d'un niveau de détail beaucoup plus fin (par exemple, étudier les relations entre la production, l'emploi, l'investissement et d'autres variables). Ainsi, ces modèles ont augmenté pour inclure des centaines ou des milliers d'équations décrivant l'évolution de centaines ou de milliers de prix et de quantités au fil du temps, rendant les ordinateurs essentiels pour leur solution. Alors que le choix des variables à inclure dans chaque équation était en partie guidé par la théorie économique (par exemple, le revenu passé comme déterminant de la consommation, comme suggéré par la théorie des anticipations adaptatives), l'inclusion variable était principalement déterminée empiriquement. Nous citons dans cette catégorie, les modèles de, Jan Tinbergen, Klein, Milton Friedman, Edmund Phelps etc...
Les modèles macroéconomiques stochastiques, ce sont des modèles plus proches de la réalité économique, que les modèles théoriques simples, malgré ses complexités. En effet, ils prennent en considérations le fait aléatoire de plusieurs agrégats économiques (par exemple : le progès technologique et l'emploi), Nous citons dans cette catégorie, les modèles VAR,EGC, DSGE, DSGE pour le B du VAR, etc...
En se basant sur les travaux antérieures, l'objectif de ce chapitre est de présenter un état d'art, autour des modèles mathématiques de la croissance économique, en essayant de donner une référence pour les nouveaux chercheurs dans ce sujet.
Ce chapitre est organisé comme suite :
Il est constitue de quatre section. Dans la première, nous donnons des notions basiques

sur les agrégats macroéconomiques et la géométrie de la production. Dans la seconde section, nous abordons les résultats de quatre modèles théoriques simple importants, dans la troisième section, nous présentons les méthodes paramétriques et non paramétriques économétrique, et nous terminerons ce chapitre par un regard bayésien sur les modèles dynamiques stochastiques de la macroéconomie.

1.2 Fonctions macroéconomiques et géométrie de la production

1.2.1 Généralité et modèles de la fonction de production

a- Généralités sur la fonction de production

Une fonction de production est une application positive de $\mathbf{IR}^n \times \mathbf{IR} \to \mathbf{IR}^+$, qui reliée à un instant $t \in \mathbf{IR}$ donné, chaque vecteur d'inputs (ou vecteur de variable explicatives) $x = (x_1, ..., x_n) \in \mathbf{IR}^n$ par l'output (ou variable à expliquée) $y(x,t) = f(x(t)) \in \mathbf{IR}^+$.
Chaque composante x_j est une fonction déterministe de $\mathbf{IR} \to \mathbf{IR}$: $t \to x_j(t)$, ou bien une variable aléatoire (v.a) sur $(\mathbf{IR}, \mathfrak{B}_{\mathbf{IR}}, P_j)$, avec $\mathfrak{B}_{\mathbf{IR}}$ la tribu borélienne de $\mathbf{IR}$ et P_j désigne la loi de la v.a x_j.
On définit et on désigne par D_f, le support de f, l'ensemble des inputs x réalisant une production strictement positive, i.e :

$$D_f = \Big\{(x,t) \in \mathbf{IR}^n \times \mathbf{IR} : f(x(t)) > 0\Big\} \tag{1.2.1}$$

Le vecteur de productivités marginales est définie par :

$$P_m = grad(f) \tag{1.2.2}$$

La productivité marginale Pm_k d'une variable x_k est donnée par :

$$Pm_k = \frac{\partial f}{\partial x_k} \tag{1.2.3}$$

La productivité total des inputs est donnée par :

$$TFP = \sum_{k=1}^{n} Pm_k \tag{1.2.4}$$

On définit le taux de croissance de la fonction f à un instant t par :

$$\gamma_f(t) = \frac{\frac{\partial f(x(t))}{\partial t}}{f(x(t))} = \frac{\partial}{\partial t} \ln(f(x(t))) \tag{1.2.5}$$

De façon similaire, le taux de croissance d'une variable explicative (input) x_j à un instant t est donnée par :

$$\gamma_{x_j}(t) = \frac{\frac{\partial x_j(t)}{\partial t}}{x_j(t)} = \frac{\partial}{\partial t} \ln(x_j(t)) \tag{1.2.6}$$

Définition 1.2.1 ***Demi- Vie***
On apelle demi- vie d'une fonction de production f, le temps nécessaire pour que f double de niveau, Autrement dit, le demi-vie de f à partir d'un instant initial t_0, est l'instant T qui vérifié :

$$f(x(T)) = 2f(x(t_0))$$

Définition 1.2.2 ***Fonction homogène***
Soit f une fonction de production. On dit que f est homogène s'il existe un réel $p \in \mathbf{IR}^+$, tel que :
$(\forall a \in \mathbf{IR})(\forall (x,t) \in D_f)$ $f(ax(t)) = a^p f(x(t))$.
p est appelé le degré de f et noté $deg(f)$, de plus on a les propriétés suivantes :
- Si $deg(f) < 1$: on dit que f a un rendement d'échelle décroissant.
- Si $deg(f) = 1$: on dit que f a un rendement d'échelle constant.
- Si $deg(f) > 1$: on dit que f a un rendement d'échelle croissant.

Définition 1.2.3 ***Fonction bien élevée***
On apelle fonction néoclassique ou bien élevée, toute fonction f vérifiant les trois conditions suivantes :

1. *Les productivités marginales sont décroissantes et positives, i.e :*
   $\forall (x,t) \in D_f \quad \frac{\partial Pm_k}{\partial x_k}(x(t)) = \frac{\partial^2 f}{\partial x_k^2}(x(t)) \leq 0$ *et* $Pm_k(x(t)) \geq 0$
2. *f est homogène et a un rendement d'échelle constant.*
3. *f vérifie les conditions d'Inada, i.e :*
   $\lim_{x_k \to 0} Pm_k = +\infty$ *et* $\lim_{x_k \to +\infty} Pm_k = 0$

b- Modèles remarquables de la fonction de production

1) Fonction Cobb douglass

On apelle fonction de production de Cobb douglass (voir [28]), toute fonction de la forme suivante :

$$f(x(t)) = c\Pi_{j=1}^{n} x_j^{\alpha_j}(t) \tag{1.2.7}$$

2) Fonction CES

On apelle fonction de production CES (Constant Elasticity of Substitution, voir [29]), toute fonction de la forme suivante :

$$f(x(t)) = c\Big(\sum_{j=1}^{n} b_j^{\rho} x_j^{\rho}(t)\Big)^{\frac{d}{\rho}} \tag{1.2.8}$$

3) Fonction de Leontief

On dit qu'une fonction de production est de Leontief (voir [30]), si elle s'écrit de la forme suivante :

$$f(x(t)) = c \min_{j=1}^{n} \left(\frac{x_j(t)}{e_j}\right) \tag{1.2.9}$$

Où :
c, d : sont des constantes réelles non nulles.
b_j sont des paramètres de partage.
$e_j > 0$, ρ et α_j représentent les élasticités pour chaque modèle de la fonction de production (1.2.7-1.2.9).

1.2.2 Propriétés géométriques de la production

Définition 1.2.4 ***Ligne isotrope- hypersurface graphique***
On apelle ligne isotrope de niveau q (et on la note $L_f(q)$) d'une fonction de production f, l'ensemble des inputs pour que la production soit égale à q, Autrement dit :

$$L_f(q) = \Big\{(x,t) \in \mathbf{IR}^n \times \mathbf{IR} : \ f(x(t)) = q\Big\}$$

-On appelle l'hypersurface graphique de f à un instant t, l'ensemble noté $\Phi_t(f)$, définie par :

$$\Phi_t(f) = \Big\{(x(t), f(x(t))) \in \mathbf{IR}^n \times \mathbf{IR}^+ : \ (x,t) \in D_f\Big\}$$

L'étude géométrique de la forme d'une fonction de production, joue un rôle important dans l'analyse macroéconomique (ainsi microéconomique), par exemple l'étude de la substitution des inputs [2]. B.Y Chen [33], a prouvé qu'une fonction de production homogène à deux inputs parfaitement substitutiables si et seulement si la surface de production est minimale. Récemment, les conditions de minimalité de surface de production de Cobb-Douglas et de CES ont été prouvées par X. Wang et Y. Fu [34], B.Y CHEN et all (voir [35]), et autres... .
Outre de la géométrie euclidienne sur $\mathbf{IR}^n$, ($n \in \mathbf{IN}^*$), on peut utiliser une autre géométrie importante sur $\mathbf{IR}^n$, appelée géométrie isotropique (voir [36,38]). D'après M.E Aydin [39] et Herrn [40], l'espace isotrope de dimension $n+1$ (noté I^{n+1}) est le produit de $\mathbf{IR}^n$ et de la ligne isotrope équipée d'une métrique de distance parabolique dégénérée (voir [20-29]).
On note :
$\mathbf{E}^{n+1}$ l'espace $\mathbf{IR}^{n+1}$ muni de la distance euclidienne.
On rappelle que la distance isotropique entre deux éléments $p, q \in \mathbf{IR}^{n+1}$ est donnée par la distance euclidienne des projections orthographiques sur $\mathbf{IR}^n \times \{0\}$ si les projections sont distinctes, sinon, il est donnée simplement par la distance euclidienne [3] .
D'après B.Y Chen [43], si $f \in \mathbf{C}^2\Big(\mathbf{IR}^n \times \mathbf{IR}\Big)$ est une fonction homogène de degré $p \neq 0$, alors on a :

$$\forall (x,t) \in Df \ \ H_f(x(t))x^{'}(t) = (p-1)grad(f(x(t))) \tag{1.2.10}$$

Où $x^{'}$ est le transposé de x et H_f est la Hessienne de f.
Si f est de rendement d'echelle constant ($p = 1$), alors $H_f(x(t))x^{'}(t) = 0$ pour tout $(x,t) \in Df$, par suite on obtient :

$$\forall (x,t) \in Df : \quad det(H_f(x(t))) = 0 \tag{1.2.11}$$

Cette équation au dessus est appelée l'équation homogène de Monge- Ampère.
Notons que (d'après le deuxième paragraphe de [44]) il y a l'équivalence entre les deux propriétés suivantes :

2. En économie, les facteurs de production ainsi les biens consommation, peuvent être substituables ou bien complémentaires. Lorsqu'il s'agit d'une production, On dit que deux inputs sont substituables si l'on peut utiliser l'un ou l'autre pour réaliser la même production. Par contre, deux inputs sont complémentaires lorsqu'il faut à la fois utiliser les deux pour produire. De point de vue du consommateur, un bien est substituable quand il peut être remplacé par un autre bien qui va répondre aux mêmes besoins, citons par exemple ; en transport, le choix entre bus et tramway [31]. Notons que, Sur le marché des biens, les biens substituables provoquent les concurrences dites indirectes. Par contre, Les biens complémentaires sont ceux qui doivent être consommés ensemble, nous donnons comme exemple des facteurs complémentaires ; en protection sociale, la consomation dans la période active et celle dans la période retraité (voir [32]).

3. Pour plus de détaille sur la métrique isotropique, on peut voir B.Y Chen [41] et M.E Aydin [42]

- L'hypersurface graphique $\Phi_t(f)$ est minimale isotropique dans I^{n+1}
- f est harmonique, (c'est-à-dire $\Delta f = 0$).
On se basant sur les travaux de B.Y CHEN et all (voir [45]), A.D. Vîlcu and all (voir [46]) et H. Pottmann [47], On a les théorèmes suivants :

Théorème 1.2.1 *Soit f une fonction de production homogène de $\mathbf{IR}^2 \times \mathbf{IR}$, alors les propriétés suivantes sont équivalentes :*

1. *f est parfaitement substituable*
2. *la surface de production est minimale dans $\mathbf{E}^3$*

Si de plus f est linéairement homogène, alors les deux propriétés (1) et (2) sont équivalentes à la propriété suivante :
- la surface de production est minimale isotropique dans $\mathbf{I}^3$

Théorème 1.2.2 *Soit f une fonction de production de $\mathbf{IR}^n$, alors :*
-Si f est une fonction CES, alors $\Phi(f)$ est isotropique minimale dans $\mathbf{I}^{n+1}$, si et selement si f est parfaitement substituable.
-Si f est une fonction de Cobb Douglas, alors $\Phi(f)$ est isotropique minimale dans $\mathbf{I}^{n+1}$, si et selement si les élasticités des inputs $\alpha_j = 1$.

Théorème 1.2.3 *Soit f une fonction de production de $\mathbf{IR}^n \times \mathbf{IR}$, et h une fonction croissante de $\mathbf{IR} \to \mathbf{IR}$.*
Soit $F = hof$ la transformation de f par h, alors :
-$\Phi(F)$ a une courbure relativement nulle, si et selement si :

- *f satisfait l'équation homogène de Monge-Ampère $det(H_f) = 0$, Ou*
- *F est linéairement homogène.*

Remarque 1.2.1 ***Cas des fonctions Cobb Dauglass et CES***
- Si f est une fonction Cobb Douglas, alors $\Phi(F)$ a une courbure relativement nulle dans $\mathbf{I}^{n+1}$, si et selement si F et f sont tous les deux linéaires.
- Si f est une fonction CES, alors $\Phi(F)$ a une courbure relativement nulle dans $\mathbf{I}^{n+1}$, si et selement si :

- *L'élasticité $\rho = 1$ ou*
- *F a un rendement d'echelle constant.*

Théorème 1.2.4 *Si $n = 2$ (i.e deux inputs), alors, $\Phi(F)$ a une courbure relativement nulle, si et selement si :*

- *f est parfaitement substituable, ou*
- *F est linéairement homogène.*

1.2.3 Méthodes de calcul du PIB et RNB

Un agrégat est une grandeur statistique synthétique et macroéconomique mesurant l'activité économique. Il existe trois grandes familles d'agrégats :

Agrégats de produit (ex. PIB)
Agrégats de revenu (ex. revenu national)
Agrégats de dépense (ex. consommation finale des ménages)
En comptabilité nationale, le PIB est le principal agrégat, car tous les autres se calcule en cascade à partir de lui. Des agrégats sont également calculés dans le domaine monétaire [4], notamment pour faciliter l'élaboration et la conduite de la politique monétaire.

a- Approches du Produit Intérieur Brut (PIB)

D'une manière très synthétique, le produit intérieur brut vise à mesurer la richesse créée par les hommes dans un pays pendant une période donnée, le plus souvent une année ou un trimestre. Cette création de richesse peut être considérée selon trois points de vue différents, appelés les trois approches du produit intérieur brut : l'approche production ; l'approche revenu ; l'approche demande (ou approche dépenses).

L'approche production

Dans l'approche production, le produit intérieur brut est calculé à partir de la valeur ajoutée, c'est-à-dire de la différence entre la production et la consommation intermédiaire. En effet, la production mesure la création de richesse et la consommation intermédiaire sa destruction au cours du processus de production. La valeur ajoutée mesure donc la richesse effectivement mise à la disposition de l'économie par le système productif. Le PIB par cet approche est calculé apartir de la relation suivante :

`Produit intérieur brut = somme des valeurs ajoutées + impôts sur les produits - subventions sur les produits`.

L'approche revenu

La production est également l'occasion d'une répartition de la richesse produite entre les salariés, les entreprises et l'Etat. L'approche revenu met en évidence cette répartition. Le calcul du produit intérieur brut selon cette approche dérive directement de la précédente, il suffit d'utiliser la décomposition de la valeur ajoutée provenant du compte d'exploitation :

Valeur ajoutée = Rémunération des salariés + autres impôts sur la production - autres subventions sur la production + excédent d'exploitation / revenu mixte

Et donc selon, cet aproche :

`Produit intérieur brut = Rémunération des salariés + autres impôts sur la production - autres subventions sur la production + excédent d'exploitation / revenu mixte + impôts sur les produits - subventions sur les produits`.

L'approche demande

L'approche demande, encore appelée approche dépenses, montre comment la richesse créée a été utilisée. Le calcul du produit intérieur brut qui lui correspond peut se déduire du compte de biens et services. Le PIB par cet approche est calculé apartir de la relation suivante :

`Produit intérieur brut = Consommation finale + Formation brute de capital fixe + Variation des stocks + Acquisitions moins cessions d'objets de valeur + Exportations - Importations`.

4. Les agrégats monétaires représentent un ensemble des moyens de paiement et des actifs financiers facilement convertibles en moyens de paiement, possédés par les agents économiques non financiers. Ils sont utilisés par les banques centrales pour orienter leur politique monétaire.

b- Revenu National Brut (RNB)

Le produit intérieur brut mesure la richesse créée par les entreprises appartenant à l'économie nationale, c'est-à-dire par les entreprises résidant à l'intérieur du territoire national, ce qui explique le terme "intérieur". Cependant, des non-résidents participent à la production des entreprises résidentes, soit par leur travail, soit par leur capital ; à l'inverse, des résidents participent à la production des entreprises non-résidentes, c'est-à-dire des entreprises appartenant au reste du monde. Pour déterminer le revenu tiré par les résidents de l'activité de production, il faut donc déduire du PIB le revenu issu de la production distribué aux non-résidents et ajouter le revenu issu de la production que les résidents ont tiré du reste du monde. Ces revenus sont les revenus du travail et du capital, c'est-à-dire des facteurs de production. Les impôts et subventions sur la production peuvent également être assimilés à des revenus des facteurs de production. L'agrégat qui mesure le revenu des résidents issus de la production prend le nom de revenu national brut, il se calcule de la manière suivante :
`Revenu national brut = Produit intérieur brut + Rémunération des salariés reçue du reste du monde - Rémunération des salariés payée au reste du monde - Impôts sur la production et les importations versés aux reste du monde + Subventions reçues du reste du monde + Revenus de la propriété reçus du reste du monde - Revenus de la propriété payés au reste du monde.`

1.3 Modèles dynamiques théoriques simples

Dans ce paragraphe, nous présentons un revue sur les importants modèles théoriques simples,des descriptions simples de manuels de la macro-économie impliquant un petit nombre d'équations ou de diagrammes sont souvent appelées «modèles». Les exemples comprennent le modèle IS-LM et le modèle Mundell-Fleming de la macroéconomie keynésienne, et le modèle de Solow de la théorie de la croissance néoclassique. Ces modèles partagent plusieurs caractéristiques. Ils sont basés sur quelques équations comportant quelques variables, qui peuvent être expliqué avec de simples schémas . La plupart de ces modèles sont statiques, mais certains sont dynamiques, décrivant l'économie sur plusieurs périodes. Les variables qui apparaissent dans ces modèles représentent souvent des agrégats macro-économiques (tels que le PIB ou le total de l'emploi) plutôt que des variables de choix individuelles et bien que les équations reliant ces variables décrivent des décisions économiques, elles ne sont généralement pas dérivées directement de modèles individuels. Ils sont assez simples pour être utilisés comme illustrations de points théoriques dans les explications introductives des idées macroéconomiques ; Par conséquent, l'application quantitative à la prévision, aux essais ou à l'évaluation des politiques est habituellement impossible sans augmenter substantiellement la structure du modèle.

1.3.1 Le modèle keynésien de Harrod (1939)

Il s'agit d'un modèle à un seul secteur de production (un seul bien) et à un seul pays. Harrod étend le modèle keynésien de base en y incluant la dynamique du capital (l'in-

vestissement) et de l'emploi (la population active)(voir [15]). Harrod s'interroge sur la capacité des des économies capitalistes à réaliser une croissance qui respecte l'équilibre du marché du bien et de celui du travail simultanément.
Il a donc deux types de problèmes :
– un problème de court terme d'existence de l'équilibre ;
– un problème de long terme de stabilité de l'équilibre.
Dans ce modèle la fonction de production est a une seul variable (n=1), et les modèles de production ; Cobb Douglass, CES et Leontief sont équivalentes.

a- Équilibre du marché du bien

Dans le cadre dynamique, et sur le marché du bien on a l'égalité entre l'épargne et l'investissement :

$$I_t = S_t \tag{1.3.1}$$

l'équation de l'accumulation du capital est donnée par :

$$K_{t+1} = K_t + I_t \quad (Cas\ discret) \tag{1.3.2}$$

$$\frac{dK_t}{dt} = I_t \quad (Cas\ continue) \tag{1.3.3}$$

L'investissement est déterminé par les anticipations des entrepreneurs sur la croissance de la demande $\left(\frac{dY_t}{dt}\right)$. De plus le coefficient de capital à l'équilibre est constant.
La production est donnée par l'équation suivante :

$$Y_t = f(K_t) = \frac{1}{\alpha} K_t \tag{1.3.4}$$

Où α est le coefficient marginal du capital.
Dans ce cas en remarque que :

$$P_m = P_{mK} = grad(Y_t) = \frac{1}{\alpha}$$

. Donc d'après l'équation (1.3.4), nous obtenons :

$$Y_{t+1} - Y_t = \frac{1}{\alpha}(K_{t+1} - K_t) \quad (Cas\ discret) \tag{1.3.5}$$

$$\frac{dY_t}{dt} = \frac{1}{\alpha}\frac{dK_t}{dt} \quad (Cas\ continue) \tag{1.3.6}$$

Donc, la variation de Y correspond ici aux anticipations des producteurs.
D'autre part, dans ce modèle le revenu total (PIB) ce réparti en deux ; l'éparne(=l'investissement) et la consommation. Ces répartition ce fait par des taux constants.
On note s le taux d'investissement et $c = 1 - s$ le taux de consommation, on a alors :

$$Y_t = I_t + C_t \tag{1.3.7}$$

$$I_t = sY_t \quad et \quad C_t = cY_t \tag{1.3.8}$$

En utilisant la définition du taux de croisance d'une fonction (voir l'équation (1.2.5)), le taux de croissance du revenu qui assure l'équilibre du marché de bien (appelé aussi taux

garanti, warranted rate) est donc donné par :

$$g_w = \gamma_Y(t) = \frac{s}{\alpha} \tag{1.3.9}$$

Nous déduisons donc, que les anticipations des producteurs et le comportement des consommateurs sont compatibles dans le temps si l'économie croît au taux g_w. Si les producteurs n'anticipent pas correctement l'évolution de la demande, le taux de croissance effectif de l'économie peut être différent du taux garanti (g_w) et impliquer ainsi un déséquilibre sur le marché des biens (le premier problème de Harrod). Puisque le taux g_w est constant dans le temps, alors l'évolution du revenu qui assure l'équilibre du marché de bien est suit le modèle malthusien, et doit suivre la trajectoire :

$$Y_t = \frac{1}{\alpha} K_0 \exp(g_w t) \tag{1.3.10}$$

b- Équilibre du marché du travail

Cela correspond à la condition :

$$Offre\ d'emploi(L_t^o) = Demande\ d'emploi(L_t^d) = population\ active \tag{1.3.11}$$

A l'équilibre (et dans le cadre keynésien), la constance de la technologie et des prix doivent conduire à un coefficient marginal du facteur-travail constant (A), c.à,d :

$$Y_t = \frac{1}{A} L_t^o = \frac{1}{A} L_t^d \tag{1.3.12}$$

Dans ce cas en remarque que :

$$P_m = P_{mL^d} = grad(Y_t) = \frac{1}{A}$$

. En utilisant la définition du taux de croisance d'une fonction, et par passage d'une dérivée logarithmique, le taux de croissance du revenu qui assure l'équilibre du marché du travail (c.à ;d de manière à ne pas créer du chômage), noté g_n est donc donné par :

$$g_n = \frac{\frac{dL_t^d}{dt}}{L_t^d} + \frac{\frac{dP_{mL^d}}{dt}}{P_{mL^d}} = n + l \tag{1.3.13}$$

où n est le taux de croissance de la population et l est le taux de croissance de la productivité du travail.
Par conséquent, nous devons avoir l'évolution suivante :

$$Y_t = \frac{1}{A} L_0^d \exp(g_n t) \tag{1.3.14}$$

c- Équilibre simultané sur les deux marchés.

L'équilibre simultané des deux marchés (du biens et du travail) est vérifié si $g_n = g_w$. D'aprés les équations (1.3.9) et 1.3.13), l'équilibre simultané des deux marchés est donc existe si et selement si :

$$\frac{s}{\alpha} = n + l \tag{1.3.15}$$

d- Critique du modèle de Harrod

Revenant à la relation (1.3.15) en dessus, On a :
– s provient du comportement de consommation des ménages ;
– α et l sont déterminés par la technologie ;
– n est déterminé par les lois démographiques.
Puisque ces quatre variables sont exogènes, alors la réalisation de la condition d'équilibre simultané (1.3.15) ne peut être que fortuite, et donc la croissance équilibrée des deux marchés simultaniment n'est pas sûrement réaliser. De plus dans le cas où l'équilibre simultané des deux marchés soit réaliser, il sera très instable. Pour montrer l'instabilité de l'équilibre simultané, nous proposons un contre exemple :
on note réspectivement Y_t^L et Y_t^K les évolutions dûs à l'équilibre de marché du travail et celle du marché du bien. On suppose que l'équilibre simultané est réalisé à un instant t_0, donc on a $Y_{t_0}^L = Y_{t_0}^K$ et d'après (1.3.15) $g_n = g_w$. on perturbe ces taux réspectivement par ε et η, et . On a donc :

$$\frac{Y_t^L}{Y_t^K} = \frac{Y_{t_0}^L \exp((g_n + \varepsilon)(t - t_0))}{Y_{t_0}^K \exp((g_w + \eta)(t - t_0))} = \exp((\varepsilon - \eta)(t - t_0))$$

Prenand par exemple $\varepsilon - \eta = 10^{-100}$, alors on a :

$$\lim_{t \to +\infty} \frac{Y_t^L}{Y_t^K} = +\infty \Rightarrow \frac{Y_t^L}{Y_t^K} \ n'est\, pas\, born\acute{e}$$

Donc nous voyons qu'une très petite différence entre les perturbations génère (à long terme) une grande différence entre les évolutions Y_t^L et Y_t^K, et par suite l'équilibre simultanie sur les marchés du bien et du travail, n'est pas stable.
En résumant ce qui précide, l'équilibre simultané des marché du bien et du travail dans le modèle de Harrod et instable et n'est pas toujours vérifié. Autrement dit, ces résultats signifient que les anticipations des producteurs et le comportement des consommateurs peuvent être incompatibles dans le temps, et le taux de chômage peut augmenter à cause de déséquilibre sur le marché de travail.(pour un exemple concret sur le modèle de Harrod, voir annexe).

1.3.2 Le modèle néo-cambridgienne de Kaldor

La réponse post-keynésienne (ou néo-cambridgienne) met l'accent sur la relation entre la répartition du revenu et l'accumulation du capital. Nicolas Kaldor explore la détermination de l'équilibre à travers un comportement différencié d'épargne. Il considère deux classes : les travailleurs et les capitalistes. La classe qui détient les capitaux a alors une position dominante dans la répartition du revenu et dans la fixation d'un taux d'épargne global compatible avec la réalisation d'un SCE (Sentier de croissance équilibré). (voir [17])

a- Répartition et équilibre

Hypothèses du modèles
— (H1) Les travailleurs reçoivent uniquement les revenus du travail, et les capitalistes reçoivent uniquement les revenus du capital.

— (H2) Le comportement d'épargne est différent entre les deux classes. Les travailleurs ont une propension à épargner s_w et les capitalistes, s_c avec $0 < s_w < s_c$.
— (H3) L'équilibre du marché des biens est assuré et il correspond à l'équation : $\frac{I_t}{Y_t} = \frac{S_t}{Y_t}$
— (H4) La part de la demande d'investissement dans le revenu est exogène et vérifie l'inégalité :

$$s_w \leq \frac{I_t}{Y_t} \leq s_c \tag{1.3.16}$$

Dans ce modèle le choix de coefficient de technologie α est ce fait de manière optimale,et il respècte l'équilibre dynamique (SCE). Or cette équilibre ne vérifie que si : $\frac{d}{dt}\left(\frac{K}{L^d}\right) = 0$. Par un simple calcul et l'utilisation des notations précédentes et la condition d'équilibre (Hypothèse(H3)), alors nous obtenons :

$$\frac{I_t}{Y_t} = n\alpha \tag{1.3.17}$$

Où n est le taux de croissance de population active L^d. Nous notons par W la masse salariale, P la masse des profit et $S = S_w + S_c = s_w W + s_c P$ l'éparne globale. Nous devons alors avoir pour chaque instant t :

$$Y_t = W_t + P_t \tag{1.3.18}$$

Donc l'équation de répartition du revenu est donnée par :

$$S_t = s_w Y_t + (s_c - s_w) P_t \tag{1.3.19}$$

En utilisant l'hypothèse (H2), alors la condition d'équilibre est donnée par :

$$\frac{I_t}{Y_t} = \frac{S_t}{Y_t} = s_w + (s_c - s_w)\frac{P_t}{Y_t} \tag{1.3.20}$$

donc on a deux types de conséquences :
-À court terme, puisque $\frac{I}{Y}$ est exogène d'après (H4) ; alors l'équilibre du marché des biens peut toujours être réalisé grâce à une répartition adaptée du revenu ;
-À long terme, quand on tient compte de la dynamique de la population, l'équation (1.3.20) ; le SCE n'est atteignable que si l'inégalité de l'Hypothèse (H4) est vérifiée. Dans ce cas, la répartition du revenu adaptée assure l'équilibre de court terme qui correspond au SCE.
La répartition d'équilibre est déterminée selon le mécanisme de multiplicateur :

$$\frac{P_t}{Y_t} = \frac{1}{s_c - s_w}\frac{I_t}{Y_t} - \frac{s_w}{s_c - s_w} \tag{1.3.21}$$

(Voir figure 1).

b- Statique comparative

Consensus social

Le consensus social peut éviter les situations extrêmes de répartition où tout le revenu est accaparé par les travailleurs $\left(\frac{P_t}{Y_t} = 0\right)$ ou par les capitalistes $\left(\frac{P_t}{Y_t} = 1\right)$. La recherche de la

paix sociale peut imposer une borne inférieure $\left(\frac{P_t}{Y_t}\right)_{min}$ et une borne supérieure $\left(\frac{P_t}{Y_t}\right)_{max}$, à la part des profits dans le revenu. (Voir figure 2)

Quel est l'impact sur l'équilibre d'une telle situation ?

L'équilibre n'existe maintenant que pour les valeurs de l'investissement vérifiant la relation :

$$s_w \leq s_{min} \leq \frac{I_t}{Y_t} \leq s_{max} \leq s_c$$

Ce qui rend l'équilibre plus difficile à atteindre.

La cruche de la veuve

On part de la condition d'équilibre du marché de bien (1.3.20), soit τ le taux de profit $\left(\tau = \frac{P_t}{K_t}\right)$, en utilisant les équations (1.3.17 - 1.3.21), alors :

$$\tau = \frac{P_t}{K_t} = \frac{1}{s_c - s_w}\left(g - \frac{s_w}{\alpha}\right) \tag{1.3.22}$$

Où $g = \frac{d\ln(K_t)}{dt} \approx g_w$

- Si l'épargne des travailleurs est négligeable $s_w \approx 0$, Alors : $\tau \approx \frac{g}{s_w}$, donc dans ce cas, plus les capitalistes consomment, plus leur taux de profit est élevé : *"les travailleurs dépensent ce qu'ils gagnent et les capitalistes gagnent ce qu'ils dépenses"* (Kalecki).(voir figure3)

Si les travailleurs n'épargnent pas et la totalité du profit est réinvestie ($s_w = 0$ et $s_c = 1$), alors $\tau = g$. C'est la règle d'or de l'accumulation : le taux de profit est égal au taux de croissance.

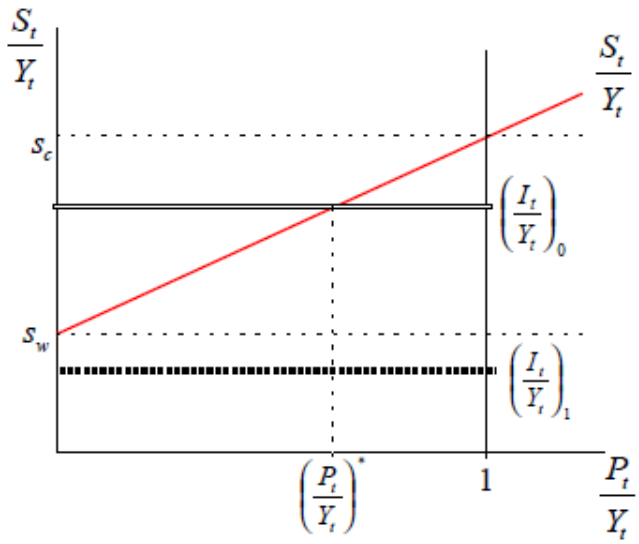


FIGURE 1.1 – Équilibre et répartition

c- Équilibre dynamique sur les deux marchés

Intègrons maintenant l'équilibre du marché du travail, nous savons que le SCE ne correspond au plein-emploi que si $g = g_w = n = g_n$ (voir Harrod). On utilise le fait que :

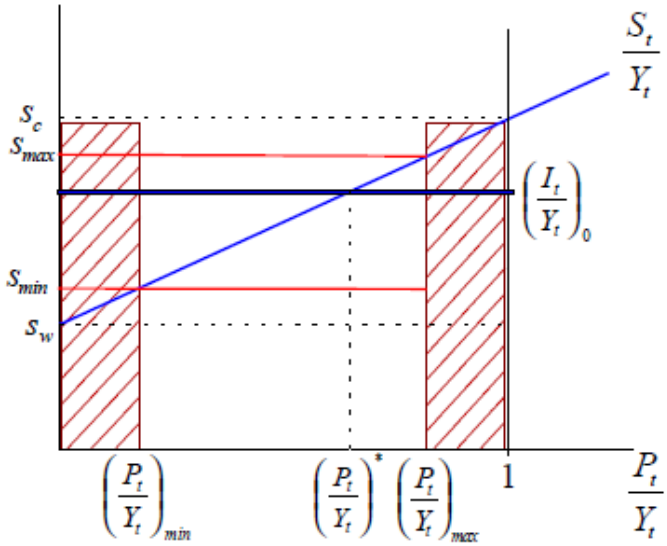


FIGURE 1.2 – Consensus social et équilibre

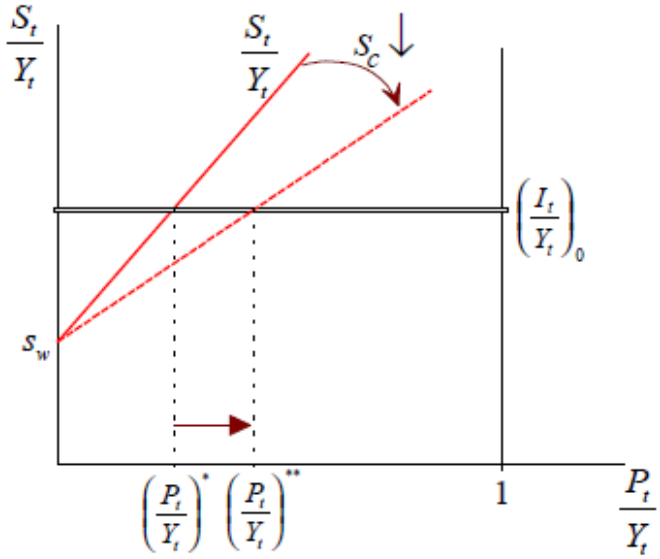


FIGURE 1.3 – La "cruche de la veuve"

$\left(\frac{P_t}{Y_t} = \frac{P_t}{K_t} \times \frac{K_t}{Y_t} = \tau\alpha\right)$, alors nous avons :

$$\frac{P_t}{Y_t} = \frac{\alpha}{s_c - s_w}\left(n - \frac{s_w}{\alpha}\right)$$

Donc

$$\frac{P_t}{Y_t} \in]0,1[\forall t \Rightarrow \frac{s_w}{\alpha} < n < \frac{s_c}{\alpha} \tag{1.3.23}$$

Cette condition rend plus facile la réalisation d'une croissance équilibrée de plein-emploi : c'est l'écart entre s_w et s_c qui détermine la possibilité d'existence de l'équilibre. De plus si tous les salaires sont consommés ($s_w = 0$) et tous les profits épargnés ($s_c = 1$) ; alors on obtient la règle d'orqui tient compte du plein-emploi du travail :

$$\tau = n \tag{1.3.24}$$

Cette dérniere égalité, exprime l'égalité entre le taux de profit et le taux de croissance naturel.

Critique de modèle

Si l'on tient compte du fait que les capitalistes ont plus de latitude pour fixer leur taux d'épargne (celui des travailleurs est plus fortement conditionné par la subsistance), le modèle de Kaldor nous donne la manière dont le comportement de cette classe détermine les grandeurs économiques : le partage du revenu, le taux de profit. Mais, il s'agit d'un modèle assez ad hoc et caricaturale, et les résultats dépendent des simplifications qui en sont à la base. Notamment la distribution et les comportements sont extrêmement schématiques et rien n'empêche les travailleurs d'avoir un capital et d'en tirer des dividendes. Le modèle de Passinetti (voir [10-11]) élimine d'ailleurs cette restriction.

1.3.3 Le modèle néoclassique de solow avec progrès technique

Il s'agit d'un modèle très simple qui fournit des intuitions fondamentales à la question initiale : "Pourquoi certains pays sont-ils si riches tandis que les autres sont appauvris ?". Dans cette section nous présentons selement le modèle de Solow avec progrès technique (pour le modèle de solow traditionnele et avec capital humain voir [13]).

Les hypothèses

Le modèle se base sur un certain nombre d'hypothèses :
(H1) : La concurrence est pure et parfaite, les agents sont preneurs de prix.
(H2) : Il y a un bien unique, (Par exemple il peut s'agir du blé). Il y a équilibre sur le marché des biens : $Y_t = C_t + I_t$.
(H3) : Il y a équilibre sur le marché des capitaux : $I_t = S_t$, et l'investissement accroît dans le temps le stock de capital. A long terme, le capital s'use et donc son stock se déprécie au taux δ. Alors l'accroissement net du stock de capital est :

$$\frac{dK_t}{dt} = I_t - \delta k_t$$

(H4) : il y a équilibre sur le marché du travail : $L_t^o = L_t^d$. et la population croît à un taux exogène constant : $\frac{dL_t}{L_t} = n$.
(H5) : la fonction de production, $Y_t = F(K_t, A_t L)$ est néoclassique, où A_t représente le progrès technique à l'instant t.
(H6) : le progrès technique est neutre au sens de Harrod, c'est-à-dire qu'il augmente l'efficacité du travail, et il croît au taux exogène constant : $\frac{dA_t}{A_t} = g$.
L'objectif de ce modèle est de trouver la solution de l'état régulier (c'-à-d une solution d'équilibre concurrentiel où toutes les variables croissent à taux constant), étudie la stabilité de l'état régulier et finalement chercher la consommation optimale (règle d'or) d'un agent dans une économie concurentielle.
La fonction de production pour ce modèle est donnée par :

$$Y_t = F(K_t, A_t L) = K_t^{\alpha}(A_t L_t)^{1-\alpha} \tag{1.3.25}$$

Puisque les rendements sont constants, on peut travailler avec la forme intensive de la fonction de production et avec les variables per capita éfficace, alors on a :

$$\hat{y} = \frac{Y_t}{A_t L_t} = \Big(\frac{K_t}{A_t L_t}\Big)^{\alpha} = \hat{k}^{\alpha} = f(\hat{k}) \tag{1.3.26}$$

D'après l'hypothèse (H6), le progrès technique est donné par : $A_t = A_0 e^{gt}$. On note r : le taux d'intérêt et w le taux de salaire. Le vecteur de productivité marginale sera :

$$Pm = \begin{pmatrix} Pm_K \\ Pm_L \end{pmatrix} = \begin{pmatrix} \frac{\partial Y_t}{\partial K_t} \\ \frac{\partial Y_t}{\partial L_t} \end{pmatrix} = \begin{pmatrix} f'(\hat{k}) \\ e^{gt}(f(\hat{k}) - \hat{k}f'(\hat{k})) \end{pmatrix} = \begin{pmatrix} r+\delta \\ w \end{pmatrix} \tag{1.3.27}$$

Où f' représente la dérivée de f par rapport à $\hat{k}$.
L'équation dynamique fondamentale du movement de capital per capita éfficace, associé à ce modèle est donnée par :

$$\frac{d\hat{k}}{dt} = sf(\hat{k}) - (n+g+\delta)\hat{k} \tag{1.3.28}$$

Où s est le taux d'investissement qui est constant dans ce modèle.
À partir des équations (1.3.28), Solow, représente le mouvement du capital per capita éfficace, dans un diagramme, appelé diagramme de Solow (figure(4)).

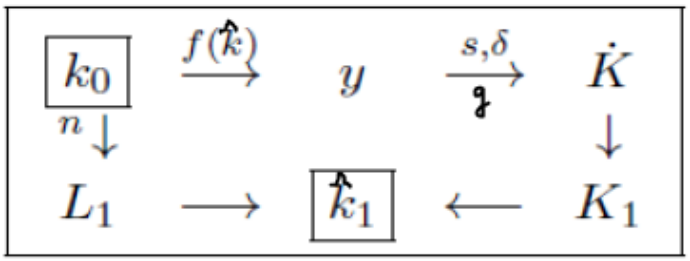


FIGURE 1.4 – Diagramme de Solow

1.3.4 Le modèle néoclassique de Cass- Koopmans (1965)

Dans le modèle de Solow, le taux d'épargne (ou propension épargner) est supposé être constant pour cause de simplicité mais également parce que Solow pensait préférable de prendre comme valeur la propension à consommer observée statistiquement. Cass [1965] et Koopmans [1965] ont remis en cause cette façon de voir les choses pour faire un pas vers le réalisme. Le consommateur dans le modèle de Solow peut être considéré comme un agent myope, incapable de voir que sa consommation actuelle influence ses possibilités de consommation future.

Modèle de croissance optimale sans progrès technique

Le modèle admet que l'horizon temporel est infini. Il existe un ménage représentatif dont la fonction d'utilité instantanée s'écrit : $u(c_t)$, avec le fait que u vérifié les conditions d'inada (voir définition 1.2.3)
La croissance de l'emploi dans ce modèle est malthusien avec un taux éxogène constant

n, c-à.d $L_t = L_0 e^{nt}$.
Le ménage représentatif est altruiste dans ce sens que les allocations en matière de consommation se font en fonction de tous ses membres et en fonction des générations à venir. Ainsi, la fonction-objectif du ménage s'écrit :

$$U_0 = \int_0^{+\infty} u(c_t) e^{-(\beta - n)t} dt \tag{1.3.29}$$

où c_t est la consommation par tête, β le taux d'escompte subjectif et $\beta - n$ le taux d'escompte effectif étant donné que le ménage dérive son utilité de la consommation de chacun de ses membres. la consommation totale C_t à un instant t est donnée par $C = c_t L_t$.
Remarque : $\beta > n$, car il s'agit d'un arbitrage entre présent et futur.
Dans ce modèle les marchés des biens et des facteurs sont compétitifs. Les possibilités de production de l'économie sont données par la fonction de production agrégée :

$$Y_t = F(K_t, L_t)$$

Comme dans le modèle de Solow, la fonction de production est « well behaved » et vérifie les conditions d'Inada. L'hypothèse de rendements d'échelle est retenue afin de travailler avec des grandeurs par tête. En terme de variable per capita, on aura :

$$y_t = f(k_t) \;\; k_t = \frac{K_t}{L_t}$$

On note R_t est la rente du capital et w_t le salaire payé à chaque travailleur. Donc, puisque les marchés sont compétitifs, on devrait vérifier que :

$$\begin{pmatrix} R_t \\ w_t \end{pmatrix} = P_{m,t} = grad(Y_t) \tag{1.3.30}$$

À la date t, le ménage représentatif dispose d'un niveau d'avoirs $B(t)$. Ainsi, au fil du temps, les avoirs du ménage changeront selon la règle suivante :

$$\frac{dB_t}{dt} = (r_t \;\; w_t - c_t) \begin{pmatrix} B_t \\ L_t \end{pmatrix} \tag{1.3.31}$$

Où $r_t = R_t - \delta$ le taux d'intérêt du marché, δ le taux d'amortissement du capital et $w_t L_t$ le flux de revenus salariaux que reçoit le ménage. Les avoirs par tête étant donné par $b_t = B_t / L_t$, alors, on établit que :

$$\frac{db_t}{dt} = (r_t - n \;\; w_t - c_t) \begin{pmatrix} b_t \\ 1 \end{pmatrix} \tag{1.3.32}$$

Il faut noter que les avoirs du ménage peuvent être composés du stock de capital K_t et des bons du Trésor. Si les bons du Trésor ne sont pas pris en compte, il vient que $a_t = k_t$. En horizon temporel infini, il doit être imposé la condition suivante (condition du jeu de No-Ponzi [5]) :

$$\lim_{t \to +\infty} a_t e^{\int_0^t (r_s - n) ds} \geq 0 \tag{1.3.33}$$

5. Un système de Ponzi est un montage financier frauduleux qui consiste à rémunérer les investissements des clients essentiellement par les fonds procurés par les nouveaux entrants. Si l'escroquerie n'est pas découverte, elle apparaît au grand jour au moment où elle s'écroule, c'est-à-dire quand les sommes procurées par les nouveaux entrants ne suffisent plus à couvrir les rémunérations des clients

Caractérisation de l'équilibre
Un équilibre compétitif consiste à un ensemble de sentiers de la consommation, du stock de capital, du taux de salaire et de rente du capital telle que le ménage représentatif maximise en tout temps son utilité étant donné un ensemble de conditions initiales.
Le problème du ménage s'écrit comme suit :

$$\begin{cases} \max_{c_t} U_0 = \int_0^{+\infty} u(c_t) e^{-(\beta-n)t} dt \\ \quad S.C \\ \frac{db_t}{dt} = (r_t - n \;\; w_t - c_t)\begin{pmatrix} b_t \\ 1 \end{pmatrix} \\ \lim_{t\to+\infty} a_t e^{\int_0^t (r_s - n)ds} \geq 0 \\ a_0 = a(0) \;\; donn\acute{e}. \end{cases} \tag{1.3.34}$$

Le Hamiltonien du problème est donné par :

$$H_t = e^{-(\beta-n)t}(u(c_t) + q_t((r_t - n)b_t + w_t - c_t)$$

On utilise le principe de maximum de Pontryagin, on obtient :

$$\frac{d}{dt}\ln(c_t) = \frac{f'(k_t) - \delta - \beta}{-c_t u''(c_t)/u'(c_t)} \tag{1.3.35}$$

Or, on a l'élasticité de l'utilité marginale de la consommation, est donnée par :

$$\theta = -c_t u''(c_t)/u'(c_t) \tag{1.3.36}$$

Donc l'équilibre compétitif est réalisable sur l'ensemble :

$$\Gamma = \{c_t / \;\frac{d}{dt}\ln(c_t) = \frac{f'(k_t) - \delta - \beta}{\theta} \;\forall t \geq 0\} \tag{1.3.37}$$

Croissance optimale Le problème consiste à définir les sentiers du capital et de la consommation afin de réaliser un équilibre Pareto-optimal au sein de l'économie. Dans le cadre de ce modèle, le problème de croissance optimale se pose en termes de maximisation de l'utilité du ménage sous une contrainte d'accumulation du capital, soit :

$$\begin{cases} \max_{c_t} U_0 = \int_0^{+\infty} u(c_t) e^{-(\beta-n)t} dt \\ \quad S.C \\ \frac{dk_t}{dt} = f(k_t) - (n+\delta)k_t - c_t \\ k_0 = k(0) \;\; donn\acute{e}. \end{cases} \tag{1.3.38}$$

Le Hamiltonien du problème s'écrit :

$$H_t = e^{-(\beta-n)t}(u(c_t) + q_t((r_t - n)b_t + w_t - c_t)$$

On utilise le principe de maximum de Pontryagin, on obtient :

$$\frac{d}{dt}\ln(c_t) = \frac{f'(k_t) - \delta - \beta}{\theta} \tag{1.3.39}$$

Modèle de Cass- Koopmans avec progrès technique

Considérons la fonction de production agrégée suivante : $Y_t = F(K_t, A_t L_t)$
où A_t qui représente le progrès technique croît au fil du temps au taux g. La fonction de production étant homogène de degré 1, on peut établir que :

$$\hat{y}_t = \frac{Y_t}{A_t L_t} = f(\hat{k}_t)$$

On définit la consommation par travailleur efficient comme suit :

$$\hat{c}_t = \frac{C_t}{A_t L_t}$$

Le taux de croissance de la consommation par travailleur efficient est donné par la différence entre le taux de croissance de la consommation par tête et le taux de croissance du progrès technique, soit :

$$\frac{d}{dt}\ln(\hat{c}_t) = \frac{d}{dt}\ln(c_t) - \frac{d}{dt}\ln(A_t) = \frac{d}{dt}\ln(c_t) - g = \frac{f'(k_t) - \delta - \beta - \theta g}{\theta} \tag{1.3.40}$$

En régime permanent le taux de croissance de la consommation par travailleur efficient devrait être nul. Par conséquent, on devrait vérifier que :

$$f'(k_t) - \delta = \beta + \theta g \tag{1.3.41}$$

Par suite, Le taux d'intérêt réel rémunère le renoncement à la consommation présente. Ce coût de l'épargne correspond la somme du facteur d'escompte subjectif (ou psychologique) et du coût du non lissage du fait de la croissance qui est mesurée par θg.

1.3.5 Modèle de Romer

D'où vient le progrès technique qui est sous-jacent à la croissance économique ? Les théories de la croissance endogène cherchent à répondre à cette question. Ces modèles introduisent notamment une idée simple : le progrès technique résulte de la recherche de profit des inventeurs et des firmes. Par conséquent, il résulte du fonctionnement même de l'économie. C'est cette idée que Paul Romer a développé dans les années 80-90. (voir [13-14])
Le modèle de Romer endogénéise le progrès technique en introduisant la recherche de nouvelles idées par des inventeurs intéressés par les profits qu'ils peuvent obtenir grâce à leur innovation.

a- Hypothèses du modèle

(H1)- Le progrès technique est endogène, et il résulte de la recherche de profit des inventeurs et des firmes.
(H2)- La fonction de production agrégée est homogène, a des rendements d'echelle croissante et elle donnée par :

$$Y_t = F(K_t, A_t, L_{t,p}) \tag{1.3.42}$$

Où $L_{t,p}$ reprèsente le travail consacré à la production.
(H3)- La population active L_t se réparti en deux ; population dans le secteur de production $L_{t,p} = s_K L_t$ et population pour le secteur de recherche et developpement (R.D)

$L_{t,A} = s_R L_t$. Où s_R et s_K représentes respectivement ; la part de la population engagée en (R.D) et la part de la population engagée dans le secteur de production.
(H4)- la concurrence est imparfaite, et l'économie et tri-sectorielles.
(H4)- la population croît avec un taux éxogène constant $n = \frac{\frac{dL_t}{dt}}{L_t} = \frac{\frac{dL_{t,p}}{dt}}{L_{t,p}} = \frac{\frac{dL_{t,A}}{dt}}{L_{t,A}}$

a- Croissance dans le modèle de Romer

Dans le sentier de croissance équilibrès, le stock de connaissance croît avec le taux :

$$\gamma_A = \frac{d}{dt}(\ln(A_t) = \frac{\lambda n}{1-\phi} \tag{1.3.43}$$

Où $(\lambda, \phi) \in [0,1[^2$ sont respectivement, l'externalité associée à la duplication (congestion) et externalité positive dans la recherche.
La production per capita qui correspond à la croissance équilibré dans ce modèle, est donnée par :

$$y_t^* = \left(\frac{s_K}{n+\gamma_A+\delta}\right)^{\frac{\alpha}{1-\alpha}} (1-s_R)\rho\frac{s_R}{\gamma_A}L_t \tag{1.3.44}$$

À partir de ces deux équations au dessus, nous constatons que :
- Le modèle suggère que si la croissance de population s'arrête, la croissance économique doit s'arrêter aussi. De plus, si l'effort de recherche reste constant, cela devrait conduire aussi à une croissance nulle. Un effort de recherche constant ne peut pas soutenir les augmentations proportionnelles du stock de connaissances nécessaires à la croissance de long terme.
- Le modèle a l'éffet d'échelle suivante : une économie mondiale plus grande est aussi une économie plus riche. Cela provient de la nature non-rivale des idées : une économie plus grande correspond à un marché plus grand pour une idée et donc à un rendement plus grand (effet de demande).
- Une croissance plus rapide est corrélée avec une proportion plus importante de la population dans le secteur de la recherche.

b- Mécanismes économiques du modèle

Romer introduit la concurrence imparfaite dans un cadre d'équilibre général (fondements microéconomiques de la macroéconomie).
Dans le modèle de Romer Il y a trois secteurs : les secteurs du bien final (secteur concurentiel), du bien intermédiaire (secteur monopoliste) et de la recherche(secteur concurentiel).

b-1 Secteur du bien final

Soientt A_t le nombre total de biens intermédiaires disponibles dans l'économie, $x_t(u)$ est le coût d'achat de l'idée u et $p(x_t(u))$ est la demande de u à l'instant t.
L'objectif des firmes du secteur final (avec la normalisation $P_Y = 1$) est :

$$\max_{p(x_t(u))} L_{t,p}^{1-\alpha} \int_0^A x_t^\alpha(u)du - w_t L_{t,p} - \int_0^A p(x_t(u))x_t(u)du \tag{1.3.45}$$

Avec $L_{t,p}$ est le travail dans le secteur du bien final.

b-2 Secteur du bien intermédiaire

Soit r le prix du capital. Alors, l'objectif du secteur de bien intermédiaire à l'instant t est :

$$\max_{x_t(u)} \pi(u) = p(x_t(u))x_t(u) - rx_t(u) \tag{1.3.46}$$

b-3 Secteur de la recherche

Une nouvelle idée correspond à une nouvelle manière de transformer une unité de capital brut en une unité de bien intermédiaire. Quand une nouvelle idée est découverte, son inventeur obtient un brevet qui lui donne l'exclusivité sur cette idée pour toujours. L'inventeur vend le brevet à une firme du secteur intermédiaire et utilise ce revenu pour consommer et épargner comme tous les autres agents dans l'économie. Mais à quel prix doit-il vendre ce brevet ?
Soit $P_{t,A}$ le prix maximal que chacun est prêt à payer et donné par la valeur actualisée des profits d'une firme du secteur intermédiaire, à l'instant t. En utilisant le principe de l'arbitrage pour ce modèle, alors ce prix doit vérifié l'équation d'arbitrage suivante :

$$rP_{t,A} = \pi + \frac{dP_{t,A}}{dt} \tag{1.3.47}$$

b-4 Résolution du modèle

- Pour le secteur du bien final, il s'agit de déterminer la demande optimale p^*, qui maximise le profit de ce secteur. après résolution du modèle (voir [14]) on obtient :

$$p_t^* = \frac{1}{\alpha} r_t \tag{1.3.48}$$

- Soit x^* le prix optimale de vendre un bien capital au secteur du bien final, et qui est au même temprs le prix optimale pour maximiser le profit du Secteur de bien intermédiaire. Alors x^* est donné par :

$$x_t^* = \Big(\frac{\alpha^2}{r_t}\Big)^{1/(1-\alpha)} L_{t,p} \tag{1.3.49}$$

- Pour le secteur de la recherche, le prix maximale d'un brevet est donné par :

$$P_{t,A} = \frac{\pi^*}{r_t - n} \tag{1.3.50}$$

Remarque : Dans ce modèle, le prix du capital est inférieur à sa productivité marginale : $r = \alpha^2 Y_t/K_t < Pm_K$.

c- Critique du modèle

le modèle de Romer souligne le rôle de la spécialisation dans la croissance, mais il néglige l'obsolescence des outils anciens du fait de l'invention de nouveaux outils. Mais si l'obsolescence avait lieu, dans ce cas les anciens outils disparaîtrait, ainsi que les gains de productivités générés par la spécialisation sur ces outils. Ce qui supprimerait la croissance de A et donc de Y. Le modèle d'Aghion- Howitt (1992) étudie la possibilité d'une croissance en présence d'un processus aléatoire d'innovation qui augmente la qualité des nouveaux biens intermédiaires tout en conduisant à une obsolescence totale des outils de production existant.

