

Louvain School of Management

Utilisation d'un Early Warning System et calcul de la probabilité de crise financière en 2022

Auteure : Gillard Camille
Promoteur : Cadelon Bertrand
Année académique 2022-2023
Travail de fin d'études (TFE) en vue d'obtenir le titre de Master en
Sciences de Gestion - International Finance

Remerciements

Je tiens tout d'abord à remercier mon promoteur, Monsieur Bertrand Cadelon, qui m'a encadré dans la réalisation de ce mémoire. Sa disponibilité et ses précieux conseils m'ont permis de mener à bien ce projet, et même de me dépasser.

Je remercie également la Louvain School of Management pour toutes les connaissances et compétences que j'ai pu développer en son sein, et pour m'avoir permis de décrocher le job de mes rêves. Je travaille actuellement dans une super boîte en tant que Financial Controller et je suis pleinement épanouie.

Je souhaite également remercier mes amis et collègues qui m'ont soutenu moralement, conseillé, relu et rassuré chaque fois que le doute a pu m'envahir. La rédaction d'un mémoire est extrêmement stressante et ils ont pu m'aider à donner le meilleur de moi-même lors de cette ultime étape de mon cursus.

Enfin, j'adresse une pensée particulière à mon Papy, décédé du Covid. Il a toujours cru en moi, même dans les moments de doute ou dans l'échec, et ce tout au long de ma scolarité. Il ne pourra assister à ma remise de diplôme mais je sais qu'il aurait été fier de moi.

Table of Contents

INTRODUCTION	7
PARTIE 1	9
CONTEXTE GÉO-ÉCONOMIQUE EN FIN D'ANNÉE 2022	9
LA GUERRE EN UKRAINE.....	10
L'INFLATION AUX USA.....	13
LE CAS DE LA CHINE	13
LA DÉPRÉCIATION DE L'EURO	14
L'IMPACT DU PRIX DES MATIÈRES PREMIÈRES SUR LE RISQUE DE CRISE	15
CONCLUSION	15
SYNTHÈSE THÉORIQUE ET REVUE DE LITTÉRATURE	17
LES DIFFÉRENTS TYPES DE CRISES FINANCIÈRES.....	17
<i>Les crises économiques</i>	17
<i>Les crises financières</i>	17
QU'EST-CE QU'UN <i>EARLY WARNING SYSTEM</i> (EWS)?.....	19
<i>Définition</i>	19
DIFFÉRENTS MODELES DE CALCULS.....	20
<i>The signaling approach - Kaminsky, Lizondo and Reinhart (1998)</i>	20
<i>The multivariate logit approach - Demirgüç-Kunt and Detragiache (1998)</i>	22
<i>The Binary Recursive Tree approach - Duttagupta and Cashin (2008)</i>	23
<i>Conclusion</i>	25
<i>MACHINE LEARNING</i> ET PRÉDICTION DES CRISES FINANCIÈRES	26
LA CRISE DE 2008	27
PARTIE 2	29
MÉTHODOLOGIE	29
DONNÉES.....	30
VARIABLES DÉPENDANTES	31
VARIABLES INDÉPENDANTES	33
ANALYSE EMPIRIQUE	36
<i>Régression bivariée</i>	36
<i>Modèle probit</i>	42
<i>Discussion des résultats</i>	45
<i>Limites et recommandations</i>	46
<i>Matrice de confusion</i>	46
<i>Résultats</i>	48
CONCLUSION	48
BIBLIOGRAPHIE	49
ANNEXES	51
<i>Graphes</i>	51
<i>Tables</i>	52
<i>Résultats statistiques</i>	54

Introduction

Il semblerait qu'au cours de ces dernières décennies, le nombre de crises financières ait augmenté. En fait, le monde financier s'accorde pour dire que l'augmentation de ces crises financières coïncideraient avec les accords de Bretton-Woods en 1971, et la fin du dollar convertible en or. Contrairement aux crises financières qu'on a pu connaître lors des siècles et des décennies passés, le risque systémique des crises actuelles est bien plus élevé en raison de la forte mondialisation et des interconnexions des économies dans le monde. Des chercheurs ont donc mis au point des modèles de calculs et d'analyses pour tenter de prédire des crises financières, c'est à dire des Early Warning System. En effet, au plus vite une crise financière est suspectée, au plus vite des solutions peuvent être mises en place pour l'éviter ou, du moins, en minimiser les conséquences. J'ai choisi ce sujet pour mon mémoire car j'ai toujours été passionnée par l'histoire des crises économiques et financières passées. Je trouvais donc intéressant et pertinent d'étudier la prévision de crise. Je voulais d'abord me concentrer sur la crise économique liée à la crise sanitaire du Covid19, mais deux raisons m'ont poussé à aller plus loin. La première étant que je trouvais difficile de pouvoir prédire cette crise économique pour la simple raison qu'elle n'a pas éclaté à la suite d'un déséquilibre financier. En effet, aucun indicateur n'aurait pu prédire cette crise dont la cause première n'est pas un choc financier mais une crise sanitaire, et un arrêt quasi total et brutal de la production et de la consommation. Ensuite, au vu de la situation socio-économique extrêmement tendue aux quatre coins du globe en ce moment, et au vu des médias qui, jours après jours, publient des gros titres parlant d'une potentielle récession mondiale et d'un potentiel krach boursier, je me suis dit qu'il pourrait être challengeant d'utiliser ces Early Warning System pour prédire ou non l'occurrence de cette récession dans les mois à venir, et plus particulièrement une récession dans la zone Euro.

Ma question de recherche est donc la suivante : "**Peut-on prédire une financière dans la zone Euro pour l'année 2022?**" Cette question va couvrir un évènement qui est en train de se passer en ce moment même et je pense qu'il est important de répondre à cette question d'un point de vue statistique. En effet, la prédiction d'une crise financière peut avoir des conséquences sur le déroulement de celle-ci car des politiques et mesures peuvent-être mises en place en amont. Cette question peut également couvrir un gap dans la littérature étant

donné qu'il s'agit d'une prédiction d'un potentiel évènement futur, dans un contexte économique tout à fait propice au déclenchement de cet évènement, et que chaque recherche à ce sujet aura comme limitation principales les données disponibles à l'instant t .

Ce mémoire sera structuré en deux parties principales. La première partie comportera un point sur l'actualité économique mondiale suivi d'une synthèse théorique et une vue historique de l'utilisation des EWS dans la prédiction de crises financières passées. La seconde partie portera sur une étude empirique qui tentera de calculer la probabilité de l'occurrence ou non d'une crise financière pour l'année 2022. Je ne peux malheureusement pas me pencher sur l'année 2023 maintenant à cause du manque de données disponibles au moment où je commence mon étude quantitative.

Pour terminer cette introduction, je dirais que ce travail de recherche comporte pas mal de limites. Pour commencer, il y a le manque de recul et le manque de données sur la situation. La littérature nous montrera qu'il est plus facile de prédire une crise à posteriori, une fois toutes les données disponibles. Ensuite, il y a le problème de récolte des données. En effet, une fois mon étude quantitative commencée, je devrai me stopper à un instant t et utiliser les données disponibles à cet instant. Il est possible que l'étude quantitative nous donne un tout autre résultats avec des données récoltées à l'instant ou les lecteurs liront ces lignes. Il est donc également possible que l'économie ait pris un autre tournant en quelques mois. J'ai donc décider de me baser sur le contexte de fin 2022 afin de ne pas trop devoir revoir mes chiffres et mes hypothèses en cours de route. Enfin, il y a également le fait que pas mal de mes sources proviennent de médias. En effet, les médias sont une source d'informations quelquefois peu fiables et il est parfois très difficile de vérifier les chiffres avancés par la presse (notamment les chiffres concernant la Chine). J'ai donc pris soins d'analyser mes sources ainsi que le contenu relaté avec attention et avec beaucoup d'esprit critique.

Partie 1

Contexte géo-économique en fin d'année 2022

Au covid et à toutes ses retombées économiques mondiales viennent s'ajouter plusieurs nouvelles situations de crise partout dans le monde. Plusieurs géants économiques mondiaux font face à des situations de crise et se trouvent dans des situations critiques. Il y a tout d'abord la guerre en Ukraine ou encore la crise du marché immobilier qui touche la Chine, toujours partiellement confinée. Ensuite, vient cette vague d'inflation énorme, particulièrement aux USA, partiellement due à une relance importante de la consommation post-covid, mais également due au prix toujours croissant de l'énergie.

Il est important de souligner que cette analyse comporte des limites assez importantes. En effet, il relate beaucoup de faits d'actualités et je n'ai bien sûr pas pu baser cette partie de mes recherches sur des articles scientifiques mais bien sur des articles de presse et autres médias. J'ai principalement utilisé des articles publiés sur des sites officiels comme le site de l'OCDE ou encore celui de la commission européenne. J'ai également utilisé plusieurs éditions du journal "The New-York Times", et les informations relatées ont toujours été vérifiées au maximum. D'autres sources médiatiques ont également été nécessaires à la rédaction de ce chapitre et il n'est donc pas à exclure qu'une erreur ait pu se glisser dans mes données, notamment pour le cas de la Chine, dont les chiffres officiels sont très compliqués à obtenir. Il faut aussi garder en tête que tous les évènements relatés évoluent jours après jours, que les chiffres évoluent eux aussi jours après jours, c'est pourquoi il y aura probablement un décalage entre les faits présentés dans ce mémoire et les faits tels qu'ils le sont au moment de la lecture de celui-ci. Cependant, le but de ce mémoire n'est pas de jouer le rôle de la presse et l'exactitude du récit de ces faits d'actualités n'est pas ce qu'il y a de plus important ici, mais je trouvais quand même important que le lecteur puisse se remettre dans le contexte des évènements en cours à l'instant où je commence mon étude empirique.

La guerre en Ukraine

Premièrement, il y a cette guerre entre l'Ukraine et la Russie, qui a éclaté le 24 février 2022, soit il y a un peu plus d'un an. L'Union Européenne a réagi instantanément en imposant toute une série de sanctions et de restrictions envers la Russie, certains membres du Conseil National de la Fédération de Russie, mais également contre le président russe Vladimir Poutine lui-même. Entre le 24 février et le mois de juin 2022, pas moins de six trains de sanctions ont été mis en place à la suite de l'invasion de l'Ukraine par la Russie. Ces sanctions touchent principalement le secteur financier, le secteur de l'énergie et des transports et les importations et exportations. D'autres secteurs sont également touchés tels que les secteurs des médias ou des cryptoactifs. L'UE a également imposé des sanctions à l'encontre de la Biélorussie et a débloqué un soutien de 500 millions d'euros pour les forces armées ukrainiennes. Toutes ces sanctions ont pour objectif de mettre la Russie au pied du mur et de faire cesser l'invasion de l'Ukraine.

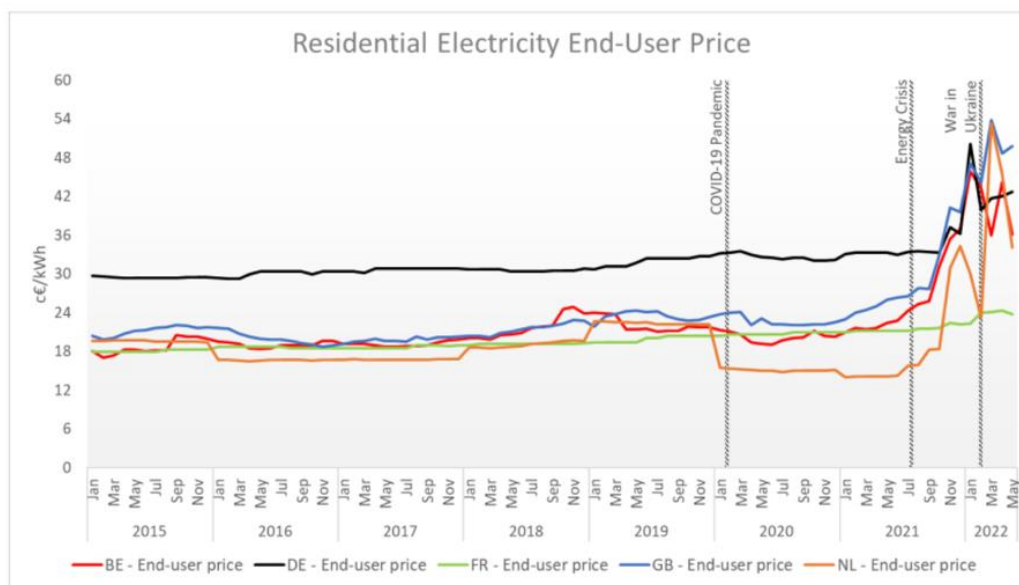
En dehors de toutes ces sanctions, la Russie et l'Ukraine ont incontestablement un impact sur l'économie mondiale via le marché des matières premières. En effet, selon le Cairn, les deux pays représentent approximativement 30% des exportations mondiales de blé, ainsi que 20% pour le maïs, le gaz naturel et les engrais minéraux, et finalement 11% pour le pétrole. Les deux pays disposent également d'une réserve importante d'uranium. La réduction des importations de céréales et d'huiles végétales touche principalement le secteur agricole ainsi que le secteur agro-alimentaire européen, et la préoccupation de l'UE est de garder les prix suffisamment abordable pour ces biens de premières nécessités.

