

**UNIVERSITE DE RENNES
FACULTÉ DES SCIENCES ÉCONOMIQUES**

-----□□□-----



PROJET FINAL

**ANALYSE DES DETERMINANTS MACRO-FINANCIERS DES CRISES
BANCAIRES : APPROCHE ACP – CLUSTERING ET REGRESSION
PANEL**

Master 1 Monnaie, Banque, Finance et Assurance parcours Ingénierie Économique et Financière
Nom de l'étudiant(e):

Nguyen Hoang Phuc PHAN

Ndeye Mbayang TALL

Promotion: 2025-2026

Rennes, 12/2025

TABLEAU DES MATIÈRES

TABLEAU DES MATIÈRES	1
LISTE DES TABLEAUX	3
LISTE DES FIGURES.....	3
RÉSUMÉ	4
INTRODUCTION	5
1. ANALYSE EN COMPOSANTES PRINCIPALES :	7
1.1. Analyse de la matrice de corrélation :	7
1.2. Analyse en composantes sans pondération :.....	8
1.3. Analyse en composantes principales pondérées :.....	10
2. CLUSTERING HIÉRARCHIQUE (HCPC / K-MEANS) :	11
2.1. Cluster 1 - Economies stables :.....	12
2.2. Cluster 2 - Économies instables, inflation élevée, forte dépréciation, forte fragilité :.....	13
2.3. Cluster 3 - Économies intermédiaires : croissance modérée, crédit modéré, mais risques bancaires faibles :.....	13
3. ANALYSE EN COMPOSANTES PRINCIPALES PRE/POST-CRISE :	15
3.1. Analyse en Composantes Principales pré-crise :.....	15
3.2. Analyse en Composantes Principales post crise :.....	17
3.3. Classement des pays :	18
4. ANALYSE ECONOMETRIQUE – TESTS ET INTERPRETATION ECONOMIQUE :	20
4.1. Hypothèses économiques et la lecture de littérature :	20
4.2. Équation de recherche de modèle:.....	21
4.3. Sélection du modèle de régression sur données de panel :.....	21
4.4. Résultat de régression panelle du modèle :.....	23
4.5. L’interprétation des variables exogènes retenues sur la présence de l’occurrence de crise:	24
4.5.1. <i>L’effet du rendement des actifs (%)</i>	24

4.5.2. <i>L'effet de balance courante / PIB (%)</i>	24
4.5.3. <i>Taux d'inflation (%)</i>	25
4.5.4. <i>L'effet de Taux de dépôt (%)</i>	25
4.6. Tests diagnostiques :	25
4.6.1. Test de normalité :.....	25
4.6.2. Test de l'homoscédasticité :.....	26
4.6.3. Test de multi colinéarité :.....	26
CONLUSION	27
RÉFÉRENCE BIBLIOGRAPHIQUES	28
ANNEXES	1
CODE	1

LISTE DES TABLEAUX

Tableau 0.1: Cadre théorique et poids attribués aux variables.	6
Tableau I.1 : Tableau de pourcentage de Variance Expliquée et pourcentage cumulé avec pondération en fonction de dimension	11
Tableau II.1: Profil des clusters : valeurs moyennes des indicateurs bancaires et macroéconomiques, et interprétation économique associée.....	15
Tableau III.1: Comparaison du CP1 pré-crise et post-crise : variations moyennes et interprétation par pays	20
Tableau IV.1: Résultat de régression du modèle sur les données pannel	23

LISTE DES FIGURES

Figure I.1: Coefficients de corrélation entre les variables utilisées dans l'ACP	7
Figure I.2: Cercle des corrélations des variables (ACP) non pondéré	9
Figure I.3 : Cercle des corrélations des variables (ACP) avec pondération	10
Figure I.4 : Le graphique de pourcentage de Variance Expliquée avec pondération en fonction de dimension	11
Figure II.1: Classification hiérarchique sur composantes principales (CHCP)	12
Figure III.1 : Cercle des corrélations des variables (ACP) pré-crise.....	16
Figure III.2 : Le graphique de pourcentage de Variance Expliquée (ACP) pré-crise	16
Figure III.3 : Cercle des corrélations des variables (ACP) post-crise	17
Figure III.4 : Le graphique de pourcentage de Variance Expliquée (ACP) post-crise....	18
Figure III.5 : Variation moyenne du CP1 (post – pré) pour les 10 pays les plus affectés	19

RÉSUMÉ

Cette étude analyse les vulnérabilités macro-financières de 14 pays émergents entre 1993 et 2002, une période marquée par plusieurs crises majeures. En mobilisant un ensemble de 13 indicateurs macroéconomiques et bancaires, l'approche combine plusieurs méthodes complémentaires. Une analyse en composantes principales, réalisée sous forme non pondérée puis pondérée, met en évidence deux dimensions majeures : la profondeur financière et la fragilité bancaire, qui structurent plus de la moitié de la variance des données. Les méthodes de clustering permettent ensuite de classifier les pays en trois groupes cohérents, distinguant les économies stables, les économies sévèrement touchées par les crises, et un groupe intermédiaire présentant une résilience partielle. Une comparaison pré-crise et post-crise révèle une reconfiguration des relations entre variables, confirmant que la crise renforce le rôle central des indicateurs bancaires. Enfin, une régression sur données de panel, sélectionnée selon la procédure de Baltagi, retient un modèle à effets aléatoires où quatre variables significatives expliquent l'occurrence de crise. L'ensemble de ces analyses offre une lecture intégrée des mécanismes de vulnérabilité financière dans les pays émergents.

INTRODUCTION

Dans un contexte d'intégration économique internationale sans cesse, les pays émergents ont fait face à une succession de crises financières à une échelle exceptionnelle depuis le début des années 90. Pendant cette période, plusieurs événements macro-financiers ont été enregistrés comme la crise mexicaine (1994), la crise d'Asie (de 1997 à 1998), la crise russe (1998) et l'effondrement argentin et brésilien (au début des années 2000). Ces derniers ont eu des problèmes macro financiers, qui ont contribué à constituer de nombreuses recherches empiriques. Ces épisodes ont profondément modifié l'allocation de crédit, de rentabilité des établissements de crédit, d'inflation, ainsi que les équilibres externes, traduisant des tensions dans des structures susceptibles d'accroître le risque systémique.

Notre étude s'inscrit dans cette période sur un panel de données annuelles allant de 1993 à 2002, se portant sur 14 pays en développement d'Amérique latine, d'Asie et d'Europe de l'Est. Tous ces pays ont été vulnérables face aux différentes crises (de change, d'instabilités macroéconomiques ou bancaires). La base des données inclut 13 indicateurs macroéconomiques et financiers axés sur l'explication de l'occurrence de crises bancaires en année T+1..

Compte tenu des circonstances particulières de cette décennie, une approche méthodologique impliquant une pondération a été appliquée aux variables pertinentes pour les dégradations de notation des banques.

Caractéristique	Variables associées	Poids attribué	Justification théorique
Fragilité bancaire	(1) ROAA (2) NPLGROSSLOANS	0.8 – 1	Variables centrales de solvabilité et de qualité des actifs. Forte contribution empirique aux crises bancaires.
Profondeur et expansion financière	(3) CREDITGDP	1	Le levier du crédit et la croissance excessive du secteur financier sont des prédicteurs robustes du « credit boom–bust cycle »
Tensions nominales	(4) INFL (5) REALRATE (6) DEPRATE	0.8	Capturent caractère instable macroéconomique, pressions inflationnistes, volatilité du change.

Dynamique réelle et investissement	(7) GDPG (8) GFCFG (9) GFCFGDP	0.6 – 0.8	Variables cycliques; reflètent l'expansion ou la contraction de l'activité, mais moins directement liées à l'éclatement de crises.
Déséquilibres externes	(10) CAGDP	0.6	Les déficits/excédents courants importants signalent vulnérabilité externe et pressions sur le change.
Consommation et demande interne	(11) HFCEG (12) HFCEGDP	0.6	Variables reflétant le cycle domestique.
Valorisation financière	(13) PER	0.1	Rôle théorique faible dans les crises bancaires ; poids réduit pour éviter sa sur-influence.

Tableau 0.1: Cadre théorique et poids attribués aux variables.

Notre démarche se déroulera en plusieurs étapes complémentaires afin de mettre en lumière l'objectif de l'étude.

Dans un premier temps, nous avons procédé à une étude de la structure interne des données via une ACP non pondérée et pondérée, afin d'examiner la question d'une amélioration de l'analyse grâce aux pondérations et les dimensions de vulnérabilité macro-financière.

Ensuite des méthodes de clustering (HCPC et k-means) sont effectuées pour donner une classification des pays selon leur profil de crises.

En troisième étape, nous avons réalisé un modèle économétrique sur des données de panel (modèles poolés, à effets fixes, puis à effets aléatoires), en suivant strictement la procédure de spécification de Baltagi (2013), afin de sélectionner le modèle le plus cohérent et robuste.

Enfin, nous avons procédé à une analyse approfondie des déterminants significatifs de l'occurrence de crise bancaire et à une série de tests diagnostiques visant à évaluer la validité et la robustesse du modèle final.

En résumé, notre étude a pour but non seulement de déterminer les facteurs exogènes qui influencent la probabilité d'occurrence des crises bancaires dans les années 90-2000, mais également d'approfondir pour trouver les facteurs qui interagissent différemment dans les périodes pré-crise et post-crise.

1. ANALYSE EN COMPOSANTES PRINCIPALES :

Afin d'étudier les liens entre les pays et leurs vulnérabilités face aux crises économiques entre 1992 et 2002, une première étape a consisté à analyser la structure interne des données macroéconomiques et financières. Nous avons suivi une démarche en deux temps. Dans un premier temps, la matrice de corrélations a permis d'identifier les relations entre les variables et de repérer les principales dimensions économiques sous-jacentes. Dans un second temps, une analyse en composantes principales (ACP) a été réalisée selon deux approches, une approche non pondérée et une pondérée afin de comparer l'influence relative des facteurs structurels, des dynamiques de croissance et des indicateurs de risque financier. La pondération a été utilisée pour réduire l'impact des épisodes extrêmes et renforcer le poids des variables reflétant la vulnérabilité bancaire.