1.4 Estimation des frontières de production et modélisation de PIB

1.4.1 Frontières paramétriques déterministes et modélisation de PIB

Notre objectif, est de touver la mielleure fonction modélisant l'output (PIB) pour m pays donnés. une telle fonction vérifie cet objectif est appellé l'output maximum réalisable par la technologie efficace. Soit f une fonction de $\mathbf{IR}^n \times \mathbf{IR} \to \mathbf{IR}$, strictement concave, continue et monotone, et soit y_i l'output observé à partir des inputs $x_i \in \mathbf{IR}^n$ sur chaque pays i. On dit que la technologie est définie par la fonction de production f, s'il existe une relation entre f et y_i, de la forme :

$$y_i = f(x_i, \beta) - u_i \tag{1.4.1}$$

Où $\beta \in \mathbf{IR}$, et u mesure l'écart entre l'output observé y et l'output maximum réalisable par la technologie efficace.

a- Estimation non statistique

Dans le prolongement des analyses pionnières de Farrell [48], Aigner et Chu [49] ont cherché la valeur de β qui minimise soit la valeur absolue, soit le carré des écarts entre les observations de l'output y et la frontière $f(.)$. Cette estimation a été faite en résolvant le programme :

$$\begin{cases} \sum_{i=1}^{m} |y_i - f(x_i, \beta)| \\ \quad S.C \\ y_i \leq f(x_i, \beta) \end{cases} \tag{1.4.2}$$

La méthode d'Aigner et Chu a le défaut de présenter les inconvénients suivants :
-premièrement, il n'est pas possible de faire de l'inférence, car les estimations calculées n'ont pas de propriétés statistiques ;
-le deuxième inconvénient est la sensibilité des estimateurs aux observations extrêmes.
Pour cette raison, on a besoin des estimations statistiques.

b- Estimation Statistique

Méthode des moindres carrés corrigés
On peut réécrire le modèle (1.3.1) sous la forme :

$$y_i = \beta_0 + \sum_{j=1}^{n} \beta_j x_{ij} - u_i \tag{1.4.3}$$

Et l'on suppose que :
- (H1) les termes u_i sont identiquement et indépendamment distribués selon une loi normale $N(\mu, \sigma_u)$
- (H2) u est sans corrélation avec les inputs.
On peut constater que sous la première hypothèse la méthode des moindres carrés n'est pas applicable $E(u_i) \neq 0$. Par cette méthode, on obtient des estimateurs sans biais pour

tous les paramètres sauf pour le terme constant, puisque le terme u a une moyenne positive dès que l'efficacité productive n'est pas parfaite pour toutes les observations. On a en effet :

$$E(\hat{\beta}_0) = \beta_0 - \mu$$

Richmond [24], qui est à l'origine de cette méthode propose de réécrire l'équation (1.4.3) sous la forme suivante :

$$y_i = \beta_0^{'} + \sum_{j=1}^{n} \beta_j x_{ij} - \delta_i \tag{1.4.4}$$

Où $\delta_i = u_i - \mu$ et $\beta_0^{'} = \beta_0 - \mu$.
L'estimation de l'équation (1.3.4) par la méthode des moindres carrés donnera des estimateurs sans biais pour tous les paramètres. Pour identifier complètement la frontière, il suffit d'ajouter à la constante estimée par les moindres carrés, la moyenne du terme u . Si par exemple, la distribution de u (on la note g) est supposée suivre une loi Gamma de paramètres (p, λ) $(p, \lambda > 0)$, alors :

$$g(x) = \frac{\lambda}{\Gamma(p)} \frac{e^{-\lambda x}}{(\lambda x)^{p-1}} \mathbf{1}_{]0,+\infty[}$$

On a $E(u) = \frac{p}{\lambda} = \mu$ et $Var(u) = \frac{p}{\lambda^2} = \frac{\mu}{\lambda}$, On estimera alors $\frac{\mu}{\lambda}$ par $\hat{\sigma}^2$ ($\hat{\sigma}^2$ représente la variance estimée des résidus des moindres carrés du modèle (1.4.3)). L'estimation des inefficacités individuelles des unités de décision s'obtient à partir des résidus de l'équation (1.3.4), soit :

$$\hat{u}_i = -(y_i - \beta_0^{'} - \sum_{j=1}^{n} \beta_j x_{ij}) = \hat{\mu} - \hat{\delta}_i \tag{1.4.5}$$

La méthode des moindres carrés corrigés présente deux inconvénients
- avec cette méthode, même après la correction, certains résidus peuvent encore être négatifs et, par conséquent, la frontière n'enveloppe pas le nuage de points ;
- la correction dépend du choix de la distribution des résidus.
Pour ces deux raisons, Thiry et Tulkens [50], notent que « on ne peut accorder une confiance très grande au niveau absolu de la valeur numérique attribuée par cette méthode à l'efficacité de chaque observation. Par contre, le rangement des observations selon les degrés d'efficacité qu'elle évalue est plus fiable, car ce rangement est évidemment indépendant de la correction du terme constant, et une observation est d'autant plus efficace que son résidu est positif ».

Estimation par le maximum de vraisemblance
Sous les deux hypothèses évoquées en a) et en spécifiant une distribution particulière pour u , le modèle (1.4.1) peut être estimé par le maximum de vraisemblance. On parle alors de frontière déterministe statistique. Schmidt [51] a par exemple montré que l'ajustement par le maximum de vraisemblance du modèle (1.3.1) avec u suivant une loi double exponentielle ou bien une loi semi normale est solution du système (5) respectivement. De son côté, Greene a montré que si la distribution de u est asymétrique, l'estimateur du maximum de vraisemblance est plus efficace que l'estimateur des moindres carrés. Comme on peut le constater, avec cette méthode, s'estompe la "frontière" entre estimations statistique et non statistique d'une frontière déterministe.

1.4.2 Frontières paramétriques stochastiques et modélisation de PIB

L'approche stochastique a été proposée simultanément par Aigner, Lovell et Schmidt et Meusen et Van Den Broeck [52]. Un terme d'erreur aléatoire v est ajoutée dans la relation (1.4.1). On obtient un modèle à erreur composée :

$$y_i = f(x_i, \beta) + v_i - u_i \tag{1.4.6}$$

Où $u \geq 0$ et $-\infty \leq v \leq +\infty$ v_i représente l'écart dû aux aléas qui influencent la production et qui ne sont pas directement sous le contrôle du gestionnaire. Par ailleurs v et u sont indépendants l'un de l'autre, ainsi que de x_i . On pose $\varepsilon = v - u$, alors, la relation (1.4.3) peut encore s'écrire sous la forme :

$$y_i = \beta_0 + \sum_{j=1}^{n} \beta_j x_{ij} + \varepsilon_i \tag{1.4.7}$$

Soit sous la forme matricielle :

$$y = X\beta + \varepsilon \tag{1.4.8}$$

Où

$$X = \begin{pmatrix} 1 & x_{11} & x_{12} & . & . & . & x_{1n} \\ 1 & x_{21} & x_{22} & . & . & . & x_{2n} \\ . & . & . & . & . & . & . \\ . & . & . & . & . & . & . \\ . & . & . & . & . & . & . \\ 1 & x_{m1} & . & . & . & . & x_{mn} \end{pmatrix}$$

$$X \in \mathfrak{M}_{mn}(\mathbf{IR})$$

Et

$$\beta = \begin{pmatrix} \beta_0 \\ \beta_1 \\ . \\ . \\ . \\ \beta_n \end{pmatrix}$$

$$\varepsilon = \begin{pmatrix} \varepsilon_0 \\ \varepsilon_1 \\ . \\ . \\ . \\ \varepsilon_n \end{pmatrix}$$

L'estimation du modèle (1.4.8) peut se faire par les moindres carrés ou par le maximum de vraisemblance si l'on spécifie les distributions des termes d'erreur u et v.

a- Le Maximum de Vraisemblance

Suivant Aigner, Lovell et Schmidt [53], on prend une distribution normale pour v, c'est à dire que $v \sim N(0, \sigma_v^2)$ et une distribution normale centrée, tronquée à gauche en zéro

pour u , c'est à dire $u \sim |N(0, \sigma_v^2)|$. La densité jointe pour v et u sachant que les deux distributions sont indépendantes s'écrit :

$$g(u,v) = \frac{1}{\pi\sigma_v\sigma_u} \exp\Big[- \frac{u^2}{2\sigma_u^2} - \frac{v^2}{2\sigma_v^2}\Big] \tag{1.4.9}$$

Si on remplace v en fonction de u, on obtient :

$$g(\varepsilon) = \frac{1}{\pi\sigma_v\sigma_u} \exp\Big[- \frac{u^2}{2\sigma_u^2} - \frac{\varepsilon^2 + u^2 + 2\varepsilon u}{2\sigma_v^2}\Big] \quad -\infty \leq \varepsilon \leq +\infty \tag{1.4.10}$$

Calculons maintenant la densité de ε, en intégrant l'équation (1.3.10) par rapport à u . On obtient :

$$g(\varepsilon) = (\frac{2}{\sigma}) g^*(\frac{\varepsilon}{\sigma})\Big[1 - G^*(\frac{\varepsilon\lambda}{\sigma})\Big] \tag{1.4.11}$$

Avec
$\sigma^2 = \sigma_v^2 + \sigma_u^2$
Et
$\lambda = \sigma_u/\sigma_v$.
G^* désigne la fonction de répartition d'une distribution normale centrée réduite et g^* sa densité. et on a :

$$E(\varepsilon) = -E(u) = -\frac{\sqrt{2}}{\sqrt{\pi}}\sigma_u \tag{1.4.12}$$

$$Var(\varepsilon) = Var(v-u) = \Big[\frac{\pi - 2}{\pi}\Big]\sigma_u^2 + \sigma_v^2 \tag{1.4.13}$$

Remarque : la paramétrisation λ joue un rôle important, en effet il représente une mésure de la variabilité relative de deux sources d'inefficacité. De plus $\lambda^2 \to 0 \Rightarrow \sigma_u^2 \to 0$ $ousigma_v^2 \to +\infty$, ce qui signifie que les chocs aléatoires dominent dans l'explication de l'inefficacité. De même, lorsque $\sigma_v^2 \to 0$, les écarts à la frontière sont alors, essentiellement dus à l'inefficacité technique.
Considérons maintenant m observation de PIB, Alors le logarithme de la vraisemblance de (1.3.7) est donné par :

$$\ln L(y/\beta, \lambda, \sigma^2) = m \ln\Big(\frac{\sqrt{2}}{\sqrt{\pi}}\Big) + m\ln(\sigma^{-1}) + \sum_{i=1}^{m} \ln\Big[1 - G^*(\frac{\varepsilon\lambda}{\sigma})\Big] - \frac{1}{2\sigma^2}\sum_{i=1}^{m}\varepsilon_i^2 \tag{1.4.14}$$

Les estimateurs correspondants au vecteur β et aux paramètres λ et σ sont les solution du systèmes des équations de vraisemblance suivant :

$$\begin{cases} \frac{\partial \ln L}{\ln \beta_i} = \frac{1}{\sigma^2}\sum_{i=1}^{m}(y_i - \beta_i x_i)x_i + \frac{\lambda}{\sigma^2}\sum_{i=1}^{m}\frac{x_i g^*(\frac{(y_i-\beta_i x_i)\lambda}{\sigma})}{1-G^*(\frac{(y_i-\beta_i x_i)\lambda}{\sigma})} \\ \frac{\partial \ln L}{\ln \lambda} = -\frac{1}{\sigma^2}\sum_{i=1}^{m}\frac{(y_i-\beta_i x_i) g^*(\frac{(y_i-\beta_i x_i)\lambda}{\sigma})}{1-G^*(\frac{(y_i-\beta_i x_i)\lambda}{\sigma})} \\ \frac{\partial \ln L}{\ln \sigma^2} = -\frac{n}{2\sigma^2} + \frac{1}{2\sigma^4}\sum_{i=1}^{m}(y_i - \beta_i x_i)^2 + \frac{\lambda}{2\sigma^3}\sum_{i=1}^{m}\frac{(y_i-\beta_i x_i) g^*(\frac{(y_i-\beta_i x_i)\lambda}{\sigma})}{1-G^*(\frac{(y_i-\beta_i x_i)\lambda}{\sigma})} \end{cases} \tag{1.4.15}$$

La résolution de ce système est peut se faire au moyen d'algorithmes d'optimisation.
Soit $\Theta = (\beta, \lambda, \sigma^2)$ le vecteur des paramètres qui maximise (1.3.14). Pour avoir les écarts types des coefficients estimés $\Theta = (\beta, \lambda, \sigma^2)$ par le maximum de vraisemblance, solution de (1.3.15), on prendra l'inverse de la quantité : $E\Big(\frac{\partial \ln L}{\partial\Theta\Theta'}\Big)$.

b- La Méthode des moindres carrés et méthode des moments

En centralisant le bruit u, la reparamétrisation du modèle (1.3.7) devient :

$$y_i = \beta_0^{'} + \sum_{j=1}^{n} \beta_j x_{ij} + \varepsilon_i^{'} \tag{1.4.16}$$

Avec $\beta_0^{'} = \beta_0 - E(u)$ et $\varepsilon_i^{'} = v_i - (u_i - E(u)) = \varepsilon + E(u) = \varepsilon - E(\varepsilon)$
L'estimation de (1.4.16) peut se faire en deux étapes :
Étape 1 : Étant donné que la distribution de $\varepsilon_i^{'}$ est symétrique, on estime (1.4.16) par la méthode des moindres carrés ordinaires ; tous les β_j seront sans biais ;
Étape 2 : On identifie complètement la frontière en estimant β_0 et donc μ. Pour ce faire, il faut spécifier une distribution particulière pour chacun des termes d'erreur. On peut alors estimer μ par la méthode des moments et, par la suite β_0 (Aigner et ali, [30]).
Comme précédemment on suppose que $v \sim N(0, \sigma_v^2)$ et $u \sim |N(0, \sigma_v^2)|$, alors les moments de $\varepsilon_i^{'}$ (moments centralisés de ε_i) d'ordre 1,2 et 3, sont données par :

$$\mu_1 = E(\varepsilon_i^{'}) = E(\varepsilon_i - E(\varepsilon)) = 0 \tag{1.4.17}$$

$$\mu_2 = Var(\varepsilon_i^{'}) = Var(\varepsilon_i - E(\varepsilon)) = \Big[\frac{\pi - 2}{\pi}\Big]\sigma_u^2 + \sigma_v^2 \tag{1.4.18}$$

$$\mu_3 = E((\varepsilon_i^{'})^3) = E((\varepsilon_i - E(\varepsilon))^3) = \Big(1 - \frac{4}{\pi}\Big)\frac{\sqrt{2}}{\sqrt{\pi}}\sigma_u^3 \tag{1.4.19}$$

À partir du vecteur résiduel $\hat{\varepsilon}^{'}$ de (1.4.16), on obtenient une estimation des moments $\hat{\mu}_2$ et $\hat{\mu}_3$ et donc de $\hat{\sigma}_u^2$ et $\hat{\sigma}_v^2$. Les variances $\hat{\sigma}_u^2$ et $\hat{\sigma}_v^2$ sont estimées de manière convergente par :

$$\hat{\sigma}_u^2 = \Big(\hat{\mu}_3 \frac{\sqrt{\pi}}{\sqrt{2}} \frac{\pi - 2}{\pi}\Big)^{2/3} \tag{1.4.20}$$

$$\hat{\sigma}_v^2 = \hat{\mu}_2 - \Big[\frac{\pi - 2}{\pi}\Big]\sigma_u^2 \tag{1.4.21}$$

Un estimateur convergent de β_0 est dans ces conditions donné par :

$$\hat{\beta}_0 = \hat{\beta}_0^{'} + \frac{\sqrt{2}}{\sqrt{\pi}}\hat{\sigma}_u^2 \tag{1.4.22}$$

Le problème ici est que l'estimateur obtenu par la méthode des moments peut ou ne pas exister (Olson, Schmidt, Waldman, [53]). Il en sera ainsi si $\hat{\mu}_3$ est positif ou quand la valeur de $\hat{\sigma}_u^2$ est négative ; ce qui est absurde. Il en sera de même, lorsqu'il est possible à partir de certaines valeurs de $\hat{\mu}_2$ et $\hat{\mu}_3$ de trouver des valeurs négatives de $\hat{\sigma}_v^2$

1.4.3 Modélisation non paramétrique du PIB

Dans le cadre de ces méthodes, il n'y a pas une spécification particulière de la fonction de production. Ces méthodes découlent directement des travaux initiaux de Farrell [54], elles permettent de construire une frontière continue par morceaux. En d'autres termes, on commence d'abord par déterminer les unités de décision efficaces, ensuite on déduit à partir de ces dernières une frontière par extrapolation linéaire ou non linéaire. Dans ce paragraphe, nous nous limiterons à décrire deux modèles les plus employés dans la littérature : le modèle CCR (Charnes, Cooper et Rhodes) et le modèle BCC (Banker, Charnes

et Cooper), qui sont deux variantes du modèle général, communément appelé DEA (data envelopment analysis), voir (Seiford et Thrall [55])et (Lewin et Seiford [56]).
Dans les deux modèles, on distingue :
- les modèles dits "orientés inputs", si l'on étudie l'efficacité en termes d'inputs, c'est à-dire si l'on s'intéresse à l'inefficacité en terme d'excès d'inputs ;
- les modèles dits "orientés outputs" si l'on veut analyser l'efficacité en termes d'outputs, c'est-à-dire si l'on souhaite appréhender l'inefficacité par l'insuffisance d'outputs.

a- Le modèle CCR

Hypothèses du modèle
- il existe une forte convexité de l'ensemble de production ;
- la technologie est à rendements constants ;
- il existe une libre disposition des inputs et des outputs.

On considère ici, le modèle dit ''orienté input". Suivant Coelli [57], on dispose de n inputs et m outputs pour chacune des k unités de décision. On note : $x_i \in \mathfrak{M}_{n,1}(\mathbf{IR})$ le vecteur des inputs pour l'unité de décision i, $y_i \in \mathfrak{M}_{m,1}(\mathbf{IR})$ le vecteur des outputs pour l'unité de décision i, $X \in \mathfrak{M}_{n,k}(\mathbf{IR})$ la matrice des inputs, $Y \in \mathfrak{M}_{m,k}(\mathbf{IR})$ la matrice des outputs et $v^{'} \in \mathfrak{M}_{n,1}(\mathbf{IR})$ et $u^{'} \in \mathfrak{M}_{m,1}(\mathbf{IR})$ sont respéctivement les vecteurs des pondérations associées aux inputs et aux outputs.
Le problème revient donc pour chaque unité de décision, à déterminer les pondérations optimales en résolvant le problème de programmation mathématique suivant :

$$\begin{cases} \max_{u,v} \left(\frac{uy_i}{vx_i}\right) \\ S.C \\ \frac{uy_j}{vx_j} \leq 1, \ j = 1, 2, ...k \\ v \geq 0_{\mathbf{IR}^n} \\ u \geq 0_{\mathbf{IR}^m} \end{cases} \tag{1.4.23}$$

Le problème au dessus, avec cette forme fractionnelle, c'est qu'elle est difficile à optimiser ; sa résolution peut admettre une infinité de solutions. Elle peut néanmoins être linéarisée si l'on définit une contrainte selon laquelle $\mu x_i = 1$. Le programme s'écrit alors :

$$\begin{cases} \max_{\mu,\vartheta}(\vartheta y_i) \\ S.C \\ \mu x_i = 1 \\ \vartheta y_j - \mu x_j \leq 0, \ j = 1, 2, ...k \\ \mu \geq 0_{\mathbf{IR}^n} \\ \vartheta \geq 0_{\mathbf{IR}^m} \end{cases} \tag{1.4.24}$$

Où μ et ϑ sont des nouvelles unités de décision obtenues à partir de u et v pour que le programme (1.4.23) soit linéaire.
Le programme dual de (1.4.24) est comme suit :

$$\begin{cases} \min_{\theta}(\lambda\theta) \\ S.C \\ -y_i + Y\lambda \geq 0, \ j = 1, 2, ...k \\ -x_i + X\lambda \geq 0, \ j = 1, 2, ...k \\ \lambda \geq 0_{\mathbf{IR}^k} \end{cases} \tag{1.4.25}$$

Où $\lambda \in \mathfrak{M}_{k,1}(\mathbf{IR})$ est un vecteur de constantes appelées multiplicateurs. θ est un scalaire qui représente le score d'efficacité technique de la i-ème unité de décision $\theta \leq 1$. Si $\theta = 1$, alors l'unité de décision observée se situe sur la frontière, c'est à dire qu'elle est efficace au sens de Farrell ; sinon $\theta < 1$, alors, on a l'existence d'une inefficacité technique.

b- Le modèle BCC

L'hypothèse des rendements constants n'est vraiment appropriée que si l'entreprise opère à une échelle optimale. Ce qui n'est pas toujours le cas (concurrence imparfaite, contraintes financières, etc.). Banker, Charnes et Cooper [58], ont proposé un modèle qui permet de déterminer, si la production se fait dans une zone de rendements croissants, constants, ou décroissants. Leur modèle conduit à la décomposition de l'efficacité technique en efficacité technique pure et en efficacité d'échelle. L'hypothèse des rendements d'échelle constants, conduit à la mesure de l'efficacité totale ; l'hypothèse de rendements d'échelle variables conduit à celle de l'efficacité technique pure. Ainsi, le modèle CCR peut être modifié en tenant compte de l'hypothèse des rendements variables à l'échelle. Il suffit pour cela d'ajouter une contrainte : $\mathbf{1}_k^{'}\lambda = 1$ au programme (1.4.24) ; avec $\mathbf{1}_k \in \mathfrak{M}_{k,1}(\mathbf{IR})$ est un vecteur unitaire. on obtient alors :

$$\begin{cases} \min_\theta(\lambda\theta) \\ S.C \\ -y_i + Y\lambda \geq 0, \ \ j = 1, 2, ...k \\ -x_i + X\lambda \geq 0, \ \ j = 1, 2, ...k \\ \mathbf{1}_k^{'}\lambda = 1 \\ \lambda \geq 0_{\mathbf{IR}^k} \end{cases} \tag{1.4.26}$$

1.4.4 Critiques des méthodes paramétriques et non paramétriques

La mesure de l'efficacité technique permet d'évaluer dans quelle mesure l'activité observée d'une unité productive se situe à la frontière, ou en deçà, de son ensemble de production. Le « degré d'efficacité » mesure alors l'importance de l'écart par rapport à la frontière. Cela suppose que la frontière de production des pays (respéctivement entreprises) soit connue ce qui n'est pas le cas en pratique. Deux méthodes différentes apportent une réponse à cette difficulté : l'approche paramétrique suppose qu'il existe une forme fonctionnelle a priori de la fonction de production (Cobb-Douglas, Translog, etc.) qui sera estimée économétriquement à partir des données de l'échantillon ; l'approche non-paramétrique ne fait pas d'hypothèse concernant la forme fonctionnelle de la fonction de production, elle construira la frontière de production en utilisant la programmation linéaire, de telle sorte qu'aucun point observé ne se trouve en dehors de celle-ci. La critique principale faite à cette seconde approche est qu'elle ne prend pas en compte l'influence d'éventuelles erreurs de mesures des données. Dans le tableaux ci dessous, nous résumons les avantages et les inconvénients de ces méthodes.

Critique	Méthodes non paramétriques	Méthodes paramétriques
Avantages	1- Technologie non spécifiée par une relation fonctionnelle ; 2- Décomposition facile des inefficiences technique, allocative et d'échelle.	1- Les inefficiences déduites peuvent avoir des propriétés statistiques ; 2- Prise en compte des aléas autres que l'inefficience (frontières stochastiques).
inconvénients	1- Les inefficiences déduites n'ont pas de propriété statistique ; 2- Les inefficiences déduites n'ont pas de propriété statistique ; Les erreurs de mesures et/ou d'oubli de variables peuvent affecter les mesures d'inefficience.	1-Nécessité de représenter la technologie par une forme paramétrique particulière ; 2- La décomposition des différentes composantes de l'inefficience n'est pas toujours possible, en particulier pour des technologies multiproduits.

TABLE 1.1 – avantages et les inconvénients des méthodes paramétriques et non paramétriques

1.5 Modèles dynamiques stochastiques en macroéconomie

L'utilisation des approches stochastiques en macrodynamique économique sont plus adaptées avec la réalité économique. En effet, vue que plusieurs chocs sur la technologie, l'emploi, etc..., l'hypothèse de certitude des prévision des agrégats macroéconomique est rejetée. Par suite l'incertain des prévision macroéconomique est fortement coherent avec la réalité économique et les faits stylisés de la croissance macroéconomique.

1.5.1 Généralité sur L'approche Bayésienne et choix des croyances a priori

a- Généralités

L'approche bayésienne propose un cadre rigoureux pour (i) formaliser nos croyances(3) a priori et (ii) déterminer comment celles-ci doivent être mises à jour une fois que les données sont observées. Les croyances, a priori ou a posteriori, sont représentées à l'aide d'une densité de probabilité jointe sur les paramètres d'un modèle. Cette densité jointe caractérise l'incertitude quant au processus générateur des données(DGP, pour Data Generating Process), en décrivant une famille (un continuum) de modèles. Nous noterons l'a priori (la densité jointe définit notre incertitude) sur un vecteur de paramètres $\theta_M = (\theta_1^M, ..., \theta_q^M)$ associé à un modèle paramétrique M, comme suite :

$$p_0(\theta_M/M) \tag{1.5.1}$$

Soit $Y_T^* = \{y_t^*\}_{i=1}^T$ un échantillon d'observations, où y_t^* est un vecteur de m variables. Nous nous limiterons au cas où t représente le temps. La vraisemblance est la densité de

l'échantillon conditionnellement au modèle et ses paramètres ; on note :

$$L(\theta_M, Y_T^*, M) = p(Y_T^*/\theta_M, M) \tag{1.5.2}$$

L'estimateur du maximum de vraisemblance (MV) sélectionne le paramètre θ_M définissant le DGP qui a le plus probablement généré les données. il représente la valeur des paramètres qui rend le plus probable l'occurrence de l'échantillon à notre disposition.
Lorsque nous disposons des densités $p_0(\theta_M/M)$ qui caractérise l'information postulée a priori, et $p(Y_T^*/\theta_M, M)$ qui caractérise l'information apportée par les données, et puisque si deux sources sont orthogonales, alors on peut trouver la densité postérieure de θ_M sachant les données Y_T^*, en utilisant le théorème de Bayes, on obtient :

$$p_1(\theta_M/Y_T^*, M) = \frac{p(Y_T^*/\theta_M, M)p_0(\theta_M/M)}{p(Y_T^*/M)} \tag{1.5.3}$$

Avec

$$p(Y_T^*/M) = \int_{\Theta_M} p(Y_T^*/\theta_M, M) \times p_0(\theta_M/M)d\theta_M \tag{1.5.4}$$

On représente nos croyances a posteriori en exhibant les propriétés de la distribution a posteriori. Par exemple, la comparaison des variances a priori et a posteriori peut nous renseigner sur l'information apportée par les données, relativement à celle contenue dans nos croyances a priori. On définit les variances a priori et a posteriori à partir des des éléments diagonaux des matrices de covariance a priori et a posteriori, alors on a :

$$Var_0(\theta_M) = \int_{\Theta_M} \theta_M\theta_M' p_0(\theta_M/M)d\theta_M - (\int_{\Theta_M} \theta_M p_0(\theta_M/M)d\theta_M)(\int_{\Theta_M} \theta_M p_0(\theta_M/M)d\theta_M)' \tag{1.5.5}$$

$$Var_1(\theta_M) = -(\int_{\Theta_M} \theta_M p_1(\theta_M/Y_T^*, M)d\theta_M)(\int_{\Theta_M} \theta_M p_1(\theta_M/Y_T^*, M)d\theta_M)' + \int_{\Theta_M} \theta_M\theta_M' p_1(\theta_M/Y_T^*, M)d\theta_M \tag{1.5.6}$$

Si $Var_1(\theta_M) << Var_0(\theta_M)$, cela signifie que les données apportent une information supplémentaire sur ce paramètre, relativement à l'information a priori.
Pour communiquer nos croyances a posteriori nous désirons souvent recourir à un média plus synthétique en résumant, à l'image de l'approche classique, la distribution postérieure par un point.On parle alors d'estimation ponctuelle. Réduire la distribution postérieure à un point s'apparente à un choix en univers incertain. Il est donc naturel de construire une estimation ponctuelle en minimisant l'espérance postérieure d'une fonction de perte :

$$\hat{\theta}_M = \arg\min_a \int_{\Theta_M} L(a, \theta_M)p_1(\theta_M/Y_T^*, M)d\theta_M \tag{1.5.7}$$

où $L(a, \theta_M)$ est une fonction associant une perte au choix a si la vraie valeur du paramètre est θ_M.
Après l'estimation, le modèle peut être utilisé pour construire des prévisions et des fonctions de réponse. À l'image de l'estimation des paramètres, le paradigme bayésien ne fournit pas des prévisions ponctuelles mais des densités prédictives. Supposons que nous souhaitions établir des prédictions sur y_{T+1}^* un vecteur d'endogènes non encore observées, le but est de construire la densité (prédictive) de ce vecteur. On a :

$$p(y_{T+1}^*/Y_T^*) = \int_{\Theta_M} p(y_{T+1}^*/\theta_M, Y_T^*)p_1(\theta_M/Y_T^*)d\theta_M \tag{1.5.8}$$

Exemple : modèle autorégressif d'ordre un (cas $\theta_M$scalaire)
Considérons le modèle autorégressif d'ordre un : $y_t^* = \theta y_{t-1}^* + \varepsilon_t \;\; t = 1, ..., T$, et $\varepsilon_t \sim_{iid} N(0, \sigma^2)$, avec σ^2 est la variance de l'innovation (connue). La distribution de y_{T+1}^* conditionnellement à θ et Y_T^* est gaussienne : $(y_{T+1}^*/Y_T^*, \theta) \sim N(\theta y_T^*, \theta, \sigma^2)$. La densité prédictive s'écrit :

$$p(y_{T+1}^*/Y_T^*) = \int_\Theta p(y_{T+1}^*/\theta, Y_T^*) p_1(\theta/Y_T^*) d\theta \tag{1.5.9}$$

À partir de cette densité prédictive, on peut construire une prédiction ponctuelle des variables en se donnant une fonction de perte, calculer un intervalle contenant $\alpha\%$ de la distribution de y_{T+1}^*. En confrontant la densité prédictive aux réalisations effectives des variables, on peut alors évaluer dans quelle mesure notre modèle tend à sur-estimer ou sous-estimer, par exemple, le taux de croissance à un trimestre du PIB par tête.
Cette comparaison peut fournir un critère d'évaluation du modèle. Si on se rend compte que les réalisations effectives d'une variable se situent systématiquement dans les queues de la densité prédictive, alors on peut conclure que le modèle est mal spécifié vis-à-vis de cette variable.

b- Le choix des croyances a priori

Le choix des croyances a priori est essentiel, dans la mesure où il détermine partiellement les résultats (surtout pour un échantillon de taille réduite). La subjectivité de l'économètre ne peut intervenir que dans la première étape d'élicitation de l'a priori, les étapes suivantes (l'évaluation de la vraisemblance,...) sont automatiques et nécessairement objectives sauf quant au choix éventuel d'un critère de perte pour l'estimation ponctuelle. La question du choix des croyances a priori est donc crucial. Il est donc important de bien comprendre le rôle de la densité a priori dans les résultats, par exemple en menant des exercices de sensibilité aux croyances a priori. Ces expériences, en donnant une idée du rôle des priors, dévoilent implicitement la forme de la vraisemblance. L'expérience la plus extrême est de considérer un a priori non informatif, c'est-à-dire le cas où nous n'avons aucune croyance a priori sur les paramètres du modèle. De façon assez surprenante, les statisticiens bayésiens ne parviennent pas à s'accorder sur une chose aussi essentielle que la caractérisation de ce non savoir. Aujourd'hui, les progrès de l'informatique nous permettent d'adopter une formulation non paramétrique plus générale. Par exemple nous pourrions caractériser nos croyances a priori sur chaque paramètre en spécifiant les quantiles de chaque distribution.
Jeffrey proposa un prior non informatif [37] basé sur la matrice d'information de Fisher : $p_0(\theta)$ est proportionnelle à $\sqrt{I(\theta)}$, avec :

$$I(\theta) = E\Big[\Big(\frac{\partial p(Y_T^*, \theta)}{\partial \theta}\Big)\Big(\frac{\partial p(Y_T^*, \theta)}{\partial \theta}\Big)'\Big]$$

La matrice d'information de Fisher quantifie l'information amenée par le modèle et les données sur le paramètre θ. En favorisant les valeurs de θ pour lesquelles l'information de Fisher est plus grande, on diminue l'influence de la loi a priori puisque l'information véhiculée par celle-ci est peu différente de l'information provenant de la vraisemblance. La définition de la densité a priori est donc liée à la courbure de la vraisemblance. Cette densité a priori est invariante à toute reparamétrisation (continue) du modèle (voir[38]).

1.5.2 Modèle VAR

Le modèle VAR gaussien, a l'avantage d'être un outil couramment utilisé en macroéconomie (voir [39]). Nous considérons un modèle $VAR(p)$ pour caractériser le vecteur $1 \times m$ de variables endogènes y_t^* observées :

$$y_t^* = \sum_{i=1}^{p} y_{t-i}^* A_i + \varepsilon_t \tag{1.5.10}$$

Où $\{A_i\}$ est une suite de matrice carrée de taille m, et ε_t est un bruit blanc gaussien, de dimension $1 \times m$ d'espérance nulle et de variance Σ. On note $Y_T^* = \{y_t^*\}_{i=-p+1}^{T}$ les données à notre disposition et on note z_t la concaténation horizontale des vecteurs lignes $y_{t-1}^*, y_{t-2}^*, ..., y_{t-p}^*$. En concaténant verticalement les vecteurs lignes y_t^*, z_t et ε_t, pour $t = 1, ..., T$ on obtient la représentation matricielle suivante du modèle $VAR(p)$:

$$Y = ZA + E$$

Où Y et E sont des matrices $T \times m$, Z est une matrice $T \times (mp)$ et $A = (A_i', A_2', ..., A_p')'$ la matrice de type $(mp) \times m$ regroupant les coefficients auto-régressifs. La vraisemblance associée à ce modèle linéaire gaussien est donnée par :

$$L(A, \Sigma, Y_T^*) = (2\pi)^{\frac{-mT}{2}} |\Sigma|^{\frac{-T}{2}} \times e^{\frac{-1}{2} tr[(Y-ZA)\Sigma^{-1}(Y-ZA)']} \tag{1.5.11}$$

L'estimateur du maximum de vraisemblance (MCO) est défini par :

$$\hat{A} = (Z'Z)^{-1}Z'Y \tag{1.5.12}$$

Et

$$\hat{\Sigma} = T^{-1}(Y - Z\hat{A})'(Y - Z\hat{A}) \tag{1.5.13}$$

Aux constantes d'intégration près on reconnaît ici les fonctions de densité de probabilité d'une gaussienne matricielle et d'une inverse Wishart (voir [40]). La vraisemblance se réécrit donc sous la forme suivante :

$$L(A, \Sigma, Y_T^*) = (2\pi)^{\frac{-m(T+k)}{2}} |Z'Z|^{\frac{-m}{2}} \times f_{MN_{k,m}}(A; \hat{A}, Z'Z, \Sigma) \times \frac{2^{\frac{\nu m}{2}} \pi^{\frac{m(m-1)}{4}} \prod_{i=1}^{m} \Gamma\left(\frac{\nu+1-i}{2}\right)}{|\hat{S}|^{\frac{\nu}{2}}} \tag{1.5.14}$$

$$\times f_{iW_m}(\Sigma; \hat{S}, \nu) \tag{1.5.15}$$

Où $k = mp$, $\nu = T - k - m - 1$ et $\hat{S} = T\hat{\Sigma}$. nous déduisons donc que la vraisemblance du VAR(p) est proportionnelle au produit de la densité d'une normale matricielle et d'une loi inverse Wishart : $L(A, \Sigma, Y_T^*) \propto f_{MN_{k,m}}(A; \hat{A}, Z'Z, \Sigma) \times f_{iW_m}(\Sigma; \hat{S}, \nu,)$
Cette propriété va nous aider à poser une forme de la densité a priori telle que nous puissions obtenir une expression analytique de la densité postérieure.

A priori non informatif
Dans ce paragraphe nous supposons que nos croyances sont non informatives, et soit :

$$p_0(A, \Sigma) = |\Sigma|^{-\frac{m+1}{2}} \tag{1.5.16}$$

La densité a posteriori satisfait donc :

$$p(A, \Sigma / Y_T^*) \propto |\Sigma|^{-\frac{m+1}{2}} \times L(A, \Sigma, Y_T^*) \tag{1.5.17}$$

La densité jointe postérieure est donc proportionnelle au produit d'une loi normale multivariée et d'une loi inverse Wishart :

$$p(A, \Sigma; Y_T^*) \propto f_{MN_{k,m}}(A; \hat{A}, Z'Z, \Sigma) \times f_{iW_m}(\Sigma; \hat{S}, \tilde{\nu}) \tag{1.5.18}$$

Avec $\tilde{\nu} = T - k$. Ainsi, la densité postérieure s'écrit sous la forme suivante :

$$A/\Sigma, Y_T^* \sim MN_{k,m}(\hat{A}, \Sigma, (Z'Z)^{-1}) \tag{1.5.19}$$

$$\Sigma/Y_T^* \sim iW_m(\hat{S}, \tilde{\nu}) \tag{1.5.20}$$

Il n'est pas surprenant de constater que la distribution postérieure de A (conditionnelle à la matrice de variance covariance) est centrée sur l'estimateur du maximum de vraisemblance, puisque notre a priori est non informatif. Nous pourrions montrer, en intégrant par rapport à Σ que la distribution postérieure (marginale) de A est une version matricielle de la loi de Student (voir [41]). On obtient la densité marginale postérieure de Y_T^* en intégrant le noyau postérieur successivement par rapport à Σ et A :

$$p(Y_T^*) = (2\pi)^{\frac{-m(T+k)}{2}} |Z'Z|^{\frac{-m}{2}} \times \frac{2^{\frac{\tilde{\nu}m}{2}} \pi^{\frac{m(m-1)}{4}} \prod_{i=1}^{m} \Gamma\left(\frac{\tilde{\nu}+1-i}{2}\right)}{|\hat{S}|^{\frac{\tilde{\nu}}{2}}} \tag{1.5.21}$$

Cette quantité nous renseigne sur la qualité d'ajustement du modèle VAR(p). On note que la densité marginale de Y_T^* est une fonction décroissante de la taille des erreurs $|\hat{S}|$. Dans cet exemple, nous pouvons caractériser la distribution postérieure analytiquement. Notons néanmoins que même si nous connaissons l'expression analytique de la distribution de Σ et A, la construction des densités prédictives nécessite une approche par simulations(voir [42]), puisque les prévisions sont des fonctions non linéaires des matrices auto-régressives (dont nous connaissons la distribution postérieure). L'intérêt pratique de l'approche bayésienne peut paraître faible dans ce cas, dans la mesure où l'espérance postérieure n'est pas différente de l'estimateur du maximum de vraisemblance.

1.5.3 Modèles DSGE

Dans cette section nous présentons de façon générale les modèles DSGE, puis soulignons les problèmes que peut poser leur estimation. Après, en présentant l'une des méthodes de simulation de Monte Carlo utilisées pour calculer les croyances postérieures.

a- Résolution et vraisemblance

Nous limitons notre attention aux modèles DSGE que nous pouvons écrire sous la forme suivante :

$$E_t[F_\theta(y_{t+1}, y_t, y_{t-1}, \varepsilon_t)] = 0 \tag{1.5.22}$$

Où $\varepsilon_t \sim iid(0, \Sigma)$ est un vecteur aléatoire dans $\mathbf{IR}^r$ des innovations structurelles, $y_t \in \Lambda \subset \mathbf{IR}^n$ est un vecteur regroupant les variables endogènes, $F_\theta : \Lambda^3 \times \mathbf{IR}^n \to \Lambda$ une fonction réelle dans $\mathbf{C}^2$ paramétrée par un vecteur réel $\theta \in \Theta \subset \mathbf{IR}^d$ regroupant l'ensemble des paramètres structurels du modèle. La fonction F est simplement l'ensemble des équations qui définissent un modèle ; on a autant d'équations que de variables endogènes. Le vecteur des variables endogènes, y_t inclut des variables d'état (endogènes ou exogènes), des

variables de choix et des variables statiques. On supposera qu'il est possible d'exhiber une unique solution stable et invariante du modèle décrit par l'équation (1.5.21) :

$$y_t = H_\theta(y_{t-1}, \varepsilon_t) \tag{1.5.23}$$

qui exprime les variables endogènes en fonction du passé et des chocs structurels contemporains. La fonction paramétrée H_θ regroupe les policy rules et les équations de transition (voir [43]). La solution (1.5.22), en décrivant une récurrence stochastique non linéaire, définit la distribution jointe d'un ensemble de variables. Pour estimer les paramètres θ du modèle, ou d'un sous ensemble des paramètres, nous devons évaluer la vraisemblance associée au modèle (1.5.21) ou à sa forme réduite (1.5.22). Afin d'amener le modèle aux données on peut l'écrire sous une forme état-mesure :

$$y_t^* = Zy_t + \eta_t = ZH_\theta(y_{t-1}, \varepsilon_t) + \eta_t \tag{1.5.24}$$

Où y_t^* est un vecteur $m \times 1$, avec $r \leq m < n$, regroupant les variables observées et Z est une matrice de sélection $m \times n$. On peut éventuellement augmenter l'équation de mesure d'un bruit blanc multivarié,η_t, représentant l'inadéquation des variables théoriques avec les variables observées, ou plus simplement une erreur de mesure. On note $Y_T^* = \{y_t^*\}_{i=-p+1}^T$ l'échantillon à notre disposition et $\psi \in \Psi \subset \mathbf{IR}^{q+\frac{n(n+1)}{2}+\frac{r(r+1)}{2}}$ le vecteur des paramètres du modèle état-mesure (θ, Σ) et éventuellement la matrice de variance-covariance de η_t. La vraisemblance est la densité de l'échantillon, conditionnellement aux paramètres ψ et au modèle défini par(1.5.23) :

$$L(\psi, Y_T^*) = p(Y_T^*/\psi) = \prod_{t=1}^{T} p(y_t^*/Y_{t-1}^*, \psi) \tag{1.5.25}$$

Or L'évaluation de la densité de y_t^* conditionnellement à Y_{t-1}^* généralement n'est pas directe, dans la mesure où y_t^* dépend de variables endogènes inobservables. Nous pouvons néanmoins utiliser la relation suivante :

$$p(y_t^*/Y_{t-1}^*, \psi) = \int_\Lambda p(y_t^*/y_t, \psi) p(y_t/Y_{t-1}^*, \psi) dy_t \tag{1.5.26}$$

La première densité sous l'intégrale est specifiée par l'équation de mesure (1.5.23). L'évaluation de la densité de la prévision des variables latentes, conditionnellement à l'information disponible en $t-1$, est moins directe, et on doit utiliser un filtre de Kalman. Il s'agit d'une procédure récursive. À chaque instant (entre 1 et T) on forme une prévision des variables latentes (y_t sachant Y_{t-1}^*), en utilisant l'équation d'état (1.5.22) et une estimation initiale des variables latentes (y_{t-1} sachant Y_{t-1}^*) puis on corrige cette prévision quand une nouvelle observation y_t^* augmente l'ensemble d'information. On peut interpréter cette démarche comme une estimation récursive bayésienne du vecteur des variables latentes. En initialisant les variables latentes avec la densité associée à la distribution ergodique des variables latentes définie par l'équation d'état (1.5.22), $p(y_0/Y_0^*, \psi) = p(y_0, \psi) = p(y_\infty, \psi)$, la récursion s'écrit de la façon suivante :

$$p(y_t/Y_{t-1}^*, \psi) = \int_\Lambda p(y_t/y_{t-1}, \psi) p(y_{t-1}/Y_{t-1}^*, \psi) dy_{t-1} \tag{1.5.27}$$

$$p(y_t/Y_t^*, \psi) = \frac{p(y_t^*/y_t, \psi) p(y_t/Y_{t-1}^*, \psi)}{\int_\Lambda p(y_t^*/y_t, \psi) p(y_t/Y_{t-1}^*, \psi) dy_t} \tag{1.5.28}$$

Puisque nous pouvons, au moins théoriquement, évaluer la vraisemblance associée au modèle DSGE, nous devrions être capable d'estimer ses paramètres. Malheureusement, les équations (1.5.25), (1.5.26) et (1.5.27) nécessitent l'évaluation d'intégrales (voir [44]) dans l'espace des variables d'état. Quand le nombre de variables latentes agmente il devient très coûteux d'évaluer ces intégrales (on parle de curse of dimensionality). En pratique,même pour des modèles de dimensions modestes, l'évaluation de la vraisemblance est difficile. Ainsi la dynamique de la distribution des variables latentes est complétement caractérisée par la dynamique de l'espérance et de la variance des variables latentes. Il y a une autre possibilité, conciste à évaluer les moments en recourant à des simulations par Monte-Carlo.

b- Les méthodes de Monte-Carlo à chaînes de Markov

Cette classe d'algorithmes permet de générer des variables aléatoires suivant approximativement la loi a posteriori, lorsque cette dernière n'est pas disponible. Elle évite donc l'appel à une fonction d'importance, h, souvent difficile à déterminer pour les modèles DSGE (voir [45]). On cherche ainsi à définir une chaîne de Markov dont la distribution ergodique est approximativement le noyau a posteriori. Si cette chaîne existe, la méthode d'échantillonnage est grossièrement définie comme suit : dans un premier temps, on initialise (arbitrairement) la chaîne de Markov. Dans un second temps, on génère les $\theta^{(s)}$ à partir de cette chaîne. À l'issue d'un certain nombre de tirages n_0, on dispose de réalisations de variables aléatoires $\theta^{(s)}, s = n_0, ..., n$ approximativement distribuées comme la distribution a posteriori.

Chaînes de Markov

Une chaîne de Markov est une suite de variables aléatoires continues à valeurs dans Θ, $\theta^{(s)}, s = n_0, ..., n$ générée par un processus de Markov. Une suite de variables aléatoires est générée par un processus de Markov (d'ordre 1) si la distribution de $\theta^{(s)}$ ne dépend que de $\theta^{(s-1)}$. Une chaîne de Markov est caractérisée par un noyau de transition qui spécifie la probabilité de passer de $\eta \in \Theta$ à $S \subset \Theta$. Nous notons $P(\eta, S)$ le noyau de transition, il vérifie $P(\eta, \Theta) = 1$, pour tout $\eta \in \Theta$. Si la chaîne de Markov définie par le noyau P converge vers une distribution invariante π, alors le noyau doit satisfaire l'identité suivante :

$$\pi(S) = \int_{\Theta} P(\eta, S)\pi(d\eta) \tag{1.5.29}$$

Plus généralement, avant d'atteindre la distribution ergodique π, si nous notons $P^{(s)}(\eta, S)$ la probabilité que $\theta^{(s)} \in S$ sachant que $\eta = \theta^{(s-1)}$, nous avons alors :

$$P^{(s)}(\eta, S) = \int_{\Theta} P(\nu, S)P^{(s-1)}(\eta, d\nu) \tag{1.5.30}$$

la distribution de θ s'ajuste d'itération en itération puis rejoint la distribution ergodique, $\lim_{s\to\infty} P^{(s)}(\eta, S) = \pi(S)$. L'idée est alors de choisir le noyau de transition qui nous amènera vers la distribution invariante désirée.

Algorithme de(Metropolis-Hastings)

(1) Se donner une condition initiale θ^0 telle que $K(\theta^0/Y_t^*) > 0$ et poser $s = 1$
(2) Générer un candidat (une proposition) θ^* à partir d'une densité $q(\theta^{(s-1)}, \theta^*)$
(3) Générer u dans une loi uniforme entre $[0, 1]$.
(4) Appliquer la règle suivante :

$$\theta^{(s)} = \begin{cases} \theta^*, \ si \ \alpha(\theta^{(s-1)}, \theta^*) > u \\ \theta^{(s-1)} \ \ sinon \end{cases}$$

Où

$$\alpha(\theta^{(s-1)}, \theta^*) = \min\left(1, \frac{K(\theta^*/Y_t^*)q(\theta^{(s-1)}, \theta^*)}{K(\theta^{(s-1)}/Y_t^*)q(\theta^*, \theta^{(s-1)})}\right)$$

(5) Effectuer (2-4) pour $s = 2, ..., n$.

1.5.4 Étude critique et comparative des modèles DSGE- VAR

Partons des critiques faits par Lucas [60], Sims [61], Wouters [62] et Fabrice et Patrick [63], l'approche VAR permet de modéliser la dynamique de différentes variables agrégées à l'aide d'un faible nombre de restrictions, la sélection du modèle ne s'effectuant que sur la base de critères statistiques. En revanche, les modèles DSGE offrent une modélisation parcimonieuse et rigoureuse de la dynamique économique et des anticipations. Les modèles DSGE de grande taille comportant des frictions informationnelles (c'est- à- dire des difficultés à récolter l'information pertinente pour prendre des décision) présentaient des bonnes qualités prédictives, qui dominent celles des modèles VAR. lorsqu'il s'agit d'identifier les chocs structurels et leurs effets sur la dynamique agrégée, il est préférable d'utilser les DSGE que les VAR

1.6 Conclusion

Dans ce chapitre, nous avons présenté une revue autour l'évolution de la théorie de la croissance macroéconomique via des modèles mathématiques dynamiques, déterministes et stochastiques. Nous avons commencé par une étude brève et concentré des propriétés de la fonction de production macroéconomique, et montré la relation entre la géométrie de la surface de production et la nature des inputs macroéconomiques (substituables ou bien complémentaires). Nous avons établi aussi selon l'ordre chronologique, une étude critique des importants modèles de la littérature, en commançant par les modèles théoriques simples, puis les modèles basés sur les méthodes paramétriques et non paramétriques, et nous avons terminé ce chapitre par des modèles stochastiques de la dynamiques macroéconomiques, tels que les modèles VAR et DSGE. Nous déduisons d'après notre étude, que les modèles théoriques simples, peuvent nous introduisent l'idée de croissance macroéconomique, mais ils sont incapables de faire des prédiction de cette croissance, non plus, d'expliquer l'impact des divers chocs sur l'économie. En revanche, les modèles basés sur les méthodes paramétriques, prennent en compte les inefficiences et les aléas stochastiques exogènes (chocs), mais, ils ne sont pas toujours éfficaces dans la décomposition des différentes composantes de l'inefficiences, surtout dans le cas d'une technologie muli- produits. Par contre, les modèles basés sur les méthodes non paramétriques, ont une décomposition facile des inefficiences et dans le cadre de ces modèles, nous n'avont pas besoin de donner une forme fonctionnelle de la technologie, l'utilité de ces modèles est lorsqu'on veut trouver une gestion optimale des inputs, pour optimiser les gains sur une courte période (mois, trimèstre ou année), mais, ils sont incapables de donner une gestion optimale sur une longue période, à cause de ; manque des propriétés statistiques des inefficiences et aussi de ne pas prendre en compte, l'influence des chocs exogènes sur la production. En dynamique stochastique, nous avons présenté deux modèles importants, qui sont basés sur le principe de maximum de vraisemblance ; le modèle VAR et le modèle DSGE. Nous avons conclu que les modèles DSGE donnent des bonne prédictions que les modèles VAR, lorsqu'il s'agit d'une grande taille des frictions informationnelles.

Chapitre 2

Méthodes des moindre carré avec mouvement brownien et application économétrique

2.1 Introduction

La méthode des moindres carrés, à été élaborée par Legendre et Gauss au début du XXe siècle [58-60]. Elle permet de comparer des observations, généralement entachées d'erreurs de mesure, à un modèle mathématique proposé pour décrire ces observations.
La méthode des moindres carrés permet alors de minimiser l'impact des erreurs expérimentales en ajoutant de l'information dans le processus de mesure.
Mais dans le cas d'existence d'un mouvement brownien dans le modèle, la méthode ne sera plus efficase, et les estimateurs obtenu à l'aide de cette méthode ne seront plus à variance minimale. Pour cette raison nous proposons d'étudier les modèles comportant des mouvements browniens, afin de trouver les estimateurs optimaux pour ces modèles selon des hypothèses bien précises.
Donc on s'intérèsse dans ce chapitre à l'étude des modèles de type :

$$(\mathcal{M}_i) \quad Y_{i,t} = F_i(X_{i,t}, \beta_i, u_{i,t}) + e_{i,t}$$

Où :
$Y_{i,t}$ est la variable endogène de l'individue i que l'on veut expliquer par les variables explicatives $X_{i,t} = (x_{i,1,t}, ..., x_{i,m,t})$ à l'instant t.
$\beta_i \in \mathbb{R}^{m+1}$ est le vecteur de paramètres à estimer.
$\left(u_{i,t}\right)_{t\geq 0}$ est un mouvement brownien unidimensionnel
F_i est une application différentiable par rapport à β_i et $x_{i,k,t}$, $k \in [1, m] \cap \mathbb{N}$.
$e_{i,t}$ est l'erreur de modélisation à l'instant t.
L'idée ici est de linéariser F_i, puis estimer le vecteur β_i selon la nature de l'erreur (hétéroscédastique, homoscédastique, autocorrélée) et la nature des variables explicatives (existence ou non d'une multicolinéarité).