Pour finir, l'importation de gaz russe est bien évidemment une des plus grosses problématiques en cette fin d'année 2022. En effet, la Russie a déjà ralenti ses exportations de gaz depuis l'année 2021. Depuis l'éclatement de la guerre, elle a déjà stoppé une partie de son approvisionnement en gaz à l'Europe, et les dirigeants européens craignent la cessation totale de ces exportations russes. Pour l'instant, une politique de réduction de la consommation a été mise en place pour contrer la hausse des prix et la baisse des réserves. Dans le cas où la Russie déciderait de stopper complètement l'approvisionnement en gaz, on se retrouverait dans une situation de pénurie en Europe. En attendant, la réduction

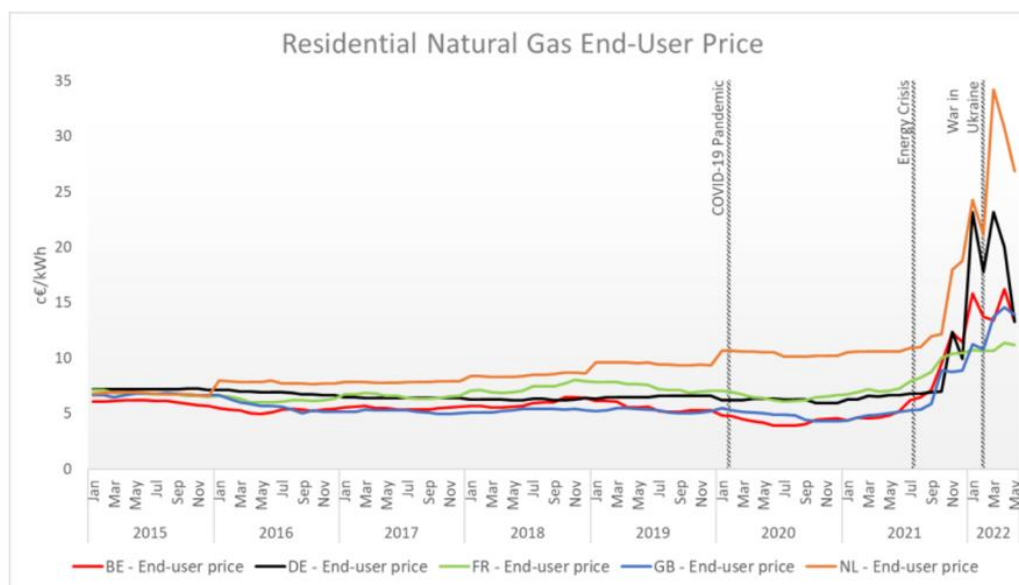
d'approvisionnement de gaz en Europe à un impact considérable sur son prix, mais également sur le prix de l'électricité. En effet, le nucléaire et le renouvelable n'étant pas suffisant pour palier à la demande, de l'électricité supplémentaire est fournie grâce à des centrales à gaz.

Graphe 1 et 2 : Évolution des prix de l'électricité et du gaz naturel pour le consommateur final, de janvier 2015 à Mai 2022. ([Source : creg.be](http://creg.be))

Graphe 1



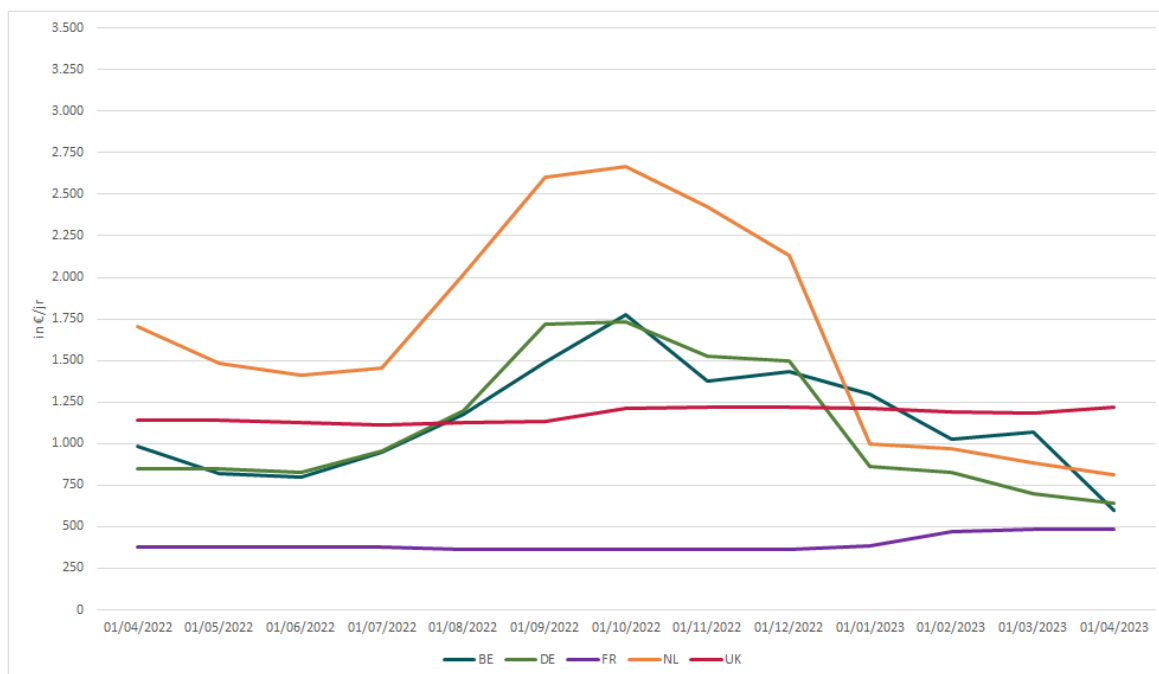
Graphe 2



Graphe 3 et 4 : Focus sur les 12 derniers mois de l'évolution de l'électricité et du gaz naturel pour un usage résidentiel ([Source : creg.be](http://creg.be)).

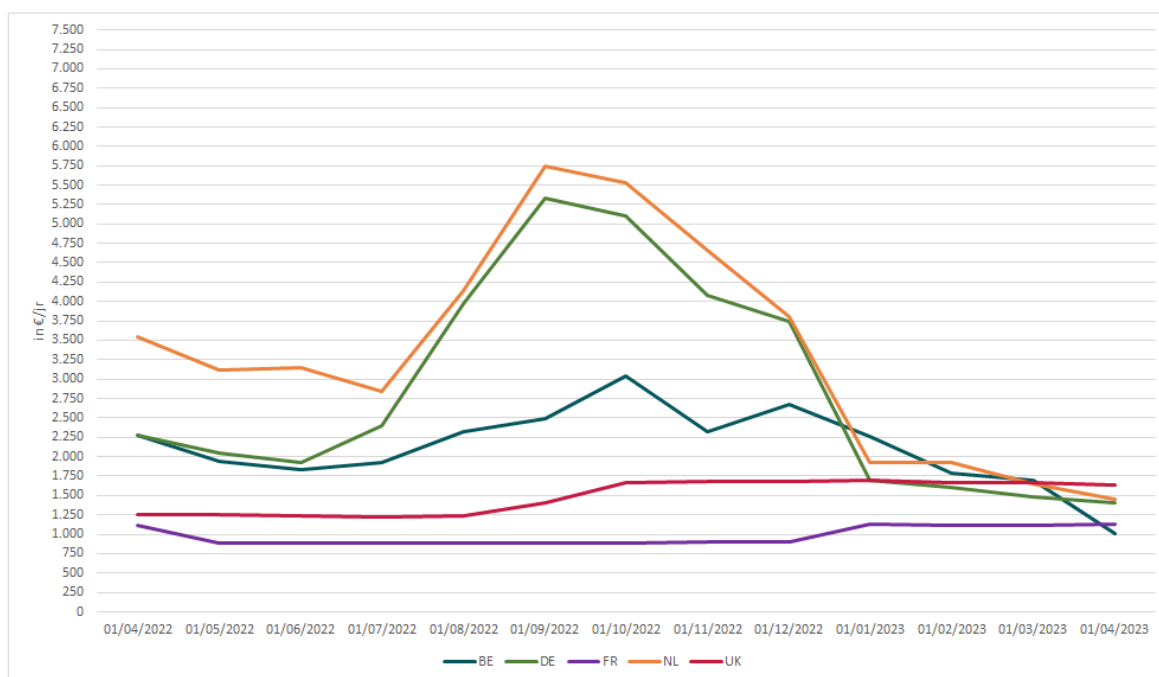
Graphe 3

Evolution prix électricité (composante énergie) pour usage résidentiel (3,5 MWh)



Graphe 4

Evolution prix gaz naturel (composante énergie) pour usage résidentiel (17,00 MWh)



Ces graphes montrent une nette évolution des prix de l'électricité et du gaz en Europe de l'Ouest, qui a commencé durant le Covid19 et qui s'est amplifié suite à l'invasion de l'Ukraine. Une baisse est cependant enregistrée à la mi-automne 2022, mais les prix actuels restent néanmoins supérieurs à ce qu'ils n'étaient avant l'été 2021.

L'inflation aux USA

Aux États-Unis aussi, la période post Covid19 a été marquée par une forte inflation et a atteint son point le plus haut en juin 2022 avec un taux d'inflation de 9,1%¹ pour redescendre juste en dessous des 8% au mois d'octobre. En effet, la période covid a énormément ralenti la croissance économique et l'effet de rareté de beaucoup de produits a eut un rôle à jouer. L'épargne accumulée pendant le covid a également permis aux consommateurs de se montrer plus conciliant envers la hausse des prix, et les salaires ont continués à être ajustés pour maintenir le pouvoir d'achat, ce qui a encore amplifié le phénomène.

Le cas de la Chine

La Chine connaît à l'heure actuelle une grave crise du secteur immobilier. Le pays a connu une importante croissance économique ces dernières années et dégage le deuxième plus gros PIB, après les USA. Selon des estimations, le secteur immobilier chinois représenterait 15 à 30% du PIB et la banque centrale chinoise, régie par le gouvernement, utilise ce secteur comme variable d'ajustement. Si le marché est en difficulté, il faut à tout prix le relancer en prêtant à des taux très faibles. Beaucoup de promoteurs immobiliers chinois pensaient que les prix continueraient de croître, mais la crise du covid et des politiques sanitaires strictes ont fait fortement ralentir la croissance et la demande et donc, baisser les prix. Beaucoup de promoteurs immobiliers surendettés et à court de liquidités se retrouvent en faillite, et sont obligés de cesser l'avancement de nombreux chantiers. Des milliers de propriétaires se retrouvent avec un bien dont la construction est suspendue et cessent de

¹ Source: U.S. bureau of labor statistics (<https://fr.tradingeconomics.com/united-states/inflation-cpi>)

payer leurs crédits par mécontentement. Le chiffre d'affaire du secteur est en chute libre et, comme cela coûte cher d'entretenir des chantiers à l'arrêt, le gouvernement a imposé la destruction pure et simple de bon nombre d'entre eux. Des millions de logements ont ainsi été détruits. Plusieurs mesures de relance du marché par la banque centrale chinoise, telle que la subvention d'achat ou l'abaissement de ses taux directeurs, sont restées sans succès. Il faut également préciser que la Chine connaît toujours, à l'heure actuelle, des confinements stricts. Début septembre 2022, les estimations indiquaient un millier de nouveaux cas chaque jour, confinant les plus grosses villes. Cette politique zéro covid coûte cher à l'état, déjà au bord de la récession. Cette crise a également des conséquences sur la zone Euro puisqu'un ralentissement de l'économie chinoise pourrait entraîner une baisse de ses importations. On ne note cependant pas vraiment de dépréciation notable du Yuan (CNY) durant l'année 2022.

La dépréciation de l'Euro

Pour la première fois depuis 2002, l'euro (EUR) est passé sous la barre du dollar (USD). En effet, le taux EUR/USD est passé à 0,96 à la fin du mois de septembre 2022, pour remonter à la hausse et atteindre la parité fin octobre (voir graphe 5). Il s'agit d'une des plus grosses baisses que l'Euro ait connu ces vingt dernières années. Il y a deux raisons principales à cette dépréciation. La première étant purement géographique. En effet, la zone Euro est bien plus exposée par la guerre en Ukraine que les USA, de par sa proximité et son importation énergétique importante, ce qui renforce l'image du dollar comme valeur refuge. De plus, les taux d'intérêts sont nettement plus intéressants aux USA qu'en Europe, et ce en raison des plans de relance massifs visant à lutter contre l'inflation américaine. Le but étant de freiner la consommation et de pousser à l'épargne, les investisseurs ont donc préféré se tourner vers l'USD, plus rentable. Cependant, l'Euro ne se déprécie pas face à certaines monnaies telles que la livre sterling (GBP) ou encore au Yen (JPY), face à lesquelles à contrario, il s'apprécie. Toutefois, le prix de certaines énergies telles que les énergies fossiles (pétrole et gaz) sont facturées en USD, ce qui contribue à augmenter encore leur prix et à accélérer l'inflation en Europe, et qui ne s'arrangera pas en cas de pénurie de gaz russe.