1.1. Analyse de la matrice de corrélation :

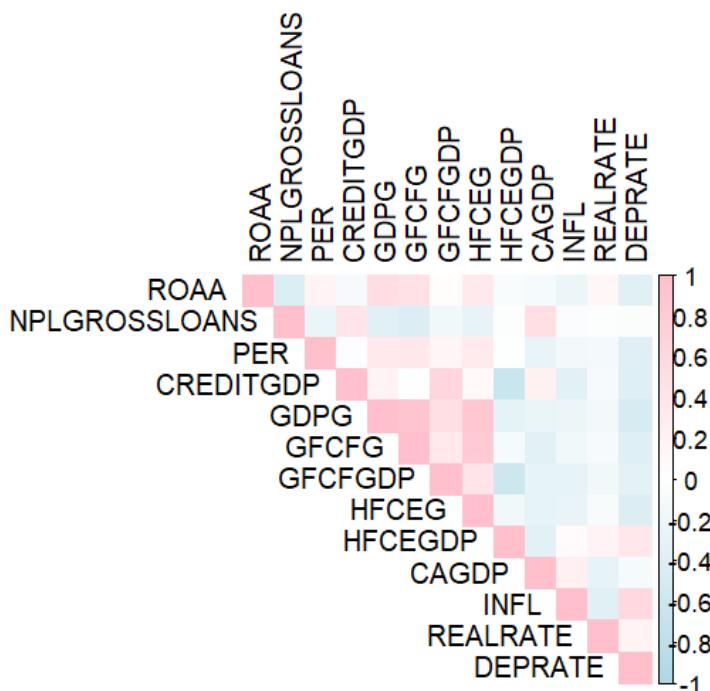


Figure I.1: Coefficients de corrélation entre les variables utilisées dans l'ACP

Tout d'abord il était important pour nous de visualiser le lien entre les variables à travers cette matrice de corrélations qui met en évidence plusieurs regroupements cohérents entre variables macroéconomiques et financières. Les indicateurs liés au crédit et au financement (CREDITGDP, GFCFG, GFCFGDP, HFCEG, HFCEGDP, CAGDP) présentent des corrélations positives importantes, reflétant une dynamique économique intégrée où l'expansion du crédit, l'investissement et la consommation évoluent conjointement. Cette configuration traduit des phases d'expansion caractérisées par une progression simultanée du crédit et de l'activité réelle.

Les variables bancaires telles que ROAA et NPLGROSSLOANS affichent des corrélations négatives avec plusieurs indicateurs macroéconomiques. Cela suggère qu'un ralentissement économique s'accompagne d'une détérioration de la qualité des actifs bancaires et d'une baisse de rentabilité, illustrant le rôle central du système bancaire dans la transmission des chocs économiques.

Les indicateurs d'instabilité nominale (INFL, REALRATE, DEPRATE) montrent des corrélations souvent négatives avec les variables d'activité réelle, ils correspondent aux tensions typiques précédant les crises: hausse des taux d'intérêt, pressions inflationnistes, dépréciation monétaire et ralentissement du crédit.

L'ensemble du graphique révèle donc une structure cohérente avec les mécanismes de vulnérabilité macro-financière. Les pays dans lesquels le crédit, la consommation et l'investissement progressent simultanément apparaissent plus sensibles aux retournements conjoncturels, tandis que les indicateurs bancaires soulignent la fragilité accrue du système financier lors des périodes de crise. Cette structure justifie le recours à une analyse en composantes principales et à des méthodes de clustering pour identifier des profils de vulnérabilité avant et après les crises.

Pour évaluer la robustesse des résultats et limiter l'influence des valeurs extrêmes caractéristiques des épisodes de crise, deux ACP ont été menées :

- ACP non pondérée : toutes les variables ont le même poids dans la construction des axes.
- ACP pondérée : des poids théoriques ont été attribués, avec une priorité donnée aux indicateurs de risque financier (NPLGROSSLOANS, CREDITGDP), afin de renforcer la représentativité des facteurs de vulnérabilité structurelle.

1.2. Analyse en composantes sans pondération :

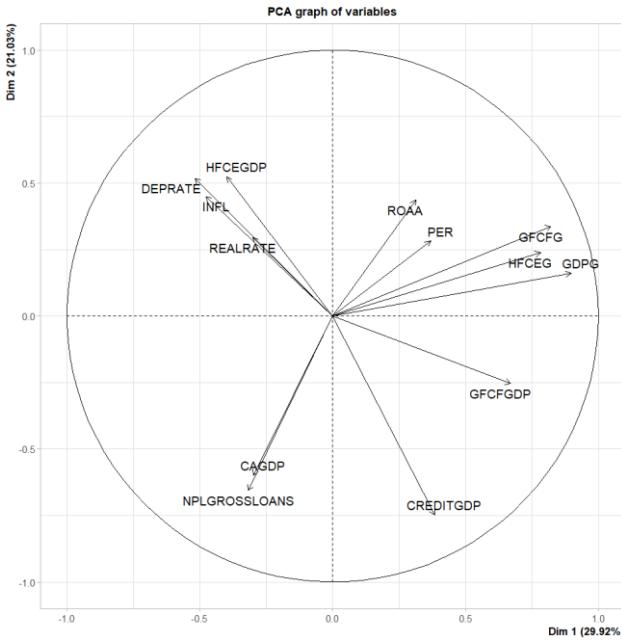


Figure I.2: Cercle des corrélations des variables (ACP) non pondéré

Dans l'ACP non pondérée, le premier axe (Dim 1) est principalement façonné par les variables reflétant la dynamique de croissance réelle et l'investissement, notamment le GDPG (0,897), GFCFG (0,821), GFCFGDP (0,669) et HFCEG (0,782) qui y contribuent fortement. Dim 1 oppose donc les pays en forte expansion économique à ceux en ralentissement, tandis que les variables d'instabilité nominale (INFL : -0,476 ; DEPRATE: -0,518) se situent à l'opposé, ce qui confirme l'opposition classique entre croissance et tensions macroéconomiques. Les valeurs associées aux variables GDPG (0,897), GFCFG (0,821), GFCFGDP (0,669) et HFCEG (0,782) correspondent à leur qualité de représentation sur la première composante principale (Dim 1), mesurée par leur \cos^2 .

La qualité de représentation indique dans quelle mesure une variable est correctement projetée sur un axe: plus le \cos^2 est élevé (proche de 1), plus l'axe explique une grande part de la variance de cette variable. Autrement dit, un \cos^2 élevé signifie que la variable est bien capturée et bien résumée par la dimension considérée (*annexe 1*).

Le second axe (Dim 2), expliquant 21,03 % de la variance, est structuré par les déséquilibres macro-financiers, avec des coordinations positives pour DEPRATE (0,518), INFL (0,449), REALRATE (0,294) et HFCEGDP (0,523). En face, les prêts non performants (NPLGROSSLOANS : -0,655) et le solde extérieur (CAGDP : -0,598) signalent des fragilités financières internes. Dim 2 oppose donc instabilité monétaire et fragilisation bancaire.

L'ACP non pondérée met en évidence une tension forte entre croissance réelle et instabilité macroéconomique, mais elle reste partiellement influencée par les épisodes

extrêmes. L'absence de pondération a pour effet d'accentuer la présence de ces chocs dans la structure factorielle.

1.3. Analyse en composantes principales pondérées :

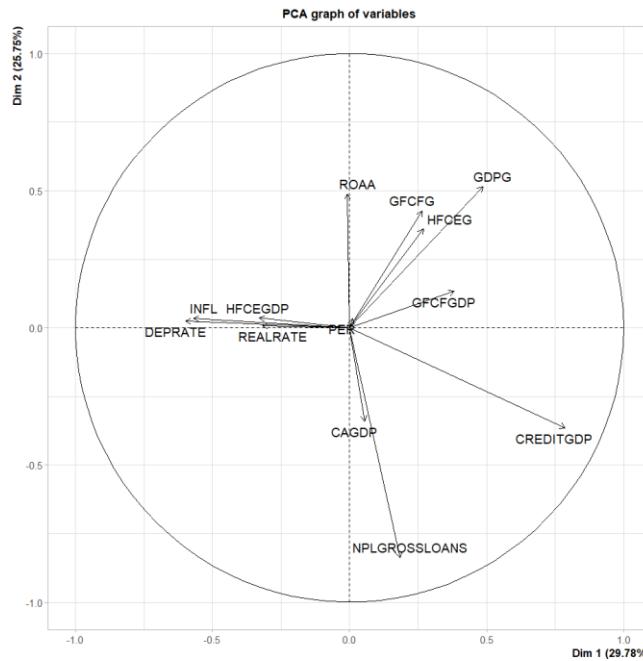


Figure I.3 : Cercle des corrélations des variables (ACP) avec pondération

L'ACP pondérée réorganise la structure factorielle en réduisant l'influence des valeurs extrêmes et en mettant davantage en avant les vulnérabilités structurelles. Le premier axe (Dim 1, 29,78 % de variance) devient dominé par CREDITGDP (0,786), désormais variable centrale de l'axe. Dim 1 reflète ainsi le cycle économique et la profondeur financière plutôt que la croissance réelle seule. Les variables GDPG, GFCFG, GFCFGDP et HFCEG conservent une contribution positive, mais plus modérée .

Le second axe (Dim 2, 25,75 % de variance) est marqué par les fragilités bancaires et les déséquilibres externes, avec NPLGROSSLOANS (-0,838) et CAGDP (-0,340) en contributions majeures. À l'opposé se trouvent des variables représentant la croissance et l'investissement (ex. GFCFG : 0,426). Dim 2 oppose donc vulnérabilité financière et stabilité bancaire (*annexe 2*).

Un élément notable est la perte d'importance des variables nominales (INFL, DEPRATE, REALRATE), dont les contributions deviennent faibles, avec des \cos^2 très bas (souvent $< 0,05$). Elles apparaissent proches de l'origine du cercle des corrélations, confirmant que la pondération atténue leur rôle dans la structure factorielle.

L'ACP pondérée met en lumière deux dimensions fortes :

- une dimension de profondeur et d'endettement financier
- une dimension de vulnérabilité bancaire et externe.

Au total, les deux premières composantes expliquent 55 % de la variance. L'analyse du scree plot ci dessous, fondée sur la règle du coude, montre que la décroissance de la variance expliquée ralentit nettement après la 2ème composante. Même si il faut quatre dimensions pour atteindre au moins 70% de variance cumulée afin de restituer l'essentiel de l'information statistique, le plan factoriel (Dim 1, Dim 2) demeure celui qui concentre les dimensions les plus pertinentes pour l'interprétation économique et pour la classification des pays. Le choix d'utiliser principalement ces deux axes repose donc à la fois sur la règle du coude et sur leur forte cohérence économique.

Le restant de notre analyse va se baser sur l'ACP avec pondération.