2.2 Linéarisation du modèle $\mathcal{M}_i$

Soit $\Big\{O_i = \{(Y_{i,t}, X_{i,1,t}, ..., X_{i,m,t})/t \in \{1, ..., T\}\}, \Big\}$ l'ensemble des observations d'un individu i pour T périodes succéssives. On suppose que les observations de l'ensemble O_i sont prélévé d'un phénomène subit par un mouvement interne des points $X_{i,k,t}$ $k \in \{1, 2, ...m\}$ à chaque instant t, et que ce mouvement peut engendrer une évolution irrégulière de la variable endogène Y_i.
On modélise ce mouvement interne au cours du temps par un mouvement brownien (noté $\Big(u_{i,t}\Big)_{t\geq 0}$).
Soient P_W la mesure de Wiener et $(\Omega, \mathfrak{F}, P_W)$ un espace probabilisé.
On note par $\mathcal{E}$ l'ensemble des applications de l'espace des fonctions continues $\mathcal{C}(\mathbf{IR}^+, \mathbf{IR})$ dans l'espace $(\Omega, \mathfrak{F}, P_W)$.
On veut modéliser $Y_{i,t}$ par une fonction $F_i : \mathbf{IR}^m \times \mathbf{IR}^{m+1} \times \mathcal{E} \rightarrow \mathbf{IR}$, avec une erreur de modélisation notée $e_{i,t}$.
Soit le modèle d'évolution suivant :

$$(\mathcal{M}_i) \quad Y_{i,t} = F_i(X_{i,t}, \beta_i, u_{i,t}) + e_{i,t}$$

Pour linéariser le modèle $(\mathcal{M}_i)$, on utilise l'une des méthodes suivantes :

2.2.1 Méthode 1

On cherche une transformation bijective h telle que $h \circ F_i$ est affine. Dans ce cas le modèle $(\mathcal{M}_i)$ peut s'écrire sous la forme :

$$(\mathcal{M}_i) \quad y_{i,t} = \beta_{i,0} + \sum_{k=1}^{m} \beta_{i,k} x_{i,k,t} + \gamma_i u_{i,t} + \psi_{i,t}$$

Avec : $y_{i,t} = h(Y_{i,t})$, $x_{i,k,t}$ sont les variables explicatives obtenues après l'application de la transformation h des observations $X_{i,k,t}$ et $\psi_{i,t}$ est l'erreur de modélisation à l'instant t.
Exemple :
Si on veut modéliser par exemple la production d'une entreprise i par une fonction Cobb Dauglass, si on note $Y_{i,t}$ est le chiffre d'affaire de l'entreprise i et $X_{i,k,t}$ $k \in \{1, ., ., m\}$ sont des m facteurs de production utilisés à l'instant t.
On suppose qu'on a un déséquilibre soit dans le marché de travail ou bien dans le marché des biens et services. On modélise l'effet de ce déséquilibre sur le chiffre d'affaire de l'entreprise par un mouvement brownien géométrique.
Si F_i est une fonction Cobb Dauglass (éq (1.2.7), chapitre 1) qui contient un mouvement brownien géométrique, autrement dit :

$$F_i(X_{i,t}, \beta_{i,k}, u_{i,t}) = c_i e^{u_{i,t}} \prod_{i=1}^{m} X_{i,k,t}^{\beta_{i,k}}$$

Où $(u_{i,t})_{t\geq 0}$ est un mouvement brownien.
Pour linéariser F_i on pose $h = \ln$, alors on a :

$$\ln(F_{i,t}) = \ln(c_i) + \sum_{i=1}^{m} \beta_{i,k} \ln(X_{i,k,t}) + u_{i,t}$$

On pose donc $\beta_{i,0} = \ln(c_i)$, $x_{i,k,t} = \ln(X_{i,k,t})$.
Donc au lieu de modéliser $Y_{i,t}$ par F_i, on modélise $y_{i,t} = \ln(Y_{i,t})$ par $\ln(F_{i,t})$ et on note $\psi_{i,t}$

l'erreur de modélisation, il s'agit donc de trouvez une bonne estimation des paramétres $\beta_{i,k},\ k \in \{1,.,.,m\}$ du modèle :

$$y_{i,t} = \beta_{i,0} + \sum_{k=1}^{m} \beta_{i,k} x_{i,k,t} + \gamma_i u_{i,t} + \psi_{i,t}$$

2.2.2 Méthode 2

On propose de suivre les étapes suivantes :

1. Pour chaque instant $t \in \{1,.,.,.,T\}$, on cherche les points stationnaires $X^*_{i,t} = (X^*_{i,1,t}, X^*_{i,t}, ..., X^*_{i,m,t})$ de la fonction $F_i(X_{i,t}, \beta_i, u_{i,t})$, en utilisant les conditions de premier ordre $\nabla_{X_{i,t}} F_i = 0$.
2. Soit $X^*_{i,t}$ un point stationnaire, on pose $d_{i,t} = X_{i,t} - X^*_{i,t}$, qui représente la différence entre l'observation $X_{i,t}$ et $X^*_{i,t}$. On pose :
$$F^*_{i,t}(d) = F_i(X^*_{i,t} + d_{i,t}, \beta_i, u_{i,t})$$
On note par $<.,.>$ le produit scalaire usuel.
On considère le développement de Taylor à l'ordre 1 de F_i, on a :
$$\begin{aligned} Y_{i,t} = F^*_{i,t}(d_{i,t}) &= F^*_{i,t}(0) + < \nabla F^*_{i,t}(0), d_{i,t} > + o(\|d_{i,t}\|) \\ &= F^*_{i,t}(0) - < \nabla_{X_{i,t}} F^*_{i,t}(0), X^*_{i,t} > + < \nabla_{X_{i,t}} F^*_{i,t}(0), X_{i,t} > + o(\|d_{i,t}\|) \end{aligned}$$
3. On pose :
$$\beta_{i,0} = F^*_{i,t}(0) - < \nabla_{X_{i,t}} F^*_{i,t}(0), X^*_{i,t} >$$
$$\sum_{k=1}^{m} \beta_{i,k} X_{i,k,t} = < \nabla_{X_{i,t}} F^*_{i,t}(0), X_{i,t} >$$
On prend :
$$o(\|d_{i,t}\|) = (u_{i,t} + \psi_{i,t})\|d_{i,t}\|$$
Avec $(u_{i,t})_{t \geq 0}$ est un mouvement brownien et $\psi_{i,t}$ est l'erreur de la modélisation.
On revient donc à la même forme du modèle $(\mathcal{M}_i)$.

2.3 Hypothèses, notations et problématique

2.3.1 Hypothèses du modèle $(\mathcal{M}_i)$:

H1 les erreurs $\psi_{i,t}$ sont centrées et hétéroscédastiques, autrement dit : $E(\psi_{i,t}) = 0$ et $V(\psi_{i,t}) = \sigma^2_{i,t}$ pour tous $t \in \{1, ..., T\}$,et $i \in \{1, ...N\}$.

H2 $cov(\psi_{i,s}, \psi_{i,t}) = \sigma^2_{i,t}\delta_{s,t}$, pour tout $(s,t) \in \{1, ..., T\}^2$. Où $\delta_{s,t}$ est le symbole de Kronecker.

H3 $cov(\psi_{i,t}, \psi_{j,t}) = \sigma^2_{i,t}\delta_{i,j}$, pour tout $(i,j) \in \{1, ..., N\}^2$. Autrement dit la décorrélation entre les erreurs des individus.

Le motif de l'hypothèse $(H1)$ est que ; dans un système contenant un mouvement brownien, le déplacement de ses particules change d'un instant à un autre. Donc la droite de régression va changer pour chaque instant t, par suite les erreurs aussi vont dépondre du temps.
L'hypothèse $(H2)$ est posé pour raison de simplification.
L'hypothèse $(H3)$ signifié que l'impact des mouvements interne d'un individu i, non pas un influence sur les autres individus.

2.3.2 Notations

Pour tout individu i, $i \in \{1, ...N\}$, le long de période $[|1, T|]$, on note :
le vecteur des variables à expliquer (vecteur des variables endogènes) :

$$y_i = \begin{pmatrix} y_{i,1} \\ . \\ . \\ . \\ y_{i,2} \end{pmatrix} \tag{2.3.1}$$

-le vecteur des coefficients du modèle :

$$\beta_i = \begin{pmatrix} \beta_{i,0} \\ . \\ . \\ . \\ \beta_{i,m} \end{pmatrix} \tag{2.3.2}$$

-la matrice des variables explicatives :

$$X_i = \begin{pmatrix} 1 & x_{i,1,1} & x_{i,2,1} & . & . & x_{i,m,1} \\ 1 & x_{i,1,2} & x_{i,2,2} & . & . & x_{i,m,2} \\ . & . & . & . & . & . \\ . & . & . & . & . & . \\ . & . & . & . & . & . \\ 1 & x_{i,1,T} & x_{i,2,T} & . & . & x_{i,m,T} \end{pmatrix} \tag{2.3.3}$$

-Le vecteur d'élements du mouvement brownien sur la gamme $[|1, T|]$:

$$u_i = \begin{pmatrix} u_{i,1} \\ . \\ . \\ . \\ u_{i,T} \end{pmatrix} \tag{2.3.4}$$

le vecteur des erreurs :

$$\psi_i = \begin{pmatrix} \psi_{i,1} \\ . \\ . \\ . \\ \psi_{i,T} \end{pmatrix} \tag{2.3.5}$$

Donc l'écriture matricielle du modèle $(\mathcal{M}_i)$ est comme suite :

$$(\mathcal{MV}_i) \;\; y_i = X_i\beta_i + \gamma_i u_i + \psi_i$$

2.3.3 Problématique

L'hétéroscédasticité du mouvement brownien et de l'erreur du modèle $(\mathcal{MV}_i)$, rend l'estimateur obtenu par les différents type de moindre carrée est inefficient. Notre contribution est de trouver une transformation inversible, qui va éliminer le mouvement brownien, et de trouver un modèle $(\mathcal{MV}_i^{'})$ équivalent au modèle $(\mathcal{MV}_i)$ telle que : si $\hat{\beta}_i$ est un

estimateur efficace pour $(\mathcal{MV}_i^{'})$, alors il est efficace aussi pour $(\mathcal{MV}_i)$.
Autrement dit :

$$(P1)\begin{cases} \textit{Chercher } \varphi_T \textit{ automorphisme de } \mathbf{R}^T, \textit{ telle que :} \\ (\mathcal{MV}_i)\ y_i = X_i\beta_i + u_i + \psi_i \ \Leftrightarrow\ (\mathcal{MV}_i^{'})\ z_i = \varphi(y_i) = M_i\beta_i + \varepsilon_i \\ O\grave{u}\ \varepsilon_i \sim BB(0, \sigma^2 I_T) \textit{ est un bruit blanc multivarié,} \\ I_T \textit{ est la matrice identité de } \mathcal{M}_T(\mathbf{IR}),\ \textit{et } M_i \in \mathcal{M}_{T,m+1}(\mathbf{IR}) \end{cases}$$

2.4 Élimination de l'hétéroscédasticité du mouvement brownien

Avant de traité le problème $(P1)$, rappelons tout d'abord quelques définitions et propriétés, dont on aura besoin dans cette section.

2.4.1 Rappel

1. On dit que le processus $(a_t)_{t\geq 0}$ est un mouvement brownien unidimensionnel, si a_t est gaussien centré et pour tout $(s,t) \in \mathbf{IR}_+^2$ la covariance $E(a_t a_s) = \min(s,t)$.
   Un mouvement brownien est dit géométrique, si est seulement si, son logarithme est un mouvement brownien.

2. Si $(a_t)_{t\geq 0}$ est un mouvement brownien unidimensionnel, alors pour tout $0 \leq t_1 \leq t_2 \leq .... \leq t_n$, le vecteur $a = (a_{t_1}, ..., a_{t_n})$ est un vecteur gaussien centré. Autrement-dit $a \sim \mathcal{N}_n(0, V(a))$.

3. le processus $(a_t)_{t\geq 0}$ est un mouvement brownien unidimensionnel, si et seulement si, il vérifie les deux propriétés suivantes :

   (a) Pour tout $(s,t) \in \mathbf{IR}_+^2$, l'acroissement $a_{t+s} - a_t$ est indépendant des variables aléatoires $a_r,\ \forall r \in [0,t]$.

   (b) Pour tout $t \geq 0$, a_t est une variable gaussienne centrée et de variance :

   $$V(a_t) = E(a_t^2) = t$$

4. On dit que le processus $B_{t\geq 0} = \Big((a_{1,t})_{t\geq 0}, ..., (a_{N,t})_{t\geq 0}\Big)$ est un mouvement brownien N-dimensionnel, si les processus $(a_{i,t})_{t\geq 0}$ sont des mouvements brownien unidimensionnels et indépendants, où $i \in \{1, ...N\}$.

5. (**Théorème spectral pour les matrices**) : Soit $A \in \mathcal{S}_n(\mathbf{IR})$ (resp hermitienne complexe), alors il existe une matrice $P \in O_n(\mathbf{IR})$ orthogonale (resp. unitaire) et une matrice D diagonale des coefficients réels, telles que $A = PDP^{'}$.

6. Si f et g deux endomorphismes qui commutent, alors $ker(f)$ et $im(f)$ sont stables par g. En générale, le noyau et l'image de tout polynôme en f est stable par g, en particulier, les sous-espaces propres de f sont stables par g, car ce sont les noyaux d'un polynôme en f de degré 1, et de même pour les sous-espaces caractéristiques. □

D'après le rappel au dessus, la matrice de covariance de u_i est donnée par :

$$\Sigma_T = V(u_i|X) = \begin{pmatrix} 1 & 1 & 1 & .... & 1 & 1 \\ 1 & 2 & 2 & .... & 2 & 2 \\ 1 & 2 & 3 & .... & 3 & 3 \\ . & . & . & .... & . & . \\ . & . & . & .... & . & . \\ 1 & 2 & 3 & .... & T-1 & T-1 \\ 1 & 2 & 3 & .... & T-1 & T \end{pmatrix} \tag{2.4.1}$$

pour tout $n \in \mathbf{IN}^*$ on pose :

$$A_n = \begin{pmatrix} 1 & 1 & 1 & .... & 1 & 1 \\ 1 & 2 & 2 & .... & 2 & 2 \\ 1 & 2 & 3 & .... & 3 & 3 \\ . & . & . & .... & . & . \\ . & . & . & .... & . & . \\ 1 & 2 & 3 & .... & n-1 & n-1 \\ 1 & 2 & 3 & .... & n-1 & n \end{pmatrix} \tag{2.4.2}$$

Proposition 2.4.1 *La forme quatratique associée à la matrice A_n, pour tout $n \geq 2$ est donnée par :*
$\forall (x_1, ..., x_n) \in \mathbf{IR^n}$:

$$q_n(x_1, ..., x_n) = n\Big[\sum_{k=1}^{n} \frac{k}{n} x_k\Big]^2 + \sum_{k=2}^{n} \frac{k-1}{k}\Big[\sum_{i=1}^{k-1} \frac{i}{k-1} x_i\Big]^2 \tag{2.4.3}$$

De plus elle est définie positive et de signature $(n, 0)$, et la matrice A_n est inversible.

Preuve 2.4.1 *Montrons la relation (2.4.3), par réccurence.*
+ On a pour $n = 2$: $A_2 = \begin{pmatrix} 1 & 1 \\ 1 & 2 \end{pmatrix}$ Donc

$$\begin{aligned} \forall (x_1, x_2) \in \mathbf{IR}^2 : \quad q_2(x_1, x_2) &= (x_1 \ x_2) \begin{pmatrix} 1 & 1 \\ 1 & 2 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} x_1 + x_2 & x_1 + 2x_2 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \\ &= x_1(x_1 + x_2) + x_2(x_1 + 2x_2) \\ &= x_1^2 + 2x_1x_2 + 2x_2^2 \end{aligned}$$

On effectue une réduction de Gauss par le plus grand pivot, qui correspond dans ce cas au coefficient de x_2^2.
On obtient Donc :

$$\begin{aligned} q_2(x_1, x_2) &= 2[x_2 + \frac{1}{2}x_1]^2 + \frac{1}{2}x_1^2 \\ &= 2\Big[\sum_{k=1}^{2} \frac{k}{2} x_k\Big]^2 + \sum_{k=2}^{2} \frac{k-1}{k}\Big[\sum_{i=1}^{k-1} \frac{i}{k-1} x_i\Big]^2 \end{aligned}$$

Donc la relation (2.4.3) est vraie pour $n = 2$. De plus la signature de q_2 est (2,0).
On suppose qu'elle est vraie pour un n fixé, telle que $n \geq 2$. Montrons qu'elle reste vraie pour l'ordre suivant $(n+1)$. Autrement dit, montrons que ; la forme quadratique associée à A_{n+1} est vérifiée :
$\forall (x_1, ..., x_{n+1}) \in \mathbf{IR^{n+1}}$:

$$q_{n+1}(x_1, ..., x_{n+1}) = (n+1)\Big[\sum_{k=1}^{n+1} \frac{k}{n+1} x_k\Big]^2 + \sum_{k=2}^{n+1} \frac{k-1}{k}\Big[\sum_{i=1}^{k-1} \frac{i}{k-1} x_i\Big]^2$$

On pose $X_n = (x_1\ x_2\ \dots x_n)$ et $p_n = (1\ 2\ 3\ \dots n)$. On a $X_{n+1} = (X_n\ x_{n+1})$.
On réecrit la matrice A_{n+1} en bloc, alors, on a :

$$\begin{aligned} q_{n+1}(x_1, ..., x_{n+1}) &= X_{n+1} A_{n+1} X_{n+1}^{'} \\ &= (X_n\ \ x_{n+1}) \begin{pmatrix} A_n & p_n^{'} \\ p_n & n+1 \end{pmatrix} \begin{pmatrix} X_n^{'} \\ x_{n+1} \end{pmatrix} \\ &= \begin{pmatrix} X_n A_n + x_{n+1} p_n & X_n p_n^{'} + (n+1) x_{n+1} \end{pmatrix} \begin{pmatrix} X_n^{'} \\ x_{n+1} \end{pmatrix} \\ &= X_n A_n X_n^{'} + x_{n+1} p_n X_n^{'} + x_{n+1} X_n p_n^{'} + (n+1) x_{n+1}^2 \end{aligned}$$

D'après l'hypothèse de réccurence, on a : $X_n A_n X_n^{'} = n\Big[\sum_{k=1}^{n} \frac{k}{n} x_k\Big]^2 + \sum_{k=2}^{n} \frac{k-1}{k}\Big[\sum_{i=1}^{k-1} \frac{i}{k-1} x_i\Big]^2$
D'autre part : $p_n X_n^{'} = X_n p_n^{'} = \sum_{k=1}^{n} k x_n$. *Donc :*

$$\begin{aligned} q_{n+1}(x_1, ..., x_{n+1}) &= n\Big[\sum_{k=1}^{n} \frac{k}{n} x_k\Big]^2 + \sum_{k=2}^{n} \frac{k-1}{k}\Big[\sum_{i=1}^{k-1} \frac{i}{k-1} x_i\Big]^2 + 2x_{n+1} \sum_{k=1}^{n} k x_n + (n+1) x_{n+1}^2 \\ &= (n+1)\Big[x_{n+1}^2 + 2x_{n+1} \sum_{k=1}^{n} \frac{k}{n+1} x_n + \frac{n}{n+1}\Big[\sum_{k=1}^{n} \frac{k}{n} x_k\Big]^2\Big] \\ &+ \sum_{k=2}^{n} \frac{k-1}{k}\Big[\sum_{i=1}^{k-1} \frac{i}{k-1} x_i\Big]^2 \\ &= (n+1)\Big[x_{n+1}^2 + 2x_{n+1} \sum_{k=1}^{n} \frac{k}{n+1} x_n + \frac{n+1}{n}\Big[\sum_{k=1}^{n+1} \frac{k}{n+1} x_k\Big]^2\Big] \\ &+ \sum_{k=2}^{n} \frac{k-1}{k}\Big[\sum_{i=1}^{k-1} \frac{i}{k-1} x_i\Big]^2 \\ &= (n+1)\Big[x_{n+1}^2 + 2x_{n+1} \sum_{k=1}^{n} \frac{k}{n+1} x_n + \Big[\sum_{k=1}^{n} \frac{k}{n+1} x_k\Big]^2\Big] + \frac{n+1}{n}\Big[\sum_{k=1}^{n} \frac{k}{n+1} x_k\Big]^2 \\ &+ \sum_{k=2}^{n} \frac{k-1}{k}\Big[\sum_{i=1}^{k-1} \frac{i}{k-1} x_i\Big]^2 \\ &= (n+1)\Big[x_{n+1} + \sum_{k=1}^{n} \frac{k}{n+1} x_k\Big]^2 + \frac{n}{n+1}\Big[\sum_{k=1}^{n} \frac{k}{n} x_k\Big]^2 \\ &+ \sum_{k=2}^{n} \frac{k-1}{k}\Big[\sum_{i=1}^{k-1} \frac{i}{k-1} x_i\Big]^2 \end{aligned}$$

Or $\frac{n}{n+1}\Big[\sum_{k=1}^{n} \frac{k}{n} x_k\Big]^2 = \sum_{k=n+1}^{n+1} \frac{k-1}{k}\Big[\sum_{i=n}^{k-1} \frac{i}{k-1} x_i\Big]^2$
et $x_{n+1} + \sum_{k=1}^{n} \frac{k}{n+1} x_k = \sum_{k=1}^{n+1} \frac{k}{n+1} x_k$ *Donc*

$$q_{n+1}(x_1,....,x_{n+1}) = (n+1)\Big[\sum_{k=1}^{n+1}\frac{k}{n+1}x_k\Big]^2 + \sum_{k=2}^{n+1}\frac{k-1}{k}\Big[\sum_{i=1}^{k-1}\frac{i}{k-1}x_i\Big]^2$$

Et la signature de q_{n+1} est $(n+1,0)$. Donc la formule (2.3.4) est vraie pour toute $n \geq 2$, et la signature de la forme quadratique q_n est $(n,0)$. Par suite $rg(A_n) = rg(q_n) = n$. Alors $ker(q_n) = 0$, et q_n est non dégénérée. On conclut donc que la forme q_n est définie positive et la matrice A_n est inversible. $\diamond$

Théorème 2.4.1 *Pour tout individu i, la matrice de covariance Σ_T admet une unique racine notée $\sqrt{\Sigma_T} \in S_T^{++}(\mathbf{IR})$ inversible, et si on pose $\varepsilon_i = \sqrt{\Sigma_T^{-1}}u_i$ alors on a :*

$$\varepsilon_i \sim \mathcal{N}_T(0, I_T) \tag{2.4.4}$$

Où I_T est la matrice identité de $\mathcal{M}_T(\mathbf{IR})$.
De plus si on pose : $\sqrt{\Sigma_T^{-1}} = (\rho_{t,l})_{(t,l)\in[|1,T|]^2}$, alors :

$$\sum_{l=1}^{T}\rho_{t,l}u_{i,l} \sim \mathcal{N}(0,1) \tag{2.4.5}$$

Preuve 2.4.2

— *On décompose la preuve en trois étapes :*

1. ***Existence de la racine carrée matricielle de Σ_i :***
Puisque les matrices Σ_i sont symétrique, alors pour tout i, d'après le théorème spéctral, $\exists P_i \in O_n(\mathbf{IR})$, et une matrice diagonale réelle $D_i = diag(\lambda_{i,1} \ldots \lambda_{i,n})$ telle que $\Sigma_i = P_iD_iP_i^{-1} = P_iD_iP_i^{'}$ (car $P_i \in O_n(\mathbf{IR})$).
Puisque les périodes t sont considérés entiers naturels, alors d'après la proposition (2.4.1), les matrices $\Sigma_T = A_T$ sont inversibles et définies positives, donc les matrices D_i sont aussi inversibles et $\forall j \in \{1,....T_i\}$ $\lambda_{i,j} > 0$.
On note $d_i = diag\Big(\sqrt{\lambda_{i,1}} \ldots \sqrt{\lambda_{i,n}}\Big)$, on a : $d_i = D_i^2$.
On pose

$$R_i = P_id_iP_i^{'} \tag{2.4.6}$$

On a alors :

$$R_i^2 = (P_id_iP_i^{'})(P_id_iP_i^{'}) = P_id_i^2P_i^{'} = P_iD_iP_i^{'} = \Sigma_T$$

On pose $\Lambda_i = diag\Big(\frac{1}{\sqrt{\lambda_{i,1}}} \ldots \frac{1}{\sqrt{\lambda_{i,n}}}\Big)$, on a $\Lambda_id_i = d_i\Lambda_i = I_{T_i}$.
Alors

$$R_i(P_i\Lambda_iP_i^{'}) = (P_id_iP_i^{'})(P_i\Lambda_iP_i^{'}) = P_id_i\Lambda_iP_i^{'} = P_iP_i^{'} = I_{T_i}$$

Et

$$(P_i\Lambda_iP_i^{'})R_i = (P_i\Lambda_iP_i^{'})(P_id_iP_i^{'}) = P_i\Lambda_id_iP_i^{'} = P_iP_i^{'} = I_{T_i}$$

Donc R_i est inversible et on a :

$$R_i^{-1} = P_i\Lambda_iP_i^{'} \tag{2.4.7}$$

Donc Σ_T admet une racine carrée inversible.

2. ***Unicité de la racine carrée :*** *On note f_i et r_i les endomorphisme associés respéctivement à la matrice Σ_T et à la matrice R_i donnée par (2.4.6), dans la base canonique. On note $v_{i,j}$ le vecteur propre associé à la valeur propre $\lambda_{i,j}$. Alors on a $\Sigma_i = P_i D_i P_i^{-1}$ avec $P_i = (v_{i,1}\ v_{i,2} \dots v_{i,n})$. Puisque chaque matrice $P_i \in O_n(\mathbf{IR})$, alors on a $P_i^{'} P_i = I_n$, par suite $v_{i,j}^2 = 1$, donc la famille $\{v_{i,1},\ v_{i,2},\ \dots v_{i,n}\}$ forme une base orthonormé des vecteurs propre. Or, toutes les valeurs propres $\lambda_{i,j} \in \mathbf{IR}_+^*$, car Σ_i est symétrique réelle définie positive(d'après la proposition (2.3.2)), donc f_i est une application autoadjoint, et $Sp(f_i) \subset \mathbf{IR}_+^*$.*
On note $E_{i,j}$ les sous-espaces propres correspondants à f_i.
Soient $H_i \in S_{T_i}^{++}(\mathbf{IR})$ telle que $H_i^2 = \Sigma_i$ et h_i son endomorphisme dans la base canonique.
Alors on a : $h_i^2 = f_i$, par suite :

$$h_i o f_i = h_i^3 = h_i^2 o h_i = f_i o h_i$$

Donc f_i et h_i commutent, par suite $E_{i,j} = ker(u - \lambda_{i,j} id|_{E_{i,j}})$ est stable par h_i
On note $h_{i,j} = h_i|_{E_{i,j}}$, alors on a : $h_{i,j}^2 = f_i|_{E_{i,j}} = \lambda_{i,j} id|_{E_{i,j}}$. Puisque $H_i \in S_{T_i}^{++}(\mathbf{IR})$, alors h_i est autoadjoint définie positif, donc $Sp(h_i) \subset \mathbf{IR}_+^*$.
Soit $\rho \in Sp(h_{i,j})$, alors $\rho^2 = \lambda_{i,j}$, ceci implique $\rho = \sqrt{\lambda_{i,j}}$. Donc $h_{i,j}$ a une unique valeur propre $\sqrt{\lambda_{i,j}}$ et il est diagonalisable. D'où $h_{i,j} = \sqrt{\lambda_{i,j}} id|_{E_{i,j}}$.
D'autre part, si on note $r_{i,j} = r_i|_{E_{i,j}}$, alors d'après (2.3.7), on a $Sp(r_i) = \left\{\sqrt{\lambda_{i,1}}, \dots \sqrt{\lambda_{i,n}}\right\}$, et par le même procedé, on peut montrer que

$$r_{i,j} = \sqrt{\lambda_{i,j}} id|_{E_{i,j}} = h_{i,j}$$

Puisque la famille $\{v_{i,1},\ v_{i,2},\ \dots v_{i,n}\}$ est une base de $\mathbf{IR}^n$, alors les espaces propres $E_{i,j}$ sont en somme directe. Donc $r_i = h_i$.
Par suite $R_i = H_i$. D'où l'unicité de la racine carrée inversible de la matrice Σ_i, on la note $\sqrt{\Sigma_T}$.
Or : Σ_i est symétrique $\Rightarrow$ $\sqrt{\Sigma_T}$ est aussi symétrique. Donc $\sqrt{\Sigma_T} \in S_{T_i}^{++}(\mathbf{IR})$.

3. ***Montrons les propriétés (2.2.4)-(2.2.5)*** *:*
Puisque $(u_{i,t})_{t\geq 0}$ est un mouvement brownien, alors $(u_{i,1}, \dots, u_{i,T_i})$ est un vecteur gaussien centré. Donc :

$$u_i \sim \mathcal{N}_T(0, \Sigma_i)$$

En revanche, on a $\sqrt{\Sigma_T}$ est inversible et symétrique, alors :

$$I_T = \sqrt{\Sigma_T}\sqrt{\Sigma_T}^{-1} = (\sqrt{\Sigma_T}\sqrt{\Sigma_T}^{-1})' = (\sqrt{\Sigma_T}^{-1})'\sqrt{\Sigma_T}^{'} = (\sqrt{\Sigma_T}^{-1})'\sqrt{\Sigma_T}$$

Donc

$$\begin{aligned} I_T = (\sqrt{\Sigma_T}^{-1})'\sqrt{\Sigma_T} &\Rightarrow I_T\sqrt{\Sigma_T}^{-1} = (\sqrt{\Sigma_T}^{-1})'\sqrt{\Sigma_T}\sqrt{\Sigma_T}^{-1} \\ &\Rightarrow \sqrt{\Sigma_T}^{-1} = (\sqrt{\Sigma_T}^{-1})' \end{aligned}$$

Par suite $\sqrt{\Sigma_T}^{-1}$ est symétrique. Donc :

$$\begin{aligned} V(\sqrt{\Sigma_T}^{-1} u_i) &= \sqrt{\Sigma_T}^{-1} V(u_i)(\sqrt{\Sigma_T}^{-1})' \\ &= \sqrt{\Sigma_T}^{-1}\Sigma_i\sqrt{\Sigma_T}^{-1} \end{aligned}$$

$$= \sqrt{\Sigma_T}^{-1}\sqrt{\Sigma_T}\sqrt{\Sigma_T}\sqrt{\Sigma_T}^{-1}$$
$$= I_T$$

$E(\sqrt{\Sigma_T}^{-1}u_i) = \sqrt{\Sigma_T}^{-1}E(u_i) = 0$
On pose $\varepsilon_i = \sqrt{\Sigma_T}^{-1}u_i$, *alors :*

$$\varepsilon_i = \sqrt{\Sigma_T}^{-1}u_i \sim \mathcal{N}_T(0, I_T)$$

De plus, puisque $(u_{i,t})_{t\geq 0}$ *est un mouvement brownien alors le vecteur* $(u_{i,1}, u_{i,2}, ..., u_{i,T})$ *est gaussien, par suite la combinaison* $\sum_{l=1}^{T}\rho_{t,l}u_{i,l}$ *est gaussien.*
Or
$\sum_{l=1}^{T}\rho_{t,l}u_{i,l} = \varepsilon_{i,t} \Rightarrow V(\sum_{l=1}^{T}\rho_{t,l}u_{i,l}) = V(\varepsilon_{i,t}) = 1$ *et* $E(\sum_{l=1}^{T}\rho_{t,l}u_{i,l}) = \sum_{l=1}^{T}\rho_{t,l}E(u_{i,l}) = 0$

La preuve est donc achevée. $\diamond$

Corollaire 2.4.1 *L'application*

$$\varphi_T : \mathbf{IR}^T \to \mathbf{IR}^T \tag{2.4.8}$$
$$y \mapsto \varphi_T(y) = \sqrt{\Sigma_T}^{-1}y \tag{2.4.9}$$

est un automorphisme de $\mathbf{IR}^T$. *Avec* Σ_T *est la matrice définit par la relation (2.4.1).*
De plus si $(\eta_t)_{t\geq 0}$ *est un mouvement brownien, et si on pose* $\eta = (\eta_1, ..., \eta_T)'$ *alors :*

$$\forall (y, A, b) \in \mathbf{IR}^T \times \mathcal{M}_{T,k}(\mathbf{IR}) \times \mathbf{IR}^k : \ \varphi_T(y - Ab - \gamma_i\eta) \sim \mathcal{N}_T(\sqrt{\Sigma_i}^{-1}(y - Ab), \gamma_i^2 I_T) \tag{2.4.10}$$

Preuve 2.4.3 *On a* φ_T *est linéaire de* $\mathbf{IR}^T \to \mathbf{IR}^T$, *puisque sa matrice* $\sqrt{\Sigma_T}^{-1}$ *dans la base canonique de* $\mathbf{IR}^T$ *est inversible, donc* φ_T *est un automorphisme.*
D'autre part, si $(\eta_t)_{t\geq 0}$ *est un mouvement brownien, alors* $\eta = (\eta_1, ..., \eta_T)'$ *est un vecteur gaussien, de plus :*

$$\eta \sim \mathcal{N}_T(0, \Sigma_T)$$

Or

$$\varphi_T(y - Ab - \gamma_i\eta) = \sqrt{\Sigma_T}^{-1}(y - Ab) - \gamma_i\sqrt{\Sigma_T}^{-1}\eta$$

Donc

$$\varphi_T(y - Ab - \gamma_i\eta) \sim \mathcal{N}_T(\sqrt{\Sigma_T}^{-1}(y - Ab), \gamma_i^2\sqrt{\Sigma_T}^{-1}V(\eta)(\sqrt{\Sigma_T}^{-1})')$$

Puisque $\sqrt{\Sigma_T}^{-1}$ *est symétrique, alors :*

$$\sqrt{\Sigma_T}^{-1}V(\eta)(\sqrt{\Sigma_T}^{-1})' = \gamma_i^2\sqrt{\Sigma_T}^{-1}\Sigma_T\sqrt{\Sigma_T}^{-1}$$
$$= \gamma_i^2 I_T$$

D'où le resultat de corollaire. $\Diamond$

Le corollaire au dessus montre que la transformation φ_T élimine l'hétéroscédasticité des variances du movement brownien. On pose : $z_i = \varphi_T(y_i)$, on a :

$$y_i = X_i\beta_i + u_i + \psi_i \Leftrightarrow \ z_i = \varphi_T(X_i\beta_i) + \varphi_T(\gamma_i u_i + \psi_i)$$

$$\Leftrightarrow \ z_i = \sqrt{\Sigma_T}^{-1} X_i\beta_i + \sqrt{\Sigma_T}^{-1}(\gamma_i u_i + \psi_i)$$

On pose

$$\nu_i = \sqrt{\Sigma_T}^{-1}(\gamma_i u_i + \psi_i) = \begin{pmatrix} \nu_{i,1} \\ \nu_{i,2} \\ . \\ . \\ \nu_{i,T} \end{pmatrix} \tag{2.4.11}$$

On considère le modèle :

$$(\mathcal{MV}_i^{'}) \ \ z_i = [\sqrt{\Sigma_T}^{-1} X_i]\beta_i + \nu_i$$

La question qui ce pose, l'élimination de l'hétéroscédasticité du mouvement brownien élimine- t- elle l'hétéroscédasticité du vecteur aléatoire ν_i ?.
En revanche l'utilisation des moindres carrée nécessite l'absence de :

— la multicolinéarité entre les variables explicatives ;
— l'autocorrélation des composantes $(\nu_{i,t})_{t\in[|1,T|]}$;
— l'hétéroscédasticité des composantes $(\nu_{i,t})_{t\in[|1,T|]}$.

Dans ce chapitre on s'intéresse à l'étude de l'existence de l'autocorrélation et de la multicolinéarité dans le modèle $(\mathcal{MV}_i)$ et dans le chapitre suivant on s'intéresse à l'étude de l'estimation optimale des paramètres du modèle dans le cas d'existence de la multicolinéarité et dans le cas d'existence de l'hétéroscédasticité.
On accepte dans le reste du chapitre l'hypothèse :
$(\mathcal{H}_0^{ht})$: le bruite stochastique ν_i est homoscédastique.

2.5 Test de la multicolinéarité et test d'autocorrélation

Considérons les modèles :

$$(\mathcal{MV})_i \ \ y_i = X_i\beta_i + \gamma_i u_i + \psi_i$$

$$(\mathcal{MV}_i^{'}) \ \ z_i = [\sqrt{\Sigma_T}^{-1} X_i]\beta_i + \nu_i$$

2.5.1 Test de la multicolinéarité : Test de Farrar et Glauber modifié

Nous proposons de commencer par cette étapes, car, la multicolinéarité à principalement troix effets indésirables :

1. Lorsqu'il y a une forte colinéarité entre quelques variables explicatives, alors la variance estimée de certains coefficients augmente. (le t de Student diminue) ;
2. la multicolinéarité engendre l'instabilité des estimations des coefficients par les moindres carrés. Autrement dit des faibles fluctuations sur les données, génèrent des variations importantes des valeurs estimées des coefficients ;
3. Si la multicolinéarité est parfaite, la matrice $X_i^{'}X_i$ est singulière, dans ce cas l'estimation des coefficients sera impossible par la méthode des moindre carré et leur variance sera infinie.

On note D la matrice de corrélation entre les vecteurs colonnes de la matrice $\sqrt{\Sigma_T}^{-1}X_i$ et on poursuit les étapes suivantes :

Étape 1 : Vérifier l'existence d'une multicolinéarité parfaite

- Si le déterminant $det(X'\Sigma_T^{-1}X) = 0$, alors il ya une multicolinéarité parfaite.

Dans ce cas, on passe au chapitre 3 pour estimer le vecteur β_i.

- Si $det(X'\Sigma_T^{-1}X) \neq 0$, On passe à l'étape suivante.

Étape 2 : Tester le degré de la multicolinéarité

On propose comme test de la multicolinéarité, le test de **Farrar et Glauber**. On calcul le déterminant $det(D)$ puis on test l'hypothèse (H_0) contre (H_1), avec :

$$\begin{cases} H_0^{mc} : \ det(D) = 1, & \text{les vecteurs explicatives sont orthogonales} \\ H_1^{mc} : \ det(D) < 1, & \text{les vecteurs explicatives sont dépendants} \end{cases} \tag{2.5.1}$$

On pose :

$$\chi^2_{emp} = -\Big[T - 1 - \frac{1}{6}(2m+7)\Big].\ln(det(D)) \tag{2.5.2}$$

Puis on compare cette valeur empirique, avec la valeur théorique $\chi_{\alpha,\frac{1}{2}m(m+1)}$ de la loi de Khi-deux.

- Si $\chi^2_{emp} < \chi_{\alpha,\frac{1}{2}m(m+1)}$ On accepte (H_0) et on rejette (H_1).

Dans ce cas les vecteurs explicatives sont considérés orthogonales, et on passe au test d'autocorrélation.

- Sinon, On accepte (H_1) et on rejette (H_0).

Dans ce cas le degré de multicolinéarité est importante, pour l'éliminer et estimer β_i on passe au troisième chapitre.

2.5.2 Étude de l'autocorrélation des erreurs

On considère le modèle $\mathcal{MV}_i^{'}$. Si l'hypothèse $(\mathcal{H}_0^{ht})$ est accépté, alors :

$$(\forall i \in [|1,N|])(\exists \sigma_i > 0) : \ \nu_{i,t} \sim \mathcal{N}(0,\sigma_i^2) \ \ \forall t \in [|1,T|] \tag{2.5.3}$$

Le problème qu'on peut confronter est dans le cas où les erreurs $(\nu_{i,t})_{t\in[|1,T|]}$ sont autocorrélés. Autrement-dit, si :

$$V(\nu_i) \neq \sigma_i^2 I_T$$

Où :

$$\nu_i = \begin{pmatrix} \nu_{i,1} \\ . \\ . \\ . \\ \nu_{i,T} \end{pmatrix}$$

Les tests qui ont existé dans la littérature n'ont pris pas en compte l'existence du mouvement brownien. Notre idée ici est de modifié le test de Breusch-Godfrey [1] fondé sur le Multiplicateur de Lagrange (LM), de tel façon de prendre en compte l'existence d'une composante brownien.

Rappelant tout d'abord la principe du test de Breusch-Godfrey avec statistique du test "Multiplicateur de Lagrange (LM)".

1. Breusch-Godfrey, On peut citer des travaux basés sur ce test [3].

a- Test de Breusch-Godfrey avec LM

On considère le modèle : $(*)$ $y_t = \beta_0 + \sum_{k=1}^{m} \beta_k x_{k,t} + \varepsilon_t$
on pose $\varepsilon_t = \mu_t + \sum_{k=1}^{p} r_k \varepsilon_{t-k}$, avec μ_t , $t \in [|0, T|]$ sont iid.
L'objectif du test est de tester l'hypothèse H_0 contre l'hypothèse H_1 :
(H_0) : $r_1 = r_2 = ... = r_p = 0$, et (H_1) : $\exists k \in [|1, p|]$: $r_k \neq 0$
Le test est composé en trois étapes :

1. Utiliser l'estimateur MCO (moindres carrés ordinaire) du modèle $(*)$ et calculer les résidus, on les notes $\hat{\varepsilon}_t$.
2. Utiliser encore une fois l'estimateur MCO, pour estimer les paramètres du modèle intermédiaire suivant :

$$\hat{\varepsilon}_t = \beta_0 + \sum_{k=1}^{m} \beta_k x_{k,t} + \sum_{i=1}^{p} r_k \hat{\varepsilon}_{t-k} + \mu_t$$

3. Calculer le coefficient de détermination R^2 et tester les hypothèses H_0 contre H_1 sur le modèle intermédiaire : on choisi comme statistique de test, la statistique "multiplicateur de Lagrange" : $(T-p)R^2$
   L'intervalle d'accepter (H_0) au niveau de confiance $1-\alpha$ est : $\mathcal{C}_\alpha = [0, \chi^2_{1-\alpha,p}]$
   Donc si $(T-p)R^2 > \chi^2_{1-\alpha,p}$, alors on rejette l'hypothèse (H_0).

Avant d'aborder la principe du test qu'on propose, qui est une version modifié du test de Breusch-Godfrey avec LM, pour tester l'autocorrélation des erreur dans le modèle $(\mathcal{MV}_i)$, on donne une définition de l'estimateur de moindres carrés ordiniaire, dans le cas d'existence d'une composante brownien, on le donne le nom "Moindres carrés ordiniaire avec composante brownien", ou encore en abréviation "BOLS" (Brownian Ordinary Least Square).

Définition 2.5.1 *On considère le modèle $(\mathcal{MV}_i)$. Soient y_i et X_i respectivement ; le vecteur d'observation de la variable à expliquer et la matrice d'observation des variables explicatives sur T périodes successives de l'individue i.*
On définit l'estimateur moindre carré ordiniaire avec composante brownien et on le note β_i^{BOLS}, associe au modèle $(\mathcal{MV}_i)$, la statistique :

$$\beta_i^{BOLS} = (X_i' \Sigma_T^{-1} X_i)^{-1} X_i' \Sigma_T^{-1} y_i \tag{2.5.4}$$

Où :
- *X_i' est la transposée de la matrice X_i*
- *Σ_T^{-1} est l'inverse de la matrice de covariance du processus brownien $(u_{i,t})_{t\in[|1,T|]}$.*

Reprenant le modèle $(\mathcal{MV}_i)$. L'idée principale de ce test est de vérifier l'existence d'une relation significative entre le résidu et ce même résidu décalé à un ordre p.
On note $\sqrt{\Sigma_T^{-1}} = (\rho_{t,l})_{(t,l)\in[|1,T|]^2}$, alors d'après l'équation (2.4.11) :

$$\forall (i,t) \in [|1,N|] \times [|1,T|] \ \nu_{i,t} = \sum_{l=1}^{T} \rho_{t,l}(\gamma_i u_{i,l} + \psi_{i,l}) \tag{2.5.5}$$

On suppose que $\nu_{i,t} \sim AR(p)$, c-à-d :

$$\nu_{i,t} = \mu_{i,t} + \sum_{k=1}^{p} r_{i,k} \nu_{i,t-k}$$

Avec $r_{i,k}$ sont des coefficients et $\mu_{i,t}$ est un bruit blanc.
Donc le modèle $(\mathcal{MV}_i)$, prend la forme :

$$z_t = \sum_{l=1}^{T} \rho_{t,l} y_l = \sum_{l=1}^{T} \rho_{t,l} \Big[\beta_{i,0} + \sum_{k=1}^{m} \beta_{i,k} x_{i,l,k}\Big] + \mu_{i,t} + \sum_{i=1}^{p} r_{i,k} \varepsilon_{i,t-k}$$

On veut tester l'hypothèse d'autocorrélation des erreurs $\nu_{i,t}$ pour $t \in [|p, T|]$.
On considère les hypothèses :

$$(H_0^{aut}) : \quad r_{i,1} = r_{i,2} = ... = r_{i,p} = 0$$

contre

$$(H_1^{aut}) : \quad \exists k \in [|1, p|] : \quad r_{i,k} \neq 0$$

On modifié le test précedent et on nomme le nouveau test ; "Test de Breusch-Godfrey avec BMLM [2]".

b- Test de Breusch-Godfrey avec BMLM

Le test est composé en trois étapes :

1. On utilise l'estimateur BOLS du modèle $(\mathcal{MV}_i)$ et on calcule les résidus, on les notes $\hat{\nu}_{i,t}$.
2. On utilise l'estimateur MCO pour estimer les paramètres du modèle intermédiaire suivant :

$$\sum_{l=1}^{T} \rho_{t,l} \hat{\varepsilon}_{i,l} = \beta_{i,0} \sum_{l=1}^{T} \rho_{t,l} + \sum_{k=1}^{m} \beta_{i,k} \sum_{l=1}^{T} \rho_{t,l} x_{i,l,k} + \mu_{i,t} + \sum_{i=1}^{p} r_{i,k} \hat{\varepsilon}_{i,t-k}$$

3. On calcule le coefficient de détermination R^2 et on teste l'hypothèse H_0^{aut} contre H_1^{aut} sur le modèle intermédiaire : On choisi comme statistique du test, la statistique "multiplicateur de Lagrange" : $(T-p)R^2$
   L'intervalle d'accepter (H_0^{aut}) au niveau de confiance $1-\alpha$ est : $\mathcal{C}_\alpha = [0, \chi^2_{1-\alpha,p}]$
   Donc si $(T-p)R^2 > \chi^2_{1-\alpha,p}$, alors on rejette l'hypothèse (H_0^{aut}).

2.6 Efficience des estimateurs des moindres carrés brownien

On suppose que l'hypothèse $(\mathcal{H}_0^{ht})$ est vérifiée. L'objectif de cette section est de trouver un estimateur optimal et sans biais (Best Linear Unbiased Estimator "BLUE") et proposer un modèle optimale après l'étude de l'autocorrélation de l'erreur.

Définition 2.6.1 *On considère le modèle $(\mathcal{MV}_i)$. Soient y_i et X_i respectivement ; le vecteur d'observation de la variable à expliquer et la matrice d'observation des variables*

2. BMLM : Browian Motion with Lagrange Multiplie

explicatives sur T périodes successives.
On pose :

$$\Omega_i = V\Big(\sqrt{\Sigma_T^{-1}}(\gamma_i u_i + \psi_i)\Big)$$

On définit et on nomme l'estimateur de Moindre Carré Généralisés avec Composante Brownien (Generalized Brownian Lesser Square) et on le note β^{GBLS}, la statistique :

$$\beta^{GBLS} = (X_i'\sqrt{\Sigma_T^{-1}}\Omega_i^{-1}\sqrt{\Sigma_T^{-1}}X_i)^{-1}X_i'\sqrt{\Sigma_T^{-1}}\Omega_i^{-1}\sqrt{\Sigma_T^{-1}}y_i \tag{2.6.1}$$

Où :

- *X_i' est la transposée de la matrice X_i*
- *Σ_T^{-1} est l'inverse de la matrice de covariance du processus brownien $(u_{i,t})_{t\in[|1,T|]}$.*

Théorème 2.6.1 *On considère le modèle $(\mathcal{MV}_i)$, et on pose :*

$$\nu_{i,t} = \sum_{l=1}^{T}\rho_{t,l}(u_{i,l} + \psi_{i,l})$$

Avec $(\rho_{t,1}, \rho_{t,2}, ..., \rho_{t,T})$ est la $t^{\text{ème}}$ ligne de la matrice $\sqrt{\Sigma_T^{-1}}$.
On suppose que :

H_1 $T > m+1$

H_2 $\chi^2_{emp} < \chi_{\alpha,\frac{1}{2}m(m+1)}$, *avec χ^2_{emp} est donné par la relation (2.5.2) et $\chi_{\alpha,\frac{1}{2}m(m+1)}$ est la valeur théorique de la loi de Khi 2.*

H_3 *La matrice $\frac{1}{T}(X_i'\Sigma_T^{-1}X_i)$ tend vers une matrice finie régulière, qu'on $T \to +\infty$.*

Alors on a les deux propriétés suivantes :

P_1 *Sous les hypothèses d'homoscédasticité (H_0^{ht}), d'absence de multicolinéarité (H_0^{mc}) et de l'absence d'autocorrélation (H_0^{aut}) des erreurs $\nu_{i,t}$, l'estimateur β^{BOLS} est BLUE.*

P_2 *Sous les hypothèses d'homoscédasticité (H_0^{ht}), d'absence de multicolinéarité (H_0^{mc}) et de l'existence de l'autocorrélation (H_1^{aut}) des erreurs $\nu_{i,t}$, l'estimateur β^{GBLS} est BLUE.*

Preuve 2.6.1 *On a :*

$$(\mathcal{MV}_i)\ \ y_i = X_i\beta_i + \gamma u_i + \psi_i$$

On pose : $\rho_{t,.} = (\rho_{t,1}, \rho_{t,2}, ..., \rho_{t,T})$, $x_{i,k,.} = (x_{i,k,1}, x_{i,k,2}, ..., x_{i,k,T})'$ et $w_{i,k,t} = \rho_{t,.}x_{i,k,.}$
On note

$$W_i = \begin{pmatrix} \sum_{l=1}^{T}\rho_{1,l} & w_{i,1,1} & w_{i,2,1} & . & . & w_{i,m,1} \\ \sum_{l=1}^{T}\rho_{2,l} & w_{i,1,2} & w_{i,2,2} & . & . & w_{i,m,2} \\ . & . & . & . & . & . \\ . & . & . & . & . & . \\ . & . & . & . & . & . \\ \sum_{l=1}^{T}\rho_{T,l} & w_{i,1,T} & w_{i,2,T} & . & . & w_{i,m,T} \end{pmatrix} \tag{2.6.2}$$

L'hypothèse H_1 implique l'absence d'une multicoléniarité parfaite entre les variables $(w_{i,k,t})_{k\in[|1,m|]}$ ce qui entraîne la régularité de la matrice $W_i'W_i$, donc $(W_i'W_i)^{-1}$ existe. Or

$$W_i = \begin{pmatrix} \sum_{l=1}^{T}\rho_{1,l} & \sum_{l=1}^{T}\rho_{1,l}x_{i,1,l} & \sum_{l=1}^{T}\rho_{1,l}x_{i,2,l} & . & . & \sum_{l=1}^{T}\rho_{1,l}x_{i,m,l} \\ \sum_{l=1}^{T}\rho_{2,l} & \sum_{l=1}^{T}\rho_{2,l}x_{i,1,l} & \sum_{l=1}^{T}\rho_{2,l}x_{i,2,l} & . & . & \sum_{l=1}^{T}\rho_{2,l}x_{i,m,l} \\ . & . & . & . & . & . \\ . & . & . & . & . & . \\ . & . & . & . & . & . \\ \sum_{l=1}^{T}\rho_{T,l} & \sum_{l=1}^{T}\rho_{T,l}x_{i,1,l} & \sum_{l=1}^{T}\rho_{T,l}x_{i,2,l} & . & . & \sum_{l=1}^{T}\rho_{T,l}x_{i,m,l} \end{pmatrix}$$

$$= \sqrt{\Sigma_T^{-1}} X_i$$

Donc

$$\begin{aligned}(W_i^{'} W_i)^{-1} &= ((\sqrt{\Sigma_T^{-1}} X_i)^{'} \sqrt{\Sigma_T^{-1}} X_i)^{-1} \\ &= (X_i^{'} \sqrt{\Sigma_T^{-1}} \sqrt{\Sigma_T^{-1}} X_i)^{-1} \quad (car\ \sqrt{\Sigma_T^{-1}}\ est\ symétrique) \\ &= (X_i^{'} \Sigma_T^{-1} X_i)^{-1}\end{aligned}$$

On déduit donc que l'estimateur β^{BOLS} *existe.*
On multiplie le modèle $(\mathcal{MV}_i)$ *Par* $\sqrt{\Sigma_T^{-1}}$*, on obtient donc :*

$$\sqrt{\Sigma_T^{-1}} y_i = W_i \beta_i + \sqrt{\Sigma_T^{-1}} (u_i + \psi_i)$$

On pose

$$\nu_i = \sqrt{\Sigma_T^{-1}} (u_i + \psi_i) = \begin{pmatrix} \nu_{i,1} \\ . \\ . \\ . \\ \nu_{i,T} \end{pmatrix}$$

Et

$$z_i = \sqrt{\Sigma_T^{-1}} y_i \tag{2.6.3}$$

Alors le modèle $(\mathcal{MV}_i)$ *est équivalent au modèle suivant :*

$$(\mathcal{MV}_i)^{'} \quad z_i = W_i \beta_i + \nu_i$$

On a :

$$\begin{aligned}E(\nu_i) &= E(\sqrt{\Sigma_T^{-1}} (u_i + \psi_i)) \\ &= \sqrt{\Sigma_T^{-1}} E(u_i + \psi_i) \\ &= 0\end{aligned}$$

On suppose que les erreurs $\nu_{i,t}$ *vérifiées les hypothèses* (H_0^{ht}) *et* (H_0^{aut})*.*
Montrons que β^{BOLS} *est BLUE :*
On a :
- $\nu_{i,t}$ *vérifiées l'hypothèses* $(H_0^{ht}) \Rightarrow \exists \sigma_i \geq 0$ *tel que :* $V(\nu_{i,t}) = \sigma_i^2\ \forall t \in [|1, T|]$ *(*)*
- $\nu_{i,t}$ *vérifiées l'hypothèses* (H_0^{aut}) *alors :*

$$V(\nu_i) = \sigma_i^2 I_T$$

D'autre part, on a :

$$\begin{aligned}\frac{1}{T}(W_i^{'} W_i) &= \frac{1}{T}((\sqrt{\Sigma_T^{-1}} X_i)^{'} \sqrt{\Sigma_T^{-1}} X_i) \\ &= \frac{1}{T}(X_i^{'} \Sigma_T^{-1} X_i)\end{aligned}$$

Alors, d'après la $3^{\text{ème}}$ *hypothèse du théorème, on a :* $\frac{1}{T}(W_i^{'} W_i)$ *tend vers une matrice finie régulière qu'on* $T \to +\infty$*.*

Donc les conditions du théorème de Gauss sont vérifiées, et l'estimateur :

$$\hat{\beta}^{W_i} = (W_i^{'} W_i)^{-1} W_i^{'} z_i$$

est BLUE pour le modèle $(\mathcal{MV}_i)^{'}$.
Or :

$$\begin{aligned}\hat{\beta}^{W_i} \in \arg\min(\|z_i - W_i\beta_i\|) &\Rightarrow \hat{\beta}^{W_i} \in \arg\min(\|\sqrt{\Sigma_T^{-1}}(y_i - X_i\beta_i)\|) \\ &\Rightarrow \hat{\beta}^{W_i} \in \arg\min(\|y_i - X_i\beta_i\|)\end{aligned}$$

Donc $\hat{\beta}^{W_i}$ *est BLUE pour le modèle* $(\mathcal{MV}_i)$. *Or*

$$\begin{aligned}\hat{\beta}^{W_i} &= (W_i^{'} W_i)^{-1} W_i^{'} z_i \\ &= (X_i^{'}\Sigma_T^{-1} X_i)^{-1}(\sqrt{\Sigma_T^{-1}} X_i)^{'} y_i \\ &= \beta^{BOLS}\end{aligned}$$

Donc, sous les hypothèses du théorème, β^{BOLS} *est BLUE pour le modèle* $(\mathcal{MV}_i)$.
Supposant maintenant que l'hypothèse d'absence d'autocorélation des erreurs $\nu_{i,t}$ *est violée. C'est-à dire ; on accepte l'hypothèse* (H_1^{aut}). *Montrons la propriété* P_2 :
-l'hypothèse d'homoscédasticité $\mathcal{H}_0^{ht}$ *des erreurs* $\nu_{i,t}$ *implique la relation (*) au dessus.*
-l'hypothèse d'existence d'autocorrélation (H_1^{aut}), *implique que* $V(\nu_i) \neq \sigma_i^2 I_T$.
On note :

$$\Omega_i = V(\nu_i) = E(\nu_i \nu_i^{'})$$

On multiplier les deux membres du modèle $(\mathcal{MV}_i)$ *par* $\sqrt{\Sigma_T^{-1}}$ *puis par* $\sqrt{\Omega_i^{-1}}$, *et on pose :*

$$\xi_i = \sqrt{\Omega_T^{-1}}\nu_i = \begin{pmatrix} \xi_{i,1} \\ . \\ . \\ . \\ \xi_{i,T} \end{pmatrix},\ q_i = \sqrt{\Omega_T^{-1}} z_i \text{ et } B_i = \sqrt{\Omega_T^{-1}} W_i$$

Alors on obtient le modèle auxilière :

$$(\mathcal{MV}_i)^{''} \quad q_i = B_i\beta_i + \xi_i$$

Or :

$$V(\xi_i) = V(\sqrt{\Omega_i^{-1}}\nu_i) = \sqrt{\Omega_i^{-1}} V(\nu_i)(\sqrt{\Omega_i^{-1}})' = \sqrt{\Omega_i^{-1}}\Omega_i\sqrt{\Omega_i^{-1}} = I_T$$

Donc $\xi_i \sim \mathcal{N}(0, I_T) \ \Rightarrow \ q_i \sim \mathcal{N}(B_i\beta_i, I_T)$.
On note $B_i = (b_{i,k,t})_{(k,t)\in[|0,m|]\times[|1,T|]}$ *alors* $q_{i,t} = \sum_{k=0}^{m} \beta_{i,k} b_{i,k,t} + \xi_{i,t}\ \forall t \in [|1,T|]$
On note f *la densité de probabilité de* q. *alors :*

$$f(q) = \frac{1}{\sqrt{2\pi}} e^{-\frac{(q - \sum_{k=0}^{m} \beta_{i,k} b_{i,k,t})^2}{2}}$$

Pour T *échantillon de la variable endogène* q *la vraisemblance de* q *est donnée par :* $L_q(q_{i,1}, ..., q_{i,T}) = \prod_{t=1}^{T} f(q_{i,t})$.
Le log vraisemblance devient donc :

$$l_q = \ln(L_q) = \sum_{k=1}^{T} \ln(f(q_{i,t})) = \sum_{k=1}^{T} \Big[-\ln(\sqrt{2\pi}) - \frac{(q_{i,t} - \sum_{k=0}^{m} \beta_{i,k} b_{i,k,t})^2}{2} \Big]$$

$$= -\frac{T}{2}\ln(2\pi) - \sum_{k=1}^{T}\Big[\frac{(q_{i,t} - \sum_{k=0}^{m}\beta_{i,k}b_{i,k,t})^2}{2}\Big]$$
$$= -\frac{T}{2}\ln(2\pi) - \frac{1}{2}(q_i - B_i\beta_i)'(q_i - B_i\beta_i) \ (\textit{écriture vectorielle})$$

Les conditions d'optimalité de premier ordre donnes :

$$\nabla_{\beta_i} l_q = 0 \Rightarrow B_i^{'}q_i - B_i^{'}B_i\beta_i = 0$$
$$\Rightarrow B_i^{'}q_i = B_i^{'}B_i\beta_i$$

Sous l'hypothèse d'abscence de multicolinéarité $det(B_i^{'}B_i) \neq 0$, donc $B_i^{'}B_i$ est inversible, par suite les conditions de premier ordre donnes :

$$\nabla_{\beta_i} l_q = 0 \Rightarrow \ \beta_i = (B_i^{'}B_i)^{-1}B_i^{'}q_i$$

On note par H_{β_i} la matrice hessienne de l_q, alors les conditions suffisante d'optimalité donnes :

$$H_{\beta_i} = -B_i^{'}B_i$$

Or $B_i^{'}B_i$ est inversible et semi définie positive, alors elle est définie positive.
Par suite H_{β_i} est définie négative, donc la statistique $\beta^{B_i} = (B_i^{'}B_i)^{-1}B_i^{'}q_i$ est l'estimateur de maximum de log vraisamblance qui correspond au modèle $(\mathcal{MV}_i)^{''}$.
D'autre part :

$$E(\beta^{B_i}) = (B_i^{'}B_i)^{-1}B_i^{'}E(q_i)$$
$$= (B_i^{'}B_i)^{-1}B_i^{'}B_i\beta_i$$
$$= \beta_i$$

Donc β^{B_i} est l'estimateur de maximum de vraisemblance sans biais, par suite il vérifié les condition de borne Cramér-Rao, donc il est efficace.
Puisque $(\mathcal{MV}_i)^{''} \Leftrightarrow (\mathcal{MV}_i)^{'} \Leftrightarrow (\mathcal{MV}_i)$, alors β^{B_i} est efficace pour le modèle $(\mathcal{MV}_i)$.
En revanche :

$$B_i^{'}B_i = (\sqrt{\Omega_T^{-1}}W_i)'\sqrt{\Omega_T^{-1}}W_i = W_i^{'}(\sqrt{\Omega_T^{-1}})'\sqrt{\Omega_T^{-1}}W_i = W_i^{'}\Omega_T^{-1}W_i$$

Alors

$$\beta^{B_i} = (W_i^{'}\Omega_i^{-1}W_i)^{-1}W_i^{'}\Omega_i^{-1}z_i$$
$$= ((\sqrt{\Sigma_T^{-1}}X_i)'\Omega_i^{-1}(\sqrt{\Sigma_T^{-1}}X_i))^{-1}(\sqrt{\Sigma_T^{-1}}X_i)'\Omega_i^{-1}\sqrt{\Sigma_T^{-1}}y_i$$
$$= (X_i^{'}\sqrt{\Sigma_T^{-1}}\Omega_i^{-1}(\sqrt{\Sigma_T^{-1}}X_i))^{-1}X_i^{'}\sqrt{\Sigma_T^{-1}}\Omega_i^{-1}\sqrt{\Sigma_T^{-1}}y_i$$
$$= \beta^{GBLS}$$

Donc dans le cas d'autocorrélation des erreurs l'estimateur β^{GBLS} est BLUE pour le modèle $(\mathcal{MV}_i)$.
D'où le théorème $\square$.