Graphe 5 : Évolution du cours de l'EUR/USD sur dix ans ([source : ecb.europa.eu](https://www.ecb.europa.eu))



L'impact du prix des matières premières sur le risque de crise

L'augmentation du prix des matières premières comme par exemple le pétrole, les métaux, les céréales ou l'huile a un impact important sur l'économie mondiale. Tout d'abord, cela ralentit la production et, comme expliqué plus haut, ça augmente l'inflation. Le PIB et l'inflation sont, nous le verrons plus tard, deux variables explicatives des crises financières. Une dégradation de ces deux facteurs augmente donc le risque de crise. Ensuite, l'augmentation du prix des matières premières incite à spéculer.

Conclusion

En règle générale, lorsque l'inflation est trop forte, les différentes banques centrales vont décider d'augmenter leurs taux directeurs pour tenter de contrôler l'inflation et d'essayer de la faire baisser à un taux plus bas. Cette augmentation des taux a pour objectif de ralentir l'économie et donc de diminuer la demande, en rendant le crédit plus cher et l'épargne plus attrayante. C'est ce que font les banques centrales en ce moment. Cependant, cette crise n'est pas aussi simple à résoudre. En effet, les principales raisons de cette inflation explosive sont l'augmentation des prix de l'énergie et du carburant, mais également des denrées alimentaires, qui sont tous trois des biens de premières nécessité. Beaucoup de

ménages ne consomment en réalité pas plus qu'à leurs habitudes et beaucoup se serrent même la ceinture pour pouvoir payer leur facture d'énergie à la fin du mois. Alors que nous faisons face à une inflation mondiale élevée depuis la guerre en Ukraine, nous assistons également à un ralentissement de l'économie et à une situation de stagflation, ainsi qu'à une dépréciation de l'Euro par rapport au dollar. Beaucoup de médias s'accordent à dire qu'une récession mondiale et un krach boursier sont à prévoir dans les prochains mois, et toutes ces rumeurs pourraient potentiellement donner lieu à des mouvements de panique et à une théorie autoréalisatrice.

Synthèse théorique et revue de littérature

Les différents types de crises financières

Dans ce mémoire, différentes crises économiques vont être abordées pour illustrer ma revue de littérature. Nous entrerons plus en profondeur sur les éléments déclencheurs de ces crises économiques et sur les modèles de calculs associés à des types de crise spécifiques. Il est donc important de bien définir dès le début quels sont les différents types de crises financières.

Les crises économiques

Une crise économique peut se définir comme un retournement de la conjoncture économique qui entraîne une dépression. Il y a un décalage entre la productivité et la consommation, ce qui entraîne une hausse du chômage et des faillites, mais également une baisse du pouvoir d'achat et du PIB. La crise économique est source de tension sociale et touche la plupart des agents économiques d'un pays. Une crise économique est bien souvent induite par une crise financière sous-jacente. Par exemple, la récession mondiale de 2008 a été induite par la crise des subprimes aux USA.

Les crises financières

Une crise financière est une perturbation économique engendrée par des problèmes liés au système financier d'un pays, à savoir le système monétaire, bancaire, aux marchés des devises, etc. Il y a différents types de crises financières, qui peuvent apparaître indépendamment ou interagir les unes avec les autres. En voici quelques exemples :

- La crise bancaire, se produit lorsqu'une ou plusieurs banques d'un même système financier se retrouvent à court de liquidités (à cause de retraits massifs de capitaux), et que les banques se retrouvent en faillite ou au bord de celle-ci.
- La crise de la dette : se produit lorsqu'un pays ne peut plus payer ses créanciers à la suite d'une autre crise financière.
- La crise boursière : Se traduit par une baisse des cours des actions plus ou moins brutale sur les marchés boursiers.

- La crise de change, se traduit par une dépréciation plus ou moins importante d'une monnaie. La difficulté étant de déterminer le seuil au-delà duquel une dépréciation sera considérée comme une crise de change. Selon les travaux de Reinhart et Rogoff sur la datation des crises financières (2008), on peut considérer qu'un *currency crash*, défini comme une dépréciation annuelle de la monnaie par rapport à l'USD, apparaît à partir d'un seuil de dépréciation d'au moins 15%. (Voir table 2 en annexe).

Il est important de ne pas confondre les crises monétaires et les crises de change. Dans un pays ou dans une union économique qui utilise un taux de change variable, il est possible que les variations d'offre et de demande étrangères aient pour conséquence une dépréciation de la monnaie, ce qui est quelque chose d'involontaire. Dans un pays utilisant un taux de change fixe, la dévaluation peut être un instrument financier servant, par exemple, à dynamiser les exportations. Ce travail s'attardera sur l'Euro, qui est une union économique utilisant un régime de change flottant.

Ces différentes crises sont le fruit d'un dérèglement dans un système financier et elles mènent quelquefois à une crise économique plus généralisée dans un pays. Nous nous concentrerons sur les deux modèles principaux de crises : les crises bancaires et monétaires. Il existe plusieurs facteurs qui sont à l'origine de ces crises financières comme par exemple des tensions financières au niveau mondial, des tensions géopolitiques, une mauvaise gestion, des politiques macro-économiques inadaptées, etc. C'est pourquoi les chercheurs se sont attardés sur des solutions pour surveiller certains facteurs et parvenir à prédire une crise financière, dans le but d'en réduire au maximum les conséquences et, dans le meilleur scénario, de l'éviter. Il existe donc des modèles de calcul, appelés Early Warning System, qui tentent de prédire ces crises, principalement les crises bancaires et monétaires.

Qu'est-ce qu'un *Early Warning System* (EWS)?

Définition

Selon Kelman et M.H. Ghantz, aucune définition universellement acceptée n'existe ou n'existera un jour. La Stratégie Internationale des Nations Unies pour la prévention des catastrophes définit un EWS comme étant "l'ensemble des capacités nécessaires pour générer et diffuser des informations d'alerte opportunes et significatives afin de permettre aux individus, aux communautés et aux organisations menacés par un danger de se préparer et d'agir de manière appropriée et suffisamment tôt pour réduire la possibilité de dommages ou de pertes". Beaucoup de définitions dans la littérature existante portant sur la réduction de risque s'accordent pour dire que l'EWS est un processus social qui a pour but d'éviter les chocs dus aux aléas, à toutes les échelles spatiales, et à tous les niveaux technologiques. Les chocs peuvent-être de toutes les natures possibles : financiers, écologiques, météorologiques, etc. Concernant les EWS de nature financière, le processus est basé sur la lecture et l'analyse de plusieurs indicateurs macroéconomiques. A titre d'exemple, Gavin et Houseman (1995), et Sachs, Torn ell, et Velasco (1996) utilisent la croissance du crédit pour prédire les boom du crédit tandis que Mish Kin (1996) préférerait la baisse du prix des actions. Calve (1996) proposait d'utiliser la *broad money* et les réserves de change. Lors de variations trop importantes de ces indicateurs, une alerte est alors donnée. Les indicateurs de crises sont des indicateurs binaires, ce qui signifie que la valeur de 1 est donnée lors de l'occurrence d'une crise tandis que la valeur de 0 est donnée sinon. La littérature s'accorde à dire qu'un ESW, quel qu'il soit, possède quatre composantes qui sont la connaissance du risque, le monitoring et l'alerte, la communication et diffusion de l'information, et enfin la capacité de réaction. Nous nous concentrerons sur les EWS dans le domaine de la finance, dont le but est de donner l'alerte en cas de suspicions de crise financière, afin que les gouvernements et banques centrales puissent désamorcer au mieux ces crises par le biais de politiques monétaires adaptées ou via d'autres solutions.

Il existe dans la littérature plusieurs modèles de calculs, qui utilisent tous des indicateurs macroéconomiques différents selon le type de crise étudiée, le contexte politique ou économique d'une région etc. Par exemple, E. Philip Davis et Dilruba Karim Brunel ont

utilisé deux approches différentes pour tenter de prédire la crise des subprimes à postériori. La première étant *the multivariate logit approach*, suivie de l'approche du *binary recursive tree*, que nous prendrons le temps d'expliquer plus tard. Leurs travaux ont montré que chacun des deux modèles ne parvenait qu'à prédire partiellement la crise des subprimes car ils n'étaient pas assez adaptés aux pays développés, mais qu'en étant mieux ajustés et utilisés de façon complémentaire, les chances de pouvoir prédire la crise auraient pu être bien meilleures.

Différents modèles de calculs

La littérature existante nous introduit deux principaux types d'EWS. Le premier est le *Signaling model*, ou *signaling approach*, (Kaminsky, & Reinhart, 1999), tandis que le second modèle est le *limited dépendant variable model*, élaboré par Berg et Pattillo (1999), ainsi que par Kumar, Moorthy et Perraudin (2003). La plupart des modèles existants et utilisés par les différentes institutions (comme par exemple l'IMF), sont des dérivés de ces deux modèles principaux. Il en existe bien sûr d'autres dont certains seront présentés plus loin.

The signaling approach - Kaminsky, Lizondo and Reinhart (1998)

Cette méthode, mise au point par Kaminsky, Lizondo et Reinhart (1998), a pour but de détecter à la fois les crises bancaires et monétaires et consiste à utiliser une quantité importante de variables qui pourraient potentiellement être utile dans la détection d'une crise. Cette méthode analyse les comportements des variables seules, des mois avant la crise et des mois après, pour déterminer leur comportement global pendant et en dehors des périodes de crise. Si par la suite, une variable présente un comportement anormal, un signal est alors émis. Cette méthode utilise des seuils estimés pour chacune de ces variables, appelés *noise-to-signal ratio*, qui minimisent le ratio entre le nombre de crises manquées et le nombre de crises détectées correctement. Le mieux serait d'adapter ce signal à chaque pays de l'échantillon étudié afin d'obtenir des ratios optimaux puisque ces ratios prendraient en compte les éléments de chacun de ceux-ci. Le seuil auquel le signal sera émis ($S_i^j = 1$) est donné par la fonction suivante :

$$\{S_i^j = 1\} = \{|X_i^j| > |X_{*i}^j|\}$$

ou

$$\{S_i^j = 0\} = \{|X_i^j| \leq |X_{*i}^j|\}$$

Où

i est un indicateur univarié

j est un pays particulier

S est la variable signal

X_i^j est l'indicateur lié à l'indicateur i et au pays j

X_{*i}^j est le seuil de normalité de l'indicateur lié à l'indicateur i et au pays j

Une fois ce modèle implémenté, il nous montre la tendance de la probabilité de l'occurrence d'une crise. Si cette tendance est positive, alors les probabilités de crise augmentent. Si cette tendance est au contraire négative, on s'éloigne alors de l'occurrence d'une crise. Une difficulté que présente ce modèle est d'interpréter le résultat. En effet, les chercheurs indiquent que ce n'est pas parce qu'une variable dépasse à peine le *noise-to-signal ratio* qu'il faut forcément décréter une crise. Pour donner un exemple, Peter J G Vlaar indique dans ses travaux qu'une valeur deux fois plus grande que le signal représente un signal fort. Il ne faut donc pas utiliser uniquement ce signal comme simple indicateur de crise. Concernant les indicateurs de crises utilisés dans cette étude, Kaminsky and Reinhart (1999) ont rapporté que les indicateurs minimisant le mieux le *noise-to-signal ratio* et avec la plus haute probabilité de crise de change seraient les suivants :

1. *The real exchange rate deviation from a deterministic trend*
2. *The occurrence of a banking crisis*
3. *The export growth rate*
4. *The stock price index rate*
5. *The M2/reserves growth rate*
6. *The output growth*
7. *Excess M1 balances*
8. *Growth of internal reserves*

9. *M2 multiplier growth*

10. *The growth rate of the domestic credit to GDP ratio*

Ce modèle est néanmoins critiqué par Demirgüç-Kunt et Detragiache (2005) sur le fait qu'il analyse chaque variable individuellement et non dans leur ensemble. En effet, cette méthode peut poser des problèmes d'interprétation des résultats lorsqu'un indicateur indique une crise et pas un autre, ou qu'un indicateur ne dépasse que très légèrement son seuil de normalité, comme mentionné plus haut. Cette méthode ne permet donc pas une lecture globale de ses résultats, ce qui est cependant le cas du modèle *logit* de ces mêmes auteurs, qui sera développé juste après.

The multivariate logit approach - Demirgüç-Kunt and Detragiache (1998)

Un des modèles les plus présents dans la littérature est *the multivariate logit approach*, mise au point par Demirgüç-Kunt et Detragiache. Ce modèle a pour but de relier la probabilité d'une *systemic banking crisis* en fonction d'un vecteur de variables explicatives. Il peut néanmoins être utilisé dans d'autres cas de figures. Les auteurs de ces travaux soulignent la corrélation d'une croissance lente du PIB, d'une forte inflation et de taux d'intérêts élevés avec les crises financières. Ce modèle est donné par la fonction suivante :

$$Prob(Y_{it} = 1) = F(\beta X_{it}) = \frac{e^{\beta'X_{it}}}{1 + e^{\beta'X_{it}}}$$

Où,

Y_{it} est la variable *dummy* pour la crise bancaire dans un pays i à un instant t .

β est le vecteur des coefficients.

X_{it} est le vecteur des variables explicatives.

$F(\beta X_{it})$ est la distribution logistique cumulative.

