



UNIVERSITÉ DE NANTES



IAE NANTES
ÉCONOMIE & MANAGEMENT

Séries Temporelles Multivariées :
Modélisation de la relation entre taux d'intérêt et production
nationale

Modèles VAR, SVAR et VECM

Aurouet Lucas, Mounier-Tebas Teodoro

Master EKAP
IAE Nantes
01/02/2021

Table des matières

1	Introduction	2
2	Théorie macroéconomique	3
2.1	La demande en économie ouverte	3
2.2	Équilibre sur le marché des biens et des devises	6
2.2.1	La croix keynésienne	6
2.2.2	la courbe IS	7
2.3	Équilibre sur le marché de la monnaie : la courbe LM	8
2.4	Un modèle à court terme d'une économie ouverte : IS-LM-FX	9
2.5	Politiques monétaires	10
3	Méthodologie économétrique	11
3.1	Modèle VAR(1)	11
3.2	Conditions de Stationnarité	12
3.3	Estimation	12
3.4	Sélection du nombre de retards	15
3.5	Diagnostic des Résidus	15
3.6	Vector Error Correction Model	16
3.7	Impulse Response Functions (IRF)	18
4	Résultats	20
4.1	Ordre d'Intégration des Variables	20
4.2	Procédure de Johansen	22
4.3	Diagnostiques des résidus	23
4.4	Interprétation des Résultats	25
5	Conclusions	28

1 Introduction

“Economics is the science of thinking in terms of models joined to the art of choosing models that are relevant to the contemporary world.” Keynes 1938.

Cette citation résume parfaitement le travail suivant. En effet, notre problématique consiste à répondre à la question suivante : quels sont les facteurs qui déterminent les variations du niveau de production et du taux d'intérêt aux États-Unis, à partir de 1950 jusqu'aux années 2000 ? La théorie macroéconomique impliquant un lien entre ces deux variables est-elle vérifiée dans la réalité ?

Pour répondre à cette problématique, nous commencerons par une révision de la théorie macroéconomique provenant principalement de la pensée keynésienne. Le modèle que nous développerons nous permettra de mettre en évidence les principaux déterminants du taux d'intérêt et de la production d'une économie ouverte, ainsi que les liens qui lient ces deux variables entre elles. Nous déterminerons donc comment le taux d'intérêt peut faire varier la production d'un pays et inversement, si la production a un impact sur le taux d'intérêt.

Dans une seconde partie, nous procéderons à la vérification des liens qui peuvent exister entre le taux d'intérêt et la production des États-Unis donc à partir de données réelles. Les données seront récupérées dans la base de données mise à disposition par la Réserve Fédérale **FED**¹. Pour simplifier la récupération des données nous avons utilisé l'API Quandl permettant le téléchargement des variables nécessaires directement depuis le script. Le taux d'intérêt sera récupéré par la variable : *Interest Rates, Discount Rate for United States* et *Real Gross Domestic Product* pour la production.

La première étape consistera à expliquer la procédure d'estimation économétrique utilisée pour les modélisations de type **VAR**² (Vector Autoregressive) qui est la généralisation d'un modèle autorégressif (AR) au cas multivarié. Puis nous procéderons aux tests et modélisations sur nos variables préalablement traitées. Enfin, nous commenterons les résultats obtenus par les modélisations en commençant par la significativité des coefficients, l'analyse des fonctions d'impulsions ainsi que la vérification des hypothèses sur les résidus.

1. <https://fred.stlouisfed.org/>

2. Modèle Vectoriel Autorégressif développé dans les années 80 par Sims.

2 Théorie macroéconomique

Avant de procéder aux modélisations économétriques pour déterminer les relations entre le taux d'intérêt et la production aux États