2.7 Méthode de moindres carrées browniens à erreurs sur les variables

2.7.1 Méthode des variables instrumentales avec mouvement browien

Dans ce qui précède, nous avons considéré que les observations sur la variable à expliquer (endogène), et les variables explicatives étaient faites sans erreur, ou que l'erreur de mesure des observations est faible par rapport à l'erreur de spécification.
Or, Dans la réalité, cette hypothèse est difficile à vérifier. Par exemple, dans le cas des observations proviennent de données d'enquêtes par sondage.
En revanche, l'existence d'un mouvement interne des variables explicatives (mouvement brownien), augmente la chance pour que les observations se fassent avec erreur. Pour cette raison , en distinguant entre les vraies variables (qui sont inconnues) et les variables observées. On note : $y^*_{i,t}$: variable endogène observée à l'instant t.
$y_{i,t}$: est la vraie variable endogène à l'instant t.
$x^*_{i,k,t}$: $k^{\text{ème}}$ variable explicative observée à l'instant t.
$x_{i,k,t}$: $k^{\text{ème}}$ vraie variable explicative à l'instant t.
On garde X_i et y_i comme ils sont définis plus haut.
On note :

$$y^*_i = \begin{pmatrix} y^*_{i,1} \\ . \\ . \\ . \\ y^*_{i,T} \end{pmatrix} \tag{2.7.1}$$

Et

$$X^*_i = \begin{pmatrix} 1 & x^*_{i,1,1} & x^*_{i,2,1} & . & . & x^*_{i,m,1} \\ 1 & x^*_{i,1,2} & x^*_{i,2,2} & . & . & x^*_{i,m,2} \\ . & . & . & . & . & . \\ . & . & . & . & . & . \\ . & . & . & . & . & . \\ 1 & x^*_{i,1,T} & x^*_{i,2,T} & . & . & x^*_{i,m,T} \end{pmatrix} \tag{2.7.2}$$

Aussi, on note respectivement les erreur des observations des vraie variables y_i et X_i par :

$$\eta_i = \begin{pmatrix} \eta_{i,1} \\ . \\ . \\ . \\ \eta_{i,T} \end{pmatrix} \tag{2.7.3}$$

Et

$$\zeta_i = \begin{pmatrix} 0 & \zeta^*_{i,1,1} & \zeta^*_{i,2,1} & . & . & \zeta^*_{i,m,1} \\ 0 & \zeta^*_{i,1,2} & \zeta^*_{i,2,2} & . & . & \zeta^*_{i,m,2} \\ . & . & . & . & . & . \\ . & . & . & . & . & . \\ . & . & . & . & . & . \\ 0 & \zeta^*_{i,1,T} & \zeta^*_{i,2,T} & . & . & \zeta^*_{i,m,T} \end{pmatrix} \tag{2.7.4}$$

On considère le modèle $(\mathcal{MV}_i)$ qu'on a défini plus haut.

$$(\mathcal{MV}_i): \quad y_i = X_i\beta_i + u_i + \psi_i$$

On a

$$X_i^* = X_i + \zeta_i$$

Et

$$y_i^* = y_i + \eta_i$$

On remplace y_i rt X_i par ses observations, alors :

$$y_i^* - \eta_i = (X_i^* - \zeta_i)\beta_i + \gamma u_i + \psi_i$$

On pose :

$$\xi_i = \eta_i + \gamma u_i + \psi_i - \zeta_i\beta_i \tag{2.7.5}$$

Donc le modèle associé au modèle $(\mathcal{MV}_i)$ dans le cas des erreurs sur les observations est comme suit :

$$(\mathcal{MV}_i^*): \quad y_i^* = X_i^*\beta_i + \xi_i$$

Pour éliminer l'hétéroscédasticité du processus brownien sur la gamme $[|1, T|]$, on multiplie le modèle $(\mathcal{MV}^*)$ par $\sqrt{\Sigma_T^{-1}}$. Rappelons que d'après l'hypothèse $(\mathcal{H}_0^{ht})$:

$$\varepsilon_i = \sqrt{\Sigma_T^{-1}}(\gamma_i u_i + \psi_i) \sim \mathcal{N}(0, \sigma_i I_T)$$

On pose :

$$z_i^* = \sqrt{\Sigma_T^{-1}} y_i^* \tag{2.7.6}$$

$$W_i^* = \sqrt{\Sigma_T^{-1}} X_i^* \tag{2.7.7}$$

$$\tilde{\zeta}_i = \sqrt{\Sigma_T^{-1}} \tilde{\zeta}_i \tag{2.7.8}$$

$$\tilde{\eta}_i = \sqrt{\Sigma_T^{-1}} \tilde{\zeta}_i \tag{2.7.9}$$

Et

$$\nu_i = \sqrt{\Sigma_T^{-1}} \xi_i \tag{2.7.10}$$

On obtient donc le modèle équivalent à $(\mathcal{MV}_i^*)$, qu'on le note $(\mathcal{MVB}_i^*)$:

$$(\mathcal{MVB}_i^*): \quad z_i^* = W_i^*\beta_i + \nu_i^*$$

Proposition 2.7.1 *Soit $W_i = \sqrt{\Sigma_T^{-1}} X_i$, avec X_i est la matrice des vraies variables. On suppose que :*

H1 *L'hypothèse $\mathcal{H}_0^{ht}$ est retenue pour les erreurs ε_i*

H2 $E(\zeta_i) = E(W_i^{'}\tilde{\zeta}_i) = E(W_i^{'}\varepsilon_i) = E(z_i^{'}\tilde{\zeta}_i) = 0$

H3 $E(\eta_i) = E(W_i^{'}\tilde{\eta}_i) = E(z_i^{'}\tilde{\eta}_i) = 0$

H4 $E(\tilde{\zeta}_i\tilde{\eta}_i) = 0$

Alors, les propriétés stochastiques de ν_i sont :

1. $E(\nu_i) = 0$
2. $E(W_i^{'}\nu_i) = 0$

3. $E((W_i^*)'\nu_i) = -E(\tilde{\zeta}_i'\zeta_i)\beta_i \neq 0$

Preuve 2.7.1 *On a :*

1.

$$\begin{aligned} E(\nu_i) &= E(\sqrt{\Sigma_T^{-1}}\xi_i) = \sqrt{\Sigma_T^{-1}}E(\eta_i + \gamma u_i + \psi_i - \zeta_i\beta_i) \\ &= \sqrt{\Sigma_T^{-1}}(E(\eta_i) + \gamma E(u_i) + E(\psi_i) - E(\zeta_i)\beta_i) \\ &= 0 \end{aligned}$$

2.

$$\begin{aligned} E(W_i'\nu_i) &= E(W_i'\sqrt{\Sigma_T^{-1}}(\eta_i + \gamma u_i + \psi_i - \zeta_i\beta_i)) \\ &= E(W_i'\sqrt{\Sigma_T^{-1}}\eta_i) + E(W_i'\varepsilon_i) - E(W_i'\sqrt{\Sigma_T^{-1}}\zeta_i)\beta_i \\ &= 0 \end{aligned}$$

3.

$$\begin{aligned} E((W_i^*)'\nu_i) &= E((X_i^*)'\sqrt{\Sigma_T^{-1}}\nu_i) \\ &= E((X_i' + \zeta_i')\sqrt{\Sigma_T^{-1}}\nu_i) \\ &= E(W_i'\nu_i) + E(\zeta_i'\sqrt{\Sigma_T^{-1}}\nu_i) \\ &= E(\zeta_i'\sqrt{\Sigma_T^{-1}}\sqrt{\Sigma_T^{-1}}(\eta_i + \gamma u_i + \psi_i - \zeta_i\beta_i)) \\ &= E\Big[(\sqrt{\Sigma_T^{-1}}\zeta_i)'(\sqrt{\Sigma_T^{-1}}\eta_i + \sqrt{\Sigma_T^{-1}}\varepsilon_i - \sqrt{\Sigma_T^{-1}}\zeta_i\beta_i)\Big] \\ &= E(\tilde{\zeta}_i'\tilde{\eta}_i) + E(\tilde{\zeta}_i'\varepsilon_i) - E(\tilde{\zeta}_i'\zeta_i)\beta_i \\ &= -E(\tilde{\zeta}_i'\zeta_i)\beta_i \end{aligned}$$

Corollaire 2.7.1 *L'estimateur β^{BOLS} est biaisé négativement est non convergent.*

Preuve 2.7.2 *La troisième propriété stochastique de ν_i (proposition 2.3.1 au dessus), montre que l'estimateur β^{BOLS} est biaisé négativement avec β_i.*
D'autre part on a :

$$\begin{aligned} z_i^* = W_i^*\beta_i + \nu_i &\Leftrightarrow ((W_i^*)'W_i^*)^{-1}(W_i^*)'z_i^* = ((W_i^*)'W_i^*)^{-1}(W_i^*)'W_i^*\beta_i + ((W_i^*)'W_i^*)^{-1}(W_i^*)'\nu_i \\ &\Leftrightarrow ((X_i^*)'\sqrt{\Sigma_T^{-1}}\sqrt{\Sigma_T^{-1}}X_i^*)^{-1}(X_i^*)'\sqrt{\Sigma_T^{-1}}\sqrt{\Sigma_T^{-1}}y_i^* = \beta_i + ((X_i^*)'\sqrt{\Sigma_T^{-1}}\sqrt{\Sigma_T^{-1}}X_i^*)^{-1}(W_i^*)'\nu_i \\ &\Leftrightarrow ((X_i^*)'\Sigma_T^{-1}X_i^*)^{-1}(X_i^*)'\Sigma_T^{-1}y_i^* = \beta_i + (\frac{(X_i^*)'\Sigma_T^{-1}X_i^*}{T})^{-1}\frac{(W_i^*)'\nu_i}{T} \\ &\Leftrightarrow \beta^{BOLS} = \beta_i + (\frac{(X_i^*)'\Sigma_T^{-1}X_i^*}{T})^{-1}\frac{(W_i^*)'\nu_i}{T} \end{aligned}$$

Puisque : $\frac{(X_i^*)'\Sigma_T^{-1}X_i^*}{T}$ tend vers une matrice régulière quand $T \to +\infty$, on la note D_i, alors :

$$\begin{aligned} Plim_{T\to+\infty}\beta^{BOLS} &= \beta_i + D_i^{-1}Plim_{T\to+\infty}\frac{(W_i^*)'\nu_i}{T} \\ &= \beta_i - D_i^{-1}E((W_i^*)'\nu_i) \end{aligned}$$

Or, d'après la 3ème proporiété stochastique de ν_i^* de la proposition (2.3.1) :

$$E((W_i^*)'\nu_i) \neq 0$$

Donc β^{BOLS} ne converge pas en probabilité vers β_i à l'infinie.□

Le but est d'adapter la méthode des variables instrumentales, dans le cas de l'existence d'un processus brownien dans le modèle.
On considère les deux modèles $(\mathcal{MV}_i)$ et $(\mathcal{MVB}_i^*)$, l'objectif est de chercher une matrice des variables instrumentales (notée θ_i), où on va projeter les donner orthogonalement sur son image $Im(\theta_i)$.
On pose

$$\theta_i = \begin{pmatrix} \theta_{i,0,1} & \theta_{i,1,1} & \theta_{i,2,1} & . & . & \theta_{i,m,1} \\ \theta_{i,0,2} & \theta_{i,1,2} & \theta_{i,2,2} & . & . & \theta_{i,m,2} \\ . & . & . & . & . & . \\ . & . & . & . & . & . \\ . & . & . & . & . & . \\ \theta_{i,0,T} & \theta_{i,1,T} & \theta_{i,2,T} & . & . & \theta_{i,m,T} \end{pmatrix} \tag{2.7.11}$$

Telle que :

$$E(\theta_i^{'}\nu_i) = 0 \qquad cov(\theta_i^{'}W_i^*) \neq 0$$

On remarque que :

$$E(\theta_i^{'}z_i^*) = E(\theta_i^{'}W_i^*)\beta_i + E(\theta_i^{'}z_i^*) = E(\theta_i^{'}W_i^*)\beta_i$$

La matrice de projection orthogonale sur $Im(\theta_i)$ est donnée par :

$$P_{\theta_i} = \theta_i(\theta_i^{'}\theta_i)^{-1}\theta_i^{'} \tag{2.7.12}$$

Remarque 2.7.1 *La matrice P_{θ_i} est idempotente et symétrique. En effet :*

$$\begin{aligned} P_{\theta_i}^2 &= \theta_i(\theta_i^{'}\theta_i)^{-1}\theta_i^{'}\theta_i(\theta_i^{'}\theta_i)^{-1}\theta_i^{'} \\ &= \theta_i(\theta_i^{'}\theta_i)^{-1}\theta_i^{'} \\ &= P_{\theta_i} \end{aligned}$$

Et

$$P_{\theta_i}^{'} = (\theta_i(\theta_i^{'}\theta_i)^{-1}\theta_i^{'})^{'} = \theta_i((\theta_i^{'}\theta_i)^{'})^{-1}\theta_i^{'} = \theta_i(\theta_i^{'}\theta_i)^{-1}\theta_i^{'} = P_{\theta_i}$$

Nous définissons l'estimateur pour les variables instrumentales dans le cas d'existence du mouvement brownien et on le note β^{BIV} (Brownian instrumental variable), comme suit :

Définition 2.7.1 *On appelle estimateur de la méthode des variables instrumentales, dans le cas d'existence d'un mouvement brownien, associé au modèle $(\mathcal{MV}_i)$, l'estimateur noté β^{BIV}, défini par :*

$$\beta^{BIV} = ((X_i^*)^{'}\sqrt{\Sigma_T^{-1}}P_{\theta_i}\sqrt{\Sigma_T^{-1}}X_i^*)^{-1}(X_i^*)^{'}\sqrt{\Sigma_T^{-1}}P_{\theta_i}\sqrt{\Sigma_T^{-1}}y_i^* \tag{2.7.13}$$

Où :
X_i^* et y_i^* sont les observations avec incertitude, qui se font pour les vraies variables X_i et y_i.

Théorème 2.7.1 *On considère le modèle $(\mathcal{MV}_i)$ et son modèle auxilière $(\mathcal{MVB})^*$.*
Soient :

— *X_i^* et y_i^* sont les observations avec incertitude, qui correspondent à les vraies variables X_i et y_i.*

— θ_i : *la matrice des instruments.*

Sous les hypothèses :

H1 $V(\nu_i|\theta_i, W_i^*) = \sigma^2 I_T$;

H2 $cov(\theta_i' W_i^*) \neq 0$;

H3 *la matrice* $\theta_i'\theta_i$ *est régulière ;*

H4 *la matrice* $\frac{\theta_i'\theta_i^*}{T}$ *tend vers une matrice finie quand* $T \to +\infty$.

L'estimateur β^{BIV} *est BLUE pour le modèle* $(\mathcal{MV}_i)$.

Preuve 2.7.3 *Soit le modèle auxiliaire*

$$(\mathcal{MVB}_i^*) : \quad z_i^* = W_i^* \beta_i + \nu_i$$

Alors on a :

$$(**) \qquad \theta_i' z_i^* = \theta_i' W_i^* \beta_i + \theta_i' \nu_i$$

Donc :

$$E(\theta_i'\nu_i|\theta_i, W_i^*) = \theta_i' E(\nu_i|\theta_i, W_i^*) = 0$$

Et

$$\begin{aligned} V(\theta_i'\nu_i|\theta_i, W_i^*) &= \theta_i' V(\nu_i|\theta_i, W_i^*)\theta_i \\ &= \sigma^2 \theta_i' I_T \theta_i \\ &= \sigma^2 \theta_i'\theta_i \end{aligned}$$

En multipliant encore une fois $(**)$ *par* $\sqrt{\theta_i'\theta_i}$, *et on pose :*

$$\tilde{z}_i^* = \sqrt{(\theta_i'\theta_i)^{-1}}\theta_i' z_i^*$$

$$\tilde{W}_i^* = \sqrt{(\theta_i'\theta_i)^{-1}}\theta_i' W_i^*$$

$$\tilde{\nu}_i = \sqrt{(\theta_i'\theta_i)^{-1}}\theta_i' \nu_i$$

Soit le nouveau modèle auxiliaire :

$$(\mathcal{M}_i^*) : \quad \tilde{z}_i^* = \tilde{W}_i^* \beta_i + \tilde{\nu}_i$$

Or

$$E(\tilde{\nu}_i|\theta_i, W_i^*) = \sqrt{(\theta_i'\theta_i)^{-1}}\theta_i' E(\nu_i|\theta_i, W_i^*) = 0$$

Et

$$\begin{aligned} V(\tilde{\nu}_i|\theta_i, W_i^*) &= \sqrt{(\theta_i'\theta_i)^{-1}}\theta_i' V(\nu_i|\theta_i, W_i^*)\theta_i \sqrt{(\theta_i'\theta_i)^{-1}} \\ &= \sigma^2 \sqrt{(\theta_i'\theta_i)^{-1}}\theta_i' I_T \theta_i \sqrt{(\theta_i'\theta_i)^{-1}} \\ &= \sigma^2 \sqrt{(\theta_i'\theta_i)^{-1}} \sqrt{\theta_i'\theta_i}^2 \sqrt{\theta_i'\theta_i} \\ &= \sigma^2 I_T \end{aligned}$$

D'autre part on a :

$$\begin{aligned} Plim_{T\to+\infty}\frac{\tilde{W}_i^*\tilde{\nu}_i}{T} &= Plim_{T\to+\infty}\frac{1}{T}\sqrt{(\theta_i'\theta_i)^{-1}}\theta_i'W_i^*\sqrt{(\theta_i'\theta_i)^{-1}}\theta_i'\nu_i \\ &= Plim_{T\to+\infty}\frac{1}{T}W_i^*\theta_i\sqrt{(\theta_i'\theta_i)^{-1}}^2\theta_i'\nu_i \\ &= Plim_{T\to+\infty}\frac{1}{T}W_i^*\theta_i(\theta_i'\theta_i)^{-1}\theta_i'\nu_i \\ &= Plim_{T\to+\infty}\frac{W_i^*\theta_i}{T}\frac{\left(\theta_i'\theta_i\right)}{T}^{-1}\frac{\theta_i'\nu_i}{T} \end{aligned}$$

Sous l'hypothèse (H3) du théorème, la matrice $\theta_i'\theta_i^*$ est inversible, et :
$Plim \to +\infty\frac{\theta_i'\theta_i^*}{T} = D_i^{-1}$, $Plim_{T\to+\infty}\frac{W_i^*\theta_i}{T} = cov(\theta_i'(W_i^*)'$ *et* $Plim_{T\to+\infty}\frac{\theta_i'\nu_i}{T} = E(\theta_i'\nu_i|\theta_i, W_i^*) = 0$.
Alors :

$$Plim_{T\to+\infty}\frac{\tilde{W}_i^*\tilde{\nu}_i}{T} = 0$$

D'autre part, on considère la matrice de projection orthogonal sur $Im(\theta_i)$:

$$P_{\theta_i} = \theta_i(\theta_i'\theta_i)^{-1}\theta_i'$$

Alors :

$$\begin{aligned} Plim_{T\to+\infty}\frac{(\tilde{W}_i^*)'\tilde{\nu}_i}{T} &= Plim_{T\to+\infty}\frac{(W_i^*)'P_{\theta_i}P_{\theta_i}\nu_i}{T} \\ &= Plim_{T\to+\infty}\frac{(W_i^*)'P_{\theta_i}'P_{\theta_i}\nu_i}{T} \\ &= Plim_{T\to+\infty}\frac{(P_{\theta_i}W_i^*)'P_{\theta_i}\nu_i}{T} \end{aligned}$$

Donc la transformation P_{θ_i} élimine la corrélation entre les variables explicatives transformées et les perturbations.
Par suite d'après le théorème de Gauss, l'etimateur de moindre carrée ordiniaire est BLUE pour le modèle $(\mathcal{M}_i^*)$ et il est donnée par :

$$\beta^* = ((\tilde{W}_i^*)'\tilde{W}_i^*)^{-1}(\tilde{W}_i^*)'\tilde{z}_i^*$$

Or

$$\begin{aligned} (\tilde{W}_i^*)'\tilde{W}_i^* &= (\sqrt{(\theta_i'\theta_i)^{-1}}\theta_i'W_i^*)'\sqrt{(\theta_i'\theta_i)^{-1}}\theta_i'W_i^* \\ &= (W_i^*)'P_{\theta_i}W_i^* \end{aligned}$$

Et

$$\begin{aligned} (\tilde{W}_i^*)'\tilde{z}_i^* &= (\sqrt{(\theta_i'\theta_i)^{-1}}\theta_i'W_i^*)'\sqrt{(\theta_i'\theta_i)^{-1}}\theta_i'z_i^* \\ &= (W_i^*)'P_{\theta_i}z_i^* \end{aligned}$$

Donc

$$\begin{aligned} \beta^* &= ((W_i^*)'P_{\theta_i}W_i^*)^{-1}(W_i^*)'P_{\theta_i}z_i^* \\ &= ((X_i^*)'\sqrt{\Sigma_T^{-1}}P_{\theta_i}\sqrt{\Sigma_T^{-1}}X_i^*)^{-1}(X_i^*)'\sqrt{\Sigma_T^{-1}}P_{\theta_i}\sqrt{\Sigma_T^{-1}}^*y_i \\ &= \beta^{BIV} \end{aligned}$$

Conclusion : L'estimateur β^{BIV} est BLUE pour le modèle $(\mathcal{MV}_i)$.

2.7.2 Méthode de moment généralisé avec mouvement brownien

L'objective de cette sous section et de trouver l'estimateur BLUE dans le cas où l'hypothèse d'autocorrélation des erreur pour le modèle $(\mathcal{MVB}_i^*)$ est violée.
On considère donc que :

$$V(\nu_i|\theta_i, W_i^*) = \Omega_i \neq \sigma^2 I_T$$

On adapte la méthode de moment généralisé existante dans la litérature avec le fait d'existence du mouvement brownien, et on défini un nouvel estimateur obtenu par l'application de la méthode des moments généralisé sur un modèle auxilière associé au modèle $(\mathcal{MV}_i)$. Puis on montre que cet estimateur converge en probabilité vers la vraie valeur β_i. Nous appelons cet estimateur ; l'estimateur de la méthode de moment généralisé sous l'existence du mouvement brownien "**Generalized moment method under existence of Brownian motion**". On le note : β^{G2MBM}. On le défini comme suit :

Définition 2.7.2 *Soient X_i^* et y_i^* les observations avec incertitude, correspondant aux vraies variables X_i et y_i et θ_i est la matrice des variables instrumentales.*
On pose :

$$\mathfrak{P}_{\theta_i} = \theta_i(\theta_i'\Omega_i\theta_i)^{-1}\theta_i'$$

On suppose que :

$$V(\nu_i|\theta_i, W_i^*) = \Omega_i \neq \sigma^2 I_T$$

On appelle estimateur de la méthode du moment généralisé sous l'existence du mouvement brownien, associe au modèle $(\mathcal{MV}_i)$, l'estimateur noté β^{G2MBM}, défini par :

$$\beta^{G2MBM} = ((X_i^*)'\sqrt{\Sigma_T^{-1}}\mathfrak{P}_{\theta_i}\sqrt{\Sigma_T^{-1}}X_i^*)^{-1}(X_i^*)'\sqrt{\Sigma_T^{-1}}\mathfrak{P}_{\theta_i}\sqrt{\Sigma_T^{-1}}y_i^* \tag{2.7.14}$$

On considère le modèle $(\mathcal{MV}_i)$ et sont modèle auxilière $(\mathcal{MVB})_i^*$.
On pose

$$\hat{z}_i^* = \sqrt{\Omega_i^{-1}}z_i^*$$
$$\hat{W}_i^* = \sqrt{\Omega_i^{-1}}W_i^*$$
$$\hat{\nu}_i = \sqrt{\Omega_i^{-1}}\nu_i$$

alors on a :

$$\begin{aligned} z_i^* = W_i^*\beta_i + \nu_i &\Leftrightarrow \sqrt{\Omega_i^{-1}}z_i^* = \sqrt{\Omega_i^{-1}}W_i^*\beta_i + \sqrt{\Omega_i^{-1}}\nu_i \\ &\Leftrightarrow \theta_i'\sqrt{\Omega_i^{-1}}z_i^* = \theta_i'\sqrt{\Omega_i^{-1}}W_i^*\beta_i + \theta_i'\sqrt{\Omega_i^{-1}}\nu_i \end{aligned}$$

Donc le modèle $(\mathcal{MVB}^*)_i$ est équivalent au modèle suivant :

$$(\mathcal{M}_i^*) : \quad \hat{z}_i^* = \hat{W}_i^*\beta_i + \hat{\nu}_i$$

Or

$$E(\hat{\nu}_i|\theta_i, W_i^*) = \sqrt{\Omega_i^{-1}}E(\nu_i|\theta_i, W_i^*) = 0$$

Et

$$V(\hat{\nu}_i|\theta_i, W_i^*) = \sqrt{\Omega_i^{-1}}V(\nu_i|\theta_i, W_i^*)\sqrt{\Omega_i^{-1}}' = \sqrt{\Omega_i^{-1}}\Omega_i\sqrt{\Omega_i^{-1}} = I_T$$

Donc on peut appliquer la méthode des variables instrumentales précedente sur le modèle auxilière $(\mathcal{M}_i^*)$ qui est équivalent à $(\mathcal{MV}_i)$ et par les mêmes étapes de la preuve du théorème (2.7.3), on montre le lemme suivant :

Lemme 2.7.2 *Soient :*

— X_i^* *et* y_i^* *les observations avec incertitude, qui correspondent aux vraies variables* X_i *et* y_i.
— θ_i *: la matrice des instruments.*

Sous les hypothèses :

H2 $cov(\theta_i' \tilde{W}_i^*) \neq 0$ *;*

H3 *la matrice* $\theta_i' \Omega^{-1} \theta_i$ *est régulière, et que* $Plim_{T \to +\infty} \frac{\theta_i' \Omega^{-1} \theta_i^*}{T} = D_i \neq 0$.

Alors l'estimateur β^{G2MBM} *est BLUE pour le modèle* $(\mathcal{MV}_i)$.

2.8 Sélection du modèle optimal

On dispose de T observations de m variables candidates pour expliquer une variable endogène, sachant que ces observations proviennent d'un système qui contient un mouvement brownien.
Considérons la problématique suivante :
"***Chercher le meilleur modèle optimal qui explique la variable endogène en fonction d'une partie des variables candidates***".
Soient :
- $X_i \in \mathcal{M}_{T,m+1}(\mathbf{IR})$ est la matrice qui contient toutes les observations des m variables candidates, telle que sa première colone est le vecteur $(1, 1, ..., 1)'$.
- Σ_T est la matrice du processus brownien $(u(i,t)_{t \in [|1,T|]})$ (définie plus haut).
La procédure que nous proposons pour selectioner le modèle optimale est comme suit :
Étape 1 : On calcul le discriminant Δ_i, qui est donné par :

$$\Delta_i = \det(X_i' \Sigma_T^{-1} X_i) \tag{2.8.1}$$

Étape 2 : On distingue ici entre deux Cas ; $\Delta_i = 0$ ou bien $\Delta_i \neq 0$.
Cas 1 :$\Delta_i = 0$
Dans ce cas, on pose :

$$w_{i,k,t} = \sum_{l=1}^{T} \rho_{t,l} x_{i,k,l}, \ \ \forall k \in [|1, m|] \tag{2.8.2}$$

Et

$$z_{i,t} = \sum_{l=1}^{T} \rho_{t,l} y_{i,l}, \ \ \forall k \in [|1, m|] \tag{2.8.3}$$

Avec $(\rho_{t,1}, \rho_{t,2}, ..., \rho_{t,T})$ est la $t^{ème}$ ligne de la matrice $\sqrt{\Sigma_T^{-1}}$.
Puis on applique la **régression par étage : (SR)** (" Stagewise Regression "), pour cela on suit la procédure suivante :

1. Calculer les coefficients de corrélation $cov(z_{i,t}, w_{i,k,t})$ $\forall k \in [|1, m|]$, puis choisir $w_{i,j,t}$, tel que :

$$cov(z_{i,t}, w_{i,j,t}) = \max\{cov(z_{i,t}, w_{i,k,t}), \ k \in [|1, m|]\}$$

2. Calculer les résidus $\hat{\nu}_{i,t,1}$ de la régression de $z_{i,t}$ sur $w_{i,j,t}$, soit :

$$\hat{\nu}_{i,t,1} = z_{i,t} - \hat{\beta}_0 - \hat{\beta}_1 w_{i,j,t}$$

3. Calculer les coefficients de corrélation : $cov(\hat{\nu}_{i,t,1}, w_{i,k,t})\ \forall k \in [|1,m|] \setminus j$. Puis choisir $w_{i,l,t}$, tel que :

$$cov(\hat{\nu}_{i,t,1}, w_{i,j,t}) = \max\{cov(\hat{\nu}_{i,t,1}, w_{i,k,t}),\ k \in [|1,m|] \setminus j\}$$

4. Revenir à (2), et Calculer les résidus $\hat{\nu}_{i,t,2}$, de la régression de $z_{i,t}$ sur $\{w_{i,j,t}, w_{i,l,t}\}$, soit ;

$$\hat{\nu}_{i,t,2} = z_{i,t} - \hat{\beta}_0 - \hat{\beta}_1 w_{i,j,t} - \hat{\beta}_2 w_{i,l,t}$$

5. La procédure est répétée, on obtient donc un nouveau résidu.
La condition d'arrêt est telle que les coefficients de corrélation ne sont plus significativement différents de 0.

Cas 2 :$\Delta_i \neq 0$
Dans ce cas on test la multicolinéarité entre les variables $w_{i,k,t},\ k \in [|1,m|]\}$, en utilisant le test de Farrar et Glauber (voir section 2.2.3). Selon le résultat du test, nous proposons :

1. Dans le cas d'absence de multicolinéarité entre les variables $w_{i,k,t},\ k \in [|1,m|]\}$, on test l'autocorrélation des erreurs du modèle auxilière :

$$z_{i,t} = \alpha_{i,0} + \sum_{l=1}^{m} \alpha_{i,k} w_{i,k,t} + \nu_{i,t}$$

- Si l'hypothèse (H_0^{aut}) est retenu, alors on estime les coefficients du modèle $(\mathcal{MV}_i)$ par l'estimateur β_i^{BOLS}.
- Si l'hypothèse (H_1^{aut}) est retenu, alors on estime les coefficients du modèle $(\mathcal{MV}_i)$ par l'estimateur β_i^{GBLS}.
- On applique le test de student selon le type de l'estimateur ($BOLS$ ou $GBLS$), pour tout $k \in [|0,m|]\}$:

$$t^*_{i,k,BOLS} = \frac{\beta^{BOLS}_{i,k}}{\sigma_{\beta^{BOLS}_{i,k}}},\quad avec:\ V(\beta^{BOLS}_{i,k}) = \sigma^2_{\beta^{BOLS}_{i,k}} \tag{2.8.4}$$

Où bien :

$$t^*_{i,k,GBLS} = \frac{\beta^{GBLS}_{i,k}}{\sigma_{\beta^{GBLS}_{i,k}}},\quad avec:\ V(\beta^{GBLS}_{i,k}) = \sigma^2_{\beta^{GBLS}_{i,k}} \tag{2.8.5}$$

- Au niveau de confiance $1-\alpha$, on élimine l'élement $w_{i,k,t}$ du modèle auxiliaire et $x_{i,k,t}$ du modèle $(\mathcal{M}_i)$, si $t^*_{i,k,BOLS} \leq t^{\alpha/2}_{T-m-1}$ (respectivement si $t^*_{i,k,GLSBM} \leq t^{\alpha/2}_{T-m-1}$, puis en réestimant le modèle auxilière, jusqu'à ce qu'on arrive à garder seulement les élements tel que la statistique de student de ces coefficients est supérieur strictement à la valeur théorique $t^{\alpha/2}_{T-m-1}$.
le modèle obtenu est le modèle linéaire optimal qui explique la variable endogène par les variables candidates sélectionnées.

2. Dans le cas d'existence d'une multicolinéarité imparfaite au niveau de confiance $1-\alpha$, on propose d'utiliser les critères de Akaike [3] ou de Schwarz [4], sur le modèle auxiliaire. Rappelons que les critères de Akaike (AIK : Akaike Information Criterion) est donné par :

$$AIC = \ln\Big(\frac{SCR}{T}\Big) + \frac{2n}{T} \tag{2.8.6}$$

Le critère de Schwarz (SC : Schwarz Criterion) est donné par :

$$SC = \ln\Big(\frac{SCR}{T}\Big) + \frac{n\ln(T)}{T} \tag{2.8.7}$$

Le modèle optimale est celui qui minimise l'un des critère "AIK" ou "SC".

Où :
T : est le nombre d'observation
SCR : somme des carrés des résidus du modèle.
n : le nombre de variables condidats utiliser dans le modèle.

2.9 Simulation des méthodes : Estimation des élasticités des facteurs d'une fonction Cobb dauglass stochastique et modélisation du PIB

L'objectif de cette section est de faire la simulation par le logiciel R des méthodes qu'on a présenté théoriquement, afin de répondre à deux questions :

1. L'évolution de certains agrégats macroéconomiques, peut elle engendrer une évolution irrégulière sur le PIB ?
2. Quelles sont les estimations des élasticités des facteurs qui contribuent dans le PIB, ainsi que l'espérence de la productivité total de ces facteurs.

Les agrégats qu'on considère ici sont :
— La formation brute du capital, on la note K.
— La population active occupée, on la note L. On a :

$$L = (1-\gamma_c)L_A$$

Où : L_A est la population active et γ_c est le taux de chômage.
Avant de répondre à ces question, on donne l'alghorithme qu'on propose pour calculer $\sqrt{\Sigma_T^{-1}}$, puis on fait la simulation de cet algorithme, pour 10 périodes.

2.9.1 Algorhithme de calcul de la matrice $\sqrt{\Sigma_T^{-1}}$

Simulation sous le logiciel R, pour 10 périodes successives :
Les figures (2.2-2.3), sont des implémentation numérique de l'algorhithme 1, pour 10 périodes succéssives :

Algorithm 1: Algorithme de calcul de $\sqrt{\Sigma_i}$ et $\sqrt{\Sigma_i}^{-1}$

Require: T Nombre de périodes.

Ensure: FUNCTION Var.brownien=function(T)

A=matrix(rep(0,T^2),ncol=T)
FOR $i = 1$ jusqu'à T
FOR $j = 1$ jusqu'à i
$A[i, j] = j$ et $A[j, i] = j$
ENDFOR
ENDFOR
RETURN A
ENDFUNCTION
FUNCTION RCI_VBROW=function(T)
P=eigen(Var.brownien(T))$vectors
D=diag(eigen(Var.brownien(T))$values)
d=diag(sqrt(eigen(Var.brownien(T))$values))
dd=diag(1/sqrt(eigen(Var.brownien(T))$values))
R=P*d*P'
RI=inverse(R)
M=list(Racine_Carree_Brownien=R,Racine_Carree_Inverse_Brownien=RI)
RETURN M
ENDFUNCTION

```
> Var.brownien(10)
      [,1] [,2] [,3] [,4] [,5] [,6] [,7] [,8] [,9] [,10]
 [1,]    1    1    1    1    1    1    1    1    1     1
 [2,]    1    2    2    2    2    2    2    2    2     2
 [3,]    1    2    3    3    3    3    3    3    3     3
 [4,]    1    2    3    4    4    4    4    4    4     4
 [5,]    1    2    3    4    5    5    5    5    5     5
 [6,]    1    2    3    4    5    6    6    6    6     6
 [7,]    1    2    3    4    5    6    7    7    7     7
 [8,]    1    2    3    4    5    6    7    8    8     8
 [9,]    1    2    3    4    5    6    7    8    9     9
[10,]    1    2    3    4    5    6    7    8    9    10
```

FIGURE 2.1 – matrice de covariance pour 10 périodes

```
> W=RCI_VBROW(10)
> W
$Racine_Carree_Brownien
            [,1]       [,2]      [,3]      [,4]      [,5]      [,6]      [,7]
 [1,] 0.85121187 0.3443394 0.2255795 0.1716131 0.1413350 0.1225935 0.1104867
 [2,] 0.34433942 1.0767913 0.5159525 0.3669145 0.2942066 0.2518217 0.2252797
 [3,] 0.22557946 0.5159525 1.2181263 0.6385461 0.4774012 0.3968928 0.3498116
 [4,] 0.17161312 0.3669145 0.6385461 1.3286130 0.7412322 0.5753911 0.4926687
 [5,] 0.14133502 0.2942066 0.4774012 0.7412322 1.4266029 0.8370082 0.6711670
 [6,] 0.12259352 0.2518217 0.3968928 0.5753911 0.8370082 1.5223789 0.9349981
 [7,] 0.11048667 0.2252797 0.3498116 0.4926687 0.6711670 0.9349981 1.6250650
 [8,] 0.10268614 0.2084766 0.3210556 0.4455876 0.5906587 0.7738532 1.0454847
 [9,] 0.09798991 0.1984621 0.3042526 0.4190455 0.5482737 0.7011453 0.8964467
[10,] 0.09577597 0.1937659 0.2964520 0.4069387 0.5295322 0.6708672 0.8424803
            [,8]       [,9]      [,10]
 [1,] 0.1026861 0.09798991 0.09577597
 [2,] 0.2084766 0.19846211 0.19376588
 [3,] 0.3210556 0.30425255 0.29645202
 [4,] 0.4455876 0.41904553 0.40693869
 [5,] 0.5906587 0.54827371 0.52953221
 [6,] 0.7738532 0.70114533 0.67086723
 [7,] 1.0454847 0.89644669 0.84248035
 [8,] 1.7476586 1.18681977 1.06805981
 [9,] 1.1868198 1.91927167 1.41239923
[10,] 1.0680598 1.41239923 2.26361109
```

```
$Racine_Carree_Inverse_Brownien
             [,1]         [,2]         [,3]         [,4]         [,5]         [,6]
 [1,]  1.35808431 -0.388112483 -0.064793620 -0.023688824 -0.011536594 -0.006634662
 [2,] -0.38811248  1.293290692 -0.411800727 -0.076330214 -0.030322906 -0.015842905
 [3,] -0.06479362 -0.411800727  1.281754098 -0.418435389 -0.080636525 -0.033427202
 [4,] -0.02368824 -0.076330214 -0.418435389  1.277447794 -0.042153969 -0.083118820
 [5,] -0.01153659 -0.030322906 -0.080636525 -0.042153969  1.274965494 -0.042375363
 [6,] -0.00663466 -0.015842905 -0.033427202 -0.083118820 -0.042375363  1.272751552
 [7,] -0.00430631 -0.009738958 -0.018325200 -0.035641142 -0.014435193 -0.085332760
 [8,] -0.00310430 -0.006788606 -0.011952898 -0.020539140 -0.014435193 -0.038123437
 [9,] -0.00248229 -0.005318236 -0.009002546 -0.014435193 -0.012106842 -0.023643436
[10,] -0.00221394 -0.004696235 -0.007800531 -0.012106842 -0.018741504 -0.030278098
             [,7]         [,8]         [,9]        [,10]
 [1,] -0.00430631 -0.003104296 -0.002482295 -0.002213940
 [2,] -0.00973896 -0.006788606 -0.005318236 -0.004696235
 [3,] -0.01832520 -0.011952898 -0.009002546 -0.007800531
 [4,] -0.03564114 -0.020539140 -0.014435193 -0.012106842
 [5,] -0.08533276 -0.038123437 -0.023643436 -0.018741504
 [6,] -0.42623592 -0.088437056 -0.042429748 -0.030278098
 [7,]  1.26964726 -0.430542231 -0.095071718 -0.053966342
 [8,] -0.43054223  1.263012594 -0.442078825 -0.118759962
 [9,] -0.09507172 -0.442078825  1.239324350 -0.506872445
[10,] -0.05396634 -0.118759962 -0.506872445  0.851211867
```

FIGURE 2.2 – Racine et racine inverse de la matrice de covariance pour 10 périodes

2.9.2 Ilustration graphique de dépendence des agrégats avec le PIB

On désigne par :
-Y le vecteur des 30 observations du PIB, et y est le vecteur des 30 observations du $\ln(PIB)$.
-K et L les vecteurs des 30 observations de leurs agrégats. Ainsi k, l sont les vecteurs des 30 observations du $\ln(Agregats)$.
-$z = \sqrt{\Sigma_T^{-1}}y$, $kk = \sqrt{\Sigma_T^{-1}}k$ et $ll = \sqrt{\Sigma_T^{-1}}l$.
On a téléchargé les données à partir du site de la Banque Mondiale [3]. Après traitements on prend les données correspondant au Maroc et au Royaume Uni, entre : (1987-2017). Les graphes illustratifs correspondent aux nuages de points de PIB et de quelques agrégats, sont données par les figures (2.3-2.14) :

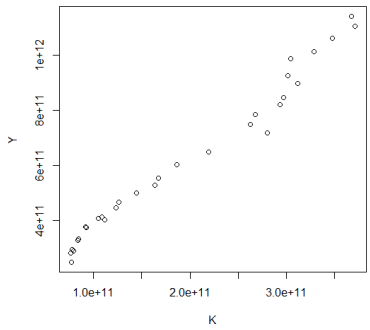


FIGURE 2.3 – Nuage de points du PIB et de Formation brute du capital (Maroc)

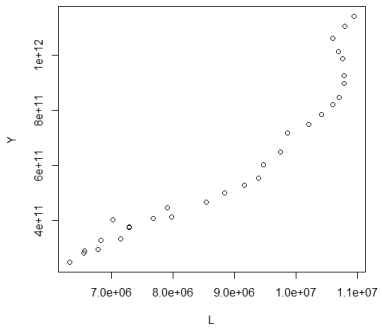


FIGURE 2.4 – Nuage de points du PIB et de population active occupée (Maroc)

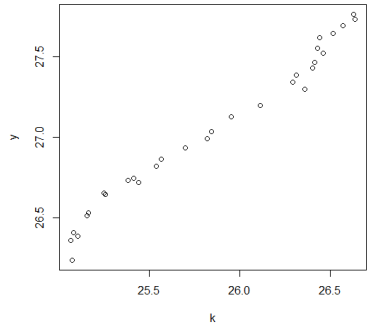


FIGURE 2.5 – Nuage de points de logarithme du PIB et de logharitme de Formation brute du capital (Maroc)

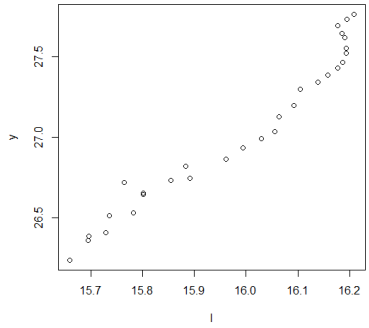


FIGURE 2.6 – Nuage de points de logarithme du PIB et de logharitme de population active occupée (Maroc)

3. Les données sont accéssibles et téléchargeables sur le site de la Banque Mondiale (World Bank Open Data) : https ://data.worldbank.org/

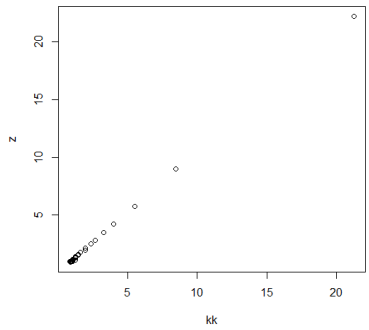


FIGURE 2.7 – Nuage de points après élimination de l'effet de mouvement brownien de logarithme du PIB et de logharitme de Formation brute du capital (Maroc)

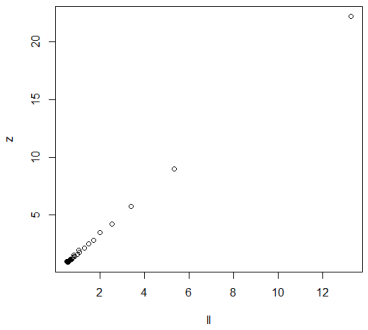


FIGURE 2.8 – Nuage de points après élimination de l'effet de mouvement brownien de logarithme du PIB et de logharitme de population active occupée (Maroc)

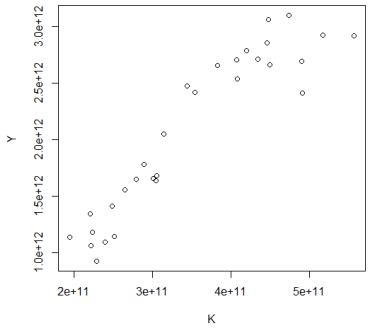


FIGURE 2.9 – Nuage de points du PIB et de Formation brute du capital (Royaume Uni)

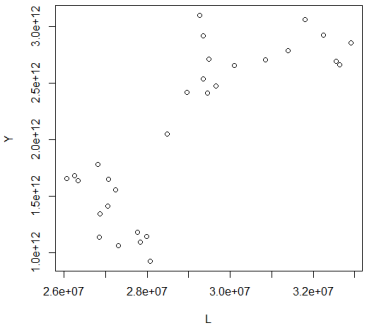


FIGURE 2.10 – Nuage de points du PIB et de population active occupée (Royaume Uni)

D'après le chapitre 1, plusieurs économistes modélisent le PIB par une fonction Cobb Dauglass (1.2.7). Les figures (2.3&2.9) peuvent illustrer ça, on peut considérer que la dépendance de PIB avec la formation brute de capital est une fonction exponentielle de base $\alpha > 0$. C'est pour cela qu'en économétrie, les staticiens expliquent le logarithme du PIB avec le logarhithmes des autres agrégats (K,L,...).
Mais si on considère les figures (2.5&2.11), on voit que même si en appliquant le logarithme sur le PIB et force du travail (L), les nuages de points ne s'approchent pas de la droite de régression.
Par contre lorsqu'on applique la transformation $\sqrt{\Sigma_T^{-1}}$ sur le logarithme des données, comme l'illustrent les figures (2.7,2.8,2.13 et 2.14), on voit que les nuages de points forme une droite, ce qui nous pousse à poser la première question (Q1).