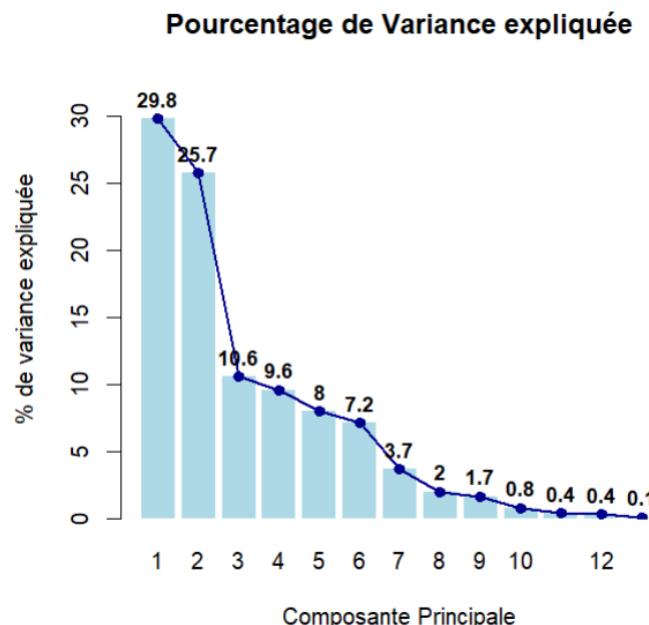


Figure I.4 : Le graphique de pourcentage de Variance Expliquée avec pondération en fonction de dimension

Composante Principale	Pourcentage de Variance Expliquée	Pourcentage Cumulé
Dim 1	29,8%	29,8%
Dim 2	25,7%	55,5%
Dim 3	10,6%	66,1%

Tableau I.2 : Tableau de pourcentage de Variance Expliquée et pourcentage cumulé avec pondération en fonction de dimension

2. CLUSTERING HIÉRARCHIQUE (HCPC / K-MEANS) :

L'analyse en composantes principales a permis de synthétiser les dimensions économiques et financières majeures et de mettre en évidence les axes de vulnérabilité et

de croissance entre pays. Ces axes fournissent un espace factoriel réduit, où chaque pays peut être positionné selon ses caractéristiques macro-financières et sa sensibilité aux crises.

Cette représentation factorielle constitue une base idéale pour la segmentation des pays à travers des méthodes de clustering. En regroupant les profils similaires sur les axes de l'ACP, on peut distinguer des clusters reflétant différents types de vulnérabilité: des pays à forte expansion mais avec une fragilité bancaire, des pays en stagnation avec des déséquilibres externes, ou des pays présentant une stabilité macro-financière relative.

Le recours au clustering permettra ainsi de compléter l'analyse exploratoire en identifiant des groupes homogènes et en facilitant la compréhension des facteurs structurants derrière la résilience ou la vulnérabilité face aux crises économiques.

Sur la base des deux axes vus précédemment avec l'ACP pondéré, une classification a été réalisée. Trois groupes se distinguent.

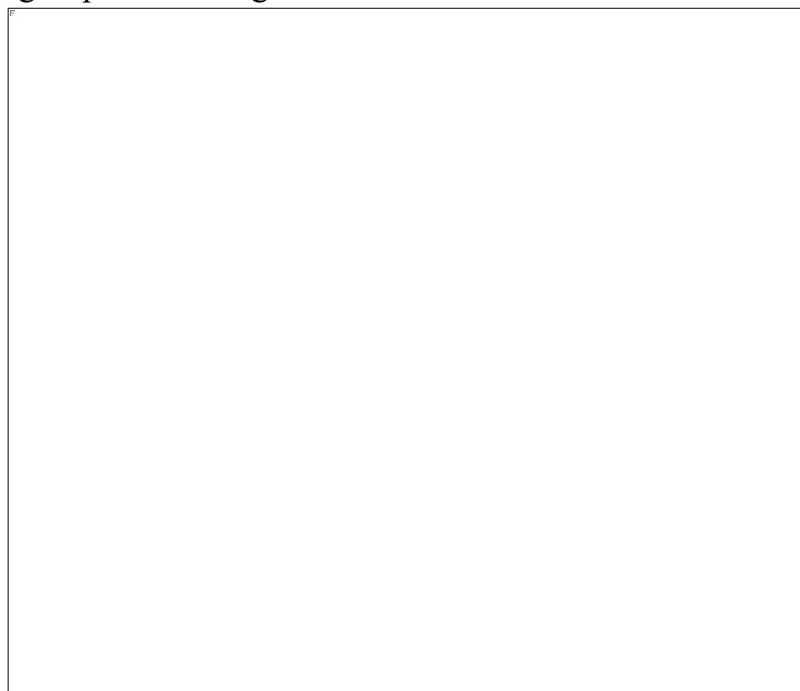


Figure II.5:Classification hiérarchique sur composantes principales (CHCP)

2.1. Cluster 1 - Economies stables :

Le premier groupe rassemble des économies où la croissance, le crédit et l'investissement semblent évoluer dans un cadre relativement stable. Le taux de croissance moyen y atteint près de 6,95 %, ce qui est cohérent avec un ratio crédit/PIB particulièrement élevé, supérieur à 107 %, traduisant un système financier capable de soutenir activement l'activité. Le niveau d'investissement, avec une moyenne d'environ 33 % du PIB, témoigne d'une dynamique d'accumulation soutenue. La rentabilité bancaire reste positive (ROAA autour de 0,54 %) et l'inflation demeure la mieux maîtrisée parmi les 3 profils, proche de 4,4 %, ce qui limite les pressions sur les marges réelles et sur la

stabilité du crédit. Malgré le fait que les prêts non performants s'établissent autour de 8,7 %, un niveau non négligeable, ils ne semblent pas compromettre l'équilibre général.

L'ensemble représente le profil d'économies en expansion où les tensions macroéconomiques restent contenues, permettant un fonctionnement relativement efficace de la finance domestique.

2.2. Cluster 2 - Économies instables, inflation élevée, forte dépréciation, forte fragilité :

Le deuxième cluster correspond à un environnement profondément altéré par des déséquilibres macro-financiers cumulatifs. Le taux de croissance moyen y est négatif (-3,67 %), reflétant une contraction économique marquée, associée à une inflation extrêmement élevée, supérieure à 60 %, et à une dépréciation moyenne de la monnaie avoisinant 52 %. Cette combinaison traduit des épisodes de crise monétaire ou budgétaire, dans lesquels la politique économique perd en crédibilité et les anticipations deviennent instables. La détérioration du cadre macroéconomique affecte directement le secteur financier: le ROAA est négatif (-1,68 %), signe de pertes régulières, et les prêts non performants dépassent 9 %. Le crédit au secteur privé, limité à environ 48 % du PIB, reflète une contraction de l'intermédiation bancaire ou une incapacité des ménages et entreprises à emprunter dans des conditions de taux réels extrêmement élevés (15,9 % en moyenne).

Ce cluster correspond donc à des économies qui se retrouvent dans des boucles d'amplification entre inflation, dépréciation, récession et fragilité bancaire, où les décisions politiques seront principalement orientées vers la stabilisation rapide de la situation macroéconomique.

2.3. Cluster 3 - Économies intermédiaires : croissance modérée, crédit modéré, mais risques bancaires faibles :

Le troisième cluster se situe dans une position intermédiaire, combinant des performances relativement solides avec quelques tensions persistantes. La croissance y est positive (4,31 %) et la rentabilité bancaire y est la plus élevée (ROAA autour de 1,22 %), soutenue par un niveau très faible de prêts non performants (5,05 %), le plus bas des trois groupes. Cette situation indique que le secteur bancaire fonctionne dans de bonnes conditions malgré un contexte marqué par une inflation moyenne d'environ 11,5 % et une dépréciation annuelle proche de 14,5 %. Le crédit représente environ 49 % du PIB, un niveau modéré, tandis que la consommation finale des ménages est particulièrement importante, dépassant 64 % du PIB, ce qui suggère une croissance tirée par la demande intérieure. Cependant, le déficit courant modéré (-1,13 %) et les taux réels relativement élevés (12,1 %) traduisent un certain degré de vulnérabilité vis-à-vis des chocs externes et des ajustements de politique monétaire.

Ces résultats illustrent trois configurations macro-financières distinctes : d'un côté, des économies en expansion soutenue où le crédit joue un rôle central, à l'opposé des pays plongés dans une dynamique de crise caractérisée par l'instabilité des prix, la dépréciation et l'effondrement de la rentabilité bancaire, enfin, un groupe intermédiaire combinant un secteur bancaire performant et une croissance solide, mais confronté à des pressions inflationnistes et à une volatilité du change. Cela montre clairement à quel point les interactions entre croissance, stabilité des prix, solidité bancaire et profondeur financière déterminent les trajectoires économiques des différents pays.

Indicateur	Cluster 1	Cluster 2	Cluster 3	Interprétation
Nombre d'observations	35	18	67	Taille des groupes
Variables Bancaires				
ROAA (Rentabilité des Actifs, %)	0.54	-1.68	1.22	Plus faible dans Cluster 2 (pertes)
NPLGROSSLOANS (Prêts non performants, %)	8.74	9.22	5.05	Plus faible dans Cluster 3 (stabilité)
PER (Prêts/Réserves, Ratio)	19.79	4.52	21.17	Faible dans Cluster 2 (fuite de capitaux)
CREDITGDP (Crédit sur PIB, %)	107.46	47.90	49.32	Très élevé dans Cluster 1
Variables Macroéconomiques				
GDPG (Croissance PIB réel, %)	6.95	-3.67	4.31	Croissance négative dans Cluster 2 (récession)
GFCFG (Croissance de la Formation Brute de Capital Fixe, %)	6.86	-15.46	5.54	Chute drastique de l'investissement dans Cluster 2
INFL(Inflation, %)	4.44	60.74	11.51	Inflation extrême dans Cluster 2 (instabilité monétaire)
CAGDP (Balance Courante/PIB, %)	1.37	2.44	-1.13	Déficit le plus marqué dans Cluster 3 (besoin de financement externe)
DEPRATE (Taux de Dépréciation, %)	6.56	51.98	14.50	Dévaluation massive dans Cluster 2 (choc de change)

Tableau II.3: Profil des clusters : valeurs moyennes des indicateurs bancaires et macroéconomiques, et interprétation économique associée

3. ANALYSE EN COMPOSANTES PRINCIPALES PRE/POST-CRISE :

Afin d'évaluer l'impact des épisodes de crise sur la structure macro-financière des pays, on a réalisé séparément une ACP sur les périodes pré-crise et post-crise (définies par la variable indicatrice *PCRISIS1*). Cette démarche permet d'identifier les changements dans les relations entre variables, ainsi que les pays les plus touchés.

On a préalablement standardisé les données via un Z-score, ce qui nous a permis de neutraliser les différences d'échelle entre variables macroéconomiques, financières et bancaires et a garanti que chaque indicateur contribue équitablement à la construction des axes factoriels.

Dans la section 1, nous avions aussi réaliser des pondérations fondées sur des considérations économiques et bancaires en donnant plus de poids aux variables reflétant la vulnérabilité bancaire (*NPLGROSSLOANS*, *ROAA*), la profondeur financière (*CREDITGDP*) ou les tensions nominales (*INFL*, *REALRATE*, *DEPRATE*) afin de renforcer l'importance des facteurs structurels de fragilité dans la construction des dimensions principales.