Y_{it} Prend la valeur de 0 lorsque aucune crise n'est détectée et la valeur de 1 lorsqu'une crise est détectée. L'avantage de cette approche paramétrique comparée aux autres est qu'elle prend en compte l'interdépendance des variables explicatives et nous donne une lecture globale des résultats de tous les indicateurs. 1 si crise, 0 si pas. On ne doit pas

interpréter le résultat de chaque indicateur de manière indépendante comme pour la méthode *signaling approach*. Il a également été prouvé que cette méthode comportait moins de risque d'erreurs de type I et de type II² que la première méthode. Un des avantages de ce modèle qui est l'interdépendance entre les variables se révèle être néanmoins aussi une difficulté. En effet, le comportement de certains paramètres est bien souvent influencé par la crise elle-même, ce qui en vient à fausser certaines relations. Pour éviter ce problème, les périodes de crises sont bien souvent retirées des échantillons lors du calibrage des seuils d'alertes.

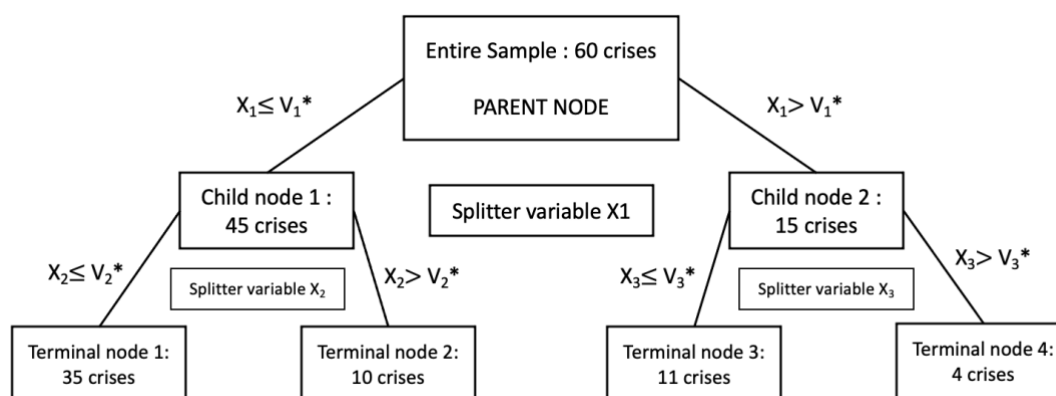
The Binary Recursive Tree approach - Duttagupta and Cashin (2008)

Cette méthode a été utilisée dans le passé par Duttagupta et Cashin (2008), pour tenter de prédire des crises bancaires. Cette procédure consiste à analyser un échantillon de données pour donner une valeur particulière à une variable explicative qui pourrait expliquer au mieux la variable dépendante. Les données représentent un ensemble de pays sur une période plus ou moins longue avec comme valeur 1 si l'année est classée comme une période de crise, et une valeur de 0 si l'année n'est pas classée comme une période de crise. Selon les chercheurs, la croissance du PIB pourrait bien être la meilleure variable explicative. Déterminer un seuil critique exact nous conduit alors à séparer les périodes de crise (*value = 1*) des périodes de calme (*value = 0*). Une fois ce premier tri effectué, on se retrouve avec à gauche, un premier groupe de données sous le seuil de croissance du PIB critique et, à droite, le second groupe de données au-dessus du seuil critique. La deuxième étape consistera à faire un second tri dans ces deux groupes et à maximiser le nombre de période de crise dans le premier groupe, et à le minimiser dans le second groupe. Notons que les seuils critiques ne sont pas toujours des maximums mais peuvent également être des minimums, il suffit de regarder le signe du coefficient ou d'interpréter l'indicateur. On déduira instinctivement qu'une crise a plus de chance de survenir dans un pays où la croissance du PIB est très lente que dans un pays où celle-ci est plus rapide. A l'inverse, on peut déduire qu'une forte inflation sera mauvaise comparé à un seuil plus bas aux alentours de 2%. Un algorithme va ensuite chercher parmi toutes les variables explicatives possibles celle qui explique le mieux les

² L'erreur de type I est le rejet de l'hypothèse nulle vraie tandis que l'erreur de type II est le non-rejet de l'hypothèse nulle fautive. L'hypothèse nulle (H_0) étant l'occurrence d'une crise financière.

périodes de crises et de calmes dans chacun des groupes, le "meilleur discriminant" X_i , ainsi que les valeurs critiques de ces discriminants V_i^* (voir figure 1). Le but est d'augmenter la "pureté" d'un nœud, à savoir le degré de concentration de crise (1) ou de calme (0) que le discriminant est capable d'apporter dans la subdivision suivante. Des nœuds et des subdivisions se créent ainsi pour arriver à une liste finale des meilleures variables explicatives et de leurs seuils critiques respectifs pour un échantillon donné.

Figure 1 : Diagramme schématique du BRT (Davis and Karim, 2008)³.



Où,

X_i est le discriminant

V_i^* est la valeur critique du seuil

Ce modèle a été appliqué aux US dans les années 2000 par Davis et Karim en 2008 pour tenter de savoir si la crise des subprimes aurait pu être anticipée. Dans leurs travaux, ils ont utilisé deux méthodes différentes, à savoir la méthode du BRT et la méthode du modèle *logit*, pour pouvoir les confronter mais également pour voir l'importance de différents

³ Les chiffres présentés dans la figure 1 sont fictifs et ne représentent pas le réel échantillon utilisé par Davis et Karim. Il s'agit simplement d'une aide à la représentation du processus.

indicateurs dans différentes études. Les deux méthodes ont décelés des variables explicatives similaires, les cinq premières étant identiques et classées dans le même ordre :

- Real GDP Growth
- Change in Term of Trade
- Nominal Depreciation
- Real Interest Rate
- Inflation

Ces indicateurs ne sont pas les mêmes que ceux utilisés dans les travaux de Demirgüç-Kunt et Detragiache (2005). Cela est expliqué par le fait que leurs travaux se concentrent sur la détection de crises bancaires tandis que cette méthode traite les crises monétaires. Au final, aucune de ces deux méthodes n'avait été capable de prédire la crise des subprimes. La conclusion en était que ces deux modèles n'étaient pas à rejeter, mais qu'ils auraient dû être mieux adaptés aux caractéristiques de cette crise. Dans ces travaux, comme dans beaucoup d'autres, il a aussi été montré que les prédictions in-sample sont en général bien plus performantes que les prédictions out-of-sample, ce qui est en fait assez intuitif.

Conclusion

En conclusion, plusieurs modèles et dérivées d'EWS financiers existent. Il serait peu prudent d'affirmer qu'un modèle est meilleur que tous les autres. En effet, ils ont chacun été inventé en vue de prédire un type de crise bien particulier, ils ont chacun leur caractéristiques et spécificités, leurs forces et leurs faiblesses. Il a même été montré que certains d'entre eux sont au final complémentaires et que dans une certaine mesure, il serait inutile d'en faire un classement, le mieux étant d'essayer de les améliorer pour les rendre encore plus précis et performants. Le problème commun à tous ces EWS est pour l'instant la difficulté à obtenir le même niveau de précision dans les prédictions out-of-sample que dans les prédictions in-sample. En effet, il est plus difficile de prédire une crise ex ante qu'à postériori. Cela est expliqué par le fait que chaque crise financière que le monde ait connu est différente, que le contexte politique et économique est lui aussi différent, et que le monde est en constante évolution et rend donc les comparaisons parfois peu pertinentes.

Machine Learning et prédiction des crises financières

Nous avons maintenant découvert certains des modèles de calculs les plus couramment utilisés et les plus présents dans la littérature. Cependant, la limite de ces modèles de calculs est qu'il faut toujours parvenir à sélectionner le bon modèle, les bonnes données, parvenir à trouver les bons tests de fiabilités, détecter les différents biais, etc. Cela peut-être notamment facilité par l'automatisation du processus et par la mise en réseau de nombreuses données à travers le monde, grâce à des intelligences artificielles par exemple.

De nos jours, le *Big Data* et les IA sont de plus en plus utilisés, et dans tous les domaines. Le *Big Data* fait référence à des masses de données trop importantes et complexes pour être traitées et analysées par les méthodes traditionnelles. Dans leurs travaux, Goldstein, Spatt et Ye (2021) attribuent trois propriétés au Big Data relatif à la recherche financière : un échantillon très large, de grande dimension, et présentant une structure complexe. Ils ajoutent également que ces données doivent pouvoir proposer de nouvelles questions de recherche, auxquelles il ne serait pas possible de répondre avec un processus de traitement de données classique.

Au vu de la complexité de ces masses de données, il est donc nécessaire d'utiliser de nouveaux outils pour pouvoir les analyser. Le concept de *Machine Learning* est le fait qu'il soit possible pour une intelligence artificielle d'apprendre automatiquement de nouvelles choses en fouillant dans toutes ces données et de créer et d'adapter de nouveaux algorithmes. Ces technologies sont utilisées dans un grand nombre de domaines tels que la finance, la médecine, l'aérospatial, la géologie, etc.

Certains chercheurs se penchent sur la question d'utiliser ces outils pour créer de nouveaux Early Warning System grâce à des intelligences artificielles qui pourraient apprendre depuis les données passées et les modèles existants. Les IA peuvent alors analyser un grand nombre de données passées et apprendre à tirer des conclusions sur les différentes tendances et scénarios possibles pour éventuellement donner à leur tour l'alerte de manière

plus dynamique. L'avantage de ces nouvelles technologies est qu'il est possible de traiter un grand nombre de données en peu de temps. Selon Forbes et Rigobon (2002), la mondialisation toujours croissante augmente considérablement le risque de "contagion" de crise d'une région géographique à une autre, ou d'un marché à un autre. A contrario, ces nombreuses "interconnexions" entre les marchés favorisent la corrélation entre un grand nombre d'indicateurs financiers. Les *Machine Learning* dans le domaine des crises financières peuvent être un moyen d'observer toutes ces tendances du même œil et de pouvoir différencier certaines corrélations de potentielles contagions que l'œil humain n'aurait pas détecté. Samitas, Kampouris et Kenourgios (2020) ont mené une étude visant à utiliser tout le système financier mondial comme un "réseau" pour créer un Early Warning Indicator pouvant prédire des périodes de vulnérabilité financières. Une autre étude (Casabianca & all, 2019) a montré que les *Machine Learning* étaient plus performants et précis que les méthodes statistiques logit traditionnelles dans la prédiction de crises financières sur un échantillon de 100 pays différents, même si les conclusions montrent que les deux modèles ont chacun leurs avantages et inconvénients. Cela vient du fait de la capacité d'une intelligence artificielle à traiter un très grand nombre de données et à automatiser des apprentissages. En conclusion, l'aire de l'intelligence artificielle ne pourra qu'être bénéfique pour ce genre d'analyse en raison d'un monde économique toujours plus interconnecté et de données toujours plus nombreuses et complexes.

La crise de 2008

Dans la littérature, on retrouve beaucoup de modèles différents et d'indicateurs différents, et pour des types de crises différentes. Jusqu'ici, chaque auteur avait sa méthode, avec ses idées et ses arguments. Dans leurs travaux, Jeffrey A. Frankel et George Saravelos (2010) expliquent que toute cette littérature manque de généralités. En effet, la plupart des indicateurs sont choisis en fonction des mécanismes économiques sous-jacent d'une crise. Bien souvent, ils sont sélectionnés après coup, ce qui selon Frankel et Saravelos, crée un biais de sélection. Ces chercheurs ont fait l'exercice inverse et ont donc cherchés à identifier les causes des crises financières en général, quel que soit le type de crise, et examiné le succès des indicateurs identifiés pour prédire la crise de 2008. Ils sont parvenu à établir un

classement des indicateurs "statistiquement significatifs" les plus fréquemment rencontrés à travers la littérature. Les cinq premiers de leurs classements sont : *foreign exchange reserves*, *the real exchange rate*, *the growth rate of credit*, *GDP*, and *the current account*. (Voir Table 1 en annexe). Cela signifie que, parmi toute l'histoire des EWS permettant de prédire une crise financière, ces cinq indicateurs sont les plus performants, peu importe l'époque, la géographie, ou le contexte économique de la crise.

La conclusion de l'analyse empirique de ces chercheurs est que les deux premiers indicateurs identifiés dans la revue de littérature, à savoir ici *the level of international reserves* et *the real exchange rate overvaluation* ont été très utiles dans la prédiction de la crise de 2008.

Partie 2

Dans la seconde partie de ce mémoire, nous allons mener une analyse quantitative sur la crise financière que nous connaissons actuellement. Nous allons répliquer des analyses faites avec des données passées mais cette fois avec les données actuelles pour comprendre où nous en sommes aujourd'hui.

Méthodologie

Nous allons dans un premier temps définir la zone géographique que nous allons étudier dans cette analyse. J'ai choisi de me concentrer sur un échantillon de six pays européens à savoir la France, la Belgique, l'Allemagne, l'Espagne, l'Italie et la Grèce. Concernant la période, les données sélectionnées s'étendent sur une période d'à peu près 10 ans. Dans un second temps, nous allons définir les variables dépendantes de cette crise, qui resteront sensiblement identiques à celles définies pour la crise de 2008. Le but étant de répliquer l'étude faite par Frankel et Saravelos, les variables indépendantes resteront les mêmes, puisqu'elles sont supposées être les plus fréquemment significatives, peu importe le type de crise. Nous les expliquerons également.