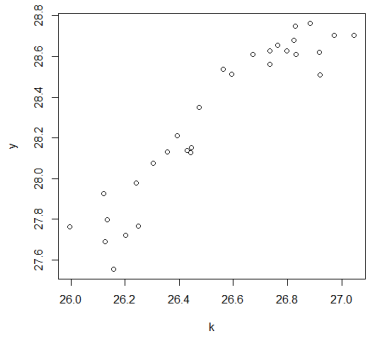


FIGURE 2.11 – Nuage de points de logarithme du PIB et de logharitme de Formation brute du capital (Royaume Uni)

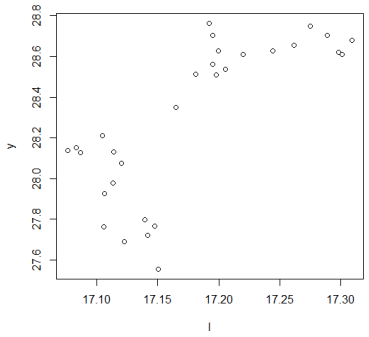


FIGURE 2.12 – Nuage de points de logarithme du PIB et de logharitme de population active occupée (Royaume Uni)

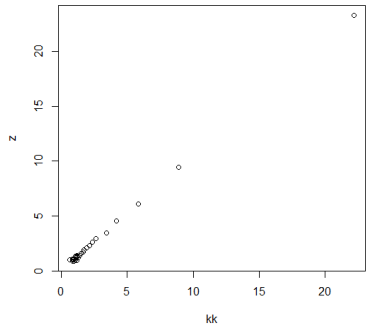


FIGURE 2.13 – Nuage de points après élimination de l'effet de mouvement brownien de logarithme du PIB et de logharitme de Formation brute du capital (Royaume Uni)

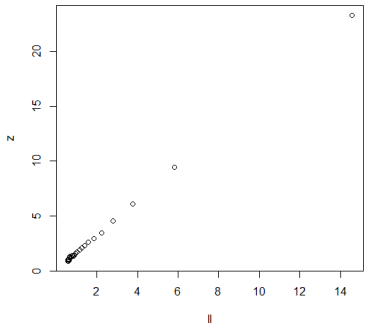


FIGURE 2.14 – Nuage de points après élimination de l'effet de mouvement brownien de logarithme du PIB et de logharitme de population active occupée (Royaume Uni)

2.9.3 Estimation par moindres carrés et moindres carrés brownien des élasticités

Pour répondre à la question (Q1), ainsi que la question, on propose de tester deux modèles, puis utiliser le critère AIC (2.3.26), pour déterminer le modèle qui explique bien le PIB en fonction des Agrégats précisés au dessus. Le premier modèle consiste à considérer, qu'il y a un mouvement brownien géométrique qui sert à une évolution irrégulière du PIB en fonctions des chocs exercés par les agrégats dans le temps. Autrement dit, on considère que le PIB est une fonction Cobb Dauglass stochastique :
On pose donc

$$PIB_{i,t} = F_i(K_{i,t}, L_{i,t}, \beta_i, u_{i,t}) = C_i K_{i,t}^{\alpha_i} (L_{i,t})^{\nu} e^{u_{i,t}} \tag{2.9.1}$$

Où :

$$\beta_i = \begin{pmatrix} \ln(C_i) \\ \alpha_i \\ \nu_i \end{pmatrix}$$

Et : $(u_{i,t})_{t\geq 0}$ est un processus brownien La transformation convenable pour la fonction (2.8.1) est $h = \ln$.
Donc :

$$\begin{aligned} y_{i,t} &= h(PIB_{i,t}) = h \circ F_i(K_{i,t}, L_{i,t}, \beta_i, u_{i,t}) \\ &= c_i + \alpha_i k_{i,t} + \nu l_{i,t} + u_{i,t} \end{aligned}$$

Avec le fait que les lettres miniscules signifient les logarithmes nipériens des agrégats qui leur correspondent.
Donc le premier modèle à tester est :

$$(\mathfrak{m}_1) \quad y_{i,t} = c_i + \alpha_i k_{i,t} + \eta_i r_{i,t} + \nu l_{i,t} + u_{i,t} + \psi_{i,t}$$

Le deuxième modèle sert à proposer que les chocs des agrégats sur le PIB, ne génère pas un mouvement très irrégulière sur l'évolution de PIB. Soit donc :

$$PIB_{i,t} = G_i(K_{i,t}, L_{i,t}, \beta_i,) = C_i K_{i,t}^{\alpha_i} (L_{i,t})^{\nu} \tag{2.9.2}$$

On pose donc :

$$(\mathfrak{m}_2) \quad g_{i,t} = c_i + \alpha_i k_{i,t} + \eta_i r_{i,t} + \nu l_{i,t} + \upsilon_{i,t}$$

Pour les données du Maroc, on applique l'estimateur β^{BOLS} pour le modèle $\mathfrak{m}_1$, car le modèle contient un mouvement brownien. Par contre, pour le modèle $\mathfrak{m}_2$, on applique l'estimateur de moindre carrée ordinaire β^{OLS} car le modèle ne contient pas un mouvement brownien.
On trouve les résultats suivants :

```
> summary(m1)

Call:
lm(formula = z ~ kk + ll)

Residuals:
      Min        1Q    Median        3Q       Max
-0.032440 -0.001931  0.001165  0.002758  0.011416

Coefficients:
             Estimate Std. Error t value Pr(>|t|)
(Intercept) -0.004045   0.001450   -2.79  0.00956 **
kk           0.310747   0.018747   16.58 1.12e-15 ***
ll           0.748589   0.016641   44.98  < 2e-16 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.007893 on 27 degrees of freedom
Multiple R-squared:  0.9998,    Adjusted R-squared:  0.9998
F-statistic: 7.589e+04 on 2 and 27 DF,  p-value: < 2.2e-16
```

TABLE 2.1 – Estimation de paramètres du modèle $\mathfrak{m}_1$ par l'estimateur β^{BOLS}

```
> summary(m2)

Call:
lm(formula = y ~ k + l)

Residuals:
     Min       1Q   Median       3Q      Max
-0.02335 -0.01498 -0.00186  0.01052  0.05825

Coefficients:
             Estimate Std. Error t value Pr(>|t|)
(Intercept) -0.031885   0.004657  -6.847 2.35e-07 ***
k            0.231254   0.042375   5.457 8.94e-06 ***
l            0.793812   0.042741  18.572  < 2e-16 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.01805 on 27 degrees of freedom
Multiple R-squared:  0.9998,    Adjusted R-squared:  0.9998
F-statistic: 6.893e+04 on 2 and 27 DF,  p-value: < 2.2e-16
```

TABLE 2.2 – Estimation de paramètres du modèle $\mathfrak{m}_2$ par l'utilisation de l'estimateur β^{OLS}

Pour choisir entre les deux modèles $\mathfrak{m}_1$ ou $\mathfrak{m}_2$, on calcul le critère de "AKAIKE", on trouve que :

$$AIC(\mathfrak{m}_1) = -200.535, \qquad AIC(\mathfrak{m}_2) = -150.8959$$

Donc on retient le modèle $\mathfrak{m}_1$, et on conclut que :

1. Les chocs des évolutions des agrégats considérés sur le PIB provoquent une évolution irrégulière du PIB.
2. L'élasticité du capital K (Formation brute du capital) est $\alpha_{Maroc} = 0.310747$.
   L'élasticité de la force de travail L est $\nu_{Maroc} = 0.748589$.
   L'espérence du logarithme de la productité total des facteur est $c_{Maroc} = -0.004045 \Rightarrow C_{Maroc} = 0.9959632$.

Pour conclure, on peut dire que le modèle qui caractérise bien l'évolution du PIB du maroc, en fonction des agrégats considérés, est le modèle :

$$PIB_{Maroc,t} = 0.9959632 e^{u_{i,t}} K_t^{0.310747} (L_t)^{0.748589} \tag{2.9.3}$$

Avec $(u_{i,t})_{t\geq 0}$ est un mouvement brownien.

2.10 Conclusion

Dans ce chapitre, nous somme arrivés à trouver les estimateurs optimaux, dans le cas d'existence d'un processus brownien dans le modèle.
nous avons commencé par l'élimination de l'effet d'hétéroscédasticité des bruits, provenant d'un mouvement brownien, en multipliant le modèle par l'inverse de la racine carré de la matrice de covariance du processus brownien sur une gamme du temps fini. Puis nous avons modifié le test d'autocorélation des erreurs pour qu'il soit convenable avec le fait d'existence du mouvement brownien. Aussi nous nous sommes basés sur les méthodes des moindres carrés, affin d'établir les nouveaux estimateurs que nous avons proposé, qu'ils tiennent en compte le fait d'existence d'un processus brownien dans le modèle.
Comme application économétrique de ce qu'on a fait, nous avons établi une nouvelle version d'une fonction Cobb Dauglass avec mouvement brownien géométrique, tout en estimant les élasticités des agrégats influençant l'évolution du PIB.
Mais dans ce chapitre nous nous somme basés sur les hypothèses de l'absence de la multicolinéarité et l'homoscédasticité de l'erreur du modèle, après remédiation de l'hétéroscédasticité du mouvement brownien. Ce qui pose la question : que ce passe-t-il si ces deux hypothèses ne sont pas vérifiées ?
Dans le chapitre suivant, nous allons répondre à cette question.

Chapitre 3

Estimation régularisée et hétéroscédasticité des erreurs

3.1 Introduction

Dans le chapitre précédent on a éliminé l'hétéroscédasticité du processus brownien sur une gamme du temps fini. On a supposé que l'élimination de l'hétéroscédasticité du processus brownien entraîne une élimination de l'hétéroscédasticité de l'erreur dans l'erreur du modèle.
Nous avons modifié la méthode des moindres carrés existant dans la littérature, ainsi que le test de l'autocorrélation des erreurs, "Test de Breusch-Godfrey avec multiplicateur de lagrange", nous avons obtenu des nouveaux estimateurs efficaces du modèle $(\mathcal{M}_i)$.
Le troisième chapitre, est un prologement du chapitre deux..
Dans ce chapitre on rejette l'hypothèse d'homoscédasticité des erreurs après l'élimination de l'hétéroscédasticité du processus brownien, ainsi que l'hypothèse de l'absence d'une multicolinéarité dans les variables explicatives. Nous commençons le chapitre par l'étude de la régression ridge dans le cas d'existence de multicolinéarité et du processus brownien dans le modèle. Puis nous présentons une version du modèle GARCH, qui prend en compte le fait d'existence du mouvement brownien dans le modèle.
Nous terminons ce chapitre par une application économétrique, en prédisant le PIB et la frontière technologique du Maroc sur une gamme du temps $[|0, T|]$.
On garde les mêmes notations que celles du chapitre 2. Rappelons les modèles :

$$(\mathcal{M}_i)\ \ y_{i,t} = \beta_{i,0} + \sum_{k=1}^{m} \beta_{i,k} x_{i,k,t} + \gamma_i u_{i,t} + \psi_{i,t}$$

$$(\mathcal{MV}_i)\ \ y_i = X_i \beta_i + \gamma_i u_i + \psi_i$$

$$(\mathcal{MVB}_i)\ \ z_i = W_i \beta_i + \nu_i$$

Où : $z_i = \sqrt{\Sigma_T^{-1}} y_i$, $W_i = \sqrt{\Sigma_T^{-1}} X_i$ et $\nu_i = \sqrt{\Sigma_T^{-1}}(\gamma_i u_{i,t} + \psi_i)$.

3.2 Régression régularisée

3.2.1 Problématique

Dans le cas d'existence d'une multicolinéarité dans le modèle, les estimation par les méthodes qu'on a proposé dans le chapitre 2 ne seront pas sinificatives. Pour illustrer cela on considère la problématique suivante :
On veut savoir quel est l'impact des dépenses en éducation et en recherche et développement sur l'évolution de PIB au Maroc, et comme dans le chapitre précédent, on propose de tester deux modèles $\mathfrak{M}_1$ et $\mathfrak{M}_2$ de l'évolution de PIB :

$$\mathfrak{M}_1 \quad Y_{i,t} = A_t K_{i,t}^{\alpha_i} (R_{i,t})^{\eta_i} (L_{i,t})^{\nu_i} (H_{i,t})^{\lambda_i}, \quad A_t = C_i e^{u_{i,t}}$$

$$\mathfrak{M}_2 \quad Y_{i,t} = C_i K_{i,t}^{\alpha_i} (R_{i,t})^{\eta_i} (L_{i,t})^{\nu_i} (H_{i,t})^{\lambda_i}$$

Où :
C est une constante. $(u_{i,t})_{t\geq 0}$ est un mouvement brownien.
$K_{i,t}$: est la formation brute du capital.
$L_{i,t}$: La force du travail.
$R_{i,t}$: représente les dépences en recherche et developpement.
$H_{i,t}$: est le capital humain qui représente les dépenses en éducation multiplié par l'indice du capital humain.
Par passage logarithmique, on trouve :

$$\mathfrak{M}_1 \quad y_{i,t} = c_i + \alpha_i k_{i,t} + \eta_i (R_{i,t}) + \nu l_{i,t} + \lambda_i h_{i,t} + u_{i,t}$$

$$\mathfrak{M}_2 \quad Y_{i,t} = c_i + \alpha_i k_{i,t} + \eta_i (R_{i,t}) + \nu l_{i,t} + \lambda_i h_{i,t}$$

Où les notations minuscules sont les logarithmes niperiens des agrégats correspondant.
Nous nous basons sur les données publiés dans le site de la Banque mondiale, après traitement de données et simulation par logiciel pour 30 périodes succéssives, on trouve que pour le Maroc :
Les estimations par moindres carrés brownien des coefficients du modèle $\mathfrak{M}_1$, ainsi que les estimations par moindres carrés des coefficients du modèle $\mathfrak{M}_2$, donnent les résultats figurés dans les tableaux (3.1)-(3.2) :

```
Call:
lm(formula = z ~ kk + ll + rr + hh)

Residuals:
       Min         1Q     Median         3Q        Max
-0.0197673 -0.0018564  0.0002906  0.0025122  0.0100726

Coefficients:
             Estimate Std. Error t value Pr(>|t|)
(Intercept) -0.002059   0.001140  -1.807 0.082880 .
kk           0.322892   0.027703  11.656 1.34e-11 ***
ll           0.381433   0.071231   5.355 1.49e-05 ***
rr          -0.114933   0.058298  -1.971 0.059833 .
hh           0.443861   0.117880   3.765 0.000903 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.005664 on 25 degrees of freedom
Multiple R-squared:  0.9999,    Adjusted R-squared:  0.9999
F-statistic: 7.369e+04 on 4 and 25 DF,  p-value: < 2.2e-16
```

TABLE 3.1 – Estimations par moindres carrés brownien : modèle $\mathfrak{M}_1$ (Maroc)

```
Call:
lm(formula = y ~ k + l + r + h)

Residuals:
       Min         1Q     Median         3Q        Max
-0.0090073 -0.0042810 -0.0008292  0.0040657  0.0165137

Coefficients:
             Estimate Std. Error t value Pr(>|t|)
(Intercept) -0.008559   0.004388  -1.950   0.0624 .
k            0.323881   0.033538   9.657 6.45e-10 ***
l            0.131825   0.064414   2.047   0.0514 .
r           -0.207074   0.042941  -4.822 5.91e-05 ***
h            0.759124   0.065328  11.620 1.43e-11 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.006556 on 25 degrees of freedom
Multiple R-squared:      1,     Adjusted R-squared:      1
F-statistic: 2.614e+05 on 4 and 25 DF,  p-value: < 2.2e-16
```

TABLE 3.2 – Estimations par moindres carrés : modèle $\mathfrak{M}_2$ (Maroc)

D'après les deux tableaux on voit, qu'on a deux coefficients non significatifs dans les deux méthodes, ce qui remet en question la fiértés des résultats obtenus. Pour voir où se trouve le problème, on teste le degré de la multicolinéarité entre les logarithmes des agrégats considérés.
Lorqu'on test la multicolinéarité des agrégats, en utilisant le test de Farrar et Glauber :
La statistique de Farrar et Glauber sur la matrice de corrélation des agrégats donne la valeur : $\chi^2_{emp} = 547.2254$.
La statistique de Farrar et Glauber sur la matrice de corrélation des agrégats avec élimination de l'hétéroscédasticité du mouvement brownien pour 30 périodes succéssives, donne la valeur : $\chi^2_{emp*} = 469.1168$.
Or la valeur théorique de fractile, pour un niveau de confiance 0.95, avec un degré de liberté $\frac{1}{2}4(4+1) = 10$ donne : $\chi^2_{0.95,10} = 18.30704 < \chi^2_{emp*} < \chi^2_{emp}$.
On conclut donc qu'il y a une multicolinéarité entre les agrégats même après élimination de l'hétéroscédasticité du mouvement brownien sur les 30 périodes.
Par conséquence, les estimations des coefficients par moindres carrés ou par moindres carrés brownien seront instables.
Pour cette raison nous nous intéressons à trouver un estimateur éfficace, qui va prendre en considérations à la fois la multicolinéarité entre les variables explicatives et l'existence d'un mouvement brownien dans le modèle.
Dans la littérature, pour résoudre le problème de multicolinéarité pour les modèles linéaires, on peut appliquer l'une des méthodes de régréssion régularisés, citons par exemple la régression ridge [80], régression LASSO [81], Elasticnet [82]. L'inconvénient de ces méthodes est qu'elles ne seront pas valables, dans le cas d'existence d'un processus stochastiques dans le modèle. Autrement dit, la régression régularisée ne tient pas en compte, les divers chocs stochastiques qui peuvent être générés par des chocs brusque sur une ou des variables explicatives. En particulier l'application de régression régularisée pour le modèle $\mathfrak{M}_1$ ne sera pas utile même sous l'hypothèse $\mathcal{H}_1^{mc}$.
Pour cette raison nous modifions la régession régularisée pour qu'elle soit compatible avec l'existence du mouvement brownien dans le modèle. On va donner le nom : "Régression régularisée brownienne", en abrévréviation "**B2R**" ("**B**rownian **R**egularized **R**egression")
.
Avant de présenter les méthodes de régression régularisée que nous allons présenter par la suite, nous donnons une nouvelle version des normes qui tiennent en compte l'existence du mouvement brownien dans le modèle.

3.2.2 Norme brownienne et programme de régularisation

Définition-Proposition 3.2.0.1 *Soit K un corps et $E = K^n$ un espace vectoriel muni d'une norme $\|.\|$.*
On définit et on appelle norme brownien associé à $\|.\|$, la norme notée $\|.\|_{\mathcal{B}}$ définie par :

$$\|v\|_{\mathcal{B}} = \|\sqrt{\Sigma_n^{-1}}v\| \tag{3.2.1}$$

Où $\sqrt{\Sigma_n^{-1}}$ est l'inverse de la racine carré de la matrice de covariance du processus $u_i = ((u_{i,t})_{t\in[|1,n|]})'$ du mouvement brownien $(u_{i,t})_{t\geq 0}$.

Preuve 3.2.1 *L'application $\|.\|_{\mathcal{B}}$ est une norme, car :*
-$\|.\|_{\mathcal{B}}$ est une semi-norme (évident).

-D'autre part On a :

$$\|v\|_{\mathcal{B}} = 0 \Rightarrow \|\sqrt{\Sigma_n^{-1}}v\| = 0 \Rightarrow \sqrt{\Sigma_n^{-1}}v = 0$$
$$\Rightarrow \sqrt{\Sigma_n}\sqrt{\Sigma_n^{-1}}v = 0 \Rightarrow v = 0$$

L'autre sens est évident. donc :

$$v = 0 \Leftrightarrow \|v\|_{\mathcal{B}} = 0$$

.
Donc $\|.\|_{\mathcal{B}}$ est une norme. □.

Définition-Proposition 3.2.0.2 *Soit K un corps et $E = K^n$ un espace vectoriel muni d'une norme $\|.\|$.*
On définit et on appelle norme brownienne générale associée à $\|.\|$, la norme notée $\|.\|_{\mathcal{GB}}$ définie par :

$$\|v\|_{\mathcal{GB}} = \|\sqrt{\Omega_i^{-1}}\sqrt{\Sigma_n^{-1}}v\| \tag{3.2.2}$$

Où : $\sqrt{\Sigma_n^{-1}}$ est l'inverse de la racine carré de la matrice de covariance du processus $u_i = ((u_{i,t})_{t\in[|1,n|]})'$ du mouvement brownien $(u_{i,t})_{t\geq 0}$.
Ω_i est une matrice symétrique définit positive.

La preuve est identique à la précédente.

Définition 3.2.1 *Soient $\|.\|$ et $|||.|||$ deux normes réspectivement de $\mathbf{IR}^T$ et $\mathbf{IR}^{m+1}$, et $\|.\|_{\mathcal{B}}$ est la norme brownienne associée à $\|.\|$.*
*On définit et on nomme : "Programme régularisé brownien (BRP)" (**B**rownian **R**egularized **P**rogram), tout programme d'optimisation, de la forme :*

$$(\mathcal{BRP}) \begin{cases} \min_{\beta\in\Theta} \|y_i - X_i\beta_i\|_{\mathcal{B}}^2 \\ S.C \\ |||\beta_i|||^2 < \tau \end{cases}$$

Avec :
$y_i \in \mathbf{IR}^T$ et $X_i \in \mathcal{M}_{T,m+1}$ sont des observations prélevées d'un phénomène contenant un mouvement brownien et $\beta_i \in \mathbf{IR}^{m+1}$ est le vecteur des paramètres du modèle $(\mathcal{M}_i)$ à estimer.
Θ : est l'ensemble des estimateurs possibles pour le modèle $(\mathcal{M}_i)$.
Toute solution du programme $(\mathcal{BRP})$ (si elle existe) est nommée : ■*estimateur régularisée brownienne*■ *"BRE".*

Définition 3.2.2 *Soient $\|.\|$ et $|||.|||$ deux normes réspectivement de $\mathbf{IR}^T$ et $\mathbf{IR}^{m+1}$.*
On pose Ω_i est la matrice variance-covariance du processus stochastique obtenu après l'élimination de l'hétéroscédasticité du mouvement brownien sur T période succéssives. Autrement- dit :
$\Omega_i = V\Big(\sqrt{\Sigma_n^{-1}}(u_i + \psi_i)\Big)$, avec $\psi_i = ((\psi_{i,t})_{t\in[|1,T|]})'$ représente les erreurs de modélisation pour T périodes succéssives.
Soit $\|.\|_{\mathcal{GB}}$ est la norme brownienne générale associée à $\|.\|$.

*On définit et on appelle Programme général régularisé brownien "BRGP" (**B**rownian **R**egularized **G**eneral **P**rogram), tout programme d'optimisation, de la forme :*

$$(\mathcal{BRGP}) \begin{cases} \min_{\beta\in\Theta} \|y_i - X_i\beta_i\|^2_{\mathcal{GB}} \\ S.C \\ |||\beta_i|||^2 < \tau \end{cases}$$

Avec :
$y_i \in \mathbf{IR}^T$ et $X_i \in \mathcal{M}_{T,m+1}$ sont des observations prélevées d'un phénomène qui contient un mouvement brownien et $\beta_i \in \mathbf{IR}^{m+1}$ est le vecteur des paramètres du modèle $(\mathcal{M}_i)$ à estimer.
Θ : est l'ensemble des estimateurs possibles pour le modèle $(\mathcal{M}_i)$.
Toute solution du programme $(\mathcal{BRGP})$ (si elle existe) est nommée : ■estimateur régularisé brownien général■ "GBRE".

Définition 3.2.3 *Le lagrangien associé au programme $(\mathcal{BRP})$ est donné par :*

$$\mathcal{L}_{\mathcal{B}}(\beta_i) = \|y_i - X_i\beta_i\|^2_{\mathcal{B}} + \lambda|||\beta_i|||^2 \quad (\lambda \geq 0)$$

Le lagrangien associé au programme $(\mathcal{BRGP})$ est donné par :

$$\mathcal{L}_{\mathcal{GB}}(\beta_i) = \|y_i - X_i\beta_i\|^2_{\mathcal{GB}} + \lambda|||\beta_i|||^2 \quad (\lambda \geq 0)$$

Ces lagrangiens sont appelés aussi "Somme des Carrés des Résidus Pénalisés".
Le multiplicateur de lagrange λ représente le coefficient de pénalité, il permet de contrôler l'impact de pénalité à fixer.
La fonction $\beta_i \to |||\beta_i|||$ est appelée fonction de pénalité.

3.3 Régréssion ridge brownien

On considères les normes $\|.\|_p$ définies par :

$$\|v\|_p = (\sum_{j=1}^{k} |v_j|^p)^{\frac{1}{p}}, \qquad v \in \mathbf{IR}^k \ k = T \ ou \ m+1 \tag{3.3.1}$$

On note : $\|.\|_{\mathcal{B},p}$ et $\|.\|_{\mathcal{GB},p}$ les normes respéctivement browniene et browniene générale associées à la norme $\|.\|_p$. On parle et on définit le programme ridge brownien (resp : le programme ridge brownien générale), lorsqu'on change la norme $\|.\|_{\mathcal{B}}$ par la norme $\|.\|_{\mathcal{B},p}$ dans la fonction objective du programme $\mathcal{BRP}$, (resp ; $\|.\|_{\mathcal{GB}}$ par la norme $\|.\|_{\mathcal{GB},2}$ dans le programme $\mathcal{BRGP}$), ainsi on change la norme $|||.|||$ par $\|.\|_2$ dans la contrainte des deux programmes ($\mathcal{BRP}$ et $\mathcal{BRGP}$).
On désigne par $\mathcal{BRP}*$ (resp $\mathcal{BRGP}*$) le programme ridge brownien (resp ; programme ridge brownien général) associé au $\mathcal{BRP}$ (resp ; $\mathcal{BRGP}$)

3.3.1 Régréssion ridge brownien

Théorème 3.3.1 *On suppose que $\lambda_i > 0$, alors, le programme $\mathcal{BRP}*$ admet une unique solution, donnée par :*

$$\hat{\beta}_i(\lambda) = (X_i'\Sigma_T^{-1}X_i + \lambda_i I_{m+1})^{-1}X_i'\Sigma_T^{-1}y_i$$

Preuve 3.3.1 *On note par $< .,. >$ le produit scalaire définit par :*

$$(a,b) \in (\mathbf{IR}^k)^2 \; k = T \; ou \; m+1, \qquad < a,b >= \sum_{j=1}^{k} a_j b_j$$

sa norme associée est la norme $\|.\|_2$.
Rapplons d'après le théorème (2.4.1- chapitre 2) que les matrices Σ_T^{-1} et $\sqrt{\Sigma_T^{-1}}$ sont symétriques.
On considère le lagrangien associé au problème $\mathcal{BRP}*$:

$$\mathcal{L}_{\mathcal{B}}(\beta_i) = \|y_i - X_i\beta_i\|^2_{\mathcal{B},2} + \lambda|||\beta_i|||_2^2 \;\; (\lambda \geq 0)$$

Alors on a :

$$\begin{aligned}
\frac{\partial \mathcal{L}_{\mathcal{B}}}{\partial \beta_i}(\beta_i) &= \frac{\partial}{\partial \beta_i}\Big(\|y_i - X_i\beta_i\|^2_{\mathcal{B}} + \lambda|||\beta_i|||^2\Big) \\
&= \frac{\partial}{\partial \beta_i}\Big(\|\sqrt{\Sigma_T^{-1}}(y_i - X_i\beta_i)\|^2 + \lambda < \beta_i, \beta_i > \Big) \\
&= \frac{\partial}{\partial \beta_i}\Big(< \sqrt{\Sigma_T^{-1}}(y_i - X_i\beta_i), \sqrt{\Sigma_T^{-1}}(y_i - X_i\beta_i) > + \lambda\beta_i'\beta_i\Big) \\
&= \frac{\partial}{\partial \beta_i}\Big((\sqrt{\Sigma_T^{-1}}(y_i - X_i\beta_i))'(\sqrt{\Sigma_T^{-1}}(y_i - X_i\beta_i))\Big) + 2\lambda\beta_i \\
&= -2(\sqrt{\Sigma_T^{-1}}X_i)'\sqrt{\Sigma_T^{-1}}(y_i - X_i\beta_i) + 2\lambda\beta_i \\
&= -2X_i'(\sqrt{\Sigma_T^{-1}})'\sqrt{\Sigma_T^{-1}}(y_i - X_i\beta_i) + 2\lambda\beta_i \\
&= -2(X_i'\Sigma_T^{-1}(y_i - X_i\beta_i) - \lambda\beta_i)
\end{aligned}$$

Les conditions de premier ordre donnent :

$$\begin{aligned}
\frac{\partial \mathcal{L}_{\mathcal{B}}}{\partial \beta_i}(\beta_i) = 0 &\Rightarrow X_i'\Sigma_T^{-1}(y_i - X_i\beta_i) - \lambda\beta_i = 0 \\
&\Rightarrow X_i'\Sigma_T^{-1}y_i = X_i'\Sigma_T^{-1}X_i\beta_i + \lambda\beta_i \\
&\Rightarrow X_i'\Sigma_T^{-1}y_i = (X_i'\Sigma_T^{-1}X_i + \lambda I_{m+1})\beta_i
\end{aligned}$$

Or $X_i'\Sigma_T^{-1}X_i = X_i'(\sqrt{\Sigma_T^{-1}})'\sqrt{\Sigma_T^{-1}}X_i = (\sqrt{\Sigma_T^{-1}}X_i)'\sqrt{\Sigma_T^{-1}}X_i$
Donc la matrice $X_i'\Sigma_T^{-1}X_i$ est positive.
Soit $v \in \mathbf{IR}^{m+1} \setminus \{0\}$, on a :

$$v'(X_i'\Sigma_T^{-1}X_i + \lambda I_{m+1})v = v'(X_i'\Sigma_T^{-1}X_i)v + \lambda v'v$$

$X_i'\Sigma_T^{-1}X_i$ est positive $\Rightarrow$ $v'(X_i'\Sigma_T^{-1}X_i)v \geq 0$.
$v \in \mathbf{IR}^{m+1} \setminus \{0\} \Rightarrow v'v = \|v\|^2 > 0$.
Donc si $\lambda > 0$, la matrice $X_i'\Sigma_T^{-1}X_i + \lambda I_{m+1}$ est définie positive, alors elle est inversible par suite :

$$\frac{\partial \mathcal{L}_{\mathcal{B}}}{\partial \beta_i}(\beta_i) = 0 \Rightarrow \beta_i = (X_i'\Sigma_T^{-1}X_i + \lambda I_{m+1})^{-1}X_i'\Sigma_T^{-1}y_i$$

On pose $\hat{\beta}_i(\lambda) = (X_i'\Sigma_T^{-1}X_i + \lambda I_{m+1})^{-1}X_i'\Sigma_T^{-1}y_i$.
Passons aux conditions du deuxième ordre, On a :

$$\frac{\partial^2 \mathcal{L}_{\mathcal{B}}}{\partial \beta_i \partial \beta_i'}(\beta_i) = -2\frac{\partial}{\partial \beta_i'}\Big((X_i'\Sigma_T^{-1}(y_i - X_i\beta_i) - \lambda\beta_i)\Big)$$

$$= -2\frac{\partial}{\partial \beta_i'}\Big(X_i'\Sigma_T^{-1}y_i - X_i'\Sigma_T^{-1}X_i\beta_i - \lambda\beta_i\Big)$$
$$= 2(X_i'\Sigma_T^{-1}X_i + \lambda I_{m+1})$$

Puisque $X_i'\Sigma_T^{-1}X_i + \lambda I_{m+1}$ est définie positive, alors la hessienne est définie positive en tout vecteur β_i, en particulier en $\hat{\beta}_i(\lambda)$.
Donc si $X_i'\Sigma_T^{-1}X_i + \lambda I_{m+1}$ est définie positive, alors $\hat{\beta}_i(\lambda)$ est l'unique minimun du programme $\mathcal{BRP}*$. D'où le théorème $\square$.

Proposition 3.3.1 *On pose :*

$$G_{\lambda,i} = (X_i'\Sigma_T^{-1}X_i + \lambda I_{m+1})^{-1}X_i'\Sigma_T^{-1}X_i \tag{3.3.2}$$

L'éspérence et la variance de l'estimateur $\hat{\beta}_i(\lambda)$ sont données par :

$$E(\hat{\beta}_i(\lambda)) = G_{\lambda,i}\beta_i \tag{3.3.3}$$

$$V(\hat{\beta}_i(\lambda)) = \sigma_i^2 G_{\lambda,i}(X_i'\Sigma_T^{-1}X_i)^{-1}G_{\lambda,i}' \tag{3.3.4}$$

Preuve 3.3.2 *D'après le théorème (3.3.1) on a :*

$$\begin{aligned} E(\hat{\beta}_i(\lambda)) &= E((X_i'\Sigma_T^{-1}X_i + \lambda_i I_{m+1})^{-1}X_i'\Sigma_T^{-1}y_i) \\ &= (X_i'\Sigma_T^{-1}X_i + \lambda_i I_{m+1})^{-1}X_i'\Sigma_T^{-1}E(y_i) \\ &= (X_i'\Sigma_T^{-1}X_i + \lambda_i I_{m+1})^{-1}X_i'\Sigma_T^{-1}X_i\beta_i \\ &= G_{\lambda,i}\beta_i \end{aligned}$$

On considère l'estimateur $\beta^{BOLS} = (X_i'\Sigma_T^{-1}X_i)^{-1}X_i'\Sigma_T^{-1}y_i$ (voir chapitre 2, définition 2.4.1), on a :

$$\begin{aligned} G_{\lambda,i}\beta^{BOLS} &= (X_i'\Sigma_T^{-1}X_i + \lambda_i I_{m+1})^{-1}X_i'\Sigma_T^{-1}X_i(X_i'\Sigma_T^{-1}X_i)^{-1}X_i'\Sigma_T^{-1}y_i \\ &= (X_i'\Sigma_T^{-1}X_i + \lambda_i I_{m+1})^{-1}X_i'\Sigma_T^{-1}y_i \\ &= \hat{\beta}_i(\lambda) \end{aligned}$$

Donc

$$\begin{aligned} V(\hat{\beta}_i(\lambda)) &= V(G_{\lambda,i}\beta^{BOLS}) \\ &= G_{\lambda,i}V(\beta^{BOLS})G_{\lambda,i}' \\ &= \sigma_i^2 G_{\lambda,i}(X_i'\Sigma_T^{-1}X_i)^{-1}G_{\lambda,i}' \end{aligned}$$

Corollaire 3.3.1 *Pour tout $\lambda > 0$ l'estimateur $\hat{\beta}_i(\lambda)$ est biaisé négativement et*

$$\hat{\beta}_i(\lambda)) \sim \mathcal{N}_{m+1}(G_{\lambda,i}\beta_i, \sigma_i^2 G_{\lambda,i}(X_i'\Sigma_T^{-1}X_i)^{-1}G_{\lambda,i}') \tag{3.3.5}$$

Preuve 3.3.3

$$\begin{aligned} E(\hat{\beta}_i(\lambda)) &= G_{\lambda,i}\beta_i \\ &= (X_i'\Sigma_T^{-1}X_i + \lambda I_{m+1})^{-1}(X_i'\Sigma_T^{-1}X_i + \lambda I_{m+1} - \lambda I_{m+1})\beta_i \end{aligned}$$

$$= \beta_i - (X_i'\Sigma_T^{-1}X_i + \lambda I_{m+1})^{-1}X_i'\Sigma_T^{-1}X_i\lambda\beta_i$$
$$= \beta_i - \lambda G_{\lambda,i}\beta_i$$

Donc $biais(\hat{\beta}_i(\lambda)) = -\lambda G_{\lambda,i}\beta_i \neq 0$
De plus d'après la proposition au dessus :

$$\hat{\beta}_i(\lambda)) \sim \mathcal{N}_{m+1}(G_{\lambda,i}\beta_i, \sigma_i^2 G_{\lambda,i}(X_i'\Sigma_T^{-1}X_i)^{-1}G_{\lambda,i}') \tag{3.3.6}$$

Proposition 3.3.2 *Soient* $y_i \in \mathbf{IR}^T$ *et* $X_i \in \mathcal{M}_{T,m+1}(\mathbf{IR})$ *les observations sur* T *périodes succéssives respéctivement de la variable endogène et des variables explicatives, avec existence d'une multicolinéarité imparfaite entre les données explicatives.*
On suppose que ces observations proviennent d'un phénomène contenant un mouvement brownien.
On considère l'estimateur de moindre carré ordinaire avec mouvement brownien β^{BOLS}*, alors :*
La matrice $V(\beta^{BOLS}) - V(\hat{\beta}_i(\lambda))$ *est définie positive pour tout* $\lambda > 0$.
De plus

$$\lim_{\lambda \to +\infty} V(\hat{\beta}_i(\lambda)) = 0 \tag{3.3.7}$$

Avec 0 *est le zéro de* $\mathcal{M}_{m+1}(\mathbf{IR})$.

Preuve 3.3.4

$$V(\beta^{BOLS}) - V(\hat{\beta}_i(\lambda)) = \sigma_i^2(X_i'\Sigma_T^{-1}X_i)^{-1} - \sigma_i^2 G_{\lambda,i}(X_i'\Sigma_T^{-1}X_i)^{-1}G_{\lambda,i}'$$
$$= \sigma_i^2[(X_i'\Sigma_T^{-1}X_i)^{-1} - G_{\lambda,i}(X_i'\Sigma_T^{-1}X_i)^{-1}G_{\lambda,i}']$$

Or $G_{\lambda,i} = (X_i'\Sigma_T^{-1}X_i + \lambda I_{m+1})^{-1}X_i'\Sigma_T^{-1}X_i$*, puisque les matrices* $X_i'\Sigma_T^{-1}X_i + \lambda I_{m+1}$ *et* $X_i'\Sigma_T^{-1}X_i$ *sont symétriques, alors* $G_{\lambda,i}' = X_i'\Sigma_T^{-1}X_i(X_i'\Sigma_T^{-1}X_i + \lambda I_{m+1})^{-1}$.
Notre objectif est d'écrire $(X_i'\Sigma_T^{-1}X_i)^{-1}$ *comme produit de* $G_{\lambda,i} \times$ *une matrice* $\times G_{\lambda,i}'$*, pour qu'on puisse voir la nature de* $V(\beta^{BOLS}) - V(\hat{\beta}_i(\lambda))$.
On pose $A_i = X_i'\Sigma_T^{-1}X_i$*, alors* $G_{\lambda,i} = (A_i + \lambda I_{m+1})^{-1}A_i$ *et* $G_{\lambda,i}' = A_i(A_i + \lambda I_{m+1})^{-1}$
Donc

$$\begin{aligned} G_{\lambda,i}(I_{m+1} + \lambda A_i^{-1})A_i^{-1}(I_{m+1} + \lambda A_i^{-1})G_{\lambda,i}' &= (A_i + \lambda I_{m+1})^{-1}A_i(I_{m+1} + \lambda A_i^{-1})A_i^{-1} \\ &\times (I_{m+1} + \lambda A_i^{-1})A_i(A_i + \lambda I_{m+1})^{-1}(A_i + \lambda I_{m+1})^{-1} \\ &= (A_i + \lambda I_{m+1})^{-1}(A_i + \lambda I_{m+1})A_i^{-1} \\ &\times (A_i + \lambda I_{m+1})(A_i + \lambda I_{m+1})^{-1}(A_i + \lambda I_{m+1})^{-1} \\ &= I_{m+1}A_i^{-1}I_{m+1} = A_i^{-1} \end{aligned}$$

Donc

$$\begin{aligned} V(\beta^{BOLS}) - V(\hat{\beta}_i(\lambda)) &= \sigma_i^2[G_{\lambda,i}(I_{m+1} + \lambda A_i^{-1})A_i^{-1}(I_{m+1} + \lambda A_i^{-1})G_{\lambda,i}' - G_{\lambda,i}A_i^{-1}G_{\lambda,i}'] \\ &= \sigma_i^2 G_{\lambda,i}[(I_{m+1} + \lambda A_i^{-1})A_i^{-1}(I_{m+1} + \lambda A_i^{-1}) - A_i^{-1}]G_{\lambda,i}' \\ &= \sigma_i^2 G_{\lambda,i}[2\lambda A_i^{-2} + \lambda^2 A_i^{-3}]G_{\lambda,i}' \\ &= \sigma_i^2(A_i + \lambda I_{m+1})^{-1}A_i[2\lambda A_i^{-2} + \lambda^2 A_i^{-3}]A_i(A_i + \lambda I_{m+1})^{-1} \end{aligned}$$

$$= \sigma_i^2(A_i + \lambda I_{m+1})^{-1}(2\lambda I_{m+1} + \lambda^2 A_i^{-1})(A_i + \lambda I_{m+1})^{-1}$$

Puisque $\lambda > 0$ *alors* $V(\beta^{BOLS}) - V(\hat{\beta}_i(\lambda))$ *est définie positive comme produit de trois matrices définies positives. De plus*

$$\lim_{\lambda \to +\infty} V(\hat{\beta}_i(\lambda)) = \lim_{\lambda \to +\infty} \sigma_i^2 G_{\lambda,i}(X_i' \Sigma_T^{-1} X_i)^{-1} G_{\lambda,i}' = 0$$

□

La proposition au dessus montre que tous les estimateurs $\hat{\beta}_i(\lambda)$ sont de variance minimale que l'estimateur β^{BOLS}.
On conclut donc que chaque régression rigde de pas $\lambda > 0$ est efficace et donnée par :

$$\hat{y}_i(\lambda) = X_i(X_i' \Sigma_T^{-1} X_i + \lambda I_{m+1})^{-1} X_i' \Sigma_T^{-1} y_i := H_i(\lambda) y_i \tag{3.3.8}$$

Pour choisir le meilleur coefficient λ, on propose d'applique soit le critère d'Akaike ou bien la validation croisée existants dans la littératture, la seule différence entre notre proposition et la littérature est que nous multiplions le vecteur de données endogènes et la matrice des données explicatives par Σ_T^{-1}.

Définition 3.3.1 *On remplace :* $\|.\|_{\mathcal{B}}$ *par* $\|.\|_{\mathcal{B},2}$ *et* $|||.||| = \|.\|_2$ *dans les normes du programme* $(\mathcal{BRP})$.
*On définit et on appelle "**estimateur ridge brownien**" associé au programme* $(\mathcal{BRP})*$ *pour les deux normes* $\|.\|_{\mathcal{B},2}$ *et* $\|.\|_2$ *, la statistique notée* β^{BRE} *(**B**rownian **R**idge **E**stimator), définie par :*

$$\beta^{BRE} = (X_i' \Sigma_T^{-1} X_i + \lambda_i^* I_T)^{-1} X_i' \Sigma_T^{-1} y_i \tag{3.3.9}$$

Avec $\lambda_i^* = \arg\min_{\lambda > 0} AIC(\hat{\beta}_i(\lambda))$.

On considère un modèle linéaire $Y = K\beta + \varepsilon$ et soit $\hat{\beta}$ un estimateur de β. Dans la littérature on appelle "facteur d'inflation de la variance" (VIF) pour chaque composante d'un estimateur $\hat{\beta}$, la quantité :

$$VIF(\hat{\beta}_j(\lambda)) := \frac{n}{\sigma^2} V(K_j) V((\hat{\beta}_j)$$

Avec K_j est la $j^{ème}$ ligne de la matrice K.
Reprenant le modèle $\mathcal{M}_i$, nous redéfinissons le VIF pour l'estimateur ridge dans le cas brownien comme suit :

Définition 3.3.2 *On appelle et on définit Le facteur d'inflation brownien de la variance ridge, pour chaque composante de l'estimateur* $\hat{\beta}(\lambda)$ *la valeur :*

$$VIF(\hat{\beta}_j(\lambda)) := \frac{n}{\sigma^2} V(w_{i,j}) V((\hat{\beta}_j(\lambda)) \tag{3.3.10}$$

avec $w_{i,j}$ *est la* $j^{ème}$ *colonne de la matrice* $\sqrt{\Sigma_T^{-1} X_i}$.

3.3.2 Régréssion ridge brownienne généralisée

Dans cette section on généralise la régression ridge brownienne précédente, pour qu'elle tienne en compte l'autocorélation de l'erreur après élimination du mouvement brownien dans le modèle $\mathcal{MV}_i$. C'est-à- dire ; on suppose que $\Omega_i = V\Big(\sqrt{\Sigma_n^{-1}}(u_i + \psi_i)\Big) \neq \sigma^2 I_T$.
Notre idée ici est d'éliminer l'hétéroscédasticité provenant du mouvement brownien dans le modèle $\mathcal{MV}_i$, puis en élimine l'autocorrélation des erreurs, en multipliant le modèle $\mathcal{MVB}_i$ par $\sqrt{\Omega_i^{-1}}$, puis en appliquant l'estimateur ridge existant dans la littérature sur le modèle $\mathcal{MVB}_i$ multiplié par $\sqrt{\Omega_i^{-1}}$.
On pose $\Lambda_i = \sqrt{\Sigma_T}\Omega_i\sqrt{\Sigma_T}$.
Donc $\Lambda_i^{-1} = \sqrt{\Sigma_T^{-1}}\Omega_i^{-1}\sqrt{\Sigma_T^{-1}}$.
On remarque que $X_i'\Lambda_i^{-1}X_i = (\sqrt{\Omega_i^{-1}}\sqrt{\Sigma_T^{-1}}X_i)'\sqrt{\Omega_i^{-1}}\sqrt{\Sigma_T^{-1}}X_i$, donc $X_i'\Lambda_i^{-1}X_i$ est semi définie positive.
De plus pour tout $v \in \mathcal{IR}^{m+1}\backslash\{0\}$ et $\lambda > 0$, on a :

$$\begin{aligned} v'(X_i'\Lambda_i^{-1}X_i + \lambda I_{m+1})v &= v'X_i'\Lambda_i^{-1}X_iv + \lambda v'v \\ &= v'X_i'\Lambda_i^{-1}X_iv + \lambda\|v\|^2 \\ &> 0 \end{aligned}$$

On suit les mêmes étapes de l'épreuve du théorème (3.3.1), on remplace la norme brownien $\|_{\mathcal{B},2}$ par $\|_{\mathcal{GB},2}$ et on utilise le faite que la matrice $X_i'\Lambda_i^{-1}X_i + \lambda I_{m+1}$ est définie positive, on obtient donc le lemme suivant :

Lemme 3.3.2 *On suppose que $\lambda > 0$ est définie positive, alors, le programme $\mathcal{BGRP}*$ admet une unique solution, donnée par :*

$$\hat{\beta}^G(\lambda) = (X_i'\sqrt{\Sigma_T^{-1}}\Omega_i^{-1}\sqrt{\Sigma_T^{-1}}X_i + \lambda_i I_{m+1})^{-1}X_i'\sqrt{\Sigma_T^{-1}}\Omega_i^{-1}\sqrt{\Sigma_T^{-1}}y_i$$

Lemme 3.3.3 *On pose :*

$$G^*_{\lambda,i} = (X_i'\Lambda_i^{-1}X_i + \lambda I_{m+1})^{-1}X_i'\Lambda_i^{-1}X_i \tag{3.3.11}$$

La variance et l'espérence de l'estimateur $\hat{\beta}_i(\lambda)$ sont données par :

$$E(\hat{\beta}_i(\lambda)) = G^*_{\lambda,i}\beta_i \tag{3.3.12}$$

$$V(\hat{\beta}_i(\lambda)) = \sigma_i^2 G^*_{\lambda,i}(X_i'\Lambda_i^{-1}X_i)^{-1}(G^*_{\lambda,i})' \tag{3.3.13}$$

La proposition au dessus, montre que :

$$\hat{\beta}_i^G(\lambda)) \sim \mathcal{N}_{m+1}(G^*_{\lambda,i}\beta_i, \sigma_i^2 G^*_{\lambda,i}(X_i'\Lambda_i^{-1}X_i)^{-1}(G^*_{\lambda,i})') \tag{3.3.14}$$

Corollaire 3.3.2 *On considère l'estimateur des moindres carrés généralisés avec mouvement brownien β^{GBLS} (définition 2.5.1), alors :*
La matrice $V(\beta^{GBLS}) - V(\hat{\beta}_i^G(\lambda))$ est définie positive pour tout $\lambda > 0$.
De plus

$$\lim_{\lambda\to+\infty} V(\hat{\beta}_i^G(\lambda)) = 0 \tag{3.3.15}$$

Avec 0 est le zéro de $\mathcal{M}_{m+1}(\mathbf{IR})$.

Le corrolaire au dessus montre que tous les estimateurs $\hat{\beta}_i^G(\lambda)$ sont de variance minimale que l'estimateur β^{GBLS}.
La régression rigde de pas $\lambda > 0$ est donnée par :

$$\hat{y}_i(\lambda) = X_i(X_i'\Lambda_i^{-1}X_i + \lambda I_{m+1})^{-1}X_i'\Lambda_i^{-1}y_i := H_i^*(\lambda)y_i \tag{3.3.16}$$

Comme dans le paragraphe précédent, pour choisir le meilleur coefficient λ on multiplie les données (endogènes et explicatives) par la matrice $\sqrt{\Lambda_i^{-1}}$.

Définition 3.3.3 *On remplace : $\|.\|_{\mathcal{GB}}$ par $\|.\|_{\mathcal{GB},2}$ et $|||.||| = \|.\|_2$ dans les normes du programme $(\mathcal{BRGP})$. On définit et on appelle "**estimateur ridge brownien avec autocorrélation des erreurs**", ou encore "**estimateur ridge brownien généralisé**" associe au programme $(\mathcal{GBRP})*$ pour les deux normes $\|.\|_{\mathcal{B},2}$ et $\|.\|_2$, la statistique notée β^{GBRE} (**G**eneralized **B**rownian **R**idge **E**stimator), définie par :*

$$\beta^{GBRE} = (X_i'\sqrt{\Sigma_T^{-1}}\Omega_i^{-1}\sqrt{\Sigma_T^{-1}}X_i + \lambda_i^* I_T)^{-1}X_i'\sqrt{\Sigma_T^{-1}}\Omega_i^{-1}\sqrt{\Sigma_T^{-1}}y_i \tag{3.3.17}$$

Avec $\lambda_i^* = \arg\min_{\lambda>0} AIC(\hat{\beta}_i(\lambda))$.

Remarque 3.3.1

1. *Dans le cas d'une multicoliéarité parfaite après élimination de l'hétéroscédasticité provenant du mouvement brownien sur la période $[|1,T|]$, les statistiques β^{BRE} et β^{GBRE} restent bien définient même si les matrices $X_i'X_i$, $X_i'\Sigma_T^{-1}X_i$ et $X_i'\Lambda_i^{-1}X_i$ ne seront pas inversibles.*

2. *Si on supprime la contrainte du programme $(\mathcal{BRP})$ (c,à-d $\lambda_i = 0$), alors on trouve que $\beta^{BRE} = \beta^{BOLS}$ et $\beta^{GBRE} = \beta^{GBLS}$.*

3. *Dans le cas où les observations sont relevé d'un phénomène non aléatoire, alors il n'y a aucun processus stochastique dans le modèle $\mathcal{M}_i$. Donc l'application φ_T (voir corollaire 2.3.1, chapitre 2) ne doit éliminer aucune hétéroscédasticité d'un tel processus stochastique. Par suite $\varphi_T = id_{\mathbf{IR}^T}$, et la matrice de φ_T dans la base canonique de $\mathbf{IR}^T$ est la matrice unité I_T. Alors il faut remplacée $\sqrt{\Sigma_T^{-1}}$ par la matrice I_T.*
   Dans ce cas on trouve l'estimateur ridge β^{Ridge} existant déja dans la littérature :

$$\beta^{BRE} = (X_i'(I_T^{-1})^2X_i + \lambda_i I_T)^{-1}X_i'(I_T^{-1})^2y_i = (X_i'X_i + \lambda_i I_T)^{-1}X_i'y_i = \beta^{Ridge}$$

   .

4. *Dans le cas d'absence de la multicolinéarité on trouve $\beta^{GBRE} = \beta^{BRE}$ Si de plus, il n'y a pas de multicolinéarité entre les variables explicatives alors on trouve que :*

$$\beta^{GBRE} = \beta^{BRE} = \beta^{Ridge} = \beta^{BOLS} = \beta^{OLS}$$

   Où : $\beta^{OLS} = (X_i'X_i)^{-1}X_i'y_i$ est l'estimateur des moindre carré ordinaire.

3.4 Régularisations Elasticnets Browniennes

Dans la littérature, la "régularisation elasticnet" (voir [80-83]) est une méthode qui combine les avantages de la régression LASSO et la régression ridge.
Pour un modèle de type $(*)$ $Y = K\beta+\varepsilon$ avec ε est un bruit blanc gaussien, si les variables explicatives figurant dans la matrice K sont multicolinéaires et si le nombre des variables explicatives m peut être strictement supérieure au nombre d'observations T, les méthodes de moindre carré ainsi que la régression ridge sont inéfficace dans l'estimation du vecteur β. Jaggi, Martin and all [83], montrent l'éfficacité de l'utilisation de la régularisation elasticnet pour ce genre de problème.
La régularisation elasticnet consiste à estimer le vecteur β par l'estimateur $\hat{\beta}$ suivant :

$$\hat{\beta} = \arg\min_{\beta\in\Theta}\{\|y - K\beta\|_2 + \lambda_1\|\beta\|_2 + \lambda_2\|\beta\|_1\}$$

L'inconvénient qu'on a dédecté pour cette méthode est qu'elle ne tient pas en compte les divers chocs qui peuvent êtres provoqués par l'une des variables explicatives. Ainsi, l'hypothèse du bruit blanc néglige l'hypothèse d'existence d'une autocorrélation des erreurs dans le modèle $(*)$.
Pour cette raison on propose dans cette section deux types de régularisation elasticnets qui vont améliorer encore la régularisation elasticnet existante dans la littérature. En modélisant les chocs qui peuvent êtres provoqués par les variables explicatives par un mouvement brownien, la première régularisation consiste à considérer qu'il y a un mouvement brownien dans le modèle mais avec abscence d'autocorrélation des erreurs après élimination de l'hétéroscédasticité de composante brownien, on la nomme "Régularisation Elasticnet Brownienne" et on la note "**BER**"("**B**rownian **E**lasticnet **R**egularization"). La deuxième régularisation consiste à considérer qu'il y a un mouvement brownien dans le modèle mais avec autocorrélation des erreurs après élimination de l'hétéroscédasticité de composante brownien, on la nomme "Régularisation Elasticnet Brownienne Généralisée" et on la note "**GBER**"("**G**eneralized **B**rownian **E**lasticnet **R**egularization").

3.4.1 Régularisation Elasticnet Brownienne

Reprenant le programme :

$$(\mathcal{BRP})\begin{cases} \min_{\beta\in\Theta} \|y_i - X_i\beta_i\|^2_{\mathcal{B}} \\ S.C \\ |||\beta_i|||^2 < \tau \end{cases}$$

On appelle et on définit la régularisation elasticnet browniene "BER"("**B**rownian **E**lasticnet **R**egularization"), lorsqu'on prend :

$$\|.\|^2_{\mathcal{B}} = \frac{1}{T}\|.\|^2_{\mathcal{B},2}$$

$$|||.||| = \frac{(1-\alpha)}{2}\|.\|_2 + \alpha\|.\|_1$$

On pose $MSE(\beta_i) = \frac{1}{T}\|y_i - X_i\beta_i\|^2_{\mathcal{B},2}$, cette fonction représente la perte quadratique moyenne "Mean Square Error".