Dans cette section, on va aller plus loin dans notre analyse en réalisant une ACP séparée pour les observations pré-crise (70 observations) et post-crise (50 observations) afin d'analyser si la structure des relations économiques demeure stable ou si elle se reconfigure après la crise, signe d'une rupture structurelle.

3.1. Analyse en Composantes Principales pré-crise :

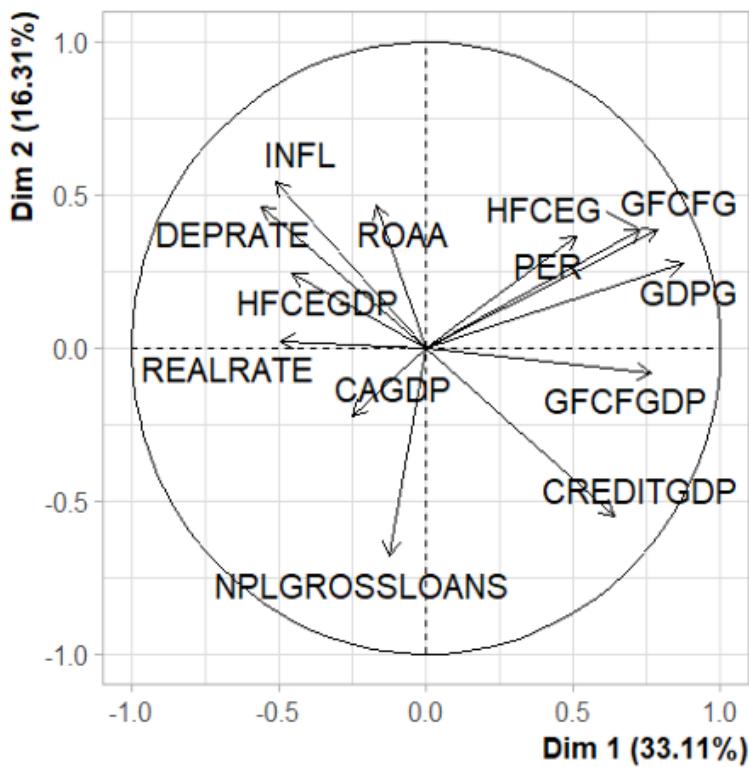


Figure III.6 : Cercle des corrélations des variables (ACP) pré-crise

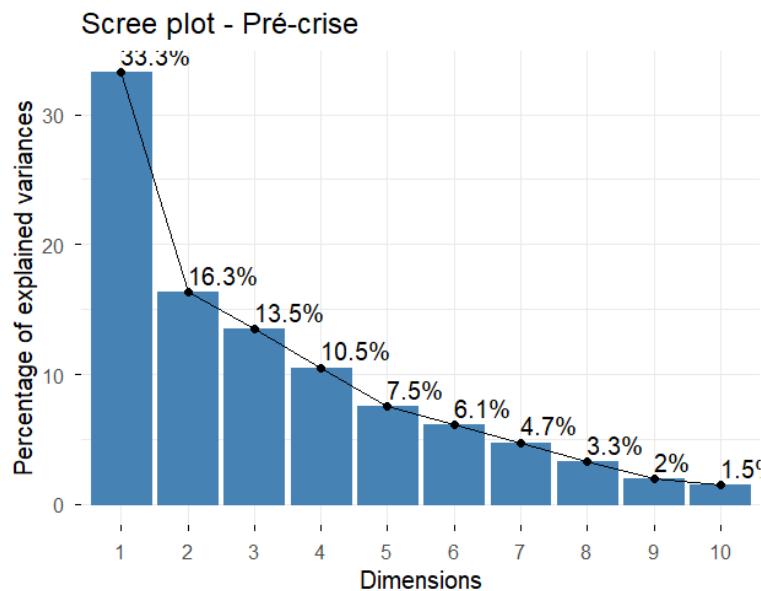


Figure III.7 : Le graphique de pourcentage de Variance Expliquée (ACP) pré-crise

Avant la crise, la variance est majoritairement expliquée par l'Axe 1 (33,3 %), traduisant un système où le cycle de croissance est le principal facteur de différenciation des

observations. La structure factorielle est donc relativement compacte: un petit nombre de variables suffit à résumer les mécanismes macro-financiers.

La première dimension oppose les variables de croissance réelle et d'investissement (telles que GDPG, GFCFG ou GFCFGDP) à des variables exprimant les tensions nominales, notamment l'inflation, le taux de change et les taux réels. Cette organisation reflète un schéma macroéconomique classique : les économies affichant une forte croissance se caractérisent par une faible instabilité monétaire et inversement. Les indicateurs de fragilité (NPLGROSSLOANS et DEPRATE) se projettent sur l'Axe 2, qui est secondaire, ce qui indique que la vulnérabilité est latente et n'est pas encore le facteur dominant qui structure la variance globale.

3.2. Analyse en Composantes Principales post crise :

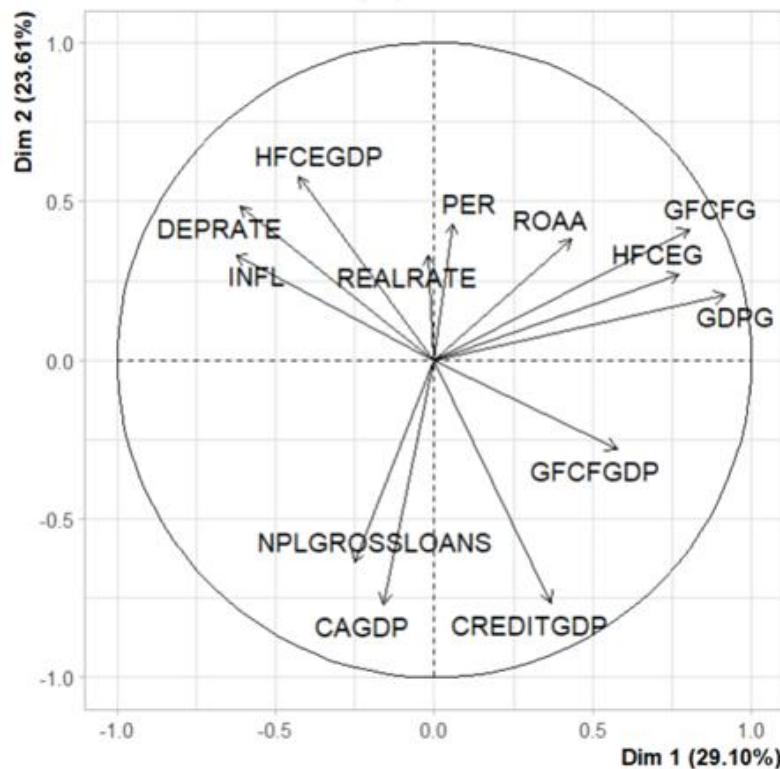


Figure III.8 : Cercle des corrélations des variables (ACP) post-crise

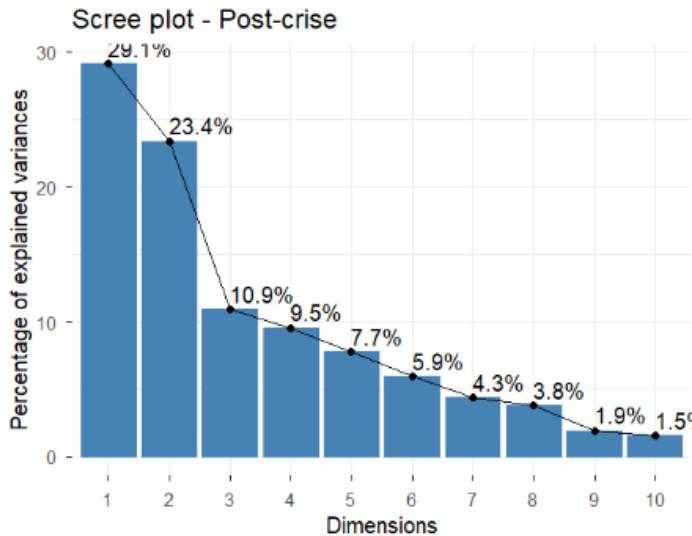


Figure III.9 : Le graphique de pourcentage de Variance Expliquée (ACP) post-crise

Dans la période post-crise, la première dimension baisse à 29,1 % et la deuxième augmente à 23,4 %, pour une variance cumulée de 52,5 %. La répartition plus équilibrée entre les deux axes montre que les relations entre variables deviennent plus complexes, ce notamment avec une recomposition des la hiérarchie des variables.

Les indicateurs bancaires prennent une importance nettement supérieure : la rentabilité des banques (ROAA) devient un contributeur plus important à la première dimension. À l'inverse, les tensions nominales deviennent marginales après la crise : l'inflation, le taux de change ou les taux réels perdent une grande partie de leur pouvoir explicatif.

Cette transition indique que la crise modifie en profondeur les mécanismes de vulnérabilité. Avant la crise, les déséquilibres macroéconomiques et monétaires dominaient. Après la crise, ce sont la solvabilité bancaire et la capacité des banques à soutenir la reprise qui deviennent centrales.

3.3. Classement des pays :

On peut également observer une évolution des scores factoriels entre les deux périodes permettant d'identifier les pays les plus touchés par la crise et ceux qui se sont améliorés. Le graphique ci contre, suivi par le tableau des récapitulation des scores moyens de la composante 1, résume cette évolution.