L'étape suivante sera de faire une régression bivariée des quatre variables dépendantes sur chaque variable indépendante. Nous pourrons ainsi déterminer si les variables indépendantes sont des indicateurs pertinents ou non, grâce au coefficient de régression et à la t-stat. Cela nous permettra de faire un tri dans nos variables indépendantes pour pouvoir passer à la suite de l'analyse.

Afin de calculer les probabilités de crise aujourd'hui, nous utiliserons la première des quatre catégories d'approches recensées par Frankel et Saravelos dans leur revue de littérature, à savoir l'utilisation d'un modèle probit/logit, nécessaire à l'analyse d'une variable dépendante binaire.

Données

Pour cette étude, nous allons comme dit plus haut analyser six pays européens, à savoir la France, la Belgique, l'Allemagne, l'Espagne, l'Italie et la Grèce. Toutes les données ont été sélectionnées sur le site de la FRED. Mon but était d'être le plus à jour possible, mais certaines données n'étaient pas encore disponibles. Il m'a donc fallu faire un compromis pour avoir un maximum de variables avec des données les plus récentes possibles. Malheureusement, les données utilisées par les auteurs ont été principalement sélectionnées sur le site World Bank World Development Indicators database, mais les données ne sont pas disponibles pour l'année 2022. Certaines données proviennent également du site de l'IMF International Financial Statistics database. J'ai donc sélectionné un panel de variables se rapprochant le plus possible de celles utilisées dans l'étude initiale. Concernant la longueur des time series, toutes les données commencent et finissent à la même date pour faciliter l'utilisation de la database. La date de début est il y a un peu plus de 8 ans soit en janvier 2015, et la date de fin est le Q3 2022. Ce n'est pas forcément une mauvaise chose que d'avoir un jeu de données s'arrêtant il y a plus de deux trimestres. En effet, cela me permettra de comparer mes résultats finaux à la conjoncture actuelle pour les confirmer ou les infirmer. Après cette date, de nombreuses données sont encore manquantes sur le site de la Fred et je ne souhaitais pas réaliser de régression avec des variables manquantes. Dans la mesure du possible, les données seront analysées de manière mensuelle pour plus de précision. Pour les variables dont les données ne sont disponibles que de façon trimestrielle, j'ai appliqué un filtre afin de transformer les données mensuelles en données trimestrielles, ce qui nous permet d'avoir des séries de données comparables. La database contient trois types de données. Il y a les montants bruts exprimés en EUR ou en USD, les taux en base 100 et les taux unitaires. Nous allons convertir ces données pour obtenir uniquement des pourcentages de variation et faciliter l'analyse. Les régressions seront faites pour chaque pays de manière individuelle, ce qui permettra de comparer la résistance des différents pays aux impacts économiques qui touchent l'Europe en ce moment.

Variables dépendantes

Les variables dépendantes définies pour la crise de 2008 dans l'étude de Frankel et Saravelos étaient les suivantes :

- *Nominal local currency percentage change versus the US dollar*
- *Equity market returns in domestic stock market benchmark indices*
- *Percentage change in the level of real GDP*
- *Percentage change in industrial production*
- *Recourse to IMF financing*

Etant donné les caractéristiques de la conjoncture économique dans laquelle nous nous trouvons actuellement et dans le but de rester fidèle à l'analyse, nous garderons les mêmes variables dépendantes. Nous ne garderons toutefois pas la dernière variable qui était le recours au financement de l'IMF car les données sont encore floues à ce sujet.

1. *Nominal local currency percentage change versus the US-dollar*

Le taux de change nominal est une variable importante dans l'analyse des crises financières. En effet, celui-ci reflète la santé financière d'un pays et notamment la confiance des investisseurs étrangers. De plus, il affecte directement les entreprises internationales. Un taux de change qui diminue signifie moins de revenus pour les entreprises exportatrices, et plus de coûts pour les entreprises importatrices, ce qui joue directement sur l'économie nationale.

2. *Equity market returns in domestic stock market benchmark indices*

Cet indice mesure le rendement des investissements dans les actions d'un pays en utilisant un indice de référence national spécifique. Cet indice a plusieurs utilités lorsque l'on analyse une crise financière. Dans un premier temps, il peut servir à identifier les secteurs les plus touchés par une crise. Ensuite, il peut évaluer l'impact de la crise sur l'investissement.

Cela permet notamment de monitorer la confiance que les investisseurs placent dans les marchés. Une méfiance de la part des investisseurs est un facteur aggravant des crises financières. Enfin, Le rendement des actions peut être utilisé pour évaluer l'efficacité des mesures prises par un gouvernement lors d'une menace ou d'une difficulté.

3. *Percentage change in the level of Real Gross Domestic Product*

Le *Real Gross Domestic Product*, or *Real GDP*, nous indique la valeur du marché des biens et services produit au sein d'un pays. A la différence du nominal GDP, le Real GDP tient compte de l'inflation et reflète donc avec plus de précision la croissance économique d'un pays. Nous convertirons la série de données en pourcentage de variation par rapport à la période précédente.

4. *Percentage change in the level of Industrial Production*

L'indice de production industriel mesure la production totale de l'industrie énergétique, minière et manufacturière d'un pays. Dans le cadre de notre analyse, l'indice de production industriel peut être utilisé pour déterminer l'impact d'une crise sur le secteur industriel. De ce fait, on peut déjà émettre l'hypothèse que cet indice a fortement diminué pendant le covid, alors qu'il est probablement resté assez stable lorsque la guerre en Ukraine a éclaté. La crise du gaz a très probablement même fait augmenter ce dernier en raison de l'urgence de trouver des sources d'énergie alternatives. Par extension, l'indice de production industriel peut aider à cibler les secteurs les plus touchés de façon plus précise. Dans notre étude, nous nous contenterons d'utiliser un indice général, reprenant toute la production industrielle des pays de l'échantillon, en excluant toutefois la construction⁴.

⁴ La construction a été exclue des données dans le but de pouvoir utiliser le même indice pour tout l'échantillon, faute de données disponibles.

Variables indépendantes

Concernant les variables indépendantes, notre but est d'utiliser les indicateurs identifiés dans la revue de littérature de l'étude pour analyser les données actuelles. Il m'a cependant été très compliqué de trouver des données récentes correspondant à chacun des indicateurs de l'étude. J'ai donc simplifié la sélection de variables en reprenant toutes les données que j'ai pu trouver pour en faire une base de donnée modulable. Certains des indicateurs sont également légèrement différents.

1. Les Consumer Price Indexes

Les *Consumer Price Indexes*, ou CPI, sont des mesures qui représentent la variance du pouvoir d'achat des consommateurs pour certains paniers de biens et services. Ces indicateurs nous permettent donc d'analyser l'augmentation du coût de la vie et le pouvoir d'achat des consommateurs. Ils nous permettent également de monitorer l'efficacité des politiques économiques mises en place. Les variables indépendantes reprennent deux différents CPI, à savoir :

- Le *CPI All Items*, qui représente l'entièreté des biens et services consommés par un ménage dans un pays.
- Le *CPI Energy*, qui nous indique la variation du prix de l'énergie, à savoir l'électricité, le gaz, le mazout et le carburant. Avoir un indicateur qui représente uniquement le prix de l'énergie semble nécessaire étant donné la crise de l'énergie qui s'estompe à peine en ce début 2023. Cela nous donne également une indication plus précise sur les composantes de l'inflation.

2. Les taux d'intérêts

Le taux d'intérêt est une variable très utile pour détecter de potentiels signes avant-coureurs d'une crise financière. En effet, ceux-ci nous indiquent ou non une tension du

secteur interbancaire ou encore du marché immobilier. Ces taux peuvent être utilisés pour mesurer la capacité des emprunteurs à rembourser leurs dettes, ou encore pour mesurer le coût d'emprunt pour les gouvernements. Dans tous les cas, au plus le taux d'intérêt est élevé, au plus le risque de défaut de paiement est accru et le risque de crise élevé. De plus, un taux d'intérêt élevé mène en général à une diminution de la consommation.

- Taux d'intérêt des obligations d'état
- Taux d'intérêt interbancaires
- Taux d'intérêt sur les prêts immobiliers
- Taux d'intérêt sur les prêts à la consommation

Pour cette analyse, nous allons utiliser le taux d'intérêt interbancaire, à savoir les taux d'intérêts que les banques se facturent les unes aux autres lorsqu'elles se prêtent de l'argent à court-terme. Encore une fois, un taux élevé peut refléter des tensions sur le marché et donc indiquer un risque de crise.

3. Les réserves et balances de paiement.

Les réserves sont un indicateur important dans l'étude d'une crise financière. En effet, ce sont les réserves qui vont permettre à un gouvernement de stabiliser l'économie en cas de crise ou de forte dépréciation de la monnaie locale. Les réserves indiquent donc à quel point une économie sera solide ou non face à un choc économique. Un pays avec de faibles réserves peut s'avérer plus sensible à un choc et cela peut faire baisser la confiance des agents économiques qui chercheront à investir en une autre monnaie, et donc à déprécier encore plus la monnaie locale. La balance de paiement nous informe des flux monétaires entre un pays et le reste du monde. On peut alors voir si un pays exporte plus qu'il n'importe ou inversement, et se rendre compte d'un éventuel déficit commercial.

4. Les valeurs d'importation et d'exportation de biens et services

La valeur d'importation et d'exportation de biens et services d'un pays est intéressant à analyser car elle nous donne une vision de la santé financière de ce pays au niveau

international. Si les exportations sont élevées, cela indique que le pays est compétitif sur le marché international et fait face à une forte demande. Les importations élevées indiquent une économie nationale dynamique. De plus, les valeurs d'importation et d'exportation peuvent servir à surveiller les flux de capitaux au niveau international. Les exportations font rentrer des devises étrangères tandis que les importations peuvent demander un paiement en devise étrangères. Ces mouvements jouent donc dans le mouvement des devises en possession de l'état, et cela à bien évidemment un impact sur la stabilité de la monnaie locale.

5. La production industrielle totale

La production industrielle totale peut nous donner une indication sur les tendances économiques à long terme d'un pays, ainsi que du flux de devises étrangères. Les exportations de biens et services génèrent des revenus en devises étrangères tandis que les importations nécessitent des paiements en devises étrangères. Cela joue sur l'appréciation ou la dépréciation de la monnaie locale.

Analyse empirique

Régression bivariée

Dans un premier temps, nous allons utiliser la méthode de la régression bivariée pour analyser les relations entre les variables indépendantes et les variables dépendantes afin de déterminer les variables les plus statistiquement significatives. Les régressions linéaires ont été faites sur chaque variable dépendante et par chaque variable indépendante, pour chaque pays. Il s'agit donc d'un total de 192 régressions, contre 305 dans l'étude initiale. Les valeurs de chaque coefficient de régression ainsi que les t-stats sont reportées dans les tableaux ci-dessous.

Légende :

Variables dépendantes

- *RGDP : Real Gross Domestic Product*
- *BY : Long-Term Government Bond Yields : 10 years*
- *EURUSD : Euro-Dollars exchange rate*
- *IP : Industrial Production*

Variables indépendantes

- *CPIA : Consumer Price Index of All items*
- *CPIE : Consumer Price Index Energy*
- *RBEER : Real Broad Effective Exchange Rate*
- *BPM6 : Reserve asset Net*
- *BPM7 : Capital Account Balance*
- *EXP : Exports of Goods and Services*
- *IMP : Imports of Goods and Services*
- *IR : Immediate rate < 24 hours : Call Money/Interbank Rate*

Tableau 1 : Résultat de l'analyse bivariée pour la France

France				
	RGDP	BY	EURUSD	IP
CPIA	0.7319	0.04912	-1.0696	0.3562
	0.664	1.695	-1.354	0.360
CPIE	0.3168	0.01875	0.2183	0.3739
	1.57	2.794	1.591	2.224
RBEER	0.4833	-0.20923	1.2971	0.3791
	0.604	-4.69	2.689	0.611
BPM6	0.000241	-0.008484	0.001241	0
	0.221	-0.824	0.856	-0.001
BPM7	0.001297	-0.00587	0.005052	0.000855
	0.593	-0.237	2.112	0.179
EXP	0	0.1692	0.16589	1.024881
	0.777	0.112	1.405	8.702
IMP	0	0.85675	0.15661	1.18086
	0.604	0.604	1.079	7.345
IR	0.0013	0.8556	0.0289	-0.006068
	0.070	1.303	2.289	-0.041

Tableau 2 : Résultat de l'analyse bivariée pour la Belgique

Belgique				
	RGDP	BY	EURUSD	IP
CPIA	0	-0.03419	0.002695	0.006359
	4.959	-1.512	1.129	2.94
CPIE	1.5226	0.197	0.03506	-0.1502
	4.025	0.115	0.198	-0.845
RBEER	0.080306	-0.8403	1.620657	0.287131
	0.136	-0.09	1.76	0.292
BPM6	0	-0.000364	0	0
	-0.283	-0.122	-0.03	-0.099
BPM7	0	-0.0006541	-0.00023	0
	0.013	-0.355	-1.258	-0.055
EXP	0.303982	-0.99674	0.276442	0.472054
	7.789	-0.918	2.718	6.377
IMP	0.22473	-0.66194	0.23584	0.267436
	4.822	-0.667	2.515	2.897
IR	0.0026644	0.4146	-0.06759	-0.002049
	0.087	0.869	-1.393	-0.04