La somme des carrées des résidus pénalisée avec mouvement brownien pour cette régularisation est donnée par :

$$\mathcal{L}_B = MSE(\beta_i) + \lambda \underbrace{[\frac{(1-\alpha)}{2}\|\beta_i\|_2 + \alpha\|\beta_i\|_1]}_{P(\beta_i,\alpha)= \, fonction \, de \, pénalité} \tag{3.4.1}$$

On voit que la fonction de pénalité $P(\beta_i, \alpha)$ n'est pas déffirentiable, car la norme $\|.\|_1$ n'est pas déffirentiable au point 0. Par suite, l'utilisation des conditions de premier ordre et de deuxième ordre ne seront pas utiles. Pour cette raison, on utilise une approximation numérique pour résoudre le programme "BER".
Toute solution numérique convergente du programme "BER", on la nomme : "Estimateur Elasticnet Brownien", en abréviation "B2E", ("**B**rownian **E**lasticnet **E**stimator").
On choisit comme méthode d'approximation la méthode de decendente du gradient stochastique. Donc la mise à jour de β_i est donnée par :

$$\beta_i = \beta_i - \eta\Big(\frac{\partial MSE(\beta_i)}{\partial \beta_i} + \lambda \frac{\partial P(\beta_i, \alpha)}{\partial \beta_i}\Big) \tag{3.4.2}$$

Où :
η ; représente le taux d'apprentissage pour moduler la convergence, peut être fixé au décroissante au fil des itérations.
L'algorithme qu'on propose pour estimer β_i est comme suit :

Algorithm 2: Algorithme BER

Require: #Importation des données
y Observations sur la variable endogène.
X Matrices des observations des variables explicatives.
Ensure: #Algorithme de la méthode de descendent stochastique "SDM".
FUNCTION SDM=function(y,X)
T=longeur de y ;
#Traitement de données.
Standardisation de y et des colonnes de X ;
$W = \sqrt{\Sigma_T^{-1}}X$; (X standardisée)
$z = \sqrt{\Sigma_T^{-1}}y$; (y standardisé)
#appel de la classe SGDRegressor dans le logiciel R.
reg=SGDRegressor(loss="squared-loss",
penality="elasticnet",λ=a,α=b,max_iter=itermax) ;
#Résolution.
β=reg.fit(W,z) ;
Afficher β
Fin Fonction.

3.4.2 Régularisation Elasticnet Brownienne Généralisée

Reprenant le programme :

$$(\mathcal{GBRP}) \begin{cases} \min_{\beta\in\Theta} \|y_i - X_i\beta_i\|^2_{\mathcal{GB}} \\ S.C \\ |||\beta_i|||^2 < \tau \end{cases}$$

On appelle et on définit la régularisation elasticnet brownienne généralisée "**GBER**"("**G**eneralized **B**rownian **E**lasticnet **R**egularization"), lorsqu'on prend :

$$\|.\|^2_{\mathcal{GB}} = \frac{1}{T}\|.\|^2_{\mathcal{GB},2}$$

$$P(.,\alpha) = |||.||| = \frac{(1-\alpha)}{2}\|.\|_2 + \alpha\|.\|_1$$

P est la fonction de pénalité.
On pose $GMSE(\beta_i) = \frac{1}{T}\|y_i - X_i\beta_i\|^2_{\mathcal{GB},2}$, cette fonction représente la perte quadratique moyenne avec autocorrélation des erreurs "Generalized Mean Square Error".
La somme des carrés des résidus pénalisés avec mouvement brownien et autocorélation des erreurs pour cette régularisation est donnée par :

$$\mathcal{L}_{GB} = GMSE(\beta_i) + \lambda P(\beta_i, \alpha) \tag{3.4.3}$$

On nomme Toute solution numérique convergente du programme "GBER", par : "Estimateur Elasticnet Brownien Généralisée", en abréviation "GB2E", ("**G**eneralized **B**rownian **E**lasticnet **E**stimator").
On choisit comme méthode d'approximation la méthode de decendente du gradient stochastique. Donc la mise à jour de β_i est donnée par :

$$\beta_i = \beta_i - \eta\Big(\frac{\partial GMSE(\beta_i)}{\partial \beta_i} + \lambda\frac{\partial P(\beta_i,\alpha)}{\partial \beta_i}\Big) \tag{3.4.4}$$

Où :
η ; représente le taux d'apprentissage pour moduler la convergence, peut être fixé au décroissante au fil des itérations.
L'algorithme qu'on propose pour estimer β_i est comme suit :

3.5 Tests de détection d'hétéroscédasticité

Parmi les importants tests des erreurs d'un modèle donné, sont les tests d'hétéroscédasticité des erreurs. En faite l'existence de l'hétéroscédasticité des erreurs dans un modèle a des conséquences identiques à celles de l'autocorrélation des erreurs, c'est-à-dire :
– estimateur sans biais ;
– les estimateurs de MCO et BOLS ne sont plus à variance minimale.
Les causes d'existence de l'hétéroscédasticité sont nombreuses, par exemple, lorsque les observations sont des moyennes calculées à partir des échantillons qui n'ont pas la même taille. Aussi, si on a une répétition d'une même valeur de l'output (variable à expliquer) pour des valeurs distinctes d'une variable explicative, (par exemple lors de regroupements en tranches de salaires ...).
L'objectif de tester l'hétéroscédasticité de l'erreur d'un modèle, est de voir si l'hypothèse de la constance de la variance de l'erreur est violée ou non. Dans la littérature il y a plusieurs tests pour détecter l'hétéroscédasticité de l'erreur, parmi eux on cite ; le test de Gleisjer [3], le test de White [4] et le test ARCH [5],... .
En revanche, ces tests sont appliqués dans le cas où les variables explicatives sont déterministes et ne contiennent pas une composante qui provient d'un mouvement brownien.
Or les variances d'un processus aléatoire brownien, changent au cours du temps. Ce qui assure l'existence de l'hétéroscédasticité des variances dans le modèle.

Algorithm 3: Algorithme GBER

Require: #Importation des données
y Observations sur la variable endogène.
X Matrices des observations des variables explicatives.
Ensure: #Algorithme de la méthode de descendent stochastique "SDM".
FUNCTION SDM=function(y,X)
T=longeur de y ;
#Traitement de données.
Standardisation de y et des colonnes de X ;
$W = \sqrt{\Sigma_T^{-1}}X$; (X standardisée)
$z = \sqrt{\Sigma_T^{-1}}y$; (y standardisé)
Ω= la matrice estimer d'autocovariance des erreurs ε du modèle $z = W\beta + \varepsilon$.
$\Lambda = \sqrt{\Omega^{-1}}W$; (X standardisée)
$v = \sqrt{\Omega^{-1}}z$; (y standardisé)
#appel de la classe SGDRegressor dans le logiciel R.
reg=SGDRegressor(loss="squared-loss",
penality="elasticnet",λ=a,α=b,max_iter=itermax) ;
#Résolution.
$\beta = reg.fit(\Lambda, v)$;
Afficher β
Fin Fonction.

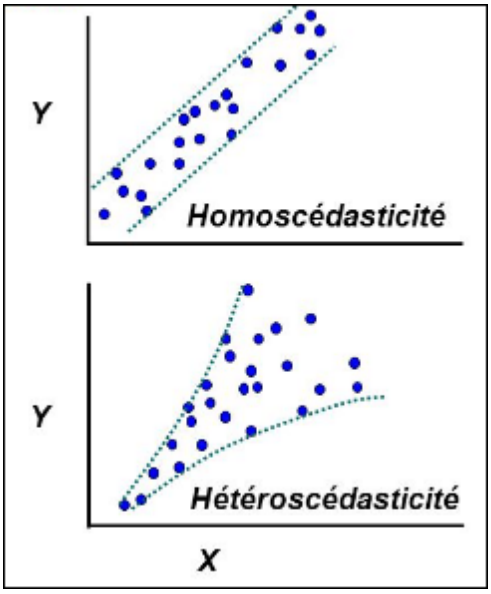


FIGURE 3.1 – Exemple d'un modèle homoscédastique et hétéroscédastique

Au dessous on donne un graphique illustratif de l'hétéroscédasticité :
Dans le chapitre 2 et les sections précedentes de ce chapitre, on a supposé que l'élimination de l'hétéroscédasticité provenante du mouvement brownien élimine l'hétéroscédasticité des erreurs dans le modèle $\mathcal{M}_i$.
L'objectif de cette section est d'étudier la violance de cette hypothèse, et de trouver un estimateur optimal du modèle $\mathcal{M}_i$.
Comme test de l'hétéroscédasticité, on propose de modifier le test ARCH ("AutoRegressive

Conditional Heteroskedasticity"), pour qu'il soit compatible avec l'existence du processus brownien dans le modèle. On nomme le test ARCH modifié : le test "B-ARCH" ("**B**rowian-**A**uto**R**egressive **C**onditional **H**eteroskedasticity").
L'idée du test B-ARCH qu'on propose est d'éliminer l'héroscédasticité du mouvement brownien sur une gamme du temps $[|1, T|]$, puis vérifier l'existence d'une hétéroscédasticité supplémentaire dans le modèle $\mathcal{M}_i$.
Avant d'aborder notre test, rappelant la définition d'un processus ARCH, ainsi que le test ARCH pour détecter l'hétéroscédasticité dans un modèle donné.

3.5.1 Test de l'hétéroscédasticité de l'erreur : Test ARCH

Définition 3.5.1 *On dit qu'un processus $(\varepsilon_t)_{t\geq 0}$ est un processus ARCH(p), "AutoRegressive Conditional Heteroskedasticity", s'il vérifie :*

$$\varepsilon_t = \eta_t \sqrt{h(t)} \tag{3.5.1}$$

Où :

- η_t *est un bruit blanc gaussien* $\mathcal{N}(0,1)$ *;*
- $h_t = \alpha_0 + \sum_{k=1}^{p} \alpha_k \varepsilon_{i,t-k}^2$
- $\alpha_i \geq 0,\ \forall i \in \{0, ..., p\}$

Définition 3.5.2 *On dit qu'un processus $(\varepsilon_t)_{t\leq 0}$ est un processus $GARCH(p,q)$ ("Generalized AutoRegressive Conditional Heteroskedasticity"), si il vérifie les propriétés suivantes :*

$$\varepsilon_t = \eta_t \sqrt{h(t)} \tag{3.5.2}$$

Où :

- η_t *est un bruit blanc gaussien* $\sim \mathcal{N}(0,1)$ *;*
- $h_t = \alpha_0 + \sum_{k=1}^{p} \alpha_k \varepsilon_{t-k}^2 + \sum_{j=1}^{q} \omega_k h_{t-j}^2$ *;*
- $\alpha_k \geq 0,\ \omega_j \geq 0,\ \forall j \in [|1, q|]\ \ \forall k \in [|0, p|]$.

On considère un modèle de type $Y = K\beta + \varepsilon$, avec Y est le vecteur d'observations de la variable endogène, K est la matrice d'observations des variables explicatives et ε est le vecteur des erreurs ε_t. On veut tester la constance des variances de ε_t dans le temps, en utilisant le test ARCH existant dans la littérature.
On pose les hypothèses altérnées ;
H_0 : il y a Homocédasticité des erreurs, contre H_1 : les erreurs sont hétéroscédastique.
La procédure d'un test ARCH(p), pour détecter l'hétéroscédasticité dans un modèle de la forme :

$$Y_t = \beta_0 + \sum_{j=1}^{m} \beta_j K_{t,j} + \varepsilon_t$$

est comme suite :

1. Soit $\hat{\varepsilon}_t$ les résidus estimés par MCO (moindre carrée ordinaire).
2. En utilise encore une fois MCO pour estimer le modèle intermédiaire suivant :

$$\hat{\varepsilon}_t = \alpha_0 + \sum_{j=1}^{p} \alpha_j \hat{\varepsilon}_{i,t-j}^2 + \mu_t$$

   Où : μ_t sont iid.

3. On calcul ensuite le coefficient de détermination R^2, et la statistique $LM = TR^2$. (LM= Lagrange Multiplier)

La région d'accepter H_0 au niveau de confiance $1-\alpha$ est $C_\alpha = [0, \chi_{1-\alpha,p}]$.
Donc si $LM > \chi_{1-\alpha,p}$, on rejette H_0 et on accepte H_1.

3.5.2 Test de l'hétéroscédasticité de l'erreur : Test B-ARCH

Revenant maintenant au modèle $(\mathcal{M}_i)$: D'après le théorème ((2.2.1)chapitre 2), on a :

$$\sum_{t=1}^{T} \rho_{l,t} u_{i,t} \sim \mathcal{N}(0, \gamma_i)$$

Ce qui élimine l'hétéroscédasticité du processus $(u_{i,t})_{t\in[|1,T|]}$.
Nous définissons le processus, qu'on l'appel **B-ARCH** ou bien ("**B**rownian- **A**uto**R**egressive **C**onditional **H**eteroskedasticity "), comme suit :

Définition 3.5.3 *On nomme et on dit qu'un processus $(\psi_{i,t})_{1\leq i\leq N 1\leq t\leq T}$ est un processus $B-ARCH(p_i)$ s'il vérifie les propriétés suivantes :*

$$\gamma_i \eta_t + \sum_{t=1}^{T} \rho_{l,t} \psi_{i,t} = \vartheta_t \sqrt{h(t)} \tag{3.5.3}$$

Où :

- $\gamma_i \geq 0, \ ,\forall i \in [|1, N|]$, *et* $(\rho_{l,t})_{t\in[|1,T|]}$ *sont les élements de la* $l^{\text{ème}}$ *ligne de la matrice* $\sqrt{\Sigma_{i,T}}^{-1}$ *;*
- ϑ_t, η_t *sont deux bruits blanc gaussiens* $\sim \mathcal{N}(0,1)$ *;*
- $h_t = \alpha_{i,0} + \sum_{k=1}^{p_i} \alpha_k \varepsilon_{i,t-k}^2$ *;*
- $\alpha_{i,k} \geq 0, \ \forall k \in \{0, ..., p_i\} \ \forall i \in [|1, N|]$.

On considère le modèle $(\mathcal{M}_i)$ et les hypothèses :

$$\begin{cases} \mathcal{H}_0^{ht}, & \text{les erreurs sont homoscédastiques} \\ \mathcal{H}_1^{ht}, & \text{les erreurs sont hétéroscédastiques} \end{cases}$$

Nous construisons le test, qu'on nomme "Test $B-ARCH(p_i)$", comme suit :
On pose :

$$\varepsilon_t = u_{i,t} + \psi_{i,t}$$

Avec $(u_{i,t})_{t\geq 0}$ est un mouvement brownien et $(\psi_{i,t})_{t\geq 0}$ sont les erreurs du modèle au cours du temps.

1. Soit $\hat{\varepsilon}_t$ les résidus estimés par BOLS (Chapitre 2, définition (2.4.1)) pour le modèle $(\mathcal{M}_i)$ sur une gamme de temps $[|1, T|]$;
2. On utilise la méthode MCO pour estimer le modèle intermédiaire suivant :

$$\hat{\varepsilon}_{i,t} = \alpha_{i,0} + \sum_{k=1}^{p_i} \alpha_{i,k} \hat{\varepsilon}_{i,t-k}^2 + \mu_{i,t}$$

Où : $\mu_{i,t}$ sont iid.

3. On calcul ensuite le coefficient de détermination R_i^2, et la statistique $LM_i = (T - p_i)R_i^2$. (LM= Lagrange Multiplier)

La région d'accepter $\mathcal{H}_0^{ht}$ au niveau de confiance $1-\alpha$ est $C_\alpha = [0, \chi_{1-\alpha,p_i}]$.
Donc si $LM > \chi_{1-\alpha,p_i}$, on rejette $\mathcal{H}_0^{ht}$ et on accepte $\mathcal{H}_1^{ht}$.

3.5.3 Estimation paramètrique avec hétéroscédasticité supplémentaire

On signifie par "hétéroscédasticité supplémentaire", lorsque l'élimination de l'hétéroscédasticité provenant d'un mouvement brownien sur une gamme du temps $[|1,T|]$ n'implique pas forcément l'élimination de l'hétéroscédasticité des erreurs du modèle $(\mathcal{M}_i)$. Autrement dit :

$$\sum_{t=1}^{T} \rho_{l,t} u_{i,t} \sim \mathcal{N}(0,\gamma_i) \;\not\Rightarrow \sum_{t=1}^{T} \rho_{l,t}(u_{i,t}+\psi_{i,t}) \sim \mathcal{N}(0,\sigma_i)$$

Avec γ_i et σ_i sont constantes et $(\rho_{l,t})_{t\in[|1,T|]}$ sont les élements de la $l^{ème}$ ligne de la matrice $\sqrt{\Sigma_T^{-1}}$.
Le test qu'on a proposé dans la section précédente, détecte l'existence d'une hétéroscédasticité supplémentaire dans le modèle avec un degré de confiance $1-\alpha$.
Si l'hypothèse $\mathcal{H}_0^{ht}$ est retenue, on peut utiliser l'un des estimateurs qu'on a proposé dans le chapitre 2 et les quatres premières sections de ce chapitre, selon l'existence d'une multicolinéarité des données explicatives ou/et l'autocorrélations des erreurs.
Si l'hypothèse $\mathcal{H}_1^{ht}$ est retenue, les estimateurs précédents ne seront pas efficaces.
Le but de cette sous section est de trouver un estimateur optimal pour le modèle $(\mathcal{M}_i)$, dans le cas où l'hypothèse $\mathcal{H}_1^{ht}$ est accéptée. Rappelant les modèles :

$$(\mathcal{M}_i)\;\; y_{i,t} = \beta_{i,0} + \sum_{k=1}^{m} \beta_{i,k} x_{i,k,t} + \gamma_i u_{i,t} + \psi_{i,t}$$

$$(\mathcal{MV}_i)\;\; y_i = X_i\beta_i + \gamma_i u_i + \psi_i$$

$$(\mathcal{MVB}_i)\;\; z_i = W_i\beta_i + \nu_i$$

Où :
$z_i = \sqrt{\Sigma_T^{-1}} y_i$, $W_i = \sqrt{\Sigma_T^{-1}} X_i$ et $\nu_i = \sqrt{\Sigma_T^{-1}}(\gamma_i u_{i,t} + \psi_i)$.
On suppose qu'on a T observation de la variable endogène, ainsi que des variables explicatives.
La procédure qu'on propose pour estimer les paramètres $\beta_{i,k},\; k \in [|0,m|]$ du modèle $(\mathcal{M}_i)$, dans le cas où l'hypothèse $\mathcal{H}_1^{ht}$ est acceptée, est comme suit :

1. On considère l'écriture vectorielle $(\mathcal{MV}_i)$ du modèle $(\mathcal{M}_i)$ sur une gamme du temps $[|1,T|]$.

2. On élimine l'hétéroscédasticité provenant du mouvement brownien, en multipliant les membres de $(\mathcal{MV}_i)$ par la matrice $\sqrt{\Sigma_T^{-1}}$, on obtient donc le modèle $(\mathcal{MVB}_i)$.

3. On pose $\nu_i = (\nu_{i,1}, \nu_{i,2}, ..., \nu_{i,T})'$, on suppose que $\nu_{i,t} \sim GARCH(p_i, q_i)$ [1].
On a : $\nu_{i,t} \sim GARCH(p_i, q_i) \Rightarrow \nu_{i,t} = \eta_t \sqrt{h_{i,t}}$, avec :

$$\eta_t \sim \mathcal{N}(0,1) \; et \; h_{i,t} = \alpha_{i,0} + \sum_{k=1}^{p_i} \alpha_{i,k} \nu_{i,t-k}^2 + \sum_{j=1}^{q_i} \omega_{i,k} h_{t-j}^2$$

On pose $\theta_i = (\beta_i, \alpha_{i,0}, ..., \alpha_{i,p_i}, \omega_{i,1}, ..., \omega_{i,q_i})'$.
On note par $\underline{z_{i,t-1}}$: l'ensemble d'information de la variable endogène z_i jusqu'à l'instant $t-1$, et $W_{i,t}$ est la $t^{ème}$ ligne de la matrice W_i ($W_i \in \mathcal{M}_{T,m+1}(\mathbf{IR})$.
On a :

$$\begin{aligned} E(z_{i,t}|\underline{z_{i,t-1}}, W_{i,t}) &= E(W_{i,t}\beta_i + \nu_{i,t}|\underline{z_{i,t-1}}, W_{i,t}) \\ &= W_{i,t}\beta_i + E(\nu_{i,t}|\underline{\nu_{i,t-1}}, W_{i,t}) \\ &= W_{i,t}\beta_i = m(\theta_i) \end{aligned}$$

Et

$$\begin{aligned} V(z_{i,t}|\underline{z_{i,t-1}}, W_{i,t}) &= V(W_{i,t}\beta_i + \nu_{i,t}|\underline{z_{i,t-1}}, W_{i,t}) \\ &= V(\nu_{i,t}|\underline{\nu_{i,t-1}}, W_{i,t}) \\ &= h_{i,t} = h_t(\theta_i) \end{aligned}$$

Sous la condition de normalité de la loi conditionnelle de $z_{i,t}$ sachant $\underline{z_{i,t-1}}$ et $W_{i,t}$, la fonction de vraisemblance de $z_{i,t}|\underline{z_{i,t-1}}, W_{i,t}$ pour T observations de la variable endogène $y_{i,t}$ est donnée par :

$$l(\theta_i) = \prod_{i=1}^{T} \frac{1}{\sqrt{2h_t(\theta_i)\pi}} e^{-\frac{(z_{i,t}-m(\theta_i))^2}{h_t(\theta_i)}} \tag{3.5.4}$$

Donc le log vraisemblance est donné par :

$$L(\theta_i) = \ln(l(\theta_i)) = -\frac{T}{2}\ln(2\pi) - \frac{1}{2}\sum_{i=1}^{T}\ln(h_t(\theta_i)) - \frac{1}{2}\sum_{i=1}^{T}\frac{(z_{i,t}-m(\theta_i))^2}{h_t(\theta_i)} \tag{3.5.5}$$

L'estimateur optimal qu'on cherche est donc :

$$\hat{\theta}_{i,max} = \arg\max_{\theta_i \in \Theta}(L(\theta_i)) \tag{3.5.6}$$

3.6 Simulation et application économétrique

Reprenant les deux modèles définis dans la section (3.2.1) :

$$\mathfrak{M}_1 \;\; y_{i,t} = c + \alpha_i k_{i,t} + \eta_i r_{i,t} + \nu l_{i,t} + \lambda_i h_{i,t} + u_{i,t}$$

$$\mathfrak{M}_2 \;\; Y_{i,t} = c + \alpha_i k_{i,t} + \eta_i r_{i,t} + \nu l_{i,t} + \lambda_i h_{i,t}$$

D'après la section (2.8- chapitre 2), la productivité globale des facteurs est stochastique. Ainsi les tests de multicolinéarité qu'on a fait au début de ce chapitre montre que le degré

1. Dans l'estimation on préfère d'utilisé le processus GARCH au lieu du processus "ARCH", car un $GARCH(p,q)$ peut s'écrire comme un processus $ARCH(+\infty)$.

de multicolinéarité dans le modèle $\mathfrak{M}_1$ est plus petit que celui du modèle $\mathfrak{M}_2$. Donc on modélise le logarhitme du PIB par le modèle $\mathfrak{M}_1$.

Simulation de la régression ridge brownienne

On applique l'estimateur ridge brownien pour estimer les élasticités des facteurs contribuant dans le PIB du maroc publiées sur le site de la banque mondiale, on trouve que :
Le meilleur pas de la régréssion ridge λ^* en utilisant la validation croisée est $\lambda^* = 0.0152154$.
L'estimation des coefficients du modèle $\mathfrak{M}_1$ par la régression ridge brownien donne les résultats suivants :

c	α_{mar}	ν_{mar}	η_{mar}	λ_{mar}
0.003092145	0.324855527	0.250870084	0.176368611	0.214785796

TABLE 3.3 – Estimation du modèle $\mathfrak{M}_1$ par β^{BRE}

D'après le tableau au dessus, on conclut que :
— l'estimation du logarithme de la frontière technologique est : $E(\ln(A_{maroc})|observation) = 0.003092145$
— l'élasticité du capital est $\alpha_{mar} = 0.324855527$
— l'élasticité de la force du travail est $\nu_{mar} = 0.250870084$
— l'élasticité des dépenses en recherche et développement est $\eta_{mar} = 0.176368611$
— l'élasticité du capital humain $\lambda_{mar} = 0.214785796$

La frontière technologique du maroc est estimé par la fonction :

$$A_{maroc,t} = e^{0.003092145+u_t} = 1.003097e^{u_t} \tag{3.6.1}$$

Avec $(u_t)_{t\geq 0}$ est un mouvement brownien.
Le PIB est modélisé par la fonction cobb Dauglass stochastique :

$$PIB_{maroc,t} = 1.003097e^{u_t}K_t^{0.324855527}R_t^{0.176368611}L_t^{0.250870084}H_t^{0.214785796} \tag{3.6.2}$$

Simulation de la méthode : Elasticnet brownienne

L'application des données elasticnet brownien aux données des Maroc donne : le pas optimal en utilisant la validation croisée est $\lambda^* = 0.003498802$.
L'estimation des paramètres du modèle $\mathfrak{M}_1$ par β^{B2E} donne les résultats figuré dans le tableau en dessous.

c	α_{mar}	ν_{mar}	η_{mar}	λ_{mar}
0.001388514	0.387665298	0.255869078	0.135311570	0.208123733

TABLE 3.4 – Estimation du modèle $\mathfrak{M}_1$ par β^{B2E}

D'après le tableau au dessus, on conclut que :
— l'estimation du logarithme de la frontière technologique est : $E(\ln(A_{maroc})|observation) = 0.001388514$
— l'élasticité du capital est $\alpha_{mar} = 0.387665298$
— l'élasticité de la force du travail est $\nu_{mar} = 0.255869078$
— l'élasticité des dépenses en recherche et développement est $\eta_{mar} = 0.135311570$

— l'élasticité du capital humain $\lambda_{mar} = 0.208123733$

La frontière technologique du maroc est estimé par la fonction :

$$A_{maroc,t} = e^{0.001388514+u_t} = 1.001389e^{u_t} \tag{3.6.3}$$

Avec $(u_t)_{t\geq 0}$ est un mouvement brownien.

Le PIB est modélisé par la fonction cobb Dauglass stochastique :

$$PIB_{maroc,t} = 1.001389e^{u_t}K_t^{0.387665298}R_t^{0.135311570}L_t^{0.255869078}H_t^{0.208123733} \tag{3.6.4}$$

Prédiction du PIB : Maroc

Puisque la méthode elasticnet brownienne combine les avantage de la régression ridge brownien et de LASSO dans la cas brownien, alors on retient l'estimation obtenue par la méthode elasticnet brownienne.

En utilisant le logiciel R, en cherche à prédire les facteurs explicatifs contribuant dans le PIB via des séries temporelles, on trouve que :

Les logarhithmes de la force du travail, de la formation brute du capital et du capital humain sont les deux des processus $ARIMA(0, 1, 0)$, des coefficients constants estimés respéctivement 0.019, $0,054$ et 0.057. Autrement dit :

$$\hat{l}_{mar,T+j} = 0.019j + l_{mar,T}$$

$$\hat{k}_{mar,T+j} = 0,054j + k_{mar,T}$$

$$\hat{h}_{mar,T+j} = 0.057j + h_{mar,T}$$

Le logarithme des dépenses en recherche et développement est un modèle de Holt-Winters provenant d'un lissage exponentiel double, avec un facteur de lissage estimé $\kappa = 0,639352$, et un coefficient de tendance estimer $\pi = 0,945145$. Autrement dit :

$$\hat{r}_{mar,T+j} = s_j + j\tau_T$$

Où : $s_2 = r_2$

$\tau_t = r_2 - r_1$

et pour $t > 2$:

$$s_t = \kappa r_t + (1-\kappa)(s_{t-1} + \tau_{t-1})$$
$$\tau_t = \pi(s_t + s_{t-1}) + (1-\pi)\tau_{t-1}$$

Donc la prédiction du logarithme du PIB_{Maroc} sur les j $(j \geq 1)$ prochaînes années est données par :

$$E(\ln(PIB_{Maroc})|information_T) = 0.001388514 + 0.387665298\hat{k}_{mar,T+j} + 0.255869078\hat{l}_{mar,T+j} + 0.135311570\hat{r}_{mar,T+j} + 0.208123733\hat{h}_{mar,T+j}$$

3.7 conclusion

Dans ce chapitre nous avons développé les méthodes de la régularisation pour qu'elles tiennent en compte l'existence d'un mouvement brownien dans le modèle. L'utilisation de ces méthodes a un rôle très important lorsque les variables explicatives sont multicolinéaires et les résultats obtenus par les moindres carrées et les moindres carrées browniennes ne sont pas significatives. les méthodes qu'on a abordé dans ce chapitre fournissent des bons estimateurs dans le cas d'existence d'une multicolinéarité entre les variables explicatives. Ainsi on a développé le test ARCH qui detecte l'existence d'une hétéroscédasticité dans le modèle, puis on a construit une procédure d'estimation des paramètres du module dans le cas d'existence d'une hétéroscédasticité supplémentaire.
Ce chapitre a été terminé par une application économétrique, en estimant les paramètres d'une fonctions Cobb Dauglass stochastiques, en utilisant les données publiées sur le site de la banque mondial concernant le Maroc.

Chapitre 4

Modélisation des données en panel

4.1 Introduction

Les données de panel sont des observations relatives aux mêmes individus (entreprises, ménages, salariés, ...) à des périodes successives. Elles comportent deux dimensions :
– Temporelle : longitudinale
– Individuelle : transversale
Dans ce chapitres on s'intérèsse à l'étude des modèles contenants plusieurs individus (variables endogènes) et plusieurs variables explicatives, avec un mouvement brownien multidimentionnel.
Ce chapitre est composé de trois section ; dans la première section nous étudions l'hypothèse d'égalité de la variance entre les individus. Dans la deuxième section nous cherchons à donner des estimateurs optimales des modèles en panel, en généralisant les estimateurs qu'on a abordé dans les chapitres 2 et 3. Dans la troisième section nous abordons le test des données en panel, pour trouver la catégorie des modèles modélisant l'ensembles des individus en fonction des variables explicatives.

4.2 Test d'égalités des variances : VETEHBM

Notre objective dans cette section est de proposer un test statistique, pour détecter l'existence de l'hétéroscédasticité dans le modèle $(\mathcal{M}_i)$, après l'élimination de l'hétéroscédasticité du mouvement brownien. Ce test qu'on propose consiste à tester l'égalité des variances dans le modèle $(\mathcal{M}_i)$, après l'élimination de l'hétéroscédasticité qui provient du mouvement brownien, On va l'appeler ; "**VETEHBM**", (Variance Equalization Test after Elimination of Heteroskedasticity from Brownian Motion).
On pose $\sqrt{\Sigma_T^{-1}} = (\rho_{l,t})_{(t,l)\in[|1,T|]^2}$
On considère le modèle $(\mathcal{MVB}_i)$, l'écriture temporelle du modèle est comme suit :

$$(\mathcal{MVB}_i)\ \ z_{i,t} = \beta_{i,0}\sum_{l=1}^{T}\rho_{t,l} + \sum_{k=1}^{m}\beta_{i,k}w_{i,k,t} + \nu_{i,t}\ \ t \in [|1,T|]$$

Avec $z_{i,t},\ w_{i,t,k}\ \nu_{i,t}$ sont les moyennes pendérées par les poids de la matrice inverse de la racine carrée de la matrice de covariance du processus brownien sur la période $t \in [|0,T|]$.
Autrement dit :

$$z_{i,t} = \sum_{l=1}^{T}\rho_{t,l}y_{i,l} \tag{4.2.1}$$

$$w_{i,k,t} = \sum_{l=1}^{T} \rho_{t,l} x_{i,k,l} \tag{4.2.2}$$

$$\nu_{i,t} = \sum_{l=1}^{T} \rho_{t,l} (\gamma_i u_{i,t} + \psi_{i,t}) \tag{4.2.3}$$

On considère les hypothèses :

$$\left(\mathcal{H}_0^{VETEHBM}\right) : \ \sigma_1^2 = \sigma_2^2 = ... = \sigma_N^2$$

$$\left(\mathcal{H}_1^{VETEHBM}\right) : \ \exists i \neq j \ (j \in [|1, N|]) : \sigma_i^2 \neq \sigma_j^2$$

Puis on suit, les étapes suivantes :
Étape 1 : Calcul de la variance empirique pour chaque variable z_i
On pose :

$$\hat{\sigma}_i^2 = \frac{\sum_{t=1}^{T} (z_{i,t} - \bar{z}_i)^2}{T-1}$$

Étape 2 : Calcul de la variance empirique total

$$\hat{\sigma}^2 = \frac{(T-1)\sum_{i=1}^{N} \hat{\sigma}_i^2}{(T-1)N} = \frac{\sum_{i=1}^{N} \hat{\sigma}_i^2}{N}$$

Étape 3 : On calcul les quantités

$$Q = N(T-1)\ln(\hat{\sigma}^2) - (T-1)\sum_{i=1}^{N} \ln(\hat{\sigma}_i^2)$$

Et

$$C = 1 + \frac{1}{3(N-1)}\Big(\frac{N}{T-1} - \frac{1}{N(T-1)}\Big)$$

Puis on compare la quantité empirique ; $Q' = \frac{Q}{C}$ avec la valeur théorique de Khi 2 ; $\chi^2_{1-\alpha,N-1}$.
-Si $Q' \leq \chi^2_{1-\alpha,N-1}$, on accepte l'hypothèse $\left(\mathcal{H}_0^{VETEHBM}\right)$, avec un niveau de confiance $1-\alpha$.
-Si $Q' > \chi^2_{1-\alpha,N-1}$, on rejette $\left(\mathcal{H}_0^{VETEHBM}\right)$ et on accepte l'hypothèse $\left(\mathcal{H}_1^{VETEHBM}\right)$ avec un niveau de confiance $1-\alpha$.

4.3 Modèle des données en panel avec mouvement brownien multidimentionnel

Dans les chapitres précédents, nous avons pésenté les estimations correspondant à T observations pour chaque modèle $(\mathcal{MV}_i)$ de l'individu i. Or dans le cas de plusieurs individus, l'étude de chaque modèle $(\mathcal{MV}_i)$ sera couteuse devant l'étude à la fois de l'ensemble des individus si cela est possible. Dans le cas où on peut modéliser l'ensemble des individus par un seul modèle, on parle alors des données en panel.
Or l'existence d'un mouvement brownien dans chaque modèle $(\mathcal{MV}_i)$, et sous l'hypothèse que le mouvement brownien de chaque individue i et indépendent de l'individu j, $(i \neq j)$,

implique l'existence d'un mouvement brownien multidimentionnel dans le modèle regroupant l'ensemble des individus. Dans ce cas les tests de Fisher pour tester l'hypothèse des données en panel ne seront pas valables.
Donc notre but est de modifier les tests de Fisher de façon à les rendre valables pour tester si les données sont en panel ou non.
On considère NT observations sur T période pour un ensemble d'individus de cardinal N.
On considère les notations (2.2.1-2.2.5), et on note encore :

$$(U_t)_{t\geq 0} = ((u_{1,t})_{t\geq 0}, (u_{2,t})_{t\geq 0}, ..., (u_{N,t})_{t\geq 0}) \tag{4.3.1}$$

$$y = \begin{pmatrix} y_1 \\ . \\ . \\ . \\ y_N \end{pmatrix} \tag{4.3.2}$$

$$X = \begin{pmatrix} X_1 & 0 & 0 & .... & 0 \\ 0 & X_2 & 0 & .... & 0 \\ 0 & 0 & X_3 & .... & 0 \\ . & . & . & .... & . \\ 0 & 0 & 0 & ....X_{N-1} & 0 \\ 0 & 0 & 0 & .... & X_N \end{pmatrix} \in \mathcal{M}_{NT}(\mathbf{IR}) \tag{4.3.3}$$

$$u = \begin{pmatrix} \gamma_1 u_1 \\ . \\ . \\ . \\ \gamma_N u_N \end{pmatrix} \tag{4.3.4}$$

$$\Psi = \begin{pmatrix} \psi_1 \\ . \\ . \\ . \\ \psi_N \end{pmatrix} \tag{4.3.5}$$

$$\beta = \begin{pmatrix} \beta_1 \\ . \\ . \\ . \\ \beta_N \end{pmatrix} \tag{4.3.6}$$

Donc le modèle coresspondant à la totalité des individus (N) pour T période est comme suit :

$$(\mathcal{MV}) \quad y = X\beta + u + \Psi$$

Puisque les mouvements unidimensionnels $(u_{i,t})_{t\geq 0}$ sont indépendants entre eux, alors le processus $(U_t)_{t\geq 0}$ est un mouvement brownien N-dimentionnel.

Sa matrice de covariance est donnée par :

$$\tilde{\Lambda}_{NT} = V(U) = \begin{pmatrix} \Sigma_T & 0 & 0 & .... & 0 \\ 0 & \Sigma_T & 0 & .... & 0 \\ 0 & 0 & \Sigma_T & .... & 0 \\ . & . & . & .... & . \\ 0 & 0 & 0 & ....\Sigma_T & 0 \\ 0 & 0 & 0 & .... & \Sigma_T \end{pmatrix} \in \mathcal{M}_{NT}(\mathbf{IR}) \tag{4.3.7}$$

Ainsi on a la matrice de covariance de u conditionnée par X est :

$$\Lambda_{NT} = V(u|X) = \begin{pmatrix} \gamma_1\Sigma_T & 0 & 0 & .... & 0 \\ 0 & \gamma_2\Sigma_T & 0 & .... & 0 \\ 0 & 0 & \gamma_3\Sigma_T & .... & 0 \\ . & . & . & .... & . \\ 0 & 0 & 0 & ....\gamma_{N-1}\Sigma_T & 0 \\ 0 & 0 & 0 & .... & \gamma_N\Sigma_T \end{pmatrix} \in \mathcal{M}_{NT}(\mathbf{IR}) \tag{4.3.8}$$

On pose :

$$\Omega_N = \begin{pmatrix} \gamma_1 & 0 & 0 & .... & 0 \\ 0 & \gamma_2 & 0 & .... & 0 \\ 0 & 0 & \gamma_3 & .... & 0 \\ . & . & . & .... & . \\ 0 & 0 & 0 & ....\gamma_{N-1} & 0 \\ 0 & 0 & 0 & .... & \gamma_N \end{pmatrix} \in \mathcal{M}_N(\mathbf{IR}) \tag{4.3.9}$$

Et

Alors on a :

$$\Lambda_{NT} = \Omega_N \bigotimes \Sigma_T \tag{4.3.10}$$

Où $\otimes$ désigne, le produit tensoriel.
Puisque les $\gamma_i \neq 0$ pour tout $i \in [|1, N|]$, alors la matrice Ω_N est inversible, son inverse est donnée par :

$$\Omega_N^{-1} = \begin{pmatrix} \frac{1}{\gamma_1} & 0 & 0 & .... & 0 \\ 0 & \frac{1}{\gamma_2} & 0 & .... & 0 \\ 0 & 0 & \frac{1}{\gamma_3} & .... & 0 \\ . & . & . & .... & . \\ 0 & 0 & 0 & ....\frac{1}{\gamma_{N-1}} & 0 \\ 0 & 0 & 0 & .... & \frac{1}{\gamma_N} \end{pmatrix} \tag{4.3.11}$$

Théorème 4.3.1 *Sous les notations ci-dessus, la matrice Λ_{NT} est symétrique inversible et admet une unique racine carrée $\sqrt{\Lambda_{NT}} = \sqrt{\Omega_N \bigotimes \Sigma_T}$ symétrique et inversible, et si on pose*

$$\varepsilon = \sqrt{\Omega_N^{-1} \bigotimes \Sigma_T^{-1}} u$$

alors on a :

$$\varepsilon|X \sim \mathcal{N}_{NT}(0, I_{NT}) \tag{4.3.12}$$

Où I_{NT} est la matrice identitée de $\mathcal{M}_{NT}(\mathbf{IR})$.

Preuve 4.3.1 *Puisque les matrices Ω_N et Σ_T sont symétriques et inversibles, alors $\Lambda_{NT} = \Omega_N \bigotimes \Sigma_T$ est symétrique et inversible. On a alors :*

$$\Lambda_{NT}^{-1} = (\Omega_N \bigotimes \Sigma_T)^{-1} = \Lambda_{NT}^{-1} = \Omega_N^{-1} \bigotimes \Sigma_T^{-1}$$

Donc on peut écrire Λ_{NT} sous la forme : $\Lambda_{NT} = PDP^{'}$, avec P est la matrice orthogonale formée par les vecteurs propres de Λ_{NT} et D est la matrice diagonale des valeurs propres de Λ_{NT}.
Or Λ_{NT} est inversible $\Rightarrow$ que les valeurs propres $\lambda_j \neq 0 \ \ \forall j \in [|0, NT|]$.
On pose :

$$d = \begin{pmatrix} sign(\lambda_1)\sqrt{|\lambda_1|} & 0 & 0\dots & 0 \\ 0 & sign(\lambda_2)\sqrt{|\lambda_2|} & 0\dots & 0 \\ . & . & \dots & . \\ 0 & 0 & \dots sign(\lambda_{NT-1})\sqrt{|\lambda_{NT-1}|} & 0 \\ 0 & 0 & 0\dots & sign(\lambda_{NT})\sqrt{|\lambda_{NT}|} \end{pmatrix}$$

Et

$$v = \begin{pmatrix} \frac{sign(\lambda_1)}{\sqrt{|\lambda_1|}} & 0 & 0 & .... & 0 \\ 0 & \frac{sign(\lambda_2)}{\sqrt{|\lambda_2|}} & 0 & .... & 0 \\ 0 & 0 & \frac{sign(\lambda_3)}{\sqrt{|\lambda_3|}} & .... & 0 \\ . & . & . & .... & . \\ 0 & 0 & 0 & ....\frac{sign(\lambda_{NT-1})}{\sqrt{|\lambda_{NT-1}|}} & 0 \\ 0 & 0 & 0 & .... & \frac{sign(\lambda_{NT})}{\sqrt{|\lambda_{NT}|}} \end{pmatrix}$$

Alors on a :

$$d^2 = D \ \ et \ \ vd = dv = I_{NT}$$

Donc D admet une racine carrée $\sqrt{D} = d$ inversible et sont inverse$\sqrt{D^{-1}} = v$.
Or :

$$(P\sqrt{D}P^{'})^2 = P\sqrt{D}P^{'}P\sqrt{D}P^{'} = PDP^{'} = \Lambda_{NT}$$

Alors Λ_{NT} admet une racine carrée $\sqrt{\Lambda_{NT}} = P\sqrt{D}P^{'}$.
Puisqu'on a :
$\sqrt{D}$ est inversible, alors $\sqrt{\Lambda_{NT}} = P\sqrt{D}P^{'}$ est aussi inversible, car :

$$P\sqrt{D}P^{'}P\sqrt{D^{-1}}P^{'} = P\sqrt{D}\sqrt{D^{-1}}P^{'} = PP^{'} = I_{NT}$$

Donc

$$\sqrt{\Lambda_{NT}}^{-1} = P\sqrt{D^{-1}}P^{'}$$

.
Conclusion : Λ_{NT} admet une racine carrée inversible.
La preuve de l'unicité de le racine carrée inversible de Λ_{NT} est identique à celle du théorème (2.2.1), dans la deuxime étape de démonstration.
Si on pose :

$$\varepsilon = \sqrt{\Omega_N^{-1} \bigotimes \Sigma_T^{-1}} u$$

Alors

$$\begin{aligned}\varepsilon = \sqrt{\Lambda_{NT}^{-1}}u \Rightarrow\ V(\varepsilon|X) &= \sqrt{\Lambda_{NT}^{-1}}V(u|X)\sqrt{\Lambda_{NT}^{-1}}' \\ &= \sqrt{\Lambda_{NT}^{-1}}\Lambda_{NT}\sqrt{\Lambda_{NT}^{-1}} \\ &= I_{NT}\end{aligned}$$

Et

$$E(\varepsilon|X) = E(\sqrt{\Lambda_{NT}^{-1}}u|X) = \sqrt{\Lambda_{NT}^{-1}}E(u|X) = 0$$

Puisque u est gaussien alors $\varepsilon = \sqrt{\Lambda_{NT}^{-1}}u$ est aussi gaussien. Donc :

$$\varepsilon|X \sim \mathcal{N}_{NT}(0, I_{NT})$$

La preuve est achevée □.

Corollaire 4.3.1 *Soit $(N,T) \in \mathbf{IN}^*$.*
On considère NT observations d'une variable endogène $y \in \mathbf{IR}^{NT}$ et des variables explicatives $X \in \mathcal{M}_{NT}(\mathcal{IR})$. On suppose que ces observations sont prises d'un système comportant un mouvement brownien $(U_t)_{t\geq 0}$.
Alors l'application définie par :

$$\Phi :\ \mathbf{IR}^{NT} \ \rightarrow\ \mathbf{IR}^{NT} \tag{4.3.13}$$

$$y \rightarrow < \sqrt{\Omega_N^{-1} \bigotimes \Sigma_T^{-1}}, y > \tag{4.3.14}$$

(où $< ., >$ désigne le produit scalaire), élimine l'effet de l'hétéroscédasticité de la variance des mouvements qui proviennent de ce mouvement brownien, dans le modèle.
De plus :

$$\arg\min_{\beta\in\Theta} ||y - X\beta|| = \arg\min_{\beta\in\Theta} ||\Phi(y) - \Phi(X\beta)||$$

Preuve 4.3.2 *On a*

$$\Phi(u) =< \sqrt{\Omega_N^{-1} \bigotimes \Sigma_T^{-1}}, u >= \varepsilon \Rightarrow V(\Phi(u)|X) = I_{NT}$$

Donc l'application Φ élimine l'effet de l'hétéroscédasticité de la variance des mouvement provenant de ce mouvement brownien.
D'autre par, d'après la construction de Φ, elle est un automorphisme, car sa matrice $\sqrt{\Omega_N^{-1} \bigotimes \Sigma_T^{-1}}$ est inversible. Aussi Φ et $||.||$ sont continues.
Donc :

$$\begin{aligned}\hat{\beta} \in \arg\min_{\beta\in\Theta} ||y - X\beta|| &\Leftrightarrow ||y - X\hat{\beta}|| = \min_{\beta\in\Theta} ||y - X\beta|| \\ &\Leftrightarrow ||\Phi(y - X\hat{\beta})|| = \min_{\beta\in\Theta} ||\Phi(y - X\beta)|| \\ &\Leftrightarrow ||\Phi(y) - \Phi(X\hat{\beta})|| = \min_{\beta\in\Theta} ||\Phi(y) - \Phi(X\beta)|| \\ &\Leftrightarrow \hat{\beta} \in \arg\min_{\beta\in\Theta} ||\Phi(y) - \Phi(X\beta)||\end{aligned}$$

Donc

$$\arg\min_{\beta\in\Theta} ||y - X\beta|| = \arg\min_{\beta\in\Theta} ||\Phi(y) - \Phi(X\beta)||$$

□.

le corrolaire au-dessus, montre qu'on peut travailler sur les nuages de points $\Phi(y)$ et $\sqrt{\Omega_N^{-1} \otimes \Sigma_T^{-1}} X$ au lieu de y et X.
Reprenons le modèle $(\mathcal{MV})$. Estimer les paramètres de ce modèle revient donc à estimer son modèle équivalent :

$$(\mathcal{MVB}) \quad \Phi(y) = \Phi(X\beta) + \Phi(u) + \Phi(\Psi)$$

Où encore :

$$(\mathcal{MVB}) \quad \sqrt{\Omega_N^{-1} \bigotimes \Sigma_T^{-1}} y = \sqrt{\Omega_N^{-1} \bigotimes \Sigma_T^{-1}} X\beta + \varepsilon + \sqrt{\Omega_N^{-1} \bigotimes \Sigma_T^{-1}} \Psi$$

On pose encore :

$$\nu = \varepsilon + \sqrt{\Omega_N^{-1} \bigotimes \Sigma_T^{-1}} \Psi \tag{4.3.15}$$

Alors

$$(\mathcal{MVB}) \quad \sqrt{\Omega_N^{-1} \bigotimes \Sigma_T^{-1}} y = \sqrt{\Omega_N^{-1} \bigotimes \Sigma_T^{-1}} X\beta + \nu$$

Donc on peut faire les tests et appliquer les estimateurs qu'on a proposés dans les sections précédentes, en changeant juste les matrices X_i par les matrice X et les vecteurs y_i par le vecteur y, ainsi que la matrice $\sqrt{\Sigma_T^{-1}}$ par la matrice $\sqrt{\Omega_N^{-1} \otimes \Sigma_T^{-1}}$.
Pour les estimateurs qu'on a proposés, dans le cas d'un mouvement brownien multi-dimensionel, il faut juste changer la matrice $\sqrt{\Sigma_T^{-1}}$ par la matrice $\sqrt{\Omega_N^{-1} \otimes \Sigma_T^{-1}}$.
Autrement- dit, on redéfinit les estimateurs β^{BOLS}, β^{GBLS}, β^{BIV} et β^{G2MBM} comme suit :

Définition 4.3.1 *Avec les notations ci dessus, on appel et on définit :*

1. *l'estimateur de moindre carré ordinaire avec composante brownienne multi-dimensionelle (Ordinary Least Square with Multi-Dimensional Brownian Component), par :*

$$\beta^{OLSMDBC} = (X' \Lambda_{NT}^{-1} X)^{-1} X' \Lambda_{NT}^{-1} y \tag{4.3.16}$$

2. *l'estimateur des moindres carrés généralisés avec composante brownienne multi-dimensionelle (Generalized Least Squares with Multi-Dimensional Brownian Component), par :*

$$\beta^{GLSMDBC} = (X' \sqrt{\Lambda_{NT}^{-1}} \Omega^{-1} \sqrt{\Lambda_{NT}^{-1}} X)^{-1} X' \sqrt{\Lambda_{NT} - 1} \Omega y \tag{4.3.17}$$

3. *l'estimateur de variables instrumentales avec composante brownienne multi-dimensionelle (Instrumental Variables with Multi-Dimensional Brownian Component), par :*

$$\beta^{IS2DBC} = ((X^*)' \sqrt{\Lambda_{NT}^{-1}} P_\theta \sqrt{\Lambda_{NT}^{-1}} X^*)^{-1} (X^*)' \sqrt{\Lambda_{NT}^{-1}} P_\theta \Lambda_{NT}^{-1} y^* \tag{4.3.18}$$

4. *l'estimateur de la méthode des moments généralisés avec composante brownienne multi-dimensionelle (Generalized Method of Moments with Multi-Dimensional Brownian Component), par :*

$$\beta^{G3MDBC} = ((X^*)' \sqrt{\Sigma^{-1}} \theta (\theta' \Omega \theta^*)^{-1} \theta' \sqrt{\Lambda_{NT}^{-1}} X^*)^{-1} (X^*)' \sqrt{\Lambda_{NT}^{-1}} \theta (\theta' \Omega \theta^*)^{-1} \Lambda_{NT}^{-1} y^* \tag{4.3.19}$$

4.4 Procédure générale de tests des donées en panel

On considère qu'on dispose des observations de N individus pour T périodes successives.
Soient : X_i est la matrice d'observations de m variables explicatives de l'individu i en T périodes, avec le fait que sa première colonne est le vecteur $\underbrace{(1, 1, \ldots 1)'}_{T}$;
y_i est le vecteur d'observations de la variable endogène de l'individu i en T périodes.
β_i : les coefficients de modèle de l'individu i. On veut tester l'hypothèse des données en panel, autrement dit, tester l'hypothèse :

$$\mathcal{H}_0 \;\; \beta_1 = \beta_2 = ... = \beta_N$$

On note encore :

$$\tilde{\beta} = \begin{pmatrix} \tilde{\beta}_1 \\ . \\ . \\ . \\ \tilde{\beta}_m \end{pmatrix} \tag{4.4.1}$$

Avec

$$\beta_i = \begin{pmatrix} \beta_{i,0} \\ \tilde{\beta}_i \end{pmatrix}$$

On a quatre Scénarios :

Scénario 1 Pour tout $i \in \{1, 2, ..., N\}$ $\beta_i = \beta$, dans ce cas on a un seul modèle expliquant les variables endogènes de tous les individus i. On parle donc d'un modèle en panel homogène.

Scénario 2 Pour tout $i \in \{1, 2, ..., N\}$ $\tilde{\beta}_i = \tilde{\beta}$ et $\forall j \in \{1, 2, ..., N\}$ $j \neq i \; , \Rightarrow \beta_{i,0} \neq \beta_{j,0}$, dans ce cas, le modèle expliquant la variable endogène de chaque individue à un autre diffère par une constante $\beta_{i,0}$. On parle donc d'un modèle à effets individuels.

Scénario 3 Pour tout $i \in \{1, 2, ..., N\}$ $\beta_{i,0} = \beta_0$ et $\forall j \in \{1, 2, ..., N\}$ $j \neq i \; , \Rightarrow \tilde{\beta}_i \neq \tilde{\beta}_j$, dans ce cas, on a N modèles différents expliquant les variables endogènes, mais avec une espérance de productivité de facteur moyenne homogène.

Scénario 4 Pour tout $(i, j) \in \{1, 2, ..., N\}^2$ $j \neq i \Rightarrow \tilde{\beta}_i \neq \tilde{\beta}_j$ et $\beta_{i,0} \neq \beta_{j,0}$, dans ce cas on rejette l'hypothèse d'une structure en panel et on a N modèles totalement différents.