Top 10 pays : baisse moyenne CP1 (post - pre)

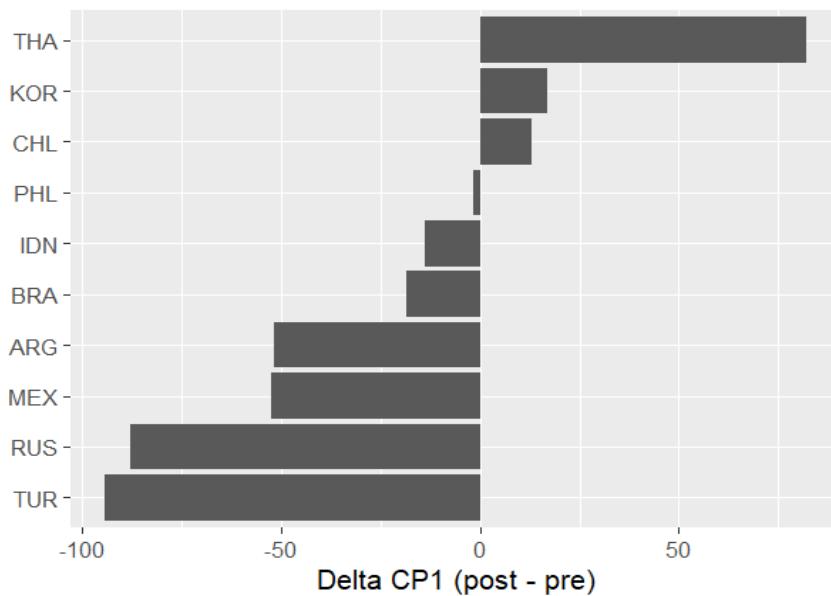


Figure III.10 : Variation moyenne du CP1 (post – pré) pour les 10 pays les plus affectés

Code Pays	CP1pre (Moyenne Avant)	CP1post (Moyenne Après)	Delta CP1 (Post–Pre)	Interprétation
TUR	59	-35	-94,10	Choc de performance le plus violent (Détérioration Sévère)
RUS	44,1	-43,80	-87,80	Choc de magnitude similaire (Perte totale de la performance)
MEX	15,10	37,40	-52,50	Forte détérioration
ARG	14,90	-36,70	-51,60	Forte détérioration
BRA	6,03	-12,60	-18,60	Détérioration modérée
IDN	0,82	-12,90	-13,70	Détérioration modérée
PHL	-2,86	-4,52	-1,66	Détérioration très faible
CHL	-12,90	0	12,90	Récupération (passage de performance négative à neutre)
KOR	-13,70	3,43	17,10	Récupération
THA	-35,80	46,10	81,90	La plus forte reprise

Tableau III.4: Comparaison du CP1 pré-crise et post-crise : variations moyennes et interprétation par pays

Cette analyse nous a permis de quantifier et de hiérarchiser l'impact de la crise sur la performance macroéconomique des pays. Elle distingue clairement les pays ayant subi un choc structurel durable sur leur santé économique (Turquie et Russie), de ceux ayant démontré une forte capacité de résilience et de rebond (Corée et Thaïlande). Cette divergence dans les évolutions illustre la variété des trajectoires de sortie de crise au sein de l'échantillon.

Les mécanismes de transmission diffèrent selon les régions. Les pays d'Asie du Sud-Est présentent une forte contagion régionale, tandis que la Russie et la Turquie suivent une dynamique propre liée à leurs fragilités structurelles. L'Amérique latine montre des trajectoires mixtes mais cohérentes avec les crises de dette et les fluctuations de capitaux observées à la même période.

Cette analyse confirme que les crises bancaires et financières provoquent non seulement des effets sur les variables macroéconomiques, mais aussi une reconfiguration de la structure des relations entre ces variables. Plusieurs limites doivent néanmoins être soulignées : l'échantillon présente un panel déséquilibré, les pondérations sont basées sur des choix a priori et l'ACP ne capte pas les relations non linéaires. Malgré ces limites, les résultats obtenus apportent un éclairage sur les mécanismes de vulnérabilité et de transmission des crises entre les années 1990 et le début des années 2000.

4. ANALYSE ECONOMETRIQUE – TESTS ET INTERPRETATION ECONOMIQUE :

4.1. Hypothèses économiques et la lecture de littérature :

En nous appuyant sur les résultats empiriques des recherches antérieures, l'analyse des différences dans l'occurrence de crise bancaire nous permet d'approfondir ce sujet. La variable endogène étudiée est donc l'occurrence de crise. Nous retenons 13 variables exogènes suivants:

(1) *ROAA – Rentabilité des actifs (%)*: Une rentabilité bancaire élevée améliore la solvabilité et la capacité d'absorption des chocs. Les banques plus profitables sont moins susceptibles d'entrer en crise.

(2) *NPLGROSSLOANS – Prêts non performants (%)* : Un niveau élevé de prêts non performants dégrade la qualité des actifs et accroît le risque de défaut systémique.

(3) *PER – Price-Earnings Ratio* : Un PER élevé peut signaler une surévaluation des marchés financiers ou une prise de risque excessive.

(4) *CREDITGDP – Crédit au secteur privé / PIB (%)* : Un développement financier sain soutient la croissance et renforce la stabilité bancaire. À un niveau normal d'expansion, le crédit améliore les conditions économiques.

(5) *GDPG – Croissance du PIB réel (%)*: Une croissance économique solide réduit les risques de défaut, améliore les revenus et renforce le système bancaire.

(6) *GFCFG – Croissance de la FBCF (%)* : Un boom de l'investissement peut parfois refléter une expansion excessive ou des bulles d'actifs, ce qui peut accroître les risques.

(7) *GFCFGDP – FBCF / PIB (%)*: Un investissement très élevé, s'il est mal alloué, peut renforcer les déséquilibres macroéconomiques.

(8) *HFCEG – Croissance de la consommation finale (%)*: Une consommation dynamique reflète une économie saine, réduisant les risques de défaut bancaire.

(9) *HFCEGDP – Consommation finale / PIB (%)* : Une part trop élevée de la consommation dans le PIB peut indiquer une faible épargne, réduisant la résilience financière.

(10) *CAGDP – Balance courante / PIB (%)* : Une détérioration du compte courant fragilise l'économie, accentuant les pressions financières externes.

(11) *INFL – Taux d'inflation (%)* : L'inflation élevée perturbe les taux d'intérêt, les marges bancaires et augmente les risques de défaut.

(12) *REALRATE – Taux d'intérêt réel (%)* : Dans certaines économies, un taux réel positif peut refléter un environnement monétaire discipliné et une inflation maîtrisée.

(13) *DEPRATE – Taux de dépôt (%)* : Un taux de dépôt élevé augmente le coût des ressources bancaires, réduisant les marges d'intermédiation et pouvant accroître la vulnérabilité du système.

4.2. Équation de recherche de modèle:

Après cadre théorique, nous avons choisi l'équation globale ci-dessous pour mieux expliquer la relation. Le modèle s'écrit sous la forme scalaire suivante:

$$\begin{aligned} \text{PCRISIS1}_{it} = & \beta_0 + \beta_1 \text{ROAA}_{it} + \beta_2 \text{NPLGROSSLOANS}_{it} + \beta_3 \text{PER}_{it} + \beta_4 \text{CREDITGDP}_{it} \\ & + \beta_5 \text{GDPG}_{it} + \beta_6 \text{GFCFG}_{it} + \beta_7 \text{GFCFGDP}_{it} + \beta_8 \text{HFCEG}_{it} + \beta_9 \text{HFCEGDP}_{it} + \beta_{10} \\ & \text{CAGDP}_{it} + \beta_{11} \text{INFL}_{it} + \beta_{12} \text{REALRATE}_{it} + \beta_{13} \text{DEPRATE}_{it} + \varepsilon_i \end{aligned}$$

Notant que i est l'indice de pays, t est l'indice de temps et ε est l'aléa

Les effets attendus sur les paramètres sont: $\beta_1 < 0$, $\beta_2 > 0$, $\beta_3 > 0$, $\beta_4 < 0$, $\beta_5 < 0$, $\beta_6 > 0$, $\beta_7 > 0$, $\beta_8 < 0$, $\beta_9 > 0$, $\beta_{10} > 0$, $\beta_{11} > 0$, $\beta_{12} < 0$, $\beta_{13} > 0$

4.3. Sélection du modèle de régression sur données de panel :

Pour les données de panel, il est évidemment primordial d'identifier quel modèle de régression le plus approprié afin d'estimer des estimateurs. Donc, nous donnons une processus incluant la comparaison entre des modèle proposé, utilisation de tests de spécification pour déterminer le modèle plus performant et sélection le modèle final.

En appuyant sur la procédure de spécification développée par Baltagi, Badi H. (2013) dans le cadre de la régression sur données de panel, nous considérons pour proposer trois types de modèles:

- Le modèle poolé (modèle 1 - Pooled OLS): les propriétés entre les pays est identique, il y a pas de différences entre eux.
- Le modèle à effets fixes (modèle 2 – Fixed Effects): raisonne que chaque pays possède ses propres identités, il n'est pas homogène l'un sur l'autre en terme de caractéristiques.
- Le modèle à effets aléatoires (modèle 3 - Random Effects) : Tous les caractéristiques entre les pays sont différer, mais cette différence est randomisée.

Dans un premier temps, nous réalisons une régression pour le modèle complet pour la spécification poolée (modèle 1) et spécification à effets fixes (modèle 2). Nous effectuons ensuite un test de Chow (F test for individuals effects) à ceux deux estimations, ce qui indique que le modèle 2 est plus adapté en termes de statistique que l'autre. Le résultat montre une p-value de 0,0141, inférieure au seuil de 5 % (*annexe 3*). C'est-à-dire, nous rejetons l'hypothèse nulle d'absence d'effets individuels ou le modèle 1 n'est pas pertinent. Donc, nous visons à tenir le modèle 2 pour effectuer les étapes postérieures.

Dans le but de respecter la contrainte structurelle du modèle à effets aléatoire (modèle 3) concernant l'égalité entre le nombre d'individus et celui de coefficients estimées, y compris le constant. Nous prenons en compte une réalisation d'élimination des variables non significative de manière progressivement descendante (Backward elimination) sur le modèle 2. Le modèle réduit permettant de réaliser la régression de modèle 3 est suivante: $PCRISIS1_{it} = \beta_0 + \beta_1 ROAA_{it} + \beta_3 PER_{it} + \beta_5 GDPG_{it} + \beta_8 HFCEG_{it} + \beta_9 HFCEGDP_{it} + \beta_{10} CAGDP_{it} + \beta_{11} INFL_{it} + \beta_{12} REALRATE_{it} + \beta_{13} DEPRATE_{it} + \varepsilon_i$

Après avoir retiré progressivement 4 variables non significatives GFCRGDP, NPLGROSSLOANS, CFCFG, GFCFGDP, nous appliquons le test de Hausman pour comparer le modèle 2 et modèle 3 de même variables explicatives. Le résultat de ce test montre $p_value = 0,78 > 5\%$ (*annexe 4*). Nous ne rejetons pas l'hypothèse nulle selon laquelle l'estimateur à effets fixés est moins performant. Autrement dit, l'estimateur à effets aléatoires est plus cohérent et contribue à constituer une spécification valable, voire préférable de manière économétrique.

En outre, nous prenons en compte le test LM de Breusch–Pagan en vue de vérifier la pertinence de conservation de modèle 1 (*annexe 5*). Toutefois, comme analyse antérieure en raison de test Chow, nous prenons la décision que le modèle retenu finale est celui à effets aléatoires.

En conséquence, dans l'objectif d'améliorer la qualité de notre estimation et de ne conserver que les variables appropriées, nous procéderons la méthode «backward

elimination». Rappelons que nous supprimons les variables en raison à la fois d'une faible significativité économétrique et d'une contribution marginale aux dimensions de vulnérabilité identifiées par l'ACP, puis estimer à nouveau via régression linéaire jusqu'à ce que toutes les variables soient toutes significatives.