Tableau 3 : Résultat de l'analyse bivariée pour l'Allemagne

Allemagne

	RGDP	BY	EURUSD	IP
CPIA	-0.158395 -0.437	4.29689 0.677	-0.45779 -0.619	-0.835968 -0.743
CPIE	0.111294 0.969	0.57804 0.281	0.05074 0.212	0.326 0.905
RBEER	-0.321125 -0.962	0.73214 0.123	0.684184 1.002	-1.113831 -1.071
BPM6	0 0.148	-0.00999 -1.194	0 0.972	0 -0.131
BPM7	-0.002648 -3.657	0.01738 1.141	-0.00165 -0.932	-0.00847 -3.788
EXP	0.4556607 14.197	0.07542 0.047	0.264183 1.455	1.449717 16.614
IMP	0.53136 8.871	0.82858 0.403	0.289712 1.242	1.697565 9.655
IR	0.010129 0.415	0.492292 1.17	-0.06759 -1.393	0.022967 0.3

Tableau 4 : Résultat de l'analyse bivariée pour l'Espagne

Espagne

	RGDP	BY	EURUSD	IP
CPIA	-0.111417 -0.215	1.922 0.28	-0.51389 -0.89	-0.7204 -0.682
CPIE	0.356164 2.83	-1.3613 -0.728	0.3613 2.643	0.667488 2.529
RBEER	-0.30126 -0.337	2.4174 0.204	1.388315 1.444	-0.479021 0.261
BPM6	0 -0.036	0.003867 0.407	-0.0007798 -0.98	0.000116 0.078
BPM7	0.004282 0.55	-0.04249 -0.412	-0.0008638 -0.098	0.00724 0.453
EXP	0.512051 17.426	-0.7609 -0.572	0.1627 1.453	0.985723 10.837
IMP	0.533268 13.793	-0.8213 -0.579	0.2019 1.73	1.046798 10.563
IR	0.002477 0.124	-0.0326 -0.412	0.0209 1.801	0.001919 0.047

Tableau 5 : Résultat de l'analyse bivariée pour l'Italie

Italie

	RGDP	BY	EURUSD	IP
CPIA	-0.058936	7.6287	-3.0392	-1.89447
	-0.091	2.309	-2.905	-0.518
CPIE	0.0679951	5.1306	-0.2006	-0.13436
	0.687	2.302	-1.374	-0.238
RBEER	-0.30324	3.9951	-0.564	-1.13738
	-0.565	1.914	-0.828	-0.374
BPM6	-0.0003	0.001524	-0.0001113	-0.001959
	-5.435	0.756	-0.949	-6.46
BPM7	0	-0.0009373	-0.000902	-0.001004
	-0.062	-0.077	-1.09	-0.269
EXP	0.44259	-1.515	0.181512	2.017872
	17.353	-0.731	1.546	6.509
IMP	0.42016	-3.2457	0.165415	1.7217
	7.926	-1.468	1.245	4.036
IR	0.002059	-0.0345	0.0165	0.01482
	0.136	-0.897	0.0118	0.173

Tableau 6 : Résultat de l'analyse bivariée pour la Grèce

Grèce

	RGDP	BY	EURUSD	IP
CPIA	0.95339	4.0362	-0.439025	-2.33629
	0.641	3.081	-1.023	-2.023
CPIE	0.59515	0.9937	0.4672	0.16501
	1.133	1.641	2.418	0.375
RBEER	0.2837	3.5855	0.3707	-3.63092
	0.09	2.769	0.853	-2.742
BPM6	-0.0003943	0.05801	-0.0008639	-0.001556
	-0.745	1.082	-0.947	-0.587
BPM7	-0.0002224	-0.01184	0.0009893	0.000114
	-0.527	-0.273	1.39	0.054
EXP	0.010628	-0.02383	0.015134	0.330527
	0.659	-0.014	0.445	6.246
IMP	0.033171	-1.3198	0.1535	0.20786
	1.236	-0.466	3.735	1.576
IR	0.001254	1.3891	0.0093	-0.11328
	0.125	1.368	0.785	-2.499

Résultats de l'analyse bivariée

Pour analyser les résultats de cette analyse, trois dimensions sont à prendre en compte. La première dimension est la significativité statistique du test, indiqué par la t-stat. La t-statistique exprime la taille du coefficient de régression par rapport à l'écart-type de l'estimation. Pour un taux de certitude de 95%, la valeur absolue de la t-stat doit être supérieure à 1.96 pour que le test soit considéré significatif et que l'hypothèse nulle qu'il n'existe pas de lien significatif entre les variables soit rejetée. En dessous de ce seuil, l'hypothèse nulle ne peut être rejetée et le coefficient, bien que pas nécessairement mauvais, ne peut être considéré comme significatif. Sur les tableaux, on peut voir quatre niveaux de significativités en fonction des quatre nuances de gris. Les niveaux représentent respectivement des *p-values* supérieures à 0.1, inférieures à 0.1, inférieure à 0.05, et enfin inférieures à 0.01. Les tests ne se sont pas montrés vraiment concluants à ce niveau. On peut observer que les couleurs ne sont pas consistantes d'un tableau à un autre, et donc que les résultats changent de significativité en fonction du pays. Les variables indépendantes CPIE, EXP et IMP sont celles dont les régressions se sont montrées significatives le plus de fois. 9 fois pour la variable CPIE, contre 14 et 15 fois sur un total de 24 régressions pour les variables EXP et IMP. Les variables qui ont montré les relations les moins significatives sont les variables BMP6 et BMP7, soit les variables relatives aux réserves. Cependant, cela ne veut pas dire qu'il n'existe pas de liens significatifs entre les variables mais simplement qu'il n'y a pas d'évidence d'une quelconque significativité. Il faut tenir compte du fait que l'échantillon n'est pas relativement élevé et qu'un échantillon plus large aurait sans doute permis d'obtenir une meilleure significativité des coefficients.

La seconde dimension est celle de l'étude des signes des coefficients. On peut voir que beaucoup d'entre eux ne sont pas consistants d'un pays à l'autre, mais les variables les plus significatives ont en général un signe consistant pour au moins cinq pays sur six.

La dernière dimension est celle des coefficients de régression, qui nous indiquent la force de la relation entre deux variables. Le coefficient de régression mesure la proportion de volatilité de la variable dépendante expliquée par la variable indépendante. Par exemple, le coefficient de régression du CPIA et du RGDP pour la France est de 0.7319, ce qui indique que

dans notre échantillon, une augmentation de 1% du CPIA provoque une augmentation de 0.7319% du RGDP. On peut voir que sur les 57 résultats avec une valeur inférieure à 0.1, 17 coefficients sont en valeur absolue supérieurs à 1 et 28 supérieurs à 0.1, ce qui indique un lien assez fort entre les variables.

Discussion

L'étape suivante va être de déterminer les variables indépendantes que nous allons utiliser pour le reste de l'étude. Pour ce faire, nous allons nous baser sur la force du coefficient de régression, sur la significativité du test et enfin, l'intuition, en se basant sur des faits purement historiques.

La paire de variables qui s'est clairement distinguée en terme de force de lien et de significativité sont les valeurs d'exportation et d'importation de biens et services. (EXP et IMP). D'un point de vue intuitif, il est clair que ce sont deux variables pertinentes. En effet, nous essayons de prédire la probabilité d'une crise de change et ces variables représentent en quelques sortes les flux de monnaies locales et étrangères au sein d'un pays. Ensuite, le *Real Broad Effective Exchange Rate* me semble être également une mesure utile, quoi que discutable. En effet, les régressions sur cette variable n'ont pas été aussi significatives que l'exportation et l'importation, mais ses coefficients de corrélations sont tout de même assez fort et dans le contexte d'une crise de change, il serait utile de garder une mesure de taux de change. Enfin, le CPIE, à savoir la variation du prix de l'énergie au sein d'une zone géographique. Bien que cette variable ne soit pas forcément intuitive dans cette analyse, il est important de noter qu'une hausse du prix de l'énergie implique une baisse du pouvoir d'achat des ménages étant donné que la demande en énergie n'est que peu élastique. Cette baisse du pouvoir d'achat se répercute directement sur la valeur de la monnaie locale par rapport aux devises étrangères. Nous écartons donc les quatre autres variables, à savoir les variables BPM6 et BPM7, relatives aux réserves, la variable CPIA ou *Consumer Price Index All Item*, et enfin la variable IR ou *Immediate Rate*, à savoir le taux interbancaire. Cette dernière variable aurait pu être intéressante d'un point de vue intuitif, mais ne présentait pas de très bons résultats statistiques lors des régressions.

Modèle probit

A l'aide des résultats statistiques obtenus plus haut qui nous ont permis de faire le tri dans les indicateurs à utiliser pour la prochaine étape, nous allons à présent modéliser une variable dépendante binaire qui représente l'occurrence ou non d'une crise financière. Cette variable prendra la valeur de 1 si le modèle prédit l'occurrence d'une crise, et la valeur de 0 si le modèle ne prédit pas de crise.

Étape 1

La première étape va être de créer une variable binaire qui deviendra notre nouvelle variable dépendante. Cette variable prendra la valeur de 1 en cas de crise, et la valeur de 0 sinon. Dans la littérature, les études se penchent sur les prédictions de seuils de décisions à appliquer afin de créer une variable dichotomique. Pour notre étude, nous allons réaliser l'exercice inverse en définissant non pas des seuils de décisions mais des années d'éclatement de crises financières, à savoir l'année 2020 (Covid19 en Europe). Étant donné que nos séries temporelles s'étendent sur 8 années, nous traiterons dans un premier temps les années 2015 à 2021, et garderons les données de l'année 2022 pour l'exercice de prédiction. Notre variable binaire prendra donc la forme suivante :

$$Y = [0, 0, 0, 0, 0, 1, 0]$$

Le but étant dans un second temps d'ajouter les données relatives à l'année 2022 pour prédire ou non l'occurrence d'une crise financière.

Étape 2

Afin d'utiliser cette variable binaire dans notre analyse, nous aurons besoin d'utiliser un modèle probit, qui nous donnera la probabilité que cette variable dépendante soit égale à 1 étant donné les variables indépendantes X . Cette probabilité est modélisée de la manière suivante :

$$P(Y = 1|X) = \Phi(X\beta)$$

Où Φ représente la fonction de distribution cumulative de la distribution normale standard, et β représente le vecteur de coefficient à estimer.

Le modèle probit fonctionne en analysant la variable latente non observée Y^* . Cette variable Y^* est en fait cette variable latente qui nous donnera la variable binaire Y . Y^* suit une distribution normale et est liée aux variables explicatives X et aux erreurs. Cette relation linéaire est définie telle que :

$$Y^* = X\beta + \varepsilon, \varepsilon \sim N(0,1)$$

La variable binaire Y est alors définie telle que :

$$Y = 0 \text{ si } Y^* \leq \text{seuil de décision}$$

$$Y = 1 \text{ si } Y^* > \text{seuil de décision}$$

Dans notre modèle, la variable Y représentera la probabilité qu'une crise financière se produise. Ce modèle probit comporte un avantage comparé à d'autres méthodes de prédiction, à savoir qu'il prend en compte l'interdépendance entre les variables.

Pour construire le modèle probit, j'ai dans un premier temps introduit mes 4 variables indépendantes sous forme de data frame ainsi que ma variable binaire dans R. J'ai ensuite créé 6 data sets (un pour chaque pays) contenant à chaque fois les 4 variables et la variable binaire "crise" et créé et ajusté 6 modèles probit. J'ai enfin créé 6 nouveaux data sets comprenant la dernière valeur de chaque variable, pour chaque pays, et calculé la probabilité de crise pour l'année 2022. Les résultats des modèles probit après ajustement sont présentés en annexes sous le titre "Résultats statistiques", tandis que les résultats des prédictions sont repris ci-dessous.

Résultats

- France

Name	Type	Value
predictions	double [1]	2.220446e-16
1	double [1]	2.220446e-16

- Allemagne

Name	Type	Value
predictions	double [1]	1
1	double [1]	1

- Belgique

Name	Type	Value
predictions	double [1]	5.581412e-16
1	double [1]	5.581412e-16

- Espagne

Name	Type	Value
predictions	double [1]	2.220446e-16
1	double [1]	2.220446e-16

- Italie

Name	Type	Value
predictions	double [1]	0.8140947
1	double [1]	0.8140947

- Grèce

Name	Type	Value
predictions	double [1]	2.220446e-16
1	double [1]	2.220446e-16

J'ai ici fais l'erreur de ne pas fixer de seuil de décision dès le départ, ce qui nous donne comme résultat des valeurs non pas binaires, mais continues entre 0 et 1. Pour remédier à cela, j'ai fixé un seuil de décision à 0,5. Au-delà de ce seuil, Y prendra la valeur de 1. En dessous, Y prendra la valeur de 0. On peut donc rapidement transformer nos résultats, présentés dans le tableau ci-dessous.