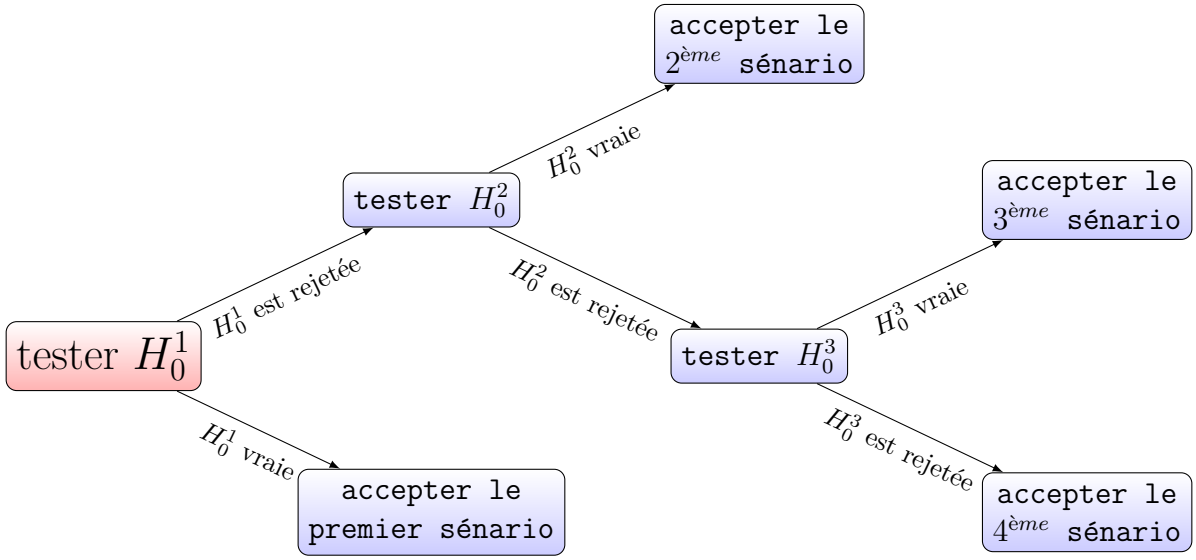
Pour tester ces scénarios, on considère les hypothèses suivantes :

$$\mathcal{H}_0^1 : \;\; i \in \{1, 2, ..., N\} \;\; \beta_i = \beta \tag{4.4.2}$$

$$\mathcal{H}_0^2 : \;\; i \in \{1, 2, ..., N\} \;\; \tilde{\beta}_i = \tilde{\beta} \tag{4.4.3}$$

$$\mathcal{H}_0^3 : \;\; i \in \{1, 2, ..., N\} \;\; \beta_{i,0} = \beta_0 \tag{4.4.4}$$

On représente la procédure générale de test d'homogénéité (test en panel) dans le schéma illustratif suivant :



On pose :

$$z_i = \sqrt{\Sigma_T^{-1}} y_i$$

$$W_i = \sqrt{\Sigma_T^{-1}} X_i = (\sum_{l=1}^{T} \rho_{l,t} x_{i,k,l})_{(k,l)\in[|0,m|]\times[|1,T|]}$$

Avec

$$x_{i,0,l} = 1, \ \forall i \in [|1,T|]$$

$$\underbrace{\bar{w}_i}_{(1,T)} = \sum_{t=1}^{T} w_{i,t}$$

$$\underbrace{\bar{z}_i}_{(1,1)} = \sum_{t=1}^{T} z_{i,t}$$

$$S_{zz,i} = \sum_{t=1}^{T} (z_{i,t} - \bar{z}_{i,t})^2$$

$$S_{zz} = \sum_{i=1}^{N} \sum_{t=1}^{T} (z_{i,t} - \bar{z}_{i,t})^2$$

$$\underbrace{S_{ww,i}}_{T,T} = \sum_{t=1}^{T} (w_{i,t} - \bar{w}_i)'(w_{i,t} - \bar{w}_i)$$

$$S_{ww} = \sum_{i=1}^{N} \sum_{t=1}^{T} (w_{i,t} - \bar{w}_i)'(w_{i,t} - \bar{w}_i)$$

$$\underbrace{S_{wz,i}}_{T,1} = \sum_{t=1}^{T} (z_{i,t} - \bar{z}_i)(w_{i,t} - \bar{w}_i)'$$

$$S_{wz} = \sum_{i=1}^{N} \sum_{t=1}^{T} (z_{i,t} - \bar{z}_i)(w_{i,t} - \bar{w}_i)'$$

$$SCR_{1,i} = S_{zz,i} - S'_{wz,i} S^{-1}_{wz,i} S_{wz,i}$$

$$SCR_{1,c} = S_{zz} - S'_{wz} S^{-1}_{wz} S_{wz}$$

$$SCR_1 = \sum_{i=1}^{N} SCR_{1,i}$$

$$SCR_{1,c'} = \sum_{i=1}^{N} S_{zz,i} - \Big(\sum_{i=1}^{N} S_{wz,i}\Big)' \Big(\sum_{i=1}^{N} S_{wz,i}\Big)^{-1} \Big(\sum_{i=1}^{N} S_{wz,i}\Big)$$

On note : F_1 est la statistique de Fisher associée à l'hypothèse H_0^1, F_2 est la statistique de Fisher associée à l'hypothèse H_0^2 et F_2 est celle associée à l'hypothèse H_0^3.
Les trois statistiques sont définies comme suit :

$$F_1 = \frac{(SCR_{1,c} - SCR_1)/(m+1)(N-1)}{SCR_1/[NT-(m+1)N]} \tag{4.4.5}$$

$$F_2 = \frac{(SCR_{1,c'} - SCR_1)/m(N-1)}{SCR_1/[NT-(m+1)N]} \tag{4.4.6}$$

$$F_3 = \frac{(SCR_{1,c} - SCR_{1,c'})/(N-1)}{SCR_{1,c'}/[(T-1)N-m]} \tag{4.4.7}$$

Où : $((m+1)(N-1), NT-(m+1)N)$, $(m(N-1), NT-(m+1)N)$ et $(N-1, (T-1)N-m)$ représentent respectivement les degrés de libertés de F_1, F_2 et F_3.
Au niveau de risque α, on pose :
$F_{théo\,1} = F_{(m+1)(N-1),NT-(m+1)N,\alpha}$, c'est le quantile d'ordre $1-\alpha$ de la loi de Fisher pour les degrés de liberté $((m+1)(N-1), NT-(m+1)N)$. De façon similaire, on note :

$$F_{théo\,1,\alpha} = F_{(m+1)(N-1),NT-(m+1)N,\alpha}$$

$$F_{théo\,2,\alpha} = F_{m(N-1),NT-(m+1)N,\alpha}$$

$$F_{théo\,3,\alpha} = F_{N-1,(T-1)N-m,\alpha}$$

On définit le p-value pour chaque hypothèse H_0^j, $(j = 1, 2, 3)$, comme suit :

$$pv_{j,\alpha} = P(F_{théo\,j,\alpha} \geq F_j) \tag{4.4.8}$$

Le p-value représente le seuil minimal pour accepter l'hypothèse nulle.
Donc pour accepter l'hypothèse H_0^j, $(j = 1, 2, 3)$, il faut que $pv_{j,\alpha} \geq \alpha$.

4.5 Conclusion

Dans ce chapitres on a étudié des modèles contenants plusieurs individus (variables endogènes) et plusieurs variables explicatives, avec un mouvement brownien multidimentionnel. En premier lieu, on a commencé par étudier l'hypothèse d'égalité de la variance entre les individus. puis on s'est intéréssé à donner des estimateurs optimaux des modèles en panel, en généralisant les estimateurs qu'on a abordés dans les chapitres 2 et 3. On a terminé ce chapitre, en abordant le test des données en panel, pour trouver la catégorie des modèles modélisant l'ensembles des individus en fonction des variables explicatives.

Deuxième partie

Modèles d'optimisation stochastiques de la croissance économique en tenant compte du rattrappage technologique et de l'équilibre écologique

Chapitre 5

Modèle stochastique de la croissance économique avec technologie hétérogène et convergence du revenu

5.1 Introduction

Les principales hypothèses du résultat de la convergence dans le modèle de Solow, sont la baisse factorielle du rendement du capital et une homogénéité de la diffusion technologique au niveau international. Cette seconde hypothèse implique une égalité des taux de croissance du progrès technique au niveau international (Phillips & Sul, [60]). Selon le travail empirique de Teheni El Ghak [77]et Di Liberto et al, l'hypothèse de l'homogénéité du taux de croissance du progrès technique est largement rejetée. Autrement dit il existe une forte hétérogénéité des niveaux technologiques entre les pays. Selon Faberger [78], les disparités des taux de croissance entre les pays sont principalement dues aux différences de niveaux technologique. [1] Selon les travaux de Seyma et Alev [63], éthier [64], McDonald [65], benhabib [66], ces disparités dans les taux de croissance proviennent du capital humain et de l'insertion de la recherche et du développement dans la production.
Nos objectifs dans ce chapitre sont :
Optimiser l'espérence du PIB par emploi, en tenant compte des équations dynamiques budgetaires et des chocs externes sur le PIB, en supposant que la technologie est un mouvement brownien géométrique.
Établir l'équation de rattrapage technologique partiel entre un pays en voix de développement et un pays avancé.
Pour attiendre ces deux objectifs, nous organisons ce chapitre comme suit :
Après cette section d'introduction, le chapitre contient cinq section et se termine par une conclusion.
Dans la première section ; nous donnons la signification des notation utilisées dans le chapitre et les hypothèses macroéconomiques de la modélisation de notre problème.
La deuxième section est consacrée à la construction du modèle.
La troisième section est consacrée à l'analyse des contraintes et l'étude de l'équilibre budgétaire.
Dans la quatrième section nous résolvons le modèle proposé, en utilisant le principe du maximum stochastique.

1. Le test de l'hypothèse d'homogénéité sur un échantillon de 102 pays non pétroliers sur la période 1960-1989 donne un taux de croissance du progrès technique compris entre -3,1 et 7,4 entre les pays.

Dans la cinquième section, nous trouvons l'équation caractérisant la condition de rattrapage technologique d'un pays en voix de développement par rapport à un pays avancé, ainsi nous faisons la simulation du modèle proposé.

5.2 Notations et hypothèses macroéconomiques modèles

5.2.1 Notations

Dans la suite de ce chapitre, on note :
t le temps.
A_t est le niveau de technologie du pays à l'instant t.
B_t Représente la frontière technologique mondiale, elle correspond à l'indice des biens et services les plus avancés.
τ La fraction du temps consacrée par l'économie à l'accumulation de savoir-faire.
K est la formation brute du capital.
H le capital humain.
L est le travail, il représente la quantité de main-d'œuvre participée dans la production.
Y Produit intérieur brute (PIB).
$\tilde{a}_i = \frac{H_i}{A_i}$ La proportion d'apprentissage par rapport à la frontière technologique mondiale.
$k = \frac{K}{L}$, $h = \frac{H}{L}$ et $y = \frac{Y}{L}$ Sont, respectivement, le capital par emploi, le capital humain par emploi et le PIB par tête.
Pour chaque aggregat F, on note et on définit :
Pour tout vecteur de variables d'entrées réelles $(x_1, x_2, ..., x_j)$, $F_{x_1,x_2,...,x_j} = F(x_1, x_2, ..., x_j)$.
$\gamma_F = \frac{\frac{\partial F}{\partial t}}{F}$ Le taux de croissance de l'agrégat F.
$f = \frac{F}{L}$ est l'agrégat par emploi
$\hat{f} = \frac{F}{AL}$ est l'agrégat par emploi efficace.
On pose $\Omega = [0, 1] \times [0, +\infty[$.
Pour des raisons de régularité, on suppose que les fonctions $H \in \mathcal{C}^1(\Omega) \bigcap \mathcal{L}^\infty(\Omega)$ et $K \in \mathcal{C}^1([0, +\infty[) \bigcap \mathcal{L}^\infty([0, +\infty[)$.
On note respéctivement s,s_K,s_H et s_R : Le taux d'investissement total, le taux d'investissement dans la production, le taux d'investissement consacré à l'accumulation de savoir-faire et le taux d'investissement dans le secteur de la recherche et développement.

5.2.2 Hypothèses macroéconomiques du modèle

1. Le revenu Y est réparti en trois ; la consommation C, les dépenses G et l'investissement I.
$$Y_t = I_t + C_t + G_t$$
2. l'investissement I_t est également divisé en trois : investissement pour la production, il est noté I_K, le deuxième noté I_H est celui pour l'accumulation de savoir-faire et le troisième dénoté I_R est celui du secteur de recherche et développement. Nous avons donc :
$$I_t = sY_t$$

$$I_K = s_K Y_t$$

$$I_H = s_H Y_t$$

$$I_R = s_R Y_t$$

Et

$$s = s_K + s_H + s_R$$

3. La technologie du pays est un processus brownien géométrique.
4. La technologie du pays leader progresse de façon Malthusianne
5. le capital humain est dépendant du niveau d'éducation de la main-d'œuvre, de la fraction du temps consacrée à l'accumulation de savoir-faire et la frontière technologique mondiale.
6. Le revenu total Y_t est une fonction de Cobb Douglas
7. La formation brute du capital et le capital humain croient respectivement par les taux s_K,s_H et diminuent par des taux de dépréciation constants δ et η. Où δ et η, sont respectivement les amortissements de la formation brute du capital et du capital humain.

5.3 Construction du modèle

Dans notre modélisation, nous supposons que le capital humain dépend des dépenses en éducation et de la qualification de main-d'œuvre.
Nous nous basons sur les travaux empiriques de Nelson & Phelps [74] et Romer avec transfert technologique [89], nous proposons une nouvelle équation dynamique d'accumulation de savoir-faire :

$$\frac{\partial H_{\tau,t}}{\partial t} = \chi_t e^{\tau\theta} H_{\tau,t}^{1-\lambda} B_t^{\lambda} \tag{5.3.1}$$

Où :
χ_t est l'indice du capital humain.
τ est la fraction du temps consacrée par l'économie à l'accumulation de savoir-faire.
λ est le coefficient de diffusion technologique.
θ est une constante strictement positive.
Selon la troisième hypothèse, la technologie $\Big(A\Big)_t$ est un processus brownien géométrique.
On pose :

$$A_t = A_{t_0} e^{\gamma_A (t-t_0)+w_t} \tag{5.3.2}$$

Avec
$(w_t)_{t\geq 0}$ est un mouvement brownien qui représente les chocs technologiques et il satisfait :

$$\frac{dw_t}{dt} = \varepsilon_t \tag{5.3.3}$$

Où ε_t est un bruits blanc, tel que :

$$E(\varepsilon_t) = 0$$

$$V(\varepsilon_t) = \sigma_a^2$$

σ_a est l'écart-type de la technologie.
Le revenu Y_t est supposé suivre le modèle d'une fonction Cobb Douglas, il est définie par :

$$Y_t = A_t K_t^{\alpha} H_{\tau,t}^{\beta_i} L_t^{\gamma} R_t^{\kappa} \tag{5.3.4}$$

Où :
$(\alpha, \beta, \gamma, \kappa) \in [0,1]^4$ sont respectivement l'élasticité de la formation brute du capital, du capital humain, du travail et des dépenses en recherche et développement.
En terme d'agrégat éfficace par emploi, nous avons :

$$\hat{y}_t = \frac{Y_t}{A_t L_t} = \hat{k}_t^{\alpha} \hat{h}_{\tau,t}^{\beta} \hat{r}_t^{\kappa} L_{i,t}^{\gamma-\alpha-\kappa-\beta-1} \tag{5.3.5}$$

Nous considérons que les équations dynamiques d'accumulation de la formation brute du capital, du capital humain et des dépenses en recherche et développement sont respectivement données par :

$$\frac{\partial K_t}{\partial t} = s_K Y_t - \delta K_t \tag{5.3.6}$$

$$\frac{\partial H_{\tau,t}}{\partial t} = s_H Y_t - \eta H_{\tau,t} \tag{5.3.7}$$

$$\frac{\partial R_t}{\partial t} = s_H Y_t - \phi R_t \tag{5.3.8}$$

Proposition 5.3.1 *On pose :*

$$I_{\hat{k},t} = \frac{I_K}{A_t L_t},\ I_{\hat{r},t} = \frac{I_R}{A_t L_t},\ I_{\hat{h},t} = \frac{I_H}{A_t L_t} \tag{5.3.9}$$

$$b_1(t, I_{\hat{k},t}, \hat{k}) = I_{\hat{k},t} - (\gamma_A + \gamma_L + \delta)\hat{k}_t \tag{5.3.10}$$

$$b_2(t, I_{\hat{r},t}, \hat{r}) = I_{\hat{r},t} - (\gamma_A + \gamma_L + \phi)\hat{r}_t \tag{5.3.11}$$

$$b_3(t, I_{\hat{h},t}, \hat{h}) = I_{\hat{h},t} - (\gamma_A + \gamma_L + \eta)\hat{h}_{\tau,t} \tag{5.3.12}$$

Les équations dynamiques d'accumulation des capitaux efficaces par emploi sont respectivement données par :

$$d\hat{k}_t = b_1(t, I_{\hat{k},t}, \hat{k})dt - \hat{k}_t dw_t \tag{5.3.13}$$

$$d\hat{r}_t = b_2(t, I_{\hat{r},t}, \hat{r})dt - \hat{r}_t dw_t \tag{5.3.14}$$

$$d\hat{h}_{\tau,t} = b_3(t, I_{\hat{h},t}, \hat{h})dt - \hat{h}_{\tau,t} dw_t \tag{5.3.15}$$

Preuve 5.3.1 *On a*

$$\frac{d\hat{k}_t}{dt} = \frac{d}{dt}\Big(\frac{K_t}{A_t L_t}\Big) = \frac{\frac{dK_t}{dt} A_t L_t - (\frac{dA_t}{dt} L_t + \frac{dL_t}{dt} A_t) K_t}{(A_t L_t)^2}$$

D'après l'équation (5.3.6), on obtient :

$$\frac{d\hat{k}_t}{dt} = \frac{(s_K Y_t - \delta K_t) A_t L_t - (\frac{dA_t}{dt} L_t + \frac{dL_t}{dt} A_t) K_t}{(A_t L_t)^2}$$

Donc

$$\frac{d\hat{k}_t}{dt} = I_{\hat{k},t} - \frac{\frac{dA_t}{dt}}{A_t}\hat{k}_t - \frac{\frac{dL_t}{dt}}{L_t}\hat{k}_t - \delta\hat{k}_t$$

D'après les équations (5.3.2)-(5.3.3), on a :

$$\frac{d}{dt}(\ln(A_t)) = \frac{\frac{dA_t}{dt}}{A_t} = \gamma_A + \frac{dw_t}{dt}$$

Et

$$\frac{\frac{dL_t}{dt}}{L_t} = \gamma_L$$

Alors

$$\frac{d\hat{k}_t}{dt} = I_{\hat{k},t} - (\gamma_A + \frac{dw_t}{dt})\hat{k}_t - \gamma_L\hat{k}_t - \delta\hat{k}_t$$

Ou encore :

$$d\hat{k}_t = I_{\hat{k},t}dt - \gamma_A\hat{k}_t dt + \hat{k}_t dw_t - \gamma_L\hat{k}_t dt - \delta\hat{k}_t dt$$

Or $b_1(t, I_{\hat{k},t}, \hat{k}) = I_{\hat{k},t} - (\gamma_A + \gamma_L + \delta)\hat{k}_t$, *alors :*

$$d\hat{k}_t = b_1(t, I_{\hat{k},t}, \hat{k})dt - \hat{k}_t dw_t$$

Pour l'équation (5.3.14) et (5.3.15) il suffit de résonner par la même méthode.□

On note par $(\Gamma_t)_{t\geq t_0}$, une filtration qui représante l'information connue à l'instant t. Nous proposons le modèle d'optimisation (noté (P)) suivant :

$$(P) = \begin{cases} \max_{I_k, I_h, I_r} E\Big[\int_{t_0}^{T} \varrho^t \hat{k}_t^{\alpha} \hat{h}_{\tau,t}^{\beta} \hat{r}_t^{\kappa} L_t^{\gamma-\alpha-\kappa-\beta-1} dt | \Gamma_{t_0}\Big] \\ S.C \\ d\hat{k}_t = b_1(t, I_{\hat{k},t}, \hat{k})dt - \hat{k}_t dw_t \\ d\hat{r}_t = b_2(t, I_{\hat{r},t}, \hat{r})dt - \hat{r}_t dw_t \\ d\hat{h}_{\tau,t} = b_3(t, I_{\hat{h},t}, \hat{h})dt - \hat{h}_{\tau,t} dw_t \\ \hat{h}_{\tau,t_0} = \hat{h}_0 \\ \hat{r}_{t_0} = \hat{r}_0 \\ \hat{k}_T = q\hat{k}_0 \\ \hat{h}_{\tau,T} = q\hat{h}_0 \\ \hat{r}_T = q\hat{r}_0 \\ q > 1 \end{cases}$$

Où
ϱ^t est le facteur d'actualisation.

5.4 Analyse des équations dynamiques budgétaires

5.4.1 Étude de trajectoire de la formation brute du capital en fonction du capital humain

Le niveau technologique mondial croît avec le rythme :

$$\gamma_{B_t} = \frac{\frac{dB_t}{dt}}{B_t} \tag{5.4.1}$$

Dans les économies en développement, ce taux est constant :

$$\gamma_{B_t} = g = constante$$

Donc

$$B_t = B_{t_0} e^{g(t-t_0)} \tag{5.4.2}$$

Utilisant l'équation (5.3.1), pour tout $t_0 \in [0, +\infty[$ et $\forall (\tau, t) \in \Omega$ on a :

$$\int_{t_0}^{t} \frac{\frac{\partial H_{\tau,s}}{\partial s}}{H_{\tau,s}^{1-\lambda}} ds = e^{\tau\theta} \int_{t_0}^{t} \chi_s B_s^{\lambda} ds$$

Donc

$$-\Big[\frac{1}{\lambda H_{\tau,s}^{-\lambda}}\Big]_{t_0}^{t} = \frac{1}{\lambda}(H_{\tau,t_0}^{\lambda} - H_{\tau,t}^{\lambda}) = B_{t_0}^{\lambda} e^{\tau\theta} \int_{t_0}^{t} \chi_s e^{g\lambda(s-t_0)} ds$$

Alors, la résolution de l'équation (1) donne :

$$H_{\tau,t} = \Big[\lambda B_{t_0}^{\lambda} e^{\tau\theta} \int_{t_0}^{t} \chi_s e^{g\lambda(s-t_0)} ds - H_{\tau,t_0}^{\lambda}\Big]^{\frac{1}{\lambda}} \tag{5.4.3}$$

dans le cas où $(\delta = \eta)$, en utilisant les équations (5.3.6) et (5.3.7) on trouve :

$$s_H \frac{\partial K_t}{\partial t} - s_K \frac{\partial H_{\tau,t}}{\partial t} = -\delta(s_H K_t - s_K H_{\tau,t}) \tag{5.4.4}$$

on pose :

$$X_{\tau,t} = s_H K_t - s_K H_{\tau,t} \tag{5.4.5}$$

Alors on obtient :

$$\frac{\partial X_{\tau,t}}{\partial t} = -\delta X_{\tau,t} \tag{5.4.6}$$

La solution de la dernière équation est la suivante :

$$X_{\tau,t} = X_{\tau,t_0} e^{-\delta(t-t_0)} \tag{5.4.7}$$

Alors les solutions des équations budgétaires (5.3.6) − (5.3.7) au cours du temps sont toutes les couples :

$$(K_t, H_{\tau,t}) \in \mathcal{C}^1(\Omega) \bigcap \mathcal{L}^{\infty}(\Omega) \times \mathcal{C}^1([0, +\infty[) \bigcap \mathcal{L}^{\infty}([0, +\infty[)$$

qui vérifient :

$$K_t = \frac{1}{s_H}[s_K H_{\tau,t} + (s_H K_{t_0} - s_K H_{\tau,t_0}) e^{-\delta(t-t_0)}] \tag{5.4.8}$$

et

$$H_{\tau,t} = \Big[\lambda \mathcal{A}_{l,t_0}^{\lambda} e^{\tau\theta} \int_{t_0}^{t} \chi_s e^{g\lambda(s-t_0)} ds - H_{\tau,t_0}^{\lambda}\Big]^{\frac{1}{\lambda}} \tag{5.4.9}$$

On a trois sénarios.
Sénario 1 : Si $\frac{K_{\tau,t_0}}{H_{\tau,t_0}} > \frac{s_K}{s_H}$
Alors la trajectoire de la formation brute du capital en fonction du capital humain, sur la période $[t_0, T]$ est donné par la figure (5.1) :

Dans ce cas en voit que la formation brute du capital humain augmente au cours du temps plus que le capital humain.

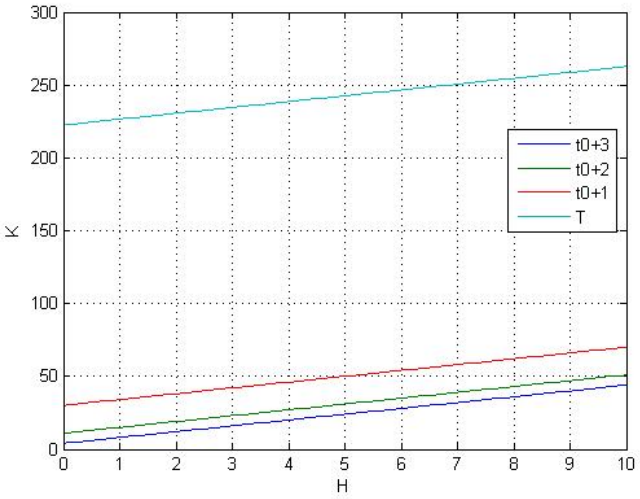


FIGURE 5.1 – Trajectoire de la formation brute du capital en fonction du capital humain

Donc si on décide à fixer les taux des investissements s_K et s_H de façon telle que : $\frac{K_{\tau,t_0}}{H_{\tau,t_0}} > \frac{s_K}{s_H}$, alors on donne la priorité d'investir dans la production plus que d'investir en capital humain.

Sénario 2 : Si $\frac{K_{\tau,t_0}}{H_{\tau,t_0}} < \frac{s_K}{s_H}$

Alors la trajectoire de la formation brute du capital en fonction du capital humain, sur la période $[t_0, T]$ est est donné par la figure (5.2) :

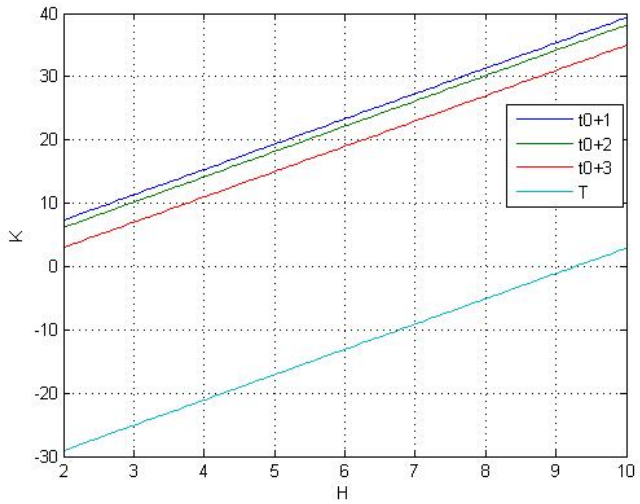


FIGURE 5.2 – Trajectoire de la formation brute du capital en fonction du capital humain

Dans ce cas en voit que la formation brute du capital humain augmente au cours du temps moins que le capital humain.

Donc si on décide à fixer les taux des investissements s_K et s_H de façon telle que : $\frac{K_{\tau,t_0}}{H_{\tau,t_0}} < \frac{s_K}{s_H}$, alors on donne la priorité d'investir dans le capital humain.

Sénario 3 : Si $\frac{K_{t_0}}{H_{\tau,t_0}} = \frac{s_K}{s_H}$

Alors la trajectoire de la formation brute du capital en fonction du capital humain, sur la

période $[t_0, T]$ est donnée par la figure (5.3).
Dans ce cas en voit que la formation brute du capital humain augmente au cours du

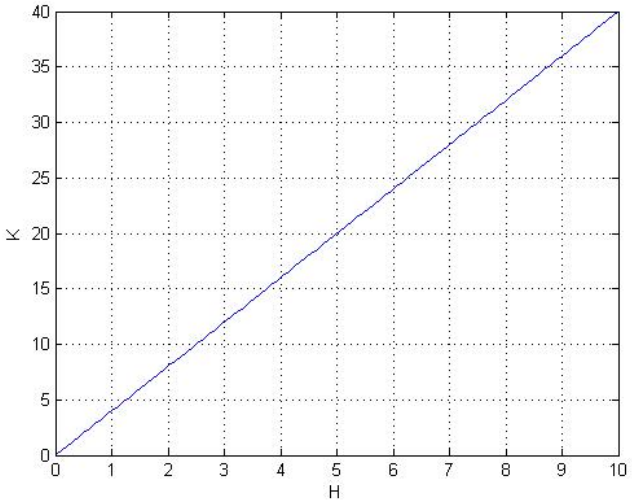


FIGURE 5.3 – Trajectoire de la formation brute du capital en fonction du capital humain

temps autant que le capital humain.
Donc si on décide à fixer les taux des investissements s_K et s_H de façon telle que : $\frac{K_{\tau,t_0}}{H_{\tau,t_0}} = \frac{s_K}{s_H}$, alors on donne la même importance d'investir dans la production aussi que dans le capital humain.

5.4.2 Analyse de l'équilibre budgétaire

L'équilibre budgétaire est réalisée lorsque les investissement et les dépreciations des capitaux sont égaux.
Durant une siècle d'équilibre économique, les taux de croissances seront nuls.
Or d'après les équations (5.3.13) et (5.4.15), les taux de croissances sont données par :

$$\frac{d\hat{k}_t}{\hat{k}_t} = \frac{b_1(t, I_{\hat{k},t}, \hat{k})}{\hat{k}_t} - \frac{dw_t}{dt} \tag{5.4.10}$$

$$\frac{d\hat{r}_t}{\hat{r}_t} = \frac{b_2(t, I_{\hat{r},t}, \hat{r})}{\hat{r}_t}dt - \frac{dw_t}{dt} \tag{5.4.11}$$

$$\frac{d\hat{h}_{\tau,t}}{\hat{h}_{\tau,t}} = \frac{b_3(t, I_{\hat{h},t}, \hat{h})}{\hat{h}_{\tau,t}} - \frac{dw_t}{dt} \tag{5.4.12}$$

Peut- on espéré cette équilibre ?
L'équilibre budgétaire est réaliser si :

$$\frac{b_1(t, I_{\hat{k},t}, \hat{k})}{\hat{k}_t} - \frac{dw_t}{dt} = 0$$

$$\frac{b_2(t, I_{\hat{r},t}, \hat{r})}{\hat{r}_t}dt - \frac{dw_t}{dt} = 0$$

$$\frac{b_3(t, I_{\hat{h},t}, \hat{h})}{\hat{h}_{\tau,t}} - \frac{dw_t}{dt} = 0$$

D'après les équations (5.3.3) et (5.3.10), on a :

$$\frac{I_{\hat{k},t} - (\gamma_A + \gamma_L + \delta)\hat{k}_t}{\hat{k}_t} - \varepsilon_t = 0$$

Donc

$$\frac{I_{\hat{k},t}}{\hat{k}_t} = \varepsilon_t + \gamma_A + \gamma_L + \delta$$

Par suite, l'équilibre budgétaire est rétablit lorsque :

$$E(\frac{I_{\hat{k},t}}{\hat{k}_t}|\Gamma_t) = \gamma_A + \gamma_L + \delta$$

On déduit donc que à l'équilibre, le ratio de l'investissement éfficace par emploi sur le capital éfficace par emploi dépend du taux de croissance technologique, du taux de croissance de l'emploi et du taux de dépréciation du capital.
De façon analogue, on peut trouver que :

$$E(\frac{I_{\hat{r},t}}{\hat{r}_t}|\Gamma_t) = \gamma_{A_t} + \gamma_{L_t} + \phi$$

$$E(\frac{I_{\hat{h},t}}{\hat{h}_{\tau,t}}|\Gamma_t) = \gamma_{A_t} + \gamma_{L_t} + \eta$$

5.5 Principe du maximum stochastique et Résolution du modèle (P)

D'après la deuxième hypothèse macroéconomique, on peut écrire :

$$\begin{aligned}\hat{k}_t^{\alpha}\hat{h}_{\tau,t}^{\beta}\hat{r}_t^{\kappa}L_t^{\gamma-\alpha-\kappa-\beta-1} = \hat{y}_{\tau,t} = \frac{Y_{\tau,t}}{A_t L_t} &= \frac{I_{K,t} + I_{R,t} + I_{H,t}}{s_i A_{i,t} L_{i,t}} \\ &= \frac{I_{\hat{k},t} + I_{\hat{r},t} + I_{\hat{h},t}}{s_i}\end{aligned}$$

On pose :

$$\hat{y}_t = F(t, \hat{k}, \hat{r}, \hat{h}, I_{\hat{k}}, I_{\hat{r}}, I_{\hat{h}}) = \frac{I_{\hat{k},t} + I_{\hat{r},t} + I_{\hat{h},t}}{s}$$

$$X_{\tau,t} = \begin{pmatrix} \hat{k}_t \\ \hat{r}_t \\ \hat{h}_{\tau,t} \end{pmatrix},\ u_t = \begin{pmatrix} I_{\hat{k},t} \\ I_{\hat{r},t} \\ I_{\hat{h},t} \end{pmatrix},\ b(t, X_{\tau,t}, u_t) = \begin{pmatrix} b_1(t, I_{\hat{k},t}, \hat{k}) \\ b_2(t, I_{\hat{r},t}, \hat{r}) \\ b_3(t, I_{\hat{h},t}, \hat{h}) \end{pmatrix}$$

$$W_t = \begin{pmatrix} w_t \\ w_t \\ w_t \end{pmatrix},\ \sigma(t, X_{\tau,t}, u_t) = \begin{pmatrix} -\hat{k}_t & 0 & 0 \\ 0 & -\hat{r}_t & 0 \\ 0 & 0 & -\hat{h}_{\tau,t} \end{pmatrix}$$

Donc le modèle (P) devient :

$$(P) = \begin{cases} \max_{u_t} E\Big[\int_{t_0}^{T} \varrho^t F(t, X_{\tau,t}, u_t)dt|\Gamma_{t_0}\Big] \\ S.C \\ dX_{\tau,t} = b(t, X_{\tau,t}, u_t)dt + \sigma(t, X_{\tau,t}, u_t)dW_t \\ \hat{X}_{\tau,t_0} = \hat{X}_0 \\ \hat{X}_{\tau,T} = c\hat{X}_0 \\ c > 1 \end{cases}$$

Pour un vecteur ou matrice v on désigne par $v^{'}$ la transposée de v et $tr(v)$ la trace de v.
Soient $p_t : [t_0, T] \to \mathbf{IR}^3$ et $q_t : [t_0, T] \to \mathbf{IR}^3$ deux multiplicateurs dynamique.
L'hamiltonien associe au problème (P) est donné par :

$$H(t, X_{\tau,t}, u_t, p_t, q_t) = \varrho^t F(t, X_{\tau,t}, u_t) + b^{'}(t, X_{\tau,t}, u_t)p_t + tr(\sigma^{'}(t, X_{\tau,t}, u_t)q_t)$$

L'équation adjointe satisfaite par les inconnus p_t et q_t

$$dp_t = -\nabla_x H(t, X_{\tau,t}, u_t, p_t, q_t)dt + q_t dW_t \tag{5.5.1}$$

On pose $p_t = (p_{1,t}, p_{2,t}, p_{3,t})^{'}$ et $q_t = (q_{1,t}, q_{2,t}, q_{3,t})^{'}$
L'équation précédente donne :

$$\begin{aligned} dp_{1,t} &= (\gamma_A + \gamma_L + \delta)p_{1,t}dt + q_{1,t}dw_t \quad (1) \\ dp_{2,t} &= (\gamma_A + \gamma_L + \phi)p_{2,t}dt + q_{2,t}dw_t \quad (2) \\ dp_{3,t} &= (\gamma_A + \gamma_L + \eta)p_{3,t}dt + q_{3,t}dw_t \quad (3) \end{aligned}$$

On essaie de chercher une solution de l'équation (1) de la forme :

$$p_{1,t} = \psi_{1,t}\hat{k}_t + \xi_{1,t} \tag{5.5.2}$$

Avec $\psi_{1,t}$ et $\xi_{1,t}$ sont deux fonctions déterministes et différentiables.
Etant donner une fonction f on note par $\dot{f} = \frac{df}{dt}$.
On applique la formule d'Itô à l'équation (5.5.2) :

$$\begin{aligned} dp_{1,t} &= (\dot{\psi}_{1,t}\hat{k}_t + \dot{\xi}_{1,t})dt + \psi_{1,t}d\hat{k}_t \\ &= (\dot{\psi}_{1,t}\hat{k}_t + \dot{\xi}_{1,t})dt + \psi_{1,t}(b_1(t, I_{\hat{k},t}, \hat{k})dt - \hat{k}_t dw_t) \\ &= [(\dot{\psi}_{1,t} - \psi_{1,t}(\gamma_A + \gamma_L + \delta))\hat{k}_t + \psi_{1,t}I_{\hat{k},t} + \dot{\xi}_{1,t}]dt - \psi_{1,t}\hat{k}_t dw_t \end{aligned}$$

On compare ce résultat avec l'équation (1), on obtient donc :

$$\begin{cases} (\gamma_A + \gamma_L + \delta)p_{1,t} = (\dot{\psi}_{1,t} - \psi_{1,t}(\gamma_A + \gamma_L + \delta))\hat{k}_t + \psi_{1,t}I_{\hat{k},t} + \dot{\xi}_{1,t} \\ q_{1,t} = -\psi_{1,t}\hat{k}_t \end{cases} \tag{5.5.3}$$

De façon analogue et par le même raisonnement on trouve :

$$\begin{cases} (\gamma_A + \gamma_L + \phi)p_{2,t} = (\dot{\psi}_{2,t} - \psi_{2,t}(\gamma_A + \gamma_L + \delta))\hat{r}_t + \psi_{2,t}I_{\hat{r},t} + \dot{\xi}_{2,t} \\ q_{2,t} = -\psi_{2,t}\hat{r}_t \end{cases} \tag{5.5.4}$$

$$\begin{cases} (\gamma_A + \gamma_L + \phi)p_{3,t} = (\dot{\psi}_{3,t} - \psi_{3,t}(\gamma_A + \gamma_L + \delta))\hat{h}_{\tau,t} + \psi_{3,t}I_{\hat{h},t} + \dot{\xi}_{3,t} \\ q_{3,t} = -\psi_{3,t}\hat{h}_{\tau,t} \end{cases} \tag{5.5.5}$$

Les conditions nécessaires d'optimalité donnent :

$$E(\nabla_u H(t, X_{\tau,t}, u_t, p_t, q_t)|\Gamma_t) = 0$$

Or

$$\begin{aligned} H(t, X_{\tau,t}, u_t, p_t, q_t) &= \varrho^t F(t, X_{\tau,t}, u_t) + b^{'}(t, X_{\tau,t}, u_t)p_t + tr(q_t^{'}\sigma(t, X_{\tau,t}, u_t)) \\ &= \varrho^t\Big[\frac{I_{\hat{k},t} + I_{\hat{r},t} + I_{\hat{h},t}}{s_i}\Big] + b_1(t, \hat{k}_t, I_{\hat{k},t})p_{1,t} + b_2(t, \hat{r}_t, I_{\hat{r},t})p_{2,t} \\ &+ b_3(t, \hat{h}_{\tau,t}, I_{\hat{h},t})p_{3,t} - \hat{k}_t q_{1,t} - \hat{r}_t q_{2,t} - \hat{h}_t q_{3,t} \end{aligned}$$

$$E(\nabla_u H(t, X_{\tau,t}, u_t, p_t, q_t)|\Gamma_t) = 0 \Rightarrow \begin{cases} E(\frac{\partial}{\partial I_{\hat{k},t}} H(t, X_{\tau,t}, u_t, p_t, q_t)|\Gamma_t) = 0 \\ E(\frac{\partial}{\partial I_{\hat{r},t}} H(t, X_{\tau,t}, u_t, p_t, q_t)|\Gamma_t) = 0 \\ E(\frac{\partial}{\partial I_{\hat{h},t}} H(t, X_{\tau,t}, u_t, p_t, q_t)|\Gamma_t) = 0 \end{cases}$$

Ceci implique

$$\begin{cases} E(\frac{\varrho^t}{s} + p_{1,t}|\Gamma_t) = 0 \Rightarrow E(p_{1,t}|\Gamma_t) = -\frac{\varrho^t}{s} \\ E(\frac{\varrho^t}{s} + p_{2,t}|\Gamma_t) = 0 \Rightarrow E(p_{2,t}|\Gamma_t) = -\frac{\varrho^t}{s} \\ E(\frac{\varrho^t}{s} + p_{3,t}|\Gamma_t) = 0 \Rightarrow E(p_{3,t}|\Gamma_t) = -\frac{\varrho^t}{s} \end{cases}$$

D'autre part en utilisant les équations (5.5.3)-(5.5.5), on obtient :

$$\begin{cases} E(p_{1,t}|\Gamma_t) = \frac{(\dot{\psi}_{1,t} - \psi_{1,t}(\gamma_A+\gamma_L+\delta))}{(\gamma_A+\gamma_L+\delta)} E(\hat{k}_t|\Gamma_t) + \psi_{1,t} I_{\hat{k},t} + \dot{\xi}_{1,t} \\ E(p_{2,t}|\Gamma_t) = \frac{(\dot{\psi}_{2,t} - \psi_{2,t}(\gamma_A+\gamma_L+\phi))}{(\gamma_A+\gamma_L+\phi)} E(\hat{r}_t|\Gamma_t) + \psi_{2,t} I_{\hat{r},t} + \dot{\xi}_{2,t} \\ E(p_{3,t}|\Gamma_t) = \frac{(\dot{\psi}_{3,t} - \psi_{3,t}(\gamma_A+\gamma_L+\eta))}{(\gamma_A+\gamma_L+\eta)} E(\hat{h}_{\tau,t}|\Gamma_t) + \psi_{3,t} I_{\hat{h},t} + \dot{\xi}_{3,t} \end{cases} \quad (5.5.6)$$

On conclut donc que si $u_t^* = (I_{\hat{k},t}^*, I_{\hat{r},t}^*, I_{\hat{h},t}^*)^{'}$ est un condidat pour le contrôle optimal, et $X_{\tau,t}^* = (\hat{k}_t^*, \hat{r}_t^*, \hat{h}_t^*)^{'}$ est la variable état qui corespond à u_t^*, alors :

$$\begin{cases} I_{\hat{k},t}^* = -\frac{\varrho^t}{s\psi_{1,t}} + (1 - \frac{\dot{\psi}_{1,t}}{(\gamma_A+\gamma_L+\delta)\psi_{1,t}}) E(\hat{k}_t^*|\Gamma_t) + \frac{\dot{\xi}_{1,t}}{\psi_{1,t}} \\ I_{\hat{r},t}^* = -\frac{\varrho^t}{s\psi_{2,t}} + (1 - \frac{\dot{\psi}_{2,t}}{(\gamma_A+\gamma_L+\phi)\psi_{2,t}}) E(\hat{r}_t^*|\Gamma_t) + \frac{\dot{\xi}_{2,t}}{\psi_{2,t}} \\ I_{\hat{k},t}^* = -\frac{\varrho^t}{s\psi_{3,t}} + (1 - \frac{\dot{\psi}_{3,t}}{(\gamma_A+\gamma_L+\eta)\psi_{3,t}}) E(\hat{h}_{\tau,t}^*|\Gamma_t) + \frac{\dot{\xi}_{3,t}}{\psi_{3,t}} \end{cases} \quad (5.5.7)$$

D'autre part on utilise l'équation (5.5.2), alors on obtient : $E(p_{1,t}|\Gamma_t) = \psi_{1,t} E(\hat{k}_t|\Gamma_t) + \xi_{1,t}$
On combine cette derniere équation avec l'équation (5.5.7), alors on obtient :

$$\begin{cases} I_{\hat{k},t}^* = (2 - \frac{\dot{\psi}_{1,t}}{(\gamma_A+\gamma_L+\delta)\psi_{1,t}}) E(\hat{k}_t^*|\Gamma_t) + \frac{\xi_{1,t} - \dot{\xi}_{1,t}}{\psi_{1,t}} \\ I_{\hat{r},t}^* = (2 - \frac{\dot{\psi}_{2,t}}{(\gamma_A+\gamma_L+\phi)\psi_{2,t}}) E(\hat{r}_t^*|\Gamma_t) + \frac{\xi_{2,t} - \dot{\xi}_{2,t}}{\psi_{2,t}} \\ I_{\hat{k},t}^* = (2 - \frac{\dot{\psi}_{3,t}}{(\gamma_A+\gamma_L+\eta)\psi_{3,t}}) E(\hat{h}_{\tau,t}^*|\Gamma_t) + \frac{\xi_{3,t} - \dot{\xi}_{3,t}}{\psi_{3,t}} \end{cases} \quad (5.5.8)$$

On soustrait (5.5.7) de (5.5.8) on obtient :

$$\begin{cases} E(\hat{k}_t^*|\Gamma_t) + \frac{\xi_{1,t} - 2\dot{\xi}_{1,t}}{\psi_{1,t}} = 0 \\ E(\hat{r}_t^*|\Gamma_t) + \frac{\xi_{2,t} - 2\dot{\xi}_{2,t}}{\psi_{2,t}} = 0 \\ E(\hat{h}_{\tau,t}^*|\Gamma_t) + \frac{\xi_{3,t} - 2\dot{\xi}_{3,t}}{\psi_{3,t}} = 0 \end{cases}$$

Donc

$$\begin{cases} \psi_{1,t}E(\hat{k}_t^*|\Gamma_t) + \xi_{1,t} - 2\dot{\xi}_{1,t} = 0 \\ \psi_{2,t}E(\hat{r}_t^*|\Gamma_t) + \xi_{2,t} - 2\dot{\xi}_{2,t} = 0 \\ \psi_{3,t}E(\hat{h}_{\tau,t}^*|\Gamma_t) + \xi_{3,t} - 2\dot{\xi}_{3,t} = 0 \end{cases}$$

D'après les équations au dessus, on trouve :

$$\begin{cases} -\frac{\varrho^t}{s} - 2\dot{\xi}_{1,t} = 0 \\ -\frac{\varrho^t}{s} - 2\dot{\xi}_{2,t} = 0 \\ -\frac{\varrho^t}{s} - 2\dot{\xi}_{3,t} = 0 \end{cases}$$

$$\begin{cases} \dot{\xi}_{1,t} = -\frac{\varrho^t}{2s} \\ \dot{\xi}_{2,t} = -\frac{\varrho^t}{2s} \\ \dot{\xi}_{3,t} = -\frac{\varrho^t}{2s} \end{cases}$$

D'où

$$\begin{cases} \xi_{1,t} = \xi_{1,t_0} - \int_{t_0}^t \frac{\varrho^l}{2s} dl \\ \xi_{2,t} = \xi_{2,t_0} - \int_{t_0}^t \frac{\varrho^l}{2s} dl \\ \xi_{3,t} = \xi_{3,t_0} - \int_{t_0}^t \frac{\varrho^l}{2s} dl \end{cases} \tag{5.5.9}$$

on déduit donc qu'on peut choisir le multiplicateur stochastique $p_{1,t}$ (resp $p_{2,t}$ et $p_{3,t}$) de la forme $p_{1,t} = \psi_{1,t}\hat{k}_t + \xi_{1,t}$, telle que $\xi_{1,t} = \xi_{1,t_0} - \int_{t_0}^t \frac{\varrho^l}{2s} dl$ et $\psi_{1,t}$ une fonction déterministe et différentiable. On prend donc $\xi_{1,t_0} = 0$ et $\psi_{1,t} = 0$, et on fait la même chose pour les autres ($p_{2,t}$ et $p_{3,t}$) Alors on obtient :

$$\begin{cases} I_{\hat{k},t}^* = -\frac{3\varrho^t}{2s} + E(\hat{k}_t^*|\Gamma_t) \\ I_{\hat{r},t}^* = -\frac{3\varrho^t}{2s} + E(\hat{r}_t^*|\Gamma_t) \\ I_{\hat{k},t}^* = -\frac{3\varrho^t}{2s} + E(\hat{h}_{\tau,t}^*|\Gamma_t) \end{cases} \tag{5.5.10}$$

Il ne reste que d'étudier la concavité de l'hamiltonien H pour la solution admissible $(X_{\tau,t}^*, u_t^*)$.
On a

$$\begin{cases} p_{1,t} = \hat{k}_t + \xi_{1,t} \\ p_{2,t} = \hat{r}_t + \xi_{2,t} \\ p_{3,t} = \hat{h}_t + \xi_{3,t} \end{cases} \tag{5.5.11}$$

On remplace ces multiplicateurs dynamique dans l'hamiltonien et on utilise les équations (5.3.10)-(5.3.12) :

$$\begin{aligned} H(t, X_{\tau,t}, u_t, p_t, q_t) = \varrho^t\Big[\frac{I_{\hat{k},t} + I_{\hat{r},t} + I_{\hat{h},t}}{s_i}\Big] + [I_{\hat{k},t} - (\gamma_A + \gamma_L + \delta)\hat{k}_t](\hat{k}_t + \xi_{1,t}) \\ + [I_{\hat{r},t} - (\gamma_A + \gamma_L + \phi)\hat{r}_t](\hat{r}_t + \xi_{2,t}) + [I_{\hat{h},t} - (\gamma_A + \gamma_L + \eta)\hat{h}_{\tau,t}](\hat{h}_t + \xi_{3,t}) \\ - \hat{k}_t^2 - \hat{r}_t^2 - \hat{h}_t^2 \end{aligned}$$

On note par $\mathcal{H}$ la matrice hessienne, donc

$$\begin{pmatrix} -2(1 + \gamma_A + \gamma_L + \delta) & 0 & 0 \\ 0 & -2(1 + \gamma_A + \gamma_L + \phi) & 0 \\ 0 & 0 & -2(\gamma_A + \gamma_L + \eta) \end{pmatrix} \tag{5.5.12}$$

La forme quadratique associé à $\mathcal{H}$ est donnée par :

$$Q(x,y,z) = (x\ \ y\ \ z)\mathcal{H}\begin{pmatrix} x \\ y \\ z \end{pmatrix} = -2(1+\gamma_A+\gamma_L+\delta)x^2 - 2(1+\gamma_A+\gamma_L+\phi)y^2$$
$$- 2(\gamma_A+\gamma_L+\eta)z^2 < 0\ (\forall(x,y,z)\neq 0)$$

Donc la matrice $\mathcal{H}$ est définie négative, par suite la fonction hamiltonien H est strictement concave. Conclusion : la solution optimale du problème (P) est donnée par l'équation (5.5.10).

5.6 Condition de rattrapage technologique et simulation du modèle (P)

5.6.1 Condition de rattrapage

On note par $\gamma_H(t)$ le taux de croissance du capital humain.
En utilisant l'équation de rattrapage technologique qui est démontrée par Nelson–Phelps [82] et développé par Benhabib et Spiegel[] dans le cas déterministe. L'équation est donnée par :

$$\frac{\frac{dA_t}{dt}}{A_t} = \gamma_H(t) + \lambda\Big(\frac{B_t}{A_t} - 1)\Big) \tag{5.6.1}$$

Benhabib et Spiegel ont montré que dans le cas où A_t est une fonction logistique déterministe avec des taux de croissances constants $\gamma_H(t) = \gamma_0$, $\gamma_B(t) = g$ alors :

$$\lim_{t\to+\infty}\frac{A_t}{B_t}\begin{cases} \frac{\gamma_0+\lambda-g}{\lambda}, & \text{Si } \gamma_0+\lambda-g>0 \\ \frac{A(0)}{B(0)}, & \text{Si } \gamma_0+\lambda-g=0 \\ 0, & \text{Si } \gamma_0+\lambda-g<0 \end{cases}$$

Donc, d'après Behabib et Spiegel le rattrapage technologique partielle entre un pays en voix de développement et un pays leader est réaliser si et seulement si $\gamma_0+\lambda-g\geq 0$.
Dans le cas où la technologie A_t est stochastique et ou le taux de croissance du capital n'est pas constant, alors le résultat de rattrapage technologique ne sera pas valable. Notre objective dans cette section est de trouver la condition de rattrapage technologique dans le cas stochastique avec un taux de croissance du capital humain variable.
Les équations (5.6.1), (5.3.2) et (5.3.3) donnes :

$$\lambda\frac{B_t}{A_t} = \gamma_A + \varepsilon_t - \gamma_H(t) + \lambda \tag{5.6.2}$$

On dérive la formule au dessus on trouve que

$$\lambda\frac{d}{dt}\Big(\frac{B_t}{A_t}\Big) = \frac{d\varepsilon_t}{dt} - \frac{d\gamma_H(t)}{dt}$$

D'après l'équation (5.3.1) on a :

$$\gamma_H(t) = \frac{\frac{\partial H_{\tau,t}}{\partial t}}{H_{\tau,t}} = \chi_t e^{\tau\theta}\Big(\frac{B_t}{H_{\tau,t}}\Big)^{\lambda}$$

Donc

$$\begin{aligned}\frac{d\gamma_H(t)}{dt} &= \lambda\chi_t e^{\tau\theta}\Big(\frac{B_t}{H_{\tau,t}}\Big)^{\lambda-1}\Big(\frac{\frac{B_t}{dt}H_{\tau,t}-\frac{\partial H_{\tau,t}}{\partial t}B_t}{H_{\tau,t}^2}\Big)\\ &= \lambda\gamma_H(t)(g-\gamma_H(t))\end{aligned}$$

Et

$$\begin{aligned}\frac{d}{dt}\Big(\frac{B_t}{A_t}\Big) &= \frac{\frac{dB_t}{dt}A_t-\frac{dA_t}{dt}B_t}{A_t^2}\\ &= g\frac{B_t}{A_t}-(\gamma_A+\varepsilon_t)\frac{B_t}{A_t}\\ &= (g-\gamma_A-\varepsilon_t)\frac{B_t}{A_t}\end{aligned}$$

Donc

$$\lambda(g-\gamma_A-\varepsilon_t)\frac{B_t}{A_t} = \frac{d\varepsilon_t}{dt}-\lambda\gamma_H(t)(g-\gamma_H(t))$$

Par suite :

$$\frac{A_t}{B_t} = \frac{\lambda(g-\gamma_A-\varepsilon_t)}{\frac{d\varepsilon_t}{dt}-\lambda\gamma_H(t)(g-\gamma_H(t))} \tag{5.6.3}$$

Donc, le seuil de confiance π_α pour qu'on peut avoir un rattrapage technologique partielle entre un pays en voix de développement et un pays leader avec niveau de confiance $1-\alpha$ est donné par l'équation :

$$P\Big(\frac{\lambda(g-\gamma_A-\varepsilon_t)}{\frac{d\varepsilon_t}{dt}-\lambda\gamma_H(t)(g-\gamma_H(t))} > \pi_\alpha\Big) = 1-\alpha \tag{5.6.4}$$

5.6.2 simulation du modèle (P)

Notre objective dans cette section est de trouver les investissements optimaux sur les années prochaînes, à partir des données des trente années précédentes (1988-2018), en utilisant les données du Maroc et les résultats de simulation par la méthode elasticnet brownien qu'on a trouver dans le chapitre 3, on trouve que :
- La prévision du capital éfficace par emploi (est donc l'espérence conditionnelle de $\hat{k}$) est un modèle ARIMA(0,1,0) de coefficient canstant 0.036962.
Donc :

$$E(\hat{k}_t^*|\Gamma_t) = 0.036962(t-2018)+\hat{k}_{2018} \tag{5.6.5}$$

- La prévision des dépences en recherche et développement éfficace par emploi (est donc l'espérence conditionnelle de $\hat{r}$) est un modèle de lissage expenontiel double de coefficient $\alpha = 0,999005$ et de coefficient de tendance $\beta = 5,648\times10^{-6}$. On voit donc que β est négligeable et α proche de 1. On considère que $\beta = 0$ et $\alpha = 1$, alors la prévision de dépenses en recherche et développement éfficace par emploi est donnée par :

$$E(\hat{r}_t^*|\Gamma_t) = \hat{r}_{2018}+(t-2018)(\hat{r}_{2018}-\hat{r}_{2017}) \tag{5.6.6}$$

- La prévision du capital humain éfficace par emploi (est donc l'espérence conditionnelle de $\hat{k}$) est un modèle de lissage expenontiel double de coefficient $\alpha = 0,299$ et de coefficient de tendance 1. Donc la prévision du capital humain éfficace par emploi est donnée par :

$$E(\hat{h}_t^*|\Gamma_t) = s_{2018}+(t-2018)(s_{2018}-s_{2017}) \tag{5.6.7}$$

Avec Où : $s_{1989} = h_{1989}$
$\tau_{1989} = r_{1989} - r_{1988}$
et pour $t > 1989$:

$$\begin{aligned} s_t &= \kappa r_t + (1-\kappa)(s_{t-1} + \tau_{t-1}) \\ \tau_t &= \pi(s_t + s_{t-1}) + (1-\pi)\tau_{t-1} \end{aligned}$$

En utilisant l'équation (5.5.10), alors on obtient :

$$\begin{cases} I^*_{\hat{k},t-2018} = -\frac{3\varrho^{t-2018}}{2\nu} + 0.036962(t-2018) + \hat{k}_{2018} \\ I^*_{\hat{r},t-2018} = -\frac{3\varrho^{t-2018}}{2\nu} + \hat{r}_{2018} + (t-2018)(\hat{r}_{2018} - \hat{r}_{2017}) \\ I^*_{\hat{k},t-2018} = -\frac{3\varrho^{t-2018}}{2\nu} + s_{2018} + (t-2018)(s_{2018} - s_{2017}) \end{cases} \tag{5.6.8}$$

Avec ν est le taux d'investissement total fixé par le gouvernement.
Pour un taux d'investissement total $\nu = 0,25$ les graphes des investissements optimaux sont donnés par les figures suivants :

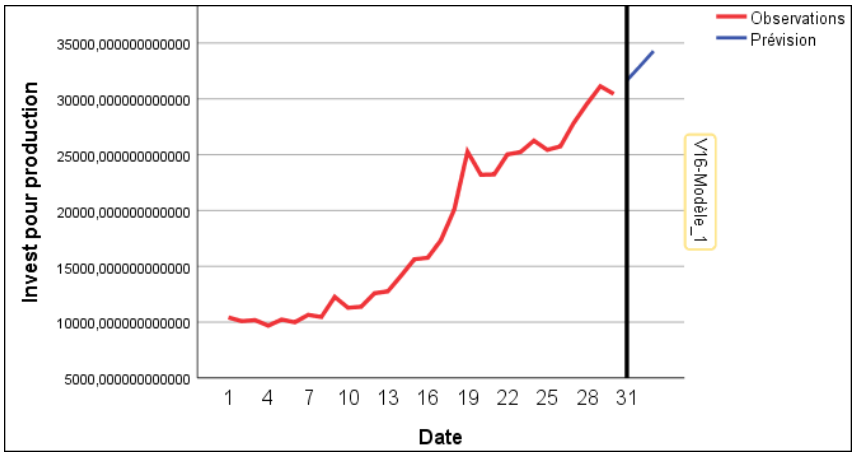


FIGURE 5.4 – Investissement optimal pour la production

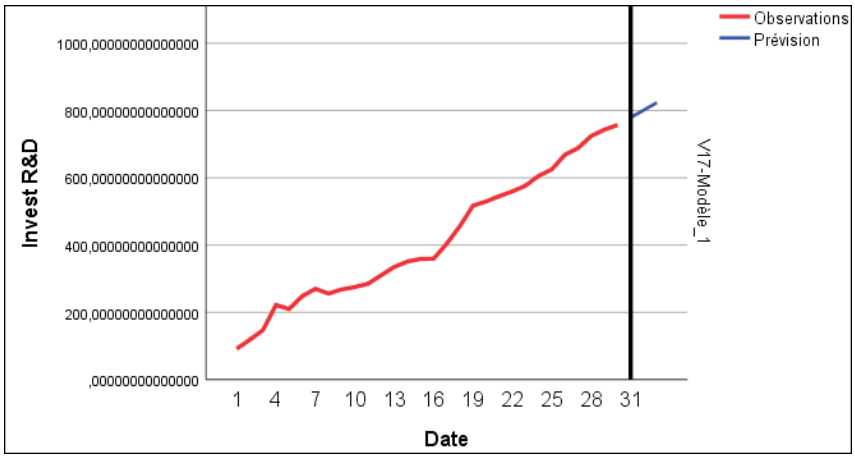


FIGURE 5.5 – Investissement optimal pour la recherche et développement

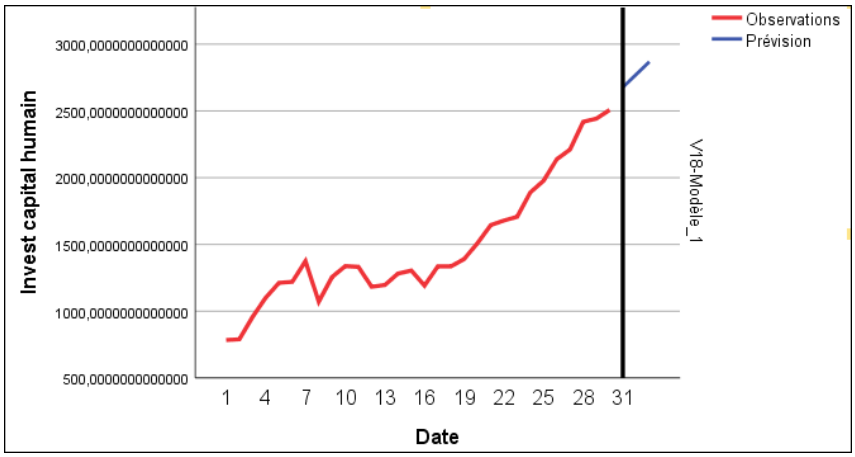


FIGURE 5.6 – Investissement optimal pour le capital humain

5.7 Conclusion

Dans ce chapitre, nous avons abordé et résolu un modèle d'optimisation stochastique qui sert à maximiser l'espérence du PIB sous contraites budgétaires, en considérant que la technologie est un mouvement brownien géométrique. Par la suite, nous avons trouvé la condition pour avoir un rattrapage technologique partiel entre un pays en voie de développement et un pays avancé. Nous avons terminé ce chapitre par une simulation des résultats théoriques du modèle, en se basant sur les données du Maroc entre les années (1988-2018).