4.4. Résultat de régression panelle du modèle :

L'utilitaire analyse conduit au tableau suivant:

	Coefficient	Écart-type	Z_value	p-value
Constant	0.3924	0.068	5.7718	7.844e-09 ***
ROAA*	-0.0436	0.0127	-3.4446	0.0005719 ***
INFL *	-0.0133	0.0045	-2.9454	0.0032250 **
CAGDP*	0.0235	0.0082	2.8491	0.0043843 **
DEPRATE*	0.0134	0.0051	2.6086	0.0090905 **
<i>R</i> ²	0.19197		Chisq-statistic	26.4122
<i>R</i> ² ajusté	0.16285		Prob(Chisq-statistic)	2.613e-05

*variables significatives au risque de première espèce 5%

Tableau IV.5: Résultat de régression du modèle sur les données pannel

Le modèle final conserve quatre variables significatives: ROAA, INFL, CAGDP, DEPRATE. Le nouveau modèle est comme suite:

$$PCRISIS1_{it} = 0.3924 - 0.0436 \text{ ROAA } \beta_{it} - 0.0133 \text{ INFL } \beta_{it} + 0.0235 \text{ CAGDP } \beta_{it} + 0.0134 \text{ DEPRATE } \beta_{it} + \varepsilon_i$$

$$(0.068) \quad (0.0127) \quad (0.0045) \quad (0.0235) \quad (0.0051)$$

À partir des données analysées, nous pouvons en déduire que le modèle renferme un coefficient de détermination $R^2 = 0.19197$, montrant que 19.197% des variations de variable dépendante sont expliqués par les variables explicatives retenues. Bien que nous constatons que cette reste relativement modeste, elle traduit pourtant d'une capacité d'explication dans le cadre de contexte d'un phénomène macro-économique complexe. Phénomènes rares et fortement non linéaires, où l'objectif principal est l'identification des déterminants significatifs plutôt que la prédiction parfaite.

En plus, quant à $R^2_{ajusté} = 0.16285$, il indique aussi que 16,285% de la variance est expliquée, après avoir corrigé pour le nombre de variables explicatives incluses. En conséquence, la différence entre R^2 et $R^2_{ajusté}$ est quasi petite (0.03). Nous pouvons arriver à la conclusion que le modèle global reste donc raisonnablement bon et tient une bonne capacité explicative.

À propos de test de significativité sur les paramètres, la statistique de Chi-caré (χ^2) affiche $\text{Chi-sq} = 26.4122$, $p\text{-value} \approx 2,61 \times 10^{-5}$ est largement significative. Ainsi, nous rejetons l'hypothèse nulle selon laquelle l'ensemble des coefficients est jointement égal à zéro et le modèle contient la qualité globalement significative.

Au sujet de la significativité individuelle des variables, effectué par le test de Wald ou de p-value, indique que les quatres variables renferme des valeurs absolues de $z_{\text{observé}}$ supérieurs à $z_{\text{critique}} = 1,96$ et ses p-value inférieurs à 0,05. Donc, ses coefficients sont également significativement différents de 0 à 95% et ont un impact significatif sur la présence crise.

À titre de test de robustesse, un modèle logit à effets aléatoires a été estimé. Les résultats confirment la stabilité des conclusions principales : les variables ROAA, CAGDP, INFL et DEPRATE conservent les mêmes signes que dans le modèle linéaire de panel. Le changement de signe du terme constant est sans interprétation économique et est inhérent à la spécification non linéaire du modèle logit (*annexe 6*). L'utilisation d'erreurs standards robustes clusterisées par pays ne modifie ni le signe ni la significativité des coefficients estimés. (*annexe 7*)

Dans la vérification de type Mundlak, les coefficients des variables ROAA, INFL, CAGDP et DEPRATE conservent leurs signes et restent significatifs, tandis que les termes moyens par pays (ROAA_mean, INFL_mean, CAGDP_mean, DEPRATE_mean) ne sont pas significatifs, ce qui va dans le sens d'une spécification à effets aléatoires acceptable. (*annexe 8*).

D'un point de vue général, nous remarquons une influence positive de la crise issus de *la balance courante / PIB* et *Taux de dépôt* contre un impact négatif venant de rendement des actifs (Return on Assets) et taux d'inflation. Ce résultat confirme à nouveau la même conclusion avec un nombre réduit de variables. L'approfondissement de l'effet de quatres variables sur l'occurrence de crise sera présenté dans la partie suivante.

4.5. L'interprétation des variables exogènes retenues sur la présence de l'occurrence de crise:

4.5.1. L'effet du rendement des actifs (%)

Le coefficient associé au ROAA (= +0.0436) est négatif et significatif (p-value = 0,00057). Économiquement, cette estimation souligne qu'une augmentation d'un point de pourcentage du ROAA entraîne une baisse moyenne de 4,36% la probabilité d'occurrence de crise. Ce résultat est conforme à la théorie de la stabilité bancaire développée par Demirguc-Kung & Huizinga (2000) selon laquelle un niveau de rentabilité plus élevé permet aux banques à renforcer sa capacité à couvrir ses pertes latentes et les aide à améliorer sa solidarité durant des chocs macroéconomiques potentiels.

4.5.2. L'effet de *balance courante / PIB* (%)

Nous trouvons qu'un effet significativement positif de la balance courante / PIB sur l'occurrence. Concrètement, une augmentation d'un point de pourcentage de solde courant accroît en moyenne 2,36 % la probabilité de crise bancaire. Ce résultat pourrait apparaître

contre-intuitif, car un solde courant excédentaire est souvent traduit par un bon équilibre macro. Cependant, nous raisonnons qu'un excédent forcément élevé contribue à masquer une appréciation de la monnaie, une sur dépendance aux activités d'exportation ou une suraccumulation de capitaux internes. Ces dernières contribuent objectivement à perturber le secteur bancaire à travers d'une diminution soit de compétitivité externe, soit de distribution du crédit.

4.5.3. Taux d'inflation (%)

Le coefficient lié à l'inflation a un effet négatif et très significatif sur la occurrence de crise ($p\text{-value} = 0.00322$). Une augmentation de 1% de taux d'inflation entraîne une diminution en moyenne de 1,33% des crises. Cet effet est inverse de notre attente qu'une augmentation de taux d'inflation puisse accroître la probabilité de présence de crise.

4.5.4. L'effet de Taux de dépôt (%)

Le coefficient associée au taux de dépôt est positif et forcément significatif ($p\text{-value} = 0.00909$). Concrètement, une augmentation de 1 % du taux de dépôt entraîne une augmentation en moyenne la probabilité de crise bancaire de 1,34 %. Cet effet est souligné dans la littérature sur la libéralisation financière (McKinnon, 1973 ; Shaw, 1973). Une augmentation de volume des crédits sera proposée par les banques suite à l'attractivité des dépôts aux taux d'intérêt élevés. Cela explique le phénomène très réputé « crédit boom – buste cycle) est considéré comme l'un des prédicteurs les plus robustes des crises bancaires.

Ces observations amènent à poser la question suivante, dans la continuité logique de notre étude :

« Est-ce que le modèle est bien adopté des propriété d'une modèle Bleu pour faire inférence?»

La série de tests de normalité, d'homoscédasticité et de multicolinéarité dans la partie suivante sera présentée pour vérifier l'hypothèse diagnostique afin de garantir la fiabilité de qualité et la robustesse du modèle à effets aléatoires réduit retenu finalement.

4.6. Tests diagnostiques :

4.6.1. Test de normalité :

Afin de vérifier si le terme de résidus suivent la loi normale, nous appliquons le test de Shapiro-Wilk aux résidus du modèle (*annexe 9*). Le résultat révèle une $p\text{-value}$ quasi nulle, ce qui mène à la conclusion de rejet de l'hypothèse de normalité.

Pourtant, ce résultat n'est pas extrêmement un grave problème en raison de binarité de variable endogène. En plus, le test de normalité est peu pertinent dans le cadre d'un modèle

non linéaire ou à effets aléatoires. En conséquence, l'absence de normalité des résidus ne génère aucune doute sur la validité de l'estimation

4.6.2. Test de l'homoscédasticité :

Le test de Breusch-Pagan appliqué au modèle indique une **p-value de 0,79**, largement supérieure au seuil de 5 % (*annexe 9*). Autrement dit, il y a une absence d'hétéroscléasticité dans le modèle où la statique de variance de résidus sur diagonal est constant

4.6.3. Test de multi colinéarité :

L'analyse de la multicolinéarité grâce à la méthode Variance Inflation Factors (VIF) met en évidence 2 côtés (*annexe 9*). D'une part, VIF de 2 variables ROAA et CAGDP sont proches de 1, une absence presque totale de corrélation entre ces variables explicatives. D'autre part, celle de 2 variables INFL et DEPRATE reste au voisinage de 6, révélant une corrélation modérée, mais il est dans intervalle plausible de 0 à 10. Cette analyse montre que le modèle représente une absence de multi colinéarité significativement.

CONCLUSION

Cette étude avait pour objectif d'examiner les déterminants structurels des crises bancaires dans les pays émergents au cours d'une décennie fortement marquée par l'instabilité financière. Grâce à une approche combinant ACP pondérée, méthodes de clustering, analyse pré/post-crise et régression de panel, elle propose une compréhension cohérente et multidimensionnelle des vulnérabilités macro-financières observées entre 1993 et 2002.

Les analyses factorielles ont permis d'identifier deux axes dominants: la profondeur financière et la fragilité bancaire, renforçant l'importance du crédit, de la solvabilité et des déséquilibres externes dans la dynamique des crises.

Le clustering a mis en évidence trois profils de pays cohérents, distinguant les économies en expansion, les économies sévèrement frappées par les crises, et un groupe intermédiaire plus résilient.

L'étude pré/post-crise a montré une transformation des mécanismes de vulnérabilité : avant la crise, les tensions nominales dominaient ; après la crise, ce sont les indicateurs bancaires qui deviennent structurants.

Sur le plan économétrique, le modèle à effets aléatoires retenu met en avant quatre variables significatives :

- ROAA, dont l'effet stabilisateur confirme la littérature sur la résilience bancaire ;
- L'inflation, dont l'effet négatif est contre-intuitif mais cohérent avec la structure de l'échantillon ;
- La balance courante et le taux de dépôt, qui renforcent la probabilité de crise en lien avec les mécanismes de surchauffe financière

Ces résultats montrent que les crises bancaires émergent de l'interaction entre fragilités structurelles, tensions externes et déséquilibres financiers internes, et soulignent l'importance d'une surveillance conjointe du système bancaire et des conditions macroéconomiques globales.