Zone géographique	Valeur de la variable Y "crisis"
France	0
Allemagne	1
Belgique	0
Espagne	0
Italie	1
Grèce	0

Discussion des résultats

En ce qui concerne les six modèles probit, on peut voir, en observant les t-stats, qu'aucun des résultats n'est significatif. Comme pour les régressions bivariées, on peut ici dire qu'aucun des résultats n'est significatif à plus de 90% et qu'on ne peut donc pas rejeter l'hypothèse nulle qu'il existe un lien entre les variables. Les liens sont présents, mais peut-être dû au hasard. Cela peut venir de deux raisons : la première étant de toute évidence la taille de l'échantillon qui est assez faible. La seconde raison est que les variables indépendantes ne présentaient déjà pas une corrélation très forte avec les variables dépendantes lors des premiers tests. Ensuite, le fait que l'échantillon soit petit et que la moitié des données soient déjà légèrement perturbée par la crise covid peut amplifier encore la confusion. Pour l'interprétation du modèle, les coefficients indiquent comme pour les régressions la part de la volatilité de la variable dépendante lorsque la variable indépendante augmente de 1%. Ici, les coefficients mesurent donc l'effet des quatre variables indépendantes sur la variable binaire "crise".

Les prédictions, quant à elles, montrent que seulement deux des six modèles ont calculés une probabilité de crise supérieur à 0.5. Néanmoins, nous aurions pu penser que

chacun des modèles aurait dû prédire une crise, au vu du contexte économique auquel nous avons dû faire face ces derniers mois. Gardons en tête que la significativité des modèles n'était pas assez élevée, et que l'échantillon était trop petit que pour tirer des conclusions trop hâtives de ces résultats. Cependant, les résultats sont encourageants. J'aurais pensé qu'au vu de la non-significativité des coefficients et au vu de la petite taille de l'échantillon, tous les résultats indiqueraient le même résultats, soit une valeur de 1 soit une valeur de 0. Les résultats extrait de R si dessus m'ont permis de me rendre compte que ce n'est pas le cas et que le modèle à tout de même donné pour l'Italie au moins un résultat qui n'était pas strictement égal à 1 ou très proche de 0.

Limites et recommandations

Les limites et recommandations de cette étude quantitatives ont déjà été citées plus haut, mais on peut souligner encore une fois la taille trop petite des échantillons utilisés. La même étude avec des times series plus longues et l'utilisation de plus de variables auraient sans doute permis de stabiliser les t-stats et auraient permis une meilleure significativité, notamment pour le modèle probit qui n'utilisaient des times series que de six valeurs. De plus, ces valeurs sont déjà perturbées par la crise covid, encore bien présente en 2021. Une autre limite est également le fait qu'on a pas encore vraiment de recul sur cette crise et qu'elle n'est pas non plus totalement terminée, même si l'économie européenne se porte mieux que ce qu'on aurait pu croire. En effet, la banque de France considère que le pays échappera bel et bien à la récession prévue en 2023, et il en est de même pour de nombreux autres pays. Nous aurons une bien meilleure visibilité sur ces derniers événements dans quelques mois, voire quelques années.

Matrice de confusion

Nous allons maintenant calculer la matrice de confusion. Cette matrice, telle que représentée ci-dessous (figure 2), nous permet de savoir si le modèle utilisé à une bonne capacité de détection ou pas. Elle nous permet en fait de comparer les résultats de prédiction d'un modèle de classification avec la réalité. La matrice est divisée en quatre parties : True Positive (TP), False Positive (FP), False Negative (FN) et True Negative (TN). FP et FN correspondant respectivement aux erreurs de type II et aux erreurs de type I, avec comme

hypothèse nulle l'occurrence d'une crise. Calculer cette matrice est importante pour évaluer la performance d'un modèle et pour décomposer le type d'erreur qu'il commet. En effet, un indicateur parfait ne ferait aucune erreur de prédiction, et le taux de FN et de FP serait égal à zéro. Ce qui est impossible dans la pratique.

Figure 2 : Matrice de confusion

		Reality	
		Crises occurrence	No crises occurrence
Prediction	Signal	True Positive	False Positive
	No signal	False Negative	True Negative

- Calcul du nombre de crises prédit par le modèle
- Calcul du nombre de non-crise prédit par le modèle
- Calcul du nombre réel de crise
- Calcul du nombre réel de non-crise

Considérons maintenant l'occurrence d'une crise dans chacun des pays pour l'année 2022, la matrice de confusion pour notre étude serait la suivante :

		Reality	
		Crises occurrence	No crises occurrence
Prediction	Signal	2	0
	No signal	4	0

Résultats

Cette matrice nous donne un taux d'erreur de type II élevé et un taux d'erreur de type I nul. Encore une fois, le faible nombre de modèle probit ajouté à la taille de l'échantillon permet de conclure que cette matrice n'est pas tout à fait significative. Il aurait fallu ajouter des modèles pour lesquels la valeur à prédire était de 0 pour obtenir un taux d'erreur de type I et pouvoir ainsi évaluer la qualité du modèle. Notons tout de même qu'on est arrivé à obtenir deux *True Positive*, ce qui montre que le modèle a correctement prédit deux des six valeurs.

Conclusion

Ce mémoire avait pour objectif de mettre en lumière les différentes méthodes de prédictions de crises financières existantes dans la littérature et de, à mon tour, utiliser ces modèles sur des données nouvelles dans le but d'observer un résultat. J'ai énormément appris sur les crises financières durant mes recherches et je me suis vraiment rendue compte de la complexité à traiter des données. Concernant la partie empirique, le but était de trouver des variations de variables qui pourraient expliquer une crise financière et d'utiliser ces mêmes variables pour créer un modèle de prédiction sur de nouvelles données et de voir si le modèle prédirait ou non l'occurrence d'une crise financière. Je pense que les variables explicatives sélectionnées étaient assez significatives mais que l'échantillon était trop petit pour tirer de réelles conclusions. Les modèles probit ont prédits une crise financière pour l'année 2022 pour l'Italie et l'Allemagne seulement, alors que les six pays étaient en difficulté suite à la crise de l'énergie et à l'inflation.

Même si mes résultats sont mitigés, je pense qu'il pourrait être intéressant de continuer à explorer la piste de crise financière en 2022, compte tenu des événements en Europe et dans le monde. Je recommanderais alors de tester plus de variables et d'utiliser des échantillons bien plus larges, surtout pour créer les modèles probit. La difficulté à trouver des données m'a un peu déstabilisé, mais je suis persuadée que les résultats auraient pu être meilleurs. Il pourrait également être intéressant de réutiliser les mêmes variables et d'utiliser un modèle de calcul différents afin de comparer les résultats et de pouvoir comparer la performance de deux ou plusieurs modèles différents sur les prédictions pour l'année 2022.

Bibliographie

- ABDELSALAM, Mamdouh Abdelmoula M., and ABDEL-LATIF, Hany. 'An Optimal Early Warning System for Currency Crises under Model Uncertainty'. *Central Bank Review* 20, no. 3 (September 2020): 99–107. <https://doi.org/10.1016/j.cbrev.2020.03.002>.
- AHN, Jae Joon, OH, Kyong Joo, KIM, Tae Yoon, and KIM, Dong Ha. 'Usefulness of Support Vector Machine to Develop an Early Warning System for Financial Crisis'. *Expert Systems with Applications* 38, no. 4 (April 2011): 2966–73. <https://doi.org/10.1016/j.eswa.2010.08.085>.
- BERG, J.J.F. van den. 'Currency Crises and Their Early Warning Systems'. Maastricht university, 2011. <https://doi.org/10.26481/dis.20110519jb>.
- BUSSIÈRE, Matthieu, and FRATZSCHER, Marcel. 'Towards a New Early Warning System of Financial Crises'. *Journal of International Money and Finance*, 2006, 21.
- CANDELON Bertrand, DIMITRESCU Elena-Ivona, and HURLIN, Christophe. 'Currency Crisis Early Warning Systems: Why They Should Be Dynamic'. *International Journal of Forecasting* 30, no. 4 (October 2014): 1016–29. <https://doi.org/10.1016/j.ijforecast.2014.03.015>.
- CANDELON Bertrand, DIMITRESCU Elena-Ivona, and HURLIN, Christophe. 'Early Warning Systems : A Comparison of Currency Crises Forecasting Methods', September 2009.
- CASABIANCA, E. J., CATALANO, M., FORNI, L., GIARDA, E., & PASSERI, S. 'An early warning system for banking crises : From regression-based analysis to machine learning techniques'. *EconPapers. Orebro: Orebro University*. (2019)
- CUMPERAYO, Phornchanok, and KOUWENBERG Roy. 'Early Warning Systems for Currency Crises: A Multivariate Extreme Value Approach'. *Journal of International Money and Finance* 36 (September 2013): 151–71. <https://doi.org/10.1016/j.jimonfin.2013.03.008>.
- DAVIS, E. Philip, and DILRUBA Karim. 'Could Early Warning Systems Have Helped To Predict the Sub-Prime Crisis?' *National Institute Economic Review* 206 (October 2008): 35–47. <https://doi.org/10.1177/0027950108099841>.
- DEMIRGUC-KUNT, Asli, and DETRAGIACHE Enrica. 'Cross-Country Empirical Studies of Systemic Bank Distress: A Survey'. *National Institute Economic Review* 192 (April 2005): 68–83. <https://doi.org/10.1177/002795010519200108>
- FRANKEL, Jeffrey, and SARAVELLOS, George. 'Are Leading Indicators of Financial Crises Useful for Assessing Country Vulnerability? Evidence from the 2008-09 Global Crisis'. Cambridge, MA: National Bureau of Economic Research, June 2010. <https://doi.org/10.3386/w16047>.
- GOLDSTEIN, I., SPATT, C. S., & YE, M. 'Big Data in Finance'. *Review of Financial Studies*, 34(7), 3213–3225. (2021) <https://doi.org/10.1093/rfs/hhab038>
- KARMARKAR, Y. " *Early Warning Signal System for Economic Crisis : A Threshold and Indicators Approach .* ", (2014)

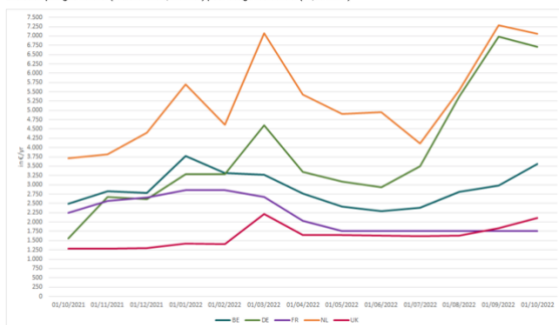
- KAMINSKY, Graciela, LIZONDO, Saul and REINHART, Carmen M. 'Leading Indicators of Currency Crises'. *Staff Papers - International Monetary Fund* 45, no. 1 (March 1998): 1. <https://doi.org/10.2307/3867328>.
- KAUPPI, Heikki, and SAIKKONEN Pentti. 'Predicting U.S. Recessions with Dynamic Binary Response Models'. *Review of Economics and Statistics* 90, no. 4 (November 2008): 777–91. <https://doi.org/10.1162/rest.90.4.777>.
- KLOPOTAN, Igor, ZOROJA, Jovana, and MESKO Maja. 'Early Warning System in Business, Finance, and Economics: Bibliometric and Topic Analysis'. *International Journal of Engineering Business Management* 10 (January 2018): 184797901879701. <https://doi.org/10.1177/1847979018797013>.
- LI, Shuangjie, and WANG, Shao. 'A Financial Early Warning Logit Model and Its Efficiency Verification Approach'. *Knowledge-Based Systems* 70 (November 2014): 78–87. <https://doi.org/10.1016/j.knosys.2014.03.017>.
- LI, Wei-Xuan, Clara, CHEN, Chia-Sheng, and FRENCH, Joseph J. 'Toward an Early Warning System of Financial Crises: What Can Index Futures and Options Tell Us?' *The Quarterly Review of Economics and Finance* 55 (February 2015): 87–99. <https://doi.org/10.1016/j.qref.2014.07.004>.
- MULDER, Christian, PERRELLI Roberto, and DUARTE ROCHA, Manuel. 'The Role of Bank and Corporate Balance Sheets on Early Warning Systems of Currency Crises—An Empirical Study'. *Emerging Markets Finance and Trade* 52, no. 7 (2 July 2016): 1542–61. <https://doi.org/10.1080/1540496X.2016.1158545>.
- REINHART, Carmen, and ROGOFF Kenneth. 'This Time Is Different: A Panoramic View of Eight Centuries of Financial Crises'. Cambridge, MA: National Bureau of Economic Research, March 2008. <https://doi.org/10.3386/w13882>.
- SAMITAS, A., KAMPOURI, E., & KENOURGIOS, D. Machine learning as an early warning system to predict financial crisis. *International Review of Financial Analysis*, 71, 101507, 2020. <https://doi.org/10.1016/j.irfa.2020.101507>
- VLAAR Peter J G. 'Early Warning Systems for Currency Crises', n.d., 253–74.