Chapitre 6

Contrôle optimal de la production économique sous condition de l'équilibre écologique

6.1 Introduction

L'objectif de toute économie est de maximiser son PIB et son RNB, autrement dit, faire une gestion optimale des investissements afin d'obtenir un niveau optimal de PIB par habitant. En revanche, une économie plus poductive est aussi plus polluante et a un impact négatif sur l'environnement et le réchauffement climatique, en effet il y a une relation forte entre l'émission de CO2, et la croissance économique.
Dans la littérature, de nombreux scientifiques ont pris en compte les facteurs environnementaux, nous citons ; David Cohen [4] qui fait valoir la relation étroite entre la croissance économique et les émissions de carbone. Des modèles similaires de croissance économique avec facteurs environnementaux ont été construits, on peut citer ; S.And, V. Capasso, H. Kunze et al [1], Zuo, H. et Ai, D. [5], Cherniwchan, J. [6], Shieh, J-Y and Chen, J-H and Chang, S-H et Lai, C-C. (Shieh, [3]).
Dans ce chapitre, nous construisons deux modèles dynamiques qui servent à maximiser le PIB par tête, en tenant compte de l'équilibre écologique. Nous considérerons que la production est divisée en deux ; production de base et production auxiliaire détruisant la pollution générée par la production de base.

6.2 Modélisation macro-dynamique et équilibre écologique

Notre objectif après formalisation du problème, est de maximiser le PIB sur une période du temps tout en essayant d'avoir un niveau minimal de pollution assurant l'équilibre écologique. Ainsi, rechercher une estimation optimale des intrants (capitaux) afin de maximiser la production par habitant(PIB/tête).

6.2.1 Notations et hyphothèses

1- Notations

Dans ce chapitre, nous notons par :
t : Le temps.
K_b : Le capital utilisé dans la production de base.
K_a : Le capital utilisé dans la production auxiliaire.
H : Le capital humain.
L : Le travail total.
A : le progrès technique de l'état.
Y : Le PIB de l'état.
Y_b : La production de base globale.
Y_a : La production auxiliaire globale.
S l'épargne de l'état et I son investissement.
$\pi \in \mathbf{C}^0(\mathbf{IR}, \mathbf{IR}^+)$: Est la quantité de pollution générée par la production de base.
$\mu \in \mathbf{C}^0(\mathbf{IR}, \mathbf{IR}^+)$: La quantité de pollution éliminée par la production auxiliaire.
$\rho \in \mathbf{C}^0(\mathbf{IR}, \mathbf{IR}^+)$: Est la quantité de pollution auto-épurante.
Avant de proposer un modèle pour atteindre notre objectif, commençons maintenant par une modélisation à priori de l'évolution économique, en tenant compte de la pollution générée par le secteur de la production.
La création d'une production auxiliaire détruisant la pollution industrielle, nécessite de réserver une partie de l'investissement qui permet la réalisation de cette production. Il s'agit donc d'un investissement tri-sectoriel ; investissement pour la production de base (noté I_b), investissement pour la production auxiliaire (notée I_a), et investissement pour le capital humain(noté I_h). Par conséquent, nous avons trois équations dynamiques d'accumulation de capital.
Pour une modélisation cohérente avec les faits stylisés de la croissance économique et de la théorie moderne de l'endogénéisation de la technologie, nous prenons en compte la technologie en tant que facteur endogène logistique de la production (voir [9-12]). Ainsi, selon les travaux de Novales, A. et de tous [2], Zuo, H. et Ai, [7], Cohen [4] et V. GRYGORKIV et tous [1], il existe une forte corrélation entre la production de base et l'augmentation de la pollution.
Pour cette raison, nous proposons les hypothèses suivantes :

2- Hypothèses macroéconomiques du modèle

(H1) L'investissement est tri-sectoriel et sur le marché des capitaux, nous avons :

$$S = I = I_a + I_b + I_h \tag{6.2.1}$$

(H2) Les capitaux sont dépréciés à taux constants ; δ_b, δ_a et δ_h :

$$\frac{dK_b(t)}{dt} = I_b - \delta_b K_b \tag{6.2.2}$$

$$\frac{dK_a(t)}{dt} = I_a - \delta_a K_a \tag{6.2.3}$$

$$\frac{dH(t)}{dt} = I_h - \delta_h H \tag{6.2.4}$$

(H3) Les fonctions π avec la production de base Y_b et μ avec le capital auxiliaire K_a sont linéairement corrélés.
(H4) La technologie et l'emploi sont endogènes et suivent le modèle logistique.
(H5) Inputs (K_b, K_a, H) et L sont dans $\mathbf{C}^1(\mathbf{IR}) \bigcap \mathbf{L}^\infty(\mathbf{IR})$.
(H6) La production de base est Cobb Douglas.
(H7) Le travail total est composé en deux ; travailler dans le secteur de la production, noté $L_p = x_p L$ et le travail dans la production auxiliaire, notée $L_a = (1 - x_p)L$. Où x_p and $1 - x_p$ sont les taux d'emploi dans chaque secteur.

6.2.2 Modèle d'évolution avec équilibre écologique

La stratégie de détérioration de la pollution industrielle nous donne deux types de production ; une base Y_b et l'autre auxiliaire (Y_a). Le revenu total (PIB) sera donc :

$$Y(t) = Y_b(t) + Y_a(t) \tag{6.2.5}$$

Selon l'hypothèse (H6), Y_b est Cobb douglas. Elle définie par :

$$Y_b(t) = F(t, K_b, H, A, L_p) = (K_b)^\alpha (H(t))^\beta (A(t)L_p(t))^{1-\alpha-\beta} \tag{6.2.6}$$

Où α, β sont respectivement les élasticités de ; capital utilisé dans la production de base et le capital humain.
Selon plusieurs travaux (see [10-13]), le progrès technique est hétérogène entre les états du monde, mais il est homogène au sein de chaque état mais avec des élasticités différentes. Nous formalisons la production auxiliaire par une fonction de Cobb Douglass, définie comme suit :

$$Y_a(t) = F(t, K_a, A, L_a) = (K_a(t))^\gamma (A(t)L_a(t))^{1-\gamma} \tag{6.2.7}$$

Où γ est l'élasticité du capital utilisé dans la production auxiliaire.
Par contre, le volume de pollution sur une période donnée $[0, t]$, est une accumulation précédente de pollution. Où la pollution en un instant t, est donnée par une différence entre la pollution industrielle et la quantité de déchets détruits par la nature elle-même (pollution auto-épurante), que :

$$dP(t) = \pi(Y_b(t)) - \rho(P(t))$$

Lorsque nous intégrons une stratégie d'investissement dans l'environnement, la pollution sera réduite à néant par la quantité de pollution industrielle éliminée par le recyclage. $(\mu(K_a))$. Le volume de pollution généré sur une période de temps $[0, t]$, est donné par :

$$P(t) = \int_0^t \pi(Y_b(s)) - \rho(P(s)) - \mu(K_a(s))ds \tag{6.2.8}$$

D'après ce qui précède, et selon les hypothèses ci-dessus, l'évolution économique en tenant compte d'une pollution industrielle, peut se formaliser selon le modèle suivant :

$$(M1) \begin{cases} Y_t = Y(t) = Y_b(t) + Y_a(t) \\ \frac{dK_b(t)}{dt} = I_b - \delta_b K_b \\ \frac{dK_a(t)}{dt} = I_a - \delta_a K_a \\ \frac{dH(t)}{dt} = I_h - \delta_h H \\ P(t) = \int_0^t \pi(Y_b(s)) - \rho(P(s)) - \mu(K_a(s))ds \\ (\delta_b, \delta_a, \delta_h, \alpha, \beta, \gamma) \in]0, 1[^6 \end{cases}$$

Selon (II.5-II.7) et l'hypothèse (H7), la production effective par habitant est donnée par :

$$\tilde{y}(t) = \frac{Y(t)}{A(t)L(t)} = x_p^{1-\alpha-\beta}(\tilde{k}_b)^{\alpha}(\tilde{h}(t))^{\beta} + (1-x_p)^{1-\gamma}(\tilde{k}_a)^{\gamma} \tag{6.2.9}$$

Selon l'hypothèse (H1), depuis l'investissement et les trois secteurs, nous avons trois taux d'investissement :
s_b : le taux d'investissement dans la production de base.
s_a : le taux d'investissement dans la production auxiliaire.
s_h : le taux d'investissement dans le capital humain.
Donc nous avons :

$$I_b = s_b Y \tag{6.2.10}$$

$$I_a = s_a Y \tag{6.2.11}$$

$$I_h = s_h Y \tag{6.2.12}$$

D'autre part, l'équation dynamique d'accumulation de capital effectif par habitant est donnée par :

$$\frac{d\tilde{k}_b(t)}{dt} = \frac{d}{dt}\Big(\frac{K_b(t)}{A(t)L(t)}\Big) = \frac{\frac{dK_b(t)}{dt}}{A(t)L(t)} - \frac{\frac{dL(t)}{dt}}{L(t)}\frac{K_b(t)}{A(t)L(t)} - \frac{\frac{dA(t)}{dt}}{A(t)}\frac{K_b(t)}{A(t)L(t)} - \delta\frac{K_b(t)}{A(t)L(t)}$$

On pose

$$n(t) = \frac{\frac{dL(t)}{dt}}{L(t)} \ (taux\ de\ croissance\ de\ l'emploi) \tag{6.2.13}$$

$$g(t) = \frac{\frac{dA(t)}{dt}}{A(t)} \ (taux\ de\ croissance\ technologique) \tag{6.2.14}$$

Alors on obtient :

$$\frac{d\tilde{k}_b(t)}{dt} = s_b\tilde{y}(t) - (\delta_b + n(t) + g(t))\tilde{k}_b(t) \tag{6.2.15}$$

Par la même façon, on obtient :

$$\frac{d\tilde{k}_a(t)}{dt} = s_b\tilde{y}(t) - (\delta_a + n(t) + g(t))\tilde{k}_a(t) \tag{6.2.16}$$

$$\frac{d\tilde{h}(t)}{dt} = s_h\tilde{y}(t) - (\delta_h + n(t) + g(t))\tilde{h}(t) \tag{6.2.17}$$

6.3 Modèle macro-dynamique optimale et équilibre écologique

Le modèle dynamique de croissance économique optimale par habitant avec condition d'équilibre écologique, sera :

$$(M2)\begin{cases} \max_{\tilde{k}_b,\tilde{k}_a,\tilde{h}} \int_0^T R^t\tilde{y}(t)dt = \int_0^T R^t x_p^{1-\alpha-\beta}(\tilde{k}_b)^\alpha(\tilde{h}(t))^\beta + (1-x_p)^{1-\gamma}(\tilde{k}_a)^\gamma dt \\ \min_{t\in[0,T]} P(t) \\ \frac{d\tilde{k}_b(t)}{dt} = s_b\tilde{y}(t) - (\delta_b + n(t) + g(t))\tilde{k}_b(t) \\ \frac{d\tilde{k}_a(t)}{dt} = s_a\tilde{y}(t) - (\delta_a + n(t) + g(t))\tilde{k}_a(t) \\ \frac{d\tilde{h}(t)}{dt} = s_h\tilde{y}(t) - (\delta_h + n(t) + g(t))\tilde{h}(t) \\ (\tilde{h}(0), \tilde{k}_b(0), \tilde{k}_a(0)) = (\tilde{h}^0, \tilde{k}_b^0, \tilde{k}_a^0) \\ (\delta_b, \delta_a, \delta_h, \alpha, \beta, \gamma, s_b, s_a, s_h) \in]0,1[^9 \end{cases}$$

Soient μ_e, K_{ea}, π_e and Y_{eb} des séries statistique temporelles sur la période [-T, 0] pour les fonctions μ, K_a, π and Y_b.
On propose le modèle suivant :

$$(M3)\begin{cases} \max_{\tilde{k}_a} \mathbf{U}_T(\tilde{k}_a) = \int_0^T R^t\Bigg(\frac{c_1}{A(t)L(t)} + c_2\tilde{k}_a(t) \\ +(1-x_p)^{1-\gamma}(\tilde{k}_a)^\gamma(t)\Bigg)dt \\ \qquad U.C \\ c_1 = \frac{\varepsilon_0 - \bar{\pi}_e + \bar{\mu}_e - \frac{cov(\mu_e,K_{ea})}{Var(K_{ea})}\bar{K}_{ea} + \frac{cov(\pi_e,Y_{eb})}{Var(Y_{eb})}\bar{Y}_{eb}}{\frac{cov(\pi_e,Y_{eb})}{Var(Y_{eb})}} \\ c_2 = \frac{\frac{cov(\pi_e,Y_{eb})}{Var(Y_{eb})}}{\frac{cov(\mu_e,K_{ea})}{Var(K_{ea})}} \\ \frac{d\tilde{k}_a(t)}{dt} = s_a\tilde{y}_a(t) - (\delta_a + n(t) + g(t))\tilde{k}_a(t) \\ \tilde{k}_a(0)) = \tilde{k}_a^0 \\ \tilde{k}_a(T) = \tilde{k}_a^T \\ \varepsilon_0 = \rho(P_{min}(t)) \\ (\delta_a, \gamma, s_a) \in]0,1[^3 \end{cases}$$

Théorème 6.3.1 *Soit $(\tilde{k}_b, \tilde{k}_a, \tilde{h})$ un vecteur des capitaux.*
On suppose que les hypothèses (H1)-(H7) sont satisfaites alors on a l'équivalance entre les deux propriétés :
(p1)-$(\tilde{k}_b, \tilde{k}_a, \tilde{h})$ est une solution de (M2).
(p2)- $\tilde{k}_a^*$ est une solution de (M3) et

$$(1-\alpha-\beta)\ln(x_p) + \alpha\ln(\tilde{k}_b) + \beta\ln(\tilde{h}) = \ln\left(\frac{c_1}{A(t)L(t)} + c_2\tilde{k}_a\right)$$

Preuve 6.3.1 *D'après les hypothèses (H3), on a :*

$$\pi_(Y_b(t)) \approx \frac{cov(\pi_e, Y_{eb})}{Var(Y_{eb})}Y_b(t) + \bar{\pi}_e\bar{Y}_{eb} - \frac{cov(\pi_e, Y_{eb})}{Var(Y_{eb})}\bar{Y}_{eb} \tag{6.3.1}$$

et

$$\mu(K_a(t)) \approx \frac{cov(\mu_e, K_{ea})}{Var(K_{ea})}K_a(t) + \bar{\mu}_e - \frac{cov(\mu_e, K_{ea})}{Var(K_{ea})}K_a(t)\bar{K}_{ea} \tag{6.3.2}$$

D'autre part, pour achever un niveau minimal de pollution, il est nécessaire que :

$$\frac{dP(t)}{dt} = 0$$

Cela implique :

$$\pi(Y_b(t)) - \rho(P_{min}(t)) - \mu(K_a(t)) = 0 \Rightarrow \pi(Y_b(t)) - \mu(K_a(t)) = \rho(P_{min}(t)) = constant \tag{6.3.3}$$

On pose : $\rho(P_{min}(t)) = \varepsilon_0$
Alors, on obtient :

$$Y_b(t) = \frac{\varepsilon_0 - \bar{\pi}_e + \bar{\mu}_e - \frac{cov(\mu_e, K_{ea})}{Var(K_{ea})}\bar{K}_{ea} + \frac{cov(\pi_e, Y_{eb})}{Var(Y_{eb})}\bar{Y}_{eb}}{\frac{cov(\pi_e, Y_{eb})}{Var(Y_{eb})}} + \frac{\frac{cov(\pi_e, Y_{eb})}{Var(Y_{eb})}}{\frac{cov(\mu_e, K_{ea})}{Var(K_{ea})}} K_a(t) \tag{6.3.4}$$

Par suite

$$Y_b(t) = c_1 + c_2 K_a(t) \tag{6.3.5}$$

On utilise les agrégats par tête :

$$\hat{y}_b(t) = \frac{c_1}{A(t)L(t)} + c_2\hat{k}_a(t) \tag{6.3.6}$$

alors :

$$\ln(\hat{y}_b(t)) = (1-\alpha-\beta)\ln(x_p) + \alpha\ln(\tilde{k}_b) + \beta\ln(\tilde{h}) = \ln\left(\frac{c_1}{A(t)L(t)} + c_2\hat{k}_a(t)\right) \tag{6.3.7}$$

Donc (p1)⇒ (p2)
Réciproquement, si (p2) est vraie. D'après (24), on a :

$$\hat{y}_b(t) = x_p^{1-\alpha-\beta}\tilde{k}_b^{\alpha}\tilde{h}^{\beta} = \frac{c_1}{A(t)L(t)} + c_2\hat{k}_a(t) \tag{6.3.8}$$

Donc

$$\hat{y}(t) = x_p^{1-\alpha-\beta}\tilde{k}_b^{\alpha}\tilde{h}^{\beta} + (1-x_p)^{1-\gamma}(\tilde{k}_a)^{\gamma}dt \quad = \frac{c_1}{A(t)L(t)} + c_2\hat{k}_a(t) + x_a^{1-\gamma}(\tilde{k}_a)^{\gamma}dt \tag{6.3.9}$$

Puisque la constante c_1 est prise de telle manière que l'équilibre écologique est maintenu, alors la contrainte $\min_{t\in[0,T]} P(t)$ est vérifiée.
Puisque les hypothèses (H1-H7) sont vérifiées, les équations d'accumulation du capital (15-17) sont vérifiées. Donc (p2)⇒ (p1)

Théorème 6.3.2 *L'estimation de la solution du modèle (M3) est donnée par :*

$$\tilde{k}_a^*(t) = \Big[\frac{c_2}{(1-\gamma)\delta_a}\Big]^{\frac{1}{\gamma}} \tag{6.3.10}$$

Où

$$c_2 = \frac{\frac{cov(\pi_e, Y_{eb})}{Var(Y_{eb})}}{\frac{cov(\mu_e, K_{ea})}{Var(K_{ea})}}$$

Preuve 6.3.2 *On pose :*

$$u(t,\tilde{k}_a) = \frac{c_1}{A(t)L(t)} + c_2\tilde{k}_a(t) + x_a^{1-\gamma}(\tilde{k}_a)^{\gamma}(t) \tag{6.3.11}$$

Et

$$f(t,\tilde{k}_a) = \frac{d\tilde{k}_a(t)}{dt} = s_a u(t,\tilde{k}_a) - (\delta_a + n(t) + g(t))\tilde{k}_a(t) \tag{6.3.12}$$

On a :

$$\frac{\partial}{\partial\tilde{k}} f(t,\tilde{k}_a) = s_a c_2 + s_a\gamma x_a^{1-\gamma}\tilde{k}_a^{\gamma-1}(t) - \delta_a$$

D'après les hypothèses (H7) :

$$\tilde{k}_a \in \mathbf{C}^1(\mathbf{IR}) \bigcap \mathbf{L}^{\infty}(\mathbf{IR})$$

Alors

$$(\exists l \in \mathbf{IR}^+) : \ \tilde{k}_a(t) \leq l$$

Donc

$$\frac{\partial}{\partial\tilde{k}_a} f(t,\tilde{k}_a) < +\infty \tag{6.3.13}$$

Donc f est lipschitzian en $\tilde{k}_a$.
Alors, D'après (H 5), f vérifie les conditions de Cauchy.
Donc, l'équation $f(t,\tilde{k}_a) = \frac{d\tilde{k}_a(t)}{dt}$, admet une unique solution. En combinant maintenant, les contraintes de la fonctions objective, et utilisant les conditions d'Euler, nous obtenons le résultat du théorème.

On pose :

$$c_3 = 2s_a s_h \delta\Big[\frac{c_2}{(1-\gamma)\delta_a}\Big]^{\frac{1}{\gamma}} \tag{6.3.14}$$

Et

$$c_4 = s_a s_h \tilde{k}_b^0 + s_a s_k \tilde{h}^0 - \frac{c_3}{\delta_a} \tag{6.3.15}$$

On suppose que les dépreciations des capitaux sont égales, i.e. $\delta_a = \delta_b = \delta_h = \delta$.
Sous les hypothèses (H1-H7), d'après les théorèmes 1 et 2 nous avons le lemme suivant :

Lemme 6.3.3 *L'estimation de la solution du modèle (M2) est donnée par le triplet $(\tilde{k}_b^*, \tilde{k}_a^*, \tilde{h}^*)$, qui vérifie :*
$\forall t \in [0,T]$

$$\tilde{k}_a^*(t) = \Big[\frac{c_2}{(1-\gamma)\delta_a}\Big]^{\frac{1}{\gamma}} \tag{6.3.16}$$

$$(1-\alpha-\beta)\ln(x_p) + \alpha\ln(\tilde{k}_b^*) + \beta\ln(\tilde{h}^*) = \ln(\frac{c_1}{A(t)L(t)}) + c_2\Big[\frac{c_2}{(1-\gamma)\delta_a}\Big]^{\frac{1}{\gamma}}) \tag{6.3.17}$$

6.4 Conclusion

Dans ce chapitre, nous avons présenté un modèle de la croissance économique, sous condition de l'équilibre environnementale. nous avons analysé l'impact de la production sur l'augementation du volume de pollution, chose qui nous a poussé à proroser un modèle d'optimisation dynamique qui sert à maximiser la production, et en même temps à minimiser la pollution. Nous somme arrivés à démontrer que pour avoir une croissance optimale, en tenant compte de l'équilibre environnemental, il faut investir dans le secteur de l'environnement, on utilise un capital auxilière fixe annuel, le long de la période $[0, T]$.

Conclusion générale

Dans cette thèse nous avons traité les points suivants :

- ♦ Présenter et critiquer les importants modèles déterministes et stochastiques existants dans la littérature qui modélisent la croissance économique ;
- ♦ Développer les méthodes des moindres carrés pour qu'elles soient bien adapter à l'existence du mouvement brownien dans le modèle qu'on veut étudier ;
- ♦ Développer les méthodes d'apprentissage statistique ridge et elasticnet pour qu'elles soient bien adaptées à l'existence du mouvement brownien dans le modèle. puis nous avons donné deux algorithmes d'apprentissage statistique et nous avons terminé l'étude par une simulation et application économétrique ;
- ♦ Généraliser les méthodes qu'on a développées lorsqu'on a plusieurs individus ainsi qu'un mouvement brownien multidimentionnel ;
- ♦ Proposé un modèle stochastique de la croissance économique sous condition de l'équilibre, en utilisant les estimations trouvées par les méthodes qu'on a développer ;
- ♦ Proposé un modèle stochastique de la croissance économique sous condition du rattrappage technologique, en utilisant les estimations trouvées par les méthodes qu'on a développer ;

La thèse à été composée en six chapitres.
Dans le premier chapitre, nous avons présenté une revue de l'évolution de la théorie de la croissance macroéconomique via des modèles mathématiques dynamiques, déterministes et stochastiques. Nous avons commencé par une étude brève et concentrée des propriétés de la fonction de production macroéconomique, et montré la relation entre la géométrie de la surface de production et la nature des inputs macroéconomiques (substituables ou bien complémentaires). Nous avons établi aussi selon l'ordre chronologique, une étude critique des importants modèles de la littérature, en commançant par les modèles théoriques simples, puis les modèles basés sur les méthodes paramétriques et non paramétriques, et nous avons terminé ce chapitre par des modèles stochastiques de la dynamiques macroéconomiques, tels que les modèles VAR et DSGE. Nous déduisons d'après notre étude, que les modèles théoriques simples, peuvent nous introduire l'idée de croissance macroéconomique, mais ils sont incapables de faire des prédiction de cette croissance, non plus, d'expliquer l'impact des divers chocs sur l'économie. En revanche, les modèles basés sur les méthodes paramétriques prennent en compte les inefficiences et les aléas stochastiques exogènes (chocs), mais, ils ne sont pas toujours éfficaces dans la décomposition des différentes composantes de l'inefficiences, surtout dans le cas d'une technologie muli- produits. Par ailleurs, les modèles basés sur les méthodes non paramétriques, ont une décomposition facile des inefficiences et dans le cadre de ces modèles, nous n'avons pas besoin de donner une forme fonctionnelle de la technologie, l'utilité de ces modèles apparaît lorsqu'on veut

trouver une gestion optimale des inputs, pour optimiser les gains sur une courte période (mois, trimèstre ou année), mais, ils sont incapables de donner une gestion optimale sur une longue période, à cause du manque des propriétés statistiques des inefficiences et la non prise en compte de l'influence des chocs exogènes sur la production. En dynamique stochastique, nous avons présenté deux modèles importants, qui sont basés sur le principe de maximum de vraisemblance ; le modèle VAR et le modèle DSGE. Nous avons conclu que les modèles DSGE donnent de meilleur prédictions que les modèles VAR, lorsqu'il s'agit d'une grande taille des frictions informationnelles.
Dans le deuxième chapitre, nous somme arrivé à trouver les estimateurs optimaux, dans le cas d'existence d'un processus brownien dans le modèle.
Nous avons commencé par l'élimination de l'effet d'hétéroscédasticité des bruits qui proviennent d'un mouvement brownien, en multipliant le modèle par la l'inverse de la racine carré de la matrice de covariance du processus brownien sur une gamme du temps fini. Puis nous avons modifié le test d'autocorélation des erreurs pour qu'il soit convenable avec le fait d'existence du mouvement brownien. Aussi nous nous somme basé sur les méthodes de moindres carré, afin d'établir les nouveaux estimateurs que nous avons proposé et qu'ils tiennent en compte le fait d'existence d'un processus brownien dans le modèle.
Comme application économétrique de ce qu'on a fait, nous avons établi une nouvelle version d'une fonction Cobb Dauglass avec mouvement brownien géométrique, tout en estimant les élasticités des agrégats influençant l'évolution du PIB.
Dans le troisième chapitre nous avons développé les méthodes de la régularisation pour qu'elles tiennent en compte l'existence d'un mouvement brownien dans le modèle. L'utilisation de ces méthodes a un rôle très important lorsque les variables explicatives sont multicolinèaires et les résultats obtenus par les moindres carrées et les moindres carrées browniennes ne sont pas significatives. Les méthodes qu'on a abordées dans ce chapitre fournissent de bons estimateurs dans le cas d'existence d'une multicolinéarité entre les variables explicatives. Ainsi on a développer le test ARCH qui detecte l'existence d'une hétéroscédasticité dans le modèle, puis on a construit une procédure d'estimation des paramètres du module dans le cas d'existence d'une hétéroscédasticité supplémentaire.
Ce chapitre a été terminé par une application économétrique, en estimant les paramètres d'une fonctions Cobb Dauglass stochastiques, en utilisant les données publiées sur le site de la banque mondial concernant le Maroc.
Dans le quatrième chapitres on a étudié des modèles contenants plusieurs individus (variables endogènes) et plusieurs variables explicatives, avec un mouvement brownien multidimentionnel. En premier lieu, on a commencé par étudier l'hypothèse d'égalité de la variance entre les individus. puis on s'est intéréssé à donner des estimateurs optimaux des modèles en panel, en généralisant les estimateurs qu'on a abordé dans les chapitres 2 et 3. On a terminé ce chapitre, en abordant le test des données en panel, pour trouver la catégorie des modèles modélisant l'ensembles des individus en fonction des variables explicatives.

Dans le cinquième chapitre, nous avons abordé un modèle de croissance hétérogène dans lequel la technologie est un mouvement brownien géométrique. Par la suite, nous avons établi que la croissance durable dépend de la qualification de la main-d'œuvre, c'est-à-dire de l'apprentissage des nouvelles technologies et du taux de diffusion technologique du pays.
On a terminé la thèse par le sixième chapitre, nous avons présenté un modèle de la croissance économique, sous condition de l'équilibre environnementale. Nous avons analysé l'impact de la production sur l'augmentation du volume de pollution, chose qui nous a

pousse à proroser un modèle d'optimisation dynamique servant à maximiser la production, et en même temps à minimiser la pollution. Nous avons arrivé à démontrer que pour avoir une croissance optimale, en tenant compte de l'équilibre environnementale, il faut investir dans le secteur de l'environnement, on utilise un capital auxilière fixe annuel, le long de la période $[0, T]$.

Perspectives

Comme tous les thèmes de recherche, il y a encore plusieurs pistes de recherches à développer.
Nos travaux de recherches nous donnent l'idée de développer dans les pistes suivantes :

1. Trouver un lien entre les méthodes à $\mathcal{M}.\mathcal{B}$ avec la statistique baysienne et la régression logistique.
2. Développer encore les modèles des données en panel dans le cas d'existence d'un $\mathcal{M}.\mathcal{B}$.
3. Exploiter et développer nos méthodes que nous venons de présenter en servant les domines de Data science : machine learning, intelligence artificielle,... .

Bibliographie

[1] Pierre-André Cornillon et Éric Matzner-Løber, Régression, Paris, Springer, 2007 (ISBN 978-2-287-39692-2).

[2] Sabine Van Huffel et Joos Vandewalle (préf. Gene H. Golub), The Total Least Squares Problem : Computational Aspects and Analysis, SIAM, coll. « Frontiers in Applied Mathematics » (no 9), 1991, xiii + 300 p. (ISBN 978-0-898712-75-9, Math Reviews 1118607,

[3] BOX G.E.P., JENKINS G.M., Time series analysis : forecasting and control, Holdenday, 1976.

[4] BOX G.E.P., PIERCE D.A., « Distribution of residual autocorrelations in autoregressiveintegrated moving average time series models », Journal of the American Statistical Association, Vol. 65, 1970.

[5] BREUSCH T., « Testing for autocorrelation in dynamic linear models », Australian Economic Papers, Vol. 17, 1978. FARRAR D.E., GLAUBER R.R., « Multicolinearity in regression analysis », Review of Economics and Statistics, Vol. 49, 1967.

[6] FULLER W.A., Introduction to statistical time series, John Wiley, 1976.

[7] GLEISJER H., « A new test for heteroscedasticity », Journal of American Statistical Association, Vol. 64, 1969.

[8] GODFREY L.G., « Testing for higher order serial correlation in regression equation when the regressors contain lagged dependant variables », Econometrica,Vol. 46, 1978.

[9] Trevor P., George C., "The Bayesian Lasso", Journal of the American Statistical Association, pages 681-686 https ://doi.org/10.1198/016214508000000337.

[10] Isabelle Couzon, Les espaces économiques de François Perroux (1950), Revue d'Histoire des Sciences Humaines, 2003/2 n° 9 | pages 81 à 102, ISBN 2859398171).

[11] A.Jamali Alaoui, M. Elkhomssi, B. Elgoumi, Stochastic Model of Economic Growth with Heterogeneous Technology and Technological UpgradingSCAMS (2017), Springer, LNNS 37, pp. 1006–1014, 2018. *https* : *//doi.org*/10.1007/978−3−319−74500−$8_9$0

[12] A.Jamali Alaoui, M. Elkhomssi, B. Elgoumi, Optimal control model of economic production under condition of ecological balance. AIP Conference Proceedings 2074, 020006 (2019) ; https ://doi.org/10.1063/1.5090623.

[13] A.Jamali Alaoui, M. Elkhomssi, B. Elgoumi, Dynamic Optimization Model of the Economic Behavior of Pension Fund. IEEE, 10.1109/ICOA.2019.8727677.

[14] Ramsey, Frank P. (1928) ; A Mathematical Theory of Saving, Appl.Anal.46 (3–4) (1992), 219–239.

[15] Harrod, Roy F. (1939). An Essay in Dynamic Theory, The Economic Journal. 49 (193) : 14-33. doi :10.2307/2225181.

[16] Targetti, Ferdinando, (1992). "Nicholas Kaldor : The Economics and Politics of Capitalism as a Dynamic System," OUP Catalogue, Oxford University Press, number 9780198283485.

[17] Solow, Robert M. (February 1956), A contribution to the theory of economic growth, Quarterly Journal of Economics. Oxford Journals. 70 (1) : 65-94.

[18] A. Krawiec & M. Szydlowski . (January 1999), The Kaldor-Kalecki business cycle model, Annals of Operations Research. Springer. DOI : https ://doi.org/10.1023/A :1018948328487.

[19] Koopmans, T. C. (1965), On the Concept of Optimal Economic Growth, The Economic Approach to Development Planning, Chicago : Rand McNally. pp. 225-287.

[20] Mankiw, N. Gregory ; Romer, David ; Weil, David N. (May 1992) ; A Contribution to the Empirics of Economic Growth, The Quarterly Journal of Economics. 107 (2) : 407-437.

[21] Aghion ; Howitt (1992), A Model of Growth Through Creative Destruction, Econometrica. 60 (2) : 323-351. doi :10.2307/2951599.

[22] J. W. Nevile, (Jun., 1962), The Mathematical Formulation of Harrod's Growth Model, The Economic Journal, Vol. 72, No. 286 , pp. 367-370. DOI : 10.2307/2228675.

[23] Soon Ryoo, (2013), Bank profitability, leverage and financial instability : a Minsky–Harrod model, Cambridge Journal of Economics, Volume 37, Issue 5, September, Pages 1127–1160, https ://doi.org/10.1093/cje/bes078

[24] Verena Halsmayer & Kevin D. Hoover , (2000), Solow's Harrod : Transforming macroeconomic dynamics into a model of long-run growth, The European Journal of the History of Economic Thought, Taylor& francis, Vol. 23, 22 - Issue 3, https ://doi.org/10.1080/01603477.2000.11490250.

[25] Khaled Hussein & A. P. Thirlwall, (2016), The AK Model of "New" Growth Theory Is the Harrod-Domar Growth Equation : Investment and Growth Revisited, Journal of Post Keynesian Economics , Taylor& francis, Vol. 23, Issue 4, https ://doi.org/10.1080/09672567.2014.1001763.

[26] Kevin Lee , M Pesaran and Ronald Smith (2004), Growth and Convergence in Multicountry Empirical Stochastic Solow Model, Journal of Applied Econometrics, vol. 12, issue 4, 357-92, 319 pp.

[27] William A. Brock & M. Scott Taylor. (May 2010), The Green Solow model, Journal of Economic Growth. Springer. DOI : https ://doi.org/10.1007/s10887-010-9051-0.

[28] Paul Douglas, Charles Cobb (1928), A Theory of Production, American Economic Review, Vol. 18.

[29] Paul K.J. Arrow, H.B. Chenery, B.S. Minhas, and R.M. Solow, (1961), Capital-labor substitution and economic efficiency, Review of Economics and Statistics (43), p. 225-250.

[30] Allen, R. G. D. (1968). Macro-economic Theory : A Mathematical Treatment. London : Macmillan. p. 35.

[31] B. Y. Chen, S. Decu, L. Verstraelen, (2013), Notes on isotropic geometry of production models, Kragujevac J. Math. 37(2), 217–220.

[32] X. Wang and Y. Fu,(2013), Some characterizations of the Cobb-Douglas and CES production functions in microeconomics, Abstr. Appl. Anal. 2013, Article ID 761832, 6 pages.

[33] B.Y. Chen, (1996), Mean curvature and shape operator of isometric immersions in real-space-forms, Glasg. Math. J. 38, 87–97.

[34] M.E. Aydin, (2016), M. Ergut, Isotropic geometry of graph surfaces associated with product production functions in economics, arXiv :1603.00222v1 [math.DG].

[35] M.E. Aydin, A.O. Ogrenmis & M. Ergut, (2015), Classification of factorable surfaces in the pseudo-Galilean space, Glas. Mat. Ser. III, 50(70), 441-451.

[36] M.E.Aydin, A.O.Ogrenmis, (2016), Homothetical and translation hypersurfaces with constant curvature in the isotropic space, B.S.G. Proceedings 23, 1–10.

[37] S. Decu, L. Verstraelen (2014), A note on the isotropical geometry of production surfaces, Kragujevac J. Math. 38(1), 23–33.

[38] Muhittin Evren Aydin & and Ion Mihai, (2017), On certain surfaces in the isotropic 4-space, Mathematical communications, 41 Math. Commun. 22, 41–51

[39] M.E Aydin & A.O. Ogrenmis, (2016), Linear weingarten factorable surfaces in isotropic spaces, http ://arxiv.org/abs/1604.01522v1.

[40] A. Tayebi, E. Peyghan & H. Sadeghi, (2012), On locally dually flat (α, β)-metrics with isotropic S-curvature, Indian Journal of Pure and Applied Mathematics volume 43, pages521–534

[41] X. Cheng and Y. Tian, (2011), Locally dually flat Finsler metrics with special curvature properties, Differential Geometry and its Applications. doi :10.1016/j.difgeo.2011.04.014.

[42] X. Cheng, Z. Shen and Y. Zhou, (2010), On a class of locally dually flat Finsler metrics, Int. J. Math., 21(11), 1–13.

[43] X. Cheng, (2008), H. Wang and M. Wang, (α, β)-metrics with relatively isotropic mean Landsberg curvature, Publ. Math. Debrecen, 72, 475–485.

[44] B.Y. Chen, (2014), Solutions to homogeneous Monge–Ampère equations of homothetic functions and their applications to production models in economics, Journal of Mathematical Analysis and Applications, 411, 223–229, http ://dx.doi.org/10.1016/j.jmaa.2013.09.029.

[45] A.D.Vîlcu, G.E.Vîlcu, (2013), On homogeneous production functions with proportional marginal rate of substitution, Math.Probl.Eng, Art.ID732643.

[46] Susan Chiona, Thomson Kalinda & Gelson Tembo, (2014), Stochastic Frontier Analysis of the Technical Efficiency of Smallholder Maize Farmers in Central Province, Zambia, Journal of Agricultural Science ; Vol. 6, No. 10, ISSN 1916-9752.

[47] Y. Liu, H. Pottmann &al (2006). Geometric modeling with conical meshes and developable surfaces. ACM SIGGRAPH 2006 Papers. Pages 681–689https ://doi.org/10.1145/1179352.1141941

[48] Farrell, M. J. (1957). The Measurement of Productive Efficiency. Journal of the Royal Statistical Society. Series A (General), 120(3), 253-290. http ://dx.doi.org/10.2307/2343100.

[49] Aigner, D. J., & Chu, S. F. (1968). On estimating the industry production function. Am. Econ. Rev., 58(4), 826-839.

[50] Thiry B., Tulkens H., (1992). Allowing for inefficiency in parametric estimation of production functions for urban transit firms, Journal of Productivity Analysis, 3, 45-65.

[51] Jondrow, J., Lovell, C. A. K., Materov, T. S., & Schmidt, P. (1982). On the estimating of technical inefficiency in the stochastic frontier production model, Journal of Econometrics, 19, 233-238. http ://dx.doi.org/10.1016/0304-4076 (82)90004-5

[52] Aigner, D. J., Lovell, C. A. K., & Schmidt, P. (1977). Formulation and estimation of stochastic frontier production models, J. Econ., 6, 21-37. http ://dx.doi.org/10.1016/0304-4076(77)90052-5

[53] Meeusen, W., & Van den Broeck, J. (1977). Efficiency Estimation from Cobb – Douglas Production Functions with Composed Error, International Economic Review, 18, 435-444. http ://dx.doi.org/10.2307/2525757

[54] Lovell C.A.K., (1993). Production frontiers and productive efficiency, in : Fried H.O., Knox Lovell C.A., Schmidt, S.S. (Eds.), The Measurement of Productive Efficiency : Techniques and Applications, Oxford University Press, New York, 3-67.

[55] Aigner, D. J., & Chu, S. F. (1968). On estimating the industry production function. Am. Econ. Rev., 58(4), 826-839.

[56] Olson, J. A, Schmidt, P & Waldman, D. M. (1980). A Monte Carlo study of estimators of stochastic frontier production functions, Journal of Econometrics, No. 13, pp 67- 82.

[57] Seiford M.L., Thrall R.M., 1990. Recent developments in DEA : The mathematical programming approach to frontier analysis, Journal of Econometrics 46, 7-38.

[58] Battese, G. E., & Coelli, T. J. (1995). A model for technical inefficiency effect in stochastic frontier production for panel data. Empirical Economics, 20, 325-332. http ://dx.doi.org/10.1007/bf01205442

[59] Rajiv D. Banker and al, (1988), A Comparison of DEA and Translog Estimates of Production Frontiers Using Simulated Observations from a Known Technology, Applications of Modern Production Theory : Efficiency and Productivity, vol 9 pp 33-55, https ://link.springer.com/chapter/10.1007/978-94-009-3253-1_2.

[60] Jeffrey and al, (2012), Issues relating to study design and risk of bias when including non-randomized studies in systematic reviews on the effects of interventions, Research Synthesis Methods, (wileyonlinelibrary.com) DOI : 10.1002/jrsm.1056.

[61] Solow, Robert M. (February 1956). ; *A contribution to the theory of economic growth*, Quarterly Journal of Economics. Oxford Journals. 70 (1) : 65-94. .

[62] Domar, Evsey (1946). ; *Capital Expansion, Rate of Growth, and Employment*. Econometrica. 14 (2) : 137-147. doi :10.2307/1905364.

[63] Koopmans, T. C. (1965). ; *On the Concept of Optimal Economic Growth*. The Economic Approach to Development Planning. Chicago : Rand McNally. pp. 225-287.

[64] Mankiw, N. Gregory ; Romer, David ; Weil, David N. (May 1992) ; *A Contribution to the Empirics of Economic Growth*, The Quarterly Journal of Economics. 107 (2) : 407-437.

[65] Aghion ; Howitt (1992) ; *A Model of Growth Through Creative Destruction*, Econometrica. 60 (2) : 323-351. doi :10.2307/2951599.

[66] Harrod, Roy F. (1939). ; *An Essay in Dynamic Theory*. The Economic Journal. 49 (193) : 14-33. doi :10.2307/2225181

[67] Acemoglu, James A. Robinson ; *Persistence of Power, Elites and Institutions*. American Economic Review, 98(1), pp267-293.

[68] Novales, Alfonso ; Fernández, Esther ; Ruiz, Jesús (2009). ; *Endogenous Growth with Accumulation of Human Capital*. Economic Growth : Theory and Numerical Solution Methods. Berlin : Springer. pp. 342-376. ISBN 978-3-540-68665-1.

[69] Novales, A., Fernandez, E., Ruiz, J. (2014) ; *Theory and Numerical Solution Methods*, Springer-Verlag Berlin Heidelberg 558p.

[70] N. Salehi Asfiji, R. Dalali Isfahane, R. Bakhshi Dastjerdi, M. Fakhar (2012) ; *Mathematical Model of Solow Economic Growth Model* Modern Sientific Press Chernivtsi, International Journal of Modern Mathematical Sciences, 4(3) : 118-125.

[71] Bernard Saby, Dominique Saby ; *Les grandes théories économiques*. Dunod, Paris, (2003), ISBN 2 10 008153 5.

[72] Kevin Lee , M Pesaran and Ronald Smith (2004) ; *Growth and Convergence in Multi-country Empirical Stochastic Solow Model*. Journal of Applied Econometrics, 1997, vol. 12, issue 4, 357-92, 319 pp.

[73] Di Liberto et al,(2008) ; *Past dominations, current institutions and the Italian regional economic performance*. Int. J. Pure Appl. Math 3(2002), no. 3, 255-273.

[74] Nelson, R. and Phelps, E. (1966) ; *Investment in humans, technological diffusion, and economic growth*, American Economic Review : Papers and Proceedings, 61, 69-75.

[75] David Coe and Elhanan Helpman ; *International R&D spillovers* , European Economic Review,1995, vol. 39, issue 5, 859-887.

[76] Jianning Kong & Peter C.B. Phillips & Donggyu Sul, (2017) ; Weak s- Convergence : Theory and Applications. Cowles Foundation Discussion Papers 2072, Cowles Foundation for Research in Economics, Yale University.

[77] Teheni El Ghak (2009) ; Productivite et clubs de convergence Productivity and convergence clubs. Volume 2, Issue 3, 2009, Pages 19-46. doi.org/10.1016/S1999-7620(09)70015-1.

[78] Filippo Ferroni, Stefano Grassi & Miguel A. León-Ledesma (Martch 2019), Selecting structural innovations in DSGE models, Journal of applied economics, Vol 34, Issue2, https ://doi.org/10.1002/jae.2664.

[79] Seyma CaliskanCavdar and Alev DilekAydin (2015) ; Empirical Analysis about Technological Development and Innovation Indicators J. Procedia - Social and Behavioral Sciences, Volume 195, 3 July 2015, Pages 1486-1495. https ://doi.org/10.1016/j.sbspro.2015.06.449.

[80] Wilfred J. Ethier, James R. Markusen (1996) ; Multinational firms, technology diffusion and trade Journal of International Economics 41 1-28. Elsevier

[81] Scott McDonald, Jennifer Robert (2002) ; Growth and multiple forms of human capital in an augmented Solow model : a panel data investigation.Economics Letters 74 (2002) 271-276.

[82] Jess Benhabib and Mark Spiegel (2005) ; Human Capital and Technology Diffusion, Chapter 13 in Handbook of Economic Growth, vol. 1, Part A, pp 935-966 from Elsevier

[83] Haralambopoulos, DA., Polatidis, H.,2003 ; *Renewable energy projects : structuring a multi-criteria group decision-making framework.* Renewable Energy (2003) ;28 :961–73.

[84] Hokkanen, J., and Salminen, P.,1997 ; *Choosing a solid waste management system using multicriteria decision analysis.* European Journal of Operational Research, 98 : 19-36, (1997).

[85] Hwang, CL., Yoon, K., 1981 ;*Multiple attribute decision making-methods and application : a state-of-the-art survey.* New York, USA : Springer-Verlag ; (1981).

[86] MarekS zydłowski & AdamKrawiec . (July 2005), The stability problem in the Kaldor–Kalecki business cycle model, Chaos, Solitons & Fractals Vol 25, Issue 2, Pages 299-305, Elsevier, https ://doi.org/10.1016/j.chaos.2004.11.012.

[87] Imrul Hossain, (2020), Extended exogenous growth model : application and investigation the long-term growth determinants of bangladesh, Asian Economic and Financial Review, Vol. 10, No. 1, 35-53. DOI : 10.18488/journal.aefr.2020.101.35.53.

[88] Ramsey, Frank P. (1928), A Mathematical Theory of Saving, Appl.Anal.46 (3–4) (1992), 219–239.

[89] Mankiw,N.G, D.Romer, and D. Weil, (1992), A contribution to the empirics of economic growth, Quarterly Journal of Economics, 107, 407-438.

[90] Erwin Dekker,2020, Entangled Economists : Ragnar Frisch and Jan Tinbergen. Erasmus University of Rotterdam.

[91] Novriana Sumarti and al, (2020), Evaluating the performance of islamic banks using a modified monti-Klein model. Journal of risk and financiel management, Vol 13 Issue 3, 13(3), 43, https ://doi.org/10.3390/jrfm13030043.

[92] Edward Nelson, (2020), Seven Fallacies Concerning Milton Friedman's "The Role of Monetary Policy", Journal of Money, Credit and Banking, Vol. 52, No. 1, https ://doi.org/10.1111/jmcb.12591.

[93] James Forder, (2019), Stabilization Policy and the Causes of Inflation. Milton Friedman pp 261-289, Springer, DOI https ://doi.org/10.1057/978-1-137-38784-4_15.

[94] Richard R. Nelson and Edmund S. Phelps (1966), Investment in Humans, Technological Diffusion, and Economic Growth. The American Economic Review Vol. 56, No. 1/2, pp. 69-75, https ://www.jstor.org/stable/1821269.

[95] Edmund S. Phelps (1972), The Statistical Theory of Racism and Sexism. The American Economic Review Vol. 62, No. 4, pp. 659-661

[96] Helmut Lütkepohl (July 2013), Vector autoregressive models, Category : Handbook Chapter DOI :https ://doi.org/10.4337/9780857931023.00012 Pages :139–164 (26 total) Collection : Economics 2013.

[97] Katerina Petrova, (2019) ; A quasi-Bayesian local likelihood approach to time varying parameter VAR models. Journal of Econometrics. Vol 212, Issue 1, https ://doi.org/10.1016/j.jeconom.2019.04.031

[98] Nicolás Rivera Garzón & Miller Rivera Lozano, (2019), A complex toolbox for the 21st century economist. Vol. 10 Núm. 2 (2019) : Revista Aglala, DOI : https ://doi.org/10.22519/22157360.1442.

[99] Michał Brzoza-Brzezina & Jacek Kotłowski ,(2020), The nonlinear nature of country risk and its implications for dsge models, Macroeconomic Dynamics, Volume 24, , pp. 601-628, DOI : https ://doi.org/10.1017/S136510051800038X.

[100] Francesco Sergi,.(2020), The standard narrative about DSGE models in central banks technical reports. Journal The European Journal of the History of Economic Thought, Volume 27, Issue 2, Pages 163-193, https ://doi.org/10.1080/09672567.2019.1651365.

[101] Adrian Pagan (2019), Implications of Partial Information for Econometric Modeling of Macroeconomic Systems, SSRN, CAMA Working Paper No. 41/2019.

[102] Jaime Martinez-Martin and al, (2019), Merging Structural and Reduced-Form Models for Forecasting : Opening the Dsge-Var Box, SSRN, ECB Working Paper No. 2335, https ://papers.ssrn.com/sol3/papers.cfm ?abstract_id=3497524.

[103] Stephen J. Cole & and Fabio Milani,(2019),The misspecification of expectations in new keynesian models : a dsge-var approach, Macroeconomic Dynamics, Vol 23, Issue 3, pp. 974-1007, DOI : https ://doi.org/10.1017/S1365100517000104.

[104] Eric Jondeau & Michael Rockinger,(2019),The misspecification of expectations in new keynesian models : a dsge-var approach, Journal of Money, Credit and Banking, Vol 51, Issue8, Pages 2239-2291, DOI : https ://doi.org/10.1111/jmcb.12586.

[105] Jia, JM., Fisher, GM., Dyer, JS.,1998. *Attribute weighting methods and decision quality in the presence of response error : a simulation study.* Journal of Behavioral Decision Making (1998) ;11 :85–105.

[106] Keeney, R., Raifa H.,1976 ; Decision with Multiobjectives, Preferences and Value Trade-Offs. New York, (1976).