RÉFÉRENCE BIBLIOGRAPHIQUES

- Baltagi, B. H. (2013). *Econometric Analysis of Panel Data* (5th ed.). John Wiley & Sons.
- Demirgüç-Kunt, A., & Huizinga, H. (2000). *Financial structure and bank profitability*. World Bank Policy Research Working Paper No. 2430.
<https://openknowledge.worldbank.org/entities/publication/1f3293c6-b003-54f4-9d03-1f27ac6e93f4>
- Eichengreen, B., & Hausmann, R. (1999). *Exchange rates and financial fragility*. NBER Working Paper No. 7418
https://www.nber.org/system/files/working_papers/w7418/w7418.pdf?utm_source=chatgpt.com
- Kaminsky, G., & Reinhart, C. (1999). *The twin crises: The causes of banking and balance-of-payments problems*. <https://ssrn.com/abstract=2882512>
- McKinnon, R. I. (1973). *Money and Capital in Economic Development*. Brookings Institution.
- Schularick, M., & Taylor, A. M. (2012). *Credit booms gone bust: Monetary policy, leverage cycles, and financial*
- Shaw, E. S. (1973). *Financial Deepening in Economic Development*. Oxford University Press.

ANNEXES

Annexe 1 : Le tableau de contributions des modalités le tableau de qualité de représentation (\cos^2) sans pondération	2
Annexe 2 : Le tableau de contributions des modalités et le tableau de qualité de représentation (\cos^2) dans ACP avec pondération.....	2
Annexe 3 : Le résultat test de Chow (F test for individuals effects	3
Annexe 4 : Le résultat de test de Hausmen	3
Annexe 5 : Le résultat de Test de Breusch-Pagan.....	3
Annexe 6 : Résultats du modèle logit à effets aléatoires (test de robustesse).....	3
Annexe 7 : Résultats du modèle à effets aléatoires avec erreurs standards robustes.....	4
Annexe 8 : Vérification de robustesse – Approche de Mundlak.....	4
Annexe 9 : Le tableau des résultats de Tests Diagnostiques	5

	Dim.1	Dim.2	Dim.3	Dim.4	Dim.5
ROAA	0.3130431	0.4370664	0.162181092	-0.28002641	0.62016177
NPLGROSSLOANS	-0.3178078	-0.6552033	-0.131368956	0.18314052	0.07950359
PER	0.3684356	0.2820878	-0.085186787	0.04783800	-0.22911833
CREDITGDP	0.3832075	-0.7478251	0.198338327	-0.17311594	-0.11015777
GDPG	0.8973087	0.1608695	0.014708901	0.15772197	0.14570119
GFCFG	0.8213202	0.3377099	-0.061046135	0.20624738	0.13158022
GFCFGDP	0.6697369	-0.2533108	0.395168298	-0.18891032	-0.43886580
HFCEG	0.7828180	0.2399288	0.081087713	0.22267147	0.08336901
HFCEGDP	-0.3992111	0.5239370	-0.571693481	0.24715611	-0.11584429
CAGDP	-0.2984146	-0.5982497	0.250867376	0.32057477	0.52953870
INFL	-0.4769146	0.4499501	0.646943653	0.29777679	-0.06618515
REALRATE	-0.2998352	0.2948104	0.007985806	-0.73624513	0.14802606
DEPRATE	-0.5186390	0.5180436	0.522072142	0.09359106	-0.13922215

Annexe 1 : Le tableau de contributions des modalités le tableau de qualité de représentation (\cos^2) sans pondération

	Dim.1	Dim.2	Dim.3	Dim.4	Dim.5
ROAA	-0.007628179	0.487824031	0.033291106	0.337535840	0.031106245
NPLGROSSLOANS	0.183352722	-0.837929111	-0.010531414	0.008523124	0.475212991
PER	0.010568198	0.033320575	-0.002632377	-0.006909907	-0.001321204
CREDITGDP	0.786186832	-0.365366688	0.327299106	0.228139259	-0.105278549
GDPG	0.486140573	0.515789287	0.079099569	-0.065799609	0.300074291
GFCFG	0.265191678	0.426343721	0.015874112	-0.076931896	0.257852759
GFCFGDP	0.382351307	0.133287089	0.263018307	0.031253108	-0.137538833
HFCEG	0.269437684	0.361204033	0.092148527	-0.077316606	0.239356456
HFCEGDP	-0.328758191	0.036044098	-0.266653540	-0.150373429	0.193720016
CAGDP	0.055116857	-0.340584659	0.115551658	-0.058889791	0.005352607
INFL	-0.569090216	0.034783650	0.506399591	-0.131620959	0.038519393
REALRATE	-0.316839731	0.008053294	-0.102414363	0.662232837	0.148934070
DEPRATE	-0.597948502	0.024873141	0.437257808	0.047288879	0.163513523

Annexe 2 : Le tableau de contributions des modalités et le tableau de qualité de représentation (\cos^2) dans ACP avec pondération

F test for individual effects

```
data: PCRISIS1 ~ ROAA + NPLGROSSLOANS + PER + CREDITGDP + GDPG + GFCFG + ...
F = 2.4626, df1 = 10, df2 = 68, p-value = 0.0141
alternative hypothesis: significant effects
```

Annexe 3 : Le résultat test de Chow (F test for individual effects)

Hausman Test

```
data: PCRISIS1 ~ ROAA + PER + GDPG + HFCEG + HFCEGDP + CAGDP + INFL + ...
chisq = 5.5972, df = 9, p-value = 0.7795
alternative hypothesis: one model is inconsistent
```

Annexe 4 : Le résultat de test de Hausmen

Lagrange Multiplier Test - (Breusch-Pagan)

```
data: PCRISIS1 ~ ROAA + PER + GDPG + HFCEG + HFCEGDP + CAGDP + INFL + ...
chisq = 2.0887, df = 1, p-value = 0.1484
alternative hypothesis: significant effects
```

Annexe 5 : Le résultat de Test de Breusch-Pagan

```
Generalized linear mixed model fit by maximum likelihood (Laplace Approximation) ['glmerMod']
Family: binomial ( logit )
Formula: PCRISIS1 ~ ROAA + CAGDP + INFL + DEPRATE + (1 | iso3c)
Data: database
      AIC      BIC      logLik -2*log(L)  df.resid
 130.5466 147.0681  -59.2733  118.5466      110
Random effects:
 Groups Name        Std.Dev.
 iso3c (Intercept) 0.4175
Number of obs: 116, groups: iso3c, 12
Fixed Effects:
(Intercept)       ROAA        CAGDP        INFL        DEPRATE
   -0.2198     -0.9361      0.1130     -0.1176      0.1285
```

Annexe 6 : Résultats du modèle logit à effets aléatoires (test de robustesse)

```
t test of coefficients:

              Estimate Std. Error t value Pr(>|t|)
(Intercept) 0.3924265 0.0598005 6.5623 1.753e-09 ***
ROAA        -0.0436270 0.0070602 -6.1793 1.091e-08 ***
CAGDP        0.0234636 0.0067872 3.4570 0.0007751 ***
INFL         -0.0133323 0.0051828 -2.5724 0.0114194 *
DEPRATE      0.0133967 0.0059864 2.2379 0.0272255 *
---
Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Annexe 7 : Résultats du modèle à effets aléatoires avec erreurs standards robustes

```
Oneway (individual) effect Random Effect Model
(Swamy-Arora's transformation)
```

```
Call:
plm(formula = PCRISIS1 ~ ROAA + INFL + CAGDP + DEPRATE + ROAA_mean +
     INFL_mean + CAGDP_mean + DEPRATE_mean, data = pdata, model = "random")
```

Balanced Panel: n = 10, T = 10, N = 100

Effects:

var	std.dev	share	
idiosyncratic	0.17630	0.41988	0.907
individual	0.01813	0.13463	0.093
theta:	0.2978		

Residuals:

Min.	1st Qu.	Median	3rd Qu.	Max.
-0.76902	-0.26755	-0.12779	0.35760	1.00906

Coefficients:

	Estimate	Std. Error	z-value	Pr(> z)
(Intercept)	0.4325979	0.1377976	3.1394	0.0016931 **
ROAA	-0.0563343	0.0189631	-2.9707	0.0029709 **
INFL	-0.0369594	0.0093730	-3.9432	8.04e-05 ***
CAGDP	0.0216858	0.0092596	2.3420	0.0191819 *
DEPRATE	0.0366941	0.0110068	3.3338	0.0008567 ***
ROAA_mean	-0.0103684	0.0681206	-0.1522	0.8790241
INFL_mean	0.0369176	0.0290194	1.2722	0.2033125
CAGDP_mean	0.0342966	0.0377496	0.9085	0.3635988
DEPRATE_mean	-0.0362205	0.0310877	-1.1651	0.2439756

```
---
Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Total Sum of Squares: 22.218

Residual Sum of Squares: 16.043

R-Squared: 0.27793

Adj. R-Squared: 0.21445

Chisq: 35.0261 on 8 DF, p-value: 2.6447e-05

Annexe 8 : Vérification de robustesse – Approche de Mundlak

Tests Diagnostiques	Résultats
Shapiro-Wilk normality test	W = 0.86747 p-value = 9.057e-09
Studentized Breusch-Pagan test	BP = 1.6847 df = 4 p-value = 0.7935
VIF	ROAA : 1.013014 CAGDP : 1.140784 INFL : 6.335399 DEPRATE : 6.342912