Annexes

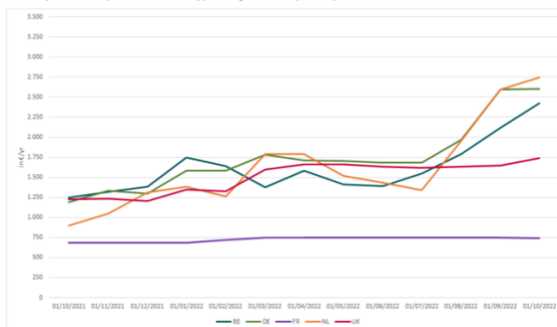
Graphes

- **Graphe 3 et 4** : Évolution des prix du gaz et de l'électricité pour un usage résidentiel durant l'année dernière. ([Source : creg.be](http://creg.be))

Evolution prix gaz naturel (facture totale, tva incl.) pour usage résidentiel (17,00 MWh)



Evolution prix électricité (facture totale, tva incl.) pour usage résidentiel (3,5 MWh)



Tables

- Table 1 : Frankel et Saravelos (2010)

Table 1 Summary of pre-2008 Early Warning Indicators

Leading Indicator ¹	KLR (1998) ²	Hawkins and Klau (2001) ³	Abiad (2003) ^{4,6}	Others ^{5,6}	Total
Reserves ^a	14	18	13	5	50
Real Exchange Rate ^b	12	22	11	3	48
GDP ^c	6	15	1	3	25
Credit ^d	5	8	6	3	22
Current Account ^e	4	10	6	2	22
Money Supply ^f	2	16	1	0	19
Exports or Imports ^{1a, g}	2	9	4	2	17
Inflation	5	7	1	2	15
Equity Returns	1	8	3	1	13
Real Interest Rate ^h	2	8	2	1	13
Debt Composition ^{1b, i}	4	4	2	0	10
Budget Balance	3	5	1	0	9
Terms of Trade	2	6	1	0	9
Contagion ^j	1	5	0	0	6
Political/Legal	3	2	1	0	6
Capital Flows ^{1c, k}	3	0	0	0	3
External Debt ^l	0	1	1	1	3
Number of Studies	28	28	20	7	83

Notes

^{1, 1a, 1b, 1c} Leading indicator categories as in Hawkins and Klau (2000), with exception of ^{1a}includes imports, ^{1b}debt composition rather than debt to international banks, ^{1c}capital flows rather than capital account.

²As reported in Hawkins and Klau (2000), but M2/reserves added to reserves, interest rate differential added to real interest rate.

³S&P, JP Morgan, IMF Indices, IMF *WEO*, IMF ICM, IMF EWS studies have been excluded due to lack of verifiability of results. The following adjustments have been made to the authors' checklist: significant credit variables reduced from 10 to 8 as Kaminsky (1999) considers level rather than growth rate of credit; significant capital account variables reduced from 1 to 0 as Honohan (1997) variable not in line with definition used here; Kaminsky (1999) significant variables for external debt reclassified to debt composition as these variables relate to short-term debt.

⁴10 out of 30 studies excluded from analysis. 7 included in Hawkins and Klau (2000) and 3 due to absence of formal testing of variables.

⁵Includes Berg, Borenstein and Pattillo (2004), Manasse and Roubini (2005), Shimpalee and Breuer (2006), Davis and Karim (2008), Bergmen et al. (2009), Obstfeld, Shambaugh and Taylor (2009), Rose and Spiegel (2009a).

⁶See App. 1 for criteria defining statistical significance in Abiad (2003) and Others studies. For rest see KLR (1998), Hawkins & Klau (2001)

Variables included in the leading indicator categories:

^aReserves: relative to GDP, M2, short-term debt, 12m change

^bReal Exchange Rate: change, over/under valuation

^cGDP: growth, level, output gap

^dCredit: nominal or real growth

^eCurrent Account: Current Account/GDP, Trade Balance/GDP

^fMoney Supply: growth rate, excess M1 balances

^gExports or Imports: relative to GDP, growth

^hReal Interest Rate: domestic or differential

ⁱDebt Composition: commercial/concess./variable-rate/debt to internat. banks/short-term/multilat./official relative to total external debt. Short-term debt relative to reserves (rather than relative to total external debt) is in the reserves category

^jContagion: dummies for crisis elsewhere

^kCapital Flows: FDI, short-term capital flows

^lExternal Debt: relative to GDP

- Table 2 : Reinhart et Rogoff (2008)

Crisis type	Threshold	Period	Maximum
Inflation	An annual inflation rate 20 percent or higher. We also examine separately the incidence of more extreme cases where inflation exceeds 40 percent per annum.	1500–1790	173.1
		1800–1913	159.6
		1914–2006	9.63E+26
Currency crashes	An annual depreciation versus the US dollar (or the relevant anchor currency—historically the UK pound, the French franc, or the German DM and presently the euro) of 15 percent or more.	1800–1913	275.7
		1914–2006	3.37E+09
Currency debasement: Type I	A reduction in the metallic content of coins in circulation of 5 percent or more.	1258–1799	–56.8
		1800–1913	–55.0
Currency debasement: Type II	A currency reform where a new currency replaces a much-depreciated earlier currency in circulation.	The most extreme episode in our sample is the 1948 Chinese conversion at a rate of 3 million to 1.	

Résultats statistiques

- Résultat 1 : Modèle probit France après ajustements

```

> glm(modelf)

Call:  glm(formula = modelf)

Coefficients:
(Intercept)      f1      f2      f3      f4
  0.12319   -0.03479   0.08119  -0.11207   0.12982

Degrees of Freedom: 6 Total (i.e. Null);  2 Residual
Null Deviance:      0.8571
Residual Deviance: 0.2234      AIC: 7.751

> summary(modelf)

Call:
glm(formula = crise ~ f1 + f2 + f3 + f4, family = binomial(link = "probit"),
    data = FR)

Deviance Residuals:
    1      2      3      4      5      6      7
-1.066e-05 -1.103e-05 -5.739e-06 -2.110e-08 -2.110e-08  9.969e-06 -1.310e-06

Coefficients:
              Estimate Std. Error z value Pr(>|z|)
(Intercept)   -9.569    21868.559      0      1
f1             -1.088     3759.792      0      1
f2              2.707     6080.568      0      1
f3             -1.826     3871.842      0      1
f4              2.963     8839.496      0      1

(Dispersion parameter for binomial family taken to be 1)

    Null deviance: 5.7416e+00  on 6  degrees of freedom
Residual deviance: 3.6936e-10  on 2  degrees of freedom
AIC: 10

Number of Fisher Scoring iterations: 23

```

Screenshot

Screenshot

- Résultat 2 : Modèle probit Allemagne après ajustements

```
> glm(modela)
```

```
Call: glm(formula = modela)
```

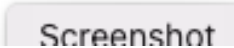
```
Coefficients:
```

```
(Intercept)          a1          a2          a3          a4
    0.31640      0.01041      0.06652      0.03386     -0.09400
```

```
Degrees of Freedom: 6 Total (i.e. Null); 2 Residual
```

```
Null Deviance:      0.8571
```

```
Residual Deviance: 0.2073      AIC: 7.228
```

```
Call: 
```

```
Call:
```

```
glm(formula = crise ~ a1 + a2 + a3 + a4, family = binomial(link = "probit"),
     data = AL)
```

```
Deviance Residuals:
```

```
      1      2      3      4      5      6      7
-2.110e-08 -7.288e-06 -1.068e-05 -2.110e-08 -9.865e-06  9.731e-06 -1.776e-07
```

```
Coefficients:
```

	Estimate	Std. Error	z value	Pr(> z)
(Intercept)	-4.274	144563.419	0	1
a1	0.677	53936.495	0	1
a2	1.186	66553.815	0	1
a3	1.555	25264.289	0	1
a4	-2.949	7564.696	0	1

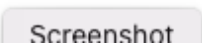
```
(Dispersion parameter for binomial family taken to be 1)
```

```
Null deviance: 5.7416e+00 on 6 degrees of freedom
```

```
Residual deviance: 3.5915e-10 on 2 degrees of freedom
```

```
AIC: 10
```

```
Number of Fisher Scoring iterations: 23
```

```
Call: 
```

- Résultat 3 : Modèle probit Belgique après ajustements

```

> glm(modelb)

Call:  glm(formula = modelb)

Coefficients:
(Intercept)      b1      b2      b3      b4
  0.11540      0.02479      0.17751      0.02861     -0.04675

Degrees of Freedom: 6 Total (i.e. Null); 2 Residual
Null Deviance:      0.8571
Residual Deviance: 0.3383      AIC: 10.66
>
> |
> summary(modelb)

Call:
glm(formula = crise ~ b1 + b2 + b3 + b4, family = binomial(link = "probit"),
    data = BE)

Deviance Residuals:
    1      2      3      4      5      6      7
-8.715e-06 -2.110e-08 -9.959e-06 -1.097e-05 -2.110e-08  1.159e-05 -3.095e-06

Coefficients:
              Estimate Std. Error z value Pr(>|z|)
(Intercept)  -11.2128  13128.3681  -0.001    0.999
b1             0.8365   1633.8961   0.001    1.000
b2             4.7220   9513.9799   0.000    1.000
b3             1.3251   2914.4203   0.000    1.000
b4            -1.7486   2648.8694  -0.001    0.999

(Dispersion parameter for binomial family taken to be 1)

    Null deviance: 5.7416e+00  on 6  degrees of freedom
Residual deviance: 4.3949e-10  on 2  degrees of freedom
AIC: 10

Number of Fisher Scoring iterations: 23

```

- Résultat 4 : Modèle probit Espagne après ajustements

```
> glm(models)
```

```
Call: glm(formula = models)
```

```
Coefficients:
```

```
(Intercept)      s1      s2      s3      s4
  0.22807    0.01193    0.14746    0.04803   -0.07018
```

```
Degrees of Freedom: 6 Total (i.e. Null); 2 Residual
```

```
Null Deviance: 0.8571
```

```
Residual Deviance: 0.4019 AIC: 11.86
```

```
> |
```

Screenshot

```
Call:
```

```
glm(formula = crise ~ s1 + s2 + s3 + s4, family = binomial(link = "probit"),
     data = ES)
```

```
Deviance Residuals:
```

```
      1      2      3      4      5      6      7
-7.355e-06 -1.266e-05 -8.962e-06 -8.063e-07 -3.531e-06  1.221e-05 -2.110e-08
```

```
Coefficients:
```

	Estimate	Std. Error	z value	Pr(> z)
(Intercept)	-2.3534	12483.5462	0	1
s1	0.1917	1072.3243	0	1
s2	2.9304	6286.7954	0	1
s3	1.4135	6350.7564	0	1
s4	-2.3767	6776.5541	0	1

```
(Dispersion parameter for binomial family taken to be 1)
```

```
Null deviance: 5.7416e+00 on 6 degrees of freedom
```

```
Residual deviance: 4.5694e-10 on 2 degrees of freedom
```

```
AIC: 10
```

```
Number of Fisher Scoring iterations: 23
```

- Résultat 5 : Modèle probit Italie après ajustements

```
> glm(modeli)
```

```
Call: glm(formula = modeli)
```

```
Coefficients:
```

(Intercept)	i1	i2	i3	i4
0.15020	-0.05771	0.02637	-0.13433	0.13436

```
Degrees of Freedom: 6 Total (i.e. Null); 2 Residual
```

```
Null Deviance: 0.8571
```

```
Residual Deviance: 0.232 AIC: 8.016
```

```
> summary(modeli)
```

```
Call:
```

```
glm(formula = crise ~ i1 + i2 + i3 + i4, family = binomial(link = "probit"),  
data = IT)
```

```
Deviance Residuals:
```

1	2	3	4	5	6	7
-1.220e-05	-4.174e-06	-7.108e-06	-2.110e-08	-2.110e-08	9.024e-06	-5.064e-06

```
Coefficients:
```

	Estimate	Std. Error	z value	Pr(> z)
(Intercept)	-7.8710	13569.5640	-0.001	1
i1	-1.3585	2621.0086	-0.001	1
i2	0.4243	6452.4692	0.000	1
i3	-3.0795	13961.6922	0.000	1
i4	3.2295	9098.6119	0.000	1

```
(Dispersion parameter for binomial family taken to be 1)
```

```
Null deviance: 5.7416e+00 on 6 degrees of freedom
```

```
Residual deviance: 3.2383e-10 on 2 degrees of freedom
```

```
AIC: 10
```

```
Number of Fisher Scoring iterations: 23
```

Screenshot

- Résultat 6 : Modèle probit Grèce après ajustements

```

> glm(modelg)

Call: glm(formula = modelg)

Coefficients:
(Intercept)      g1      g2      g3      g4
  0.12821    -0.09354   -0.16007   -0.14975    0.18270

Degrees of Freedom: 6 Total (i.e. Null);  2 Residual
Null Deviance:      0.8571
Residual Deviance: 0.3718      AIC: 11.32
>
> summary(modelg)

Call:
glm(formula = crise ~ g1 + g2 + g3 + g4, family = binomial(link = "probit"),
    data = GR)

Deviance Residuals:
    1      2      3      4      5      6      7
-6.469e-06 -2.110e-08 -8.728e-06 -1.260e-05 -2.110e-08  1.182e-05 -4.824e-06

Coefficients:
              Estimate Std. Error z value Pr(>|z|)
(Intercept)   -3.547   14729.402   0.000   1.000
g1             -2.536   3393.440  -0.001   0.999
g2             -2.704   16419.668   0.000   1.000
g3             -6.987   9340.892  -0.001   0.999
g4              7.147   9479.600   0.001   0.999

(Dispersion parameter for binomial family taken to be 1)

    Null deviance: 5.7416e+00  on 6  degrees of freedom
Residual deviance: 4.3999e-10  on 2  degrees of freedom
AIC: 10

Number of Fisher Scoring iterations: 23

```