Annexe 8 : Le tableau des résultats de Tests Diagnostiques

CODE

```
\install.packages("corrplot") #Matrice de covariance
install.packages("FactoMineR") #ACP
install.packages("dplyr")
install.packages("tidyverse")
library(dplyr)
library(tidyverse)
library(FactoMineR)
library(corrplot)

#=====
# Partie 1: Télécharger des données et prendre la pondération
#=====

# Import file ews
database <- read.csv("ews.csv")

# Voir les 6 premières lignes
head(database)

# Vérifier la structure de database
str(database)

# Nettoyage de données
colSums(is.na(database))
database_clean <- database %>% mutate(across(where(is.numeric), ~ ifelse(is.na(.),
mean(., na.rm = TRUE), .)))

summary(database)
# Standardisation (z-score)
vars <- c("ROAA", "NPLGROSSLOANS", "PER", "CREDITGDP", "GDPG", "GFCFG",
"GFCFGDP", "HFCEG", "HFCEGDP", "CAGDP", "INFL", "REALRATE", "DEPRATE")
database_z <- database_clean %>% mutate(across(all_of(vars), scale))

# Pondération
weights <- c(ROAA = 0.8, NPLGROSSLOANS = 1, CREDITGDP = 1,
```

```

GDPG = 0.8, INFL = 0.8, CAGDP = 0.6,
GFCFG = 0.6, GFCFGDP = 0.6, HFCEG = 0.6, HFCEGDP = 0.6,
PER = 0.1, REALRATE = 0.8, DEPRATE = 0.8)
database_w <- database_z %>% mutate(across(all_of(names(weights)), ~ . * 
weights[cur_column()]))
```

```

#=====#
# Partie 2: ANALYSE EN COMPOSANTES PRINCIPALES
#=====#
# Partie 2.1: Analyse de la matrice de corrélation
#=====#
corr <- cor(database_w[,vars], use = "complete.obs")
corrplot(corr,
  method = "color",
  col = colorRampPalette(c("lightblue", "white", "pink"))(200),
  tl.col = "black",
  tl.cex = 1,
  tl.font = 1,
  type = "upper",
  diag = FALSE
)
```

```

#=====#
# Partie 2.2: Analyse en composantes sans pondération
#=====#
res.pca.simple <- PCA(database_z[, vars], scale.unit = FALSE, graph = FALSE)
plot(res.pca.simple, choix = "var")
res.pca.simple$var$coord
```

```

#=====#
# Partie 2.3: Analyse en composantes avec pondération
#=====#
res.pca.w <- PCA(database_w[, vars], scale.unit = FALSE, graph = FALSE)
plot(res.pca.w, choix = "var")
res.pca.w$var$coord
```

```
#Scree plot individuel  
plot(res.pca.w, choix = "ind")
```

```
# Visualisation avec un barplot des pourcentage de Variance expliquée  
eig <- res.pca.w$eig[, "percentage of variance"]  
barplot <- barplot(eig,  
                     main = "Pourcentage de Variance expliquée",  
                     names.arg = 1:nrow(res.pca.w$eig),  
                     col = "lightblue",  
                     border = NA,  
                     xlab = "Composante Principale",  
                     ylab = "% de variance expliquée",  
                     ylim = c(0, max(eig) * 1.1))  
lines(barplot, eig, type = "o", pch = 19, col = "darkblue", lwd = 2)  
text(x = barplot,  
      y = eig + max(eig)*0.05,  
      labels = round(eig, 1),  
      cex = 0.9, font = 2)
```

```
#Visualisation avec un barplot des pourcentage de Variance expliquée cumul  
cum_eig <- cumsum(eig)  
barplot(cum_eig,  
        main = "Variance Cumulée Expliquée",  
        names.arg = 1:length(cum_eig),  
        col = "lightblue",  
        xlab = "Composante Principale",  
        ylab = "% de variance cumulée",  
        ylim = c(0, max(cum_eig) * 1.15)  
)  
text(x = barplot,  
      y = cum_eig + max(cum_eig)*0.03,  
      labels = round(cum_eig, 1),  
      cex = 0.9, font = 2)
```

```

#=====
#  Partie 3: CLUSTERING HIÉRARCHIQUE (HCPC/KMEANS)
#=====

#Clustering (k-means)
pc_scores <- res.pca.w$ind$coord[, 1:3]

#k-means
set.seed(123)
km <- kmeans(pc_scores, centers = 3)
database_clean$cluster <- km$cluster

#Profil des clusters
aggregate(database_clean[, vars],
          by = list(database_clean$cluster),
          FUN = mean)
HACP = HCPC(res.pca.w, nb.clust = 3)

#=====
#  Partie 4: ANALYSE EN COMPOSANTES PRINCIPALES PRE/POST-CRISE
#=====

#  Partie 4.1: Analyse en Composantes Principales post-crise
#=====

required <-
c("dplyr","tidyverse","FactoMineR","factoextra","corrplot","strucchange","ggplot2","reshape2")
for(pkg in required){
  if(!requireNamespace(pkg, quietly = TRUE)) install.packages(pkg)
  library(pkg, character.only = TRUE)
}
post <- database_w %>% filter(PCRISIS1 == 1)
pca_post <- PCA(post[, vars], graph = FALSE)
plot(pca_post, choix = "var")
pca_post$var$coord

```

```

#=====
# Partie 4.2: Analyse en Composantes Principales pre-crise
#=====

pre <- database_w %>% filter(PCRISIS1 == 0)
pca_pre <- PCA(pre[, vars], graph = FALSE)
plot(pca_pre, choix = "var")
pca_pre$var$coord

#=====
# Partie 4.3: Classement de pays
#=====

pre_df <- database_w %>% filter(PCRISIS1 == 0)
post_df <- database_w %>% filter(PCRISIS1 == 1)

cat("Observations pré-crise :", nrow(pre_df), "\n")
cat("Observations post-crise:", nrow(post_df), "\n")

# Exécuter PCA sur subsets pré & post
res.pca.pre <- PCA(pre_df[, vars], scale.unit = FALSE, graph = FALSE)
res.pca.post <- PCA(post_df[, vars], scale.unit = FALSE, graph = FALSE)

#SCREE PLOT pre/post crise
png("scree_pre.png", width = 800, height = 600)
fviz_eig(pca_pre, addlabels = TRUE, main = "Scree plot - Pré-crise")
dev.off()
png("scree_post.png", width = 800, height = 600)
fviz_eig(pca_post, addlabels = TRUE, main = "Scree plot - Post-crise")
dev.off()

scores_pre <- as.data.frame(res.pca.pre$ind$coord)
scores_pre$iso3c <- pre_df$iso3c
scores_pre$obs_id <- rownames(pre_df)

scores_post <- as.data.frame(res.pca.post$ind$coord)
scores_post$iso3c <- post_df$iso3c
scores_post$obs_id <- rownames(post_df)

```

```

mean_pre_by_country <- scores_pre %>% group_by(iso3c) %>%
summarise(mean_CP1_pre = mean(Dim.1, na.rm = TRUE), n_pre = n())
mean_post_by_country <- scores_post %>% group_by(iso3c) %>%
summarise(mean_CP1_post = mean(Dim.1, na.rm = TRUE), n_post = n())

country_cmp <- full_join(mean_pre_by_country, mean_post_by_country, by = "iso3c")
%>%
mutate(mean_CP1_pre = ifelse(is.na(mean_CP1_pre), 0, mean_CP1_pre),
mean_CP1_post = ifelse(is.na(mean_CP1_post), 0, mean_CP1_post),
delta_CP1 = mean_CP1_post - mean_CP1_pre) %>%
arrange(delta_CP1) # tri asc : les plus négatifs = plus forte baisse

# Exporter top 20
write.csv(country_cmp, "country_cp1_pre_post_comparison.csv", row.names = FALSE)

# Top 10 pays les plus durement touchés (plus grande baisse de CP1)
top10_hurt <- head(country_cmp, 10)

```

```

#=====
# Partie 5: ANALYSE ECONOMETRIQUE – TESTS ET INTERPRETATION
ECONOMIQUE :Regression panel
#=====

install.packages("plm")
library(plm)
pdata <- pdata.frame(database, index = c("iso3c", "date"))
summary(pdata)

#=====
# Partie 5.1: Sélection du modèle de régression sur données de panel
#=====

#Modele 1: Pooled OLS
#Modèle complet:
pooling <- plm(PCRISIS1 ~ ROAA + NPLGROSSLOANS + PER + CREDITGDP +
GDPG + GFCFG + GFCFGDP + HFCEG + HFCEGDP + CAGDP + INFL +
REALRATE + DEPRATE,
      data = database,
      model = "pooling")
summary(pooling)

#Modele 2:FIXED EFFECTS
#Modèle complet:
fe <- plm(PCRISIS1 ~ ROAA + NPLGROSSLOANS + PER + CREDITGDP + GDPG +
GFCFG + GFCFGDP + HFCEG + HFCEGDP + CAGDP + INFL + REALRATE +
DEPRATE,
      data = pdata,
      model = "within")

#Test de chow
pFtest(fe, pooling)

#Modele 3: RANDOM EFFECTS
#Pour analyser cette modèle, nous devrons assurer que le nombre de coefficients, y
compris le constant, est égal au celui de individus. Comme le modèle pooling MCO est
meilleur que celui Fixed Effet. Nous visons à réaliser une élimination des variables

```

explicative via méthode Backward jusqu'à avoir le même nombre de variables explicatives et des individus.

#Via cette dernière, nous retirons 4 variables le plus non significatives où p_value le plus élevé, ce sont respectivement NPLGROSSLOANS, GFCRGDP,CFCFG, GFCFGDP

#Modele FE de 9 variables

```
fe1 <- plm(PCRISIS1 ~ ROAA + PER + GDPG + HFCEG + HFCEGDP + CAGDP +  
INFL + REALRATE + DEPRATE,  
            data = pdata,  
            model = "within")  
summary(fe1)
```

#Modèle complet de RE de 9 variables

```
re <- plm(PCRISIS1 ~ ROAA + PER + GDPG + HFCEG + HFCEGDP + CAGDP +  
INFL + REALRATE + DEPRATE,  
           data = pdata,  
           model = "random")
```

#Modele de pooled de 9 varaibles

```
pooling1 <- plm(PCRISIS1 ~ ROAA + PER + GDPG + HFCEG + HFCEGDP + CAGDP  
+ INFL + REALRATE + DEPRATE,  
                  data = pdata,  
                  model = "pooling")
```

#Le test de Hausman

```
phptest(fe1, re)
```

#Test de Breusch-Pagan

```
plmttest(pooling1, type = "bp")
```

```

#=====
# Partie 5.2: Modèle final et interprétation des données
#=====

#Modèle final:
re <- plm(PCRISIS1 ~ ROAA + CAGDP + INFL + DEPRATE,
           data = pdata,
           model = "random")
summary(re)
print(re)

#robustness check - Résultats du modèle logit à effets aléatoires
install.packages("lme4")
library(lme4)
logit_re <- glmer( PCRISIS1 ~ ROAA + CAGDP + INFL + DEPRATE + (1 | iso3c),
                   data = database,
                   family = binomial(link = "logit"))
print(logit_re)

# Test de robustesse – Modèle à effets aléatoires avec erreurs standards robustes
coeftest(re, vcov = vcovHC(re, type = "HC1", cluster = "group"))

#Approche de Mundlak
library(dplyr)
pdata <- pdata %>%
  group_by(iso3c) %>%
  mutate(
    ROAA_mean = mean(ROAA),
    INFL_mean = mean(INFL),
    CAGDP_mean = mean(CAGDP),
    DEPRATE_mean = mean(DEPRATE) )
re_mundlak <- plm( PCRISIS1 ~ ROAA + INFL + CAGDP + DEPRATE +
                    ROAA_mean + INFL_mean + CAGDP_mean + DEPRATE_mean,
                    data = pdata,
                    model = "random" )
summary(re_mundlak)

```

```
#=====
# Partie 5.3: Tests diagnostiques
#=====
#Test de normalité
res_re <- residuals(re)
shapiro.test(res_re)

#Test de homoscadasticité
install.packages("lmtest")
install.packages("car")
library(lmtest)
library(car)
bptest(re)

#Test de multi colinéarité
df <- as.data.frame(pdata)
lm_for_vif <- lm(formula(re), data = df)
vif(lm_for_vif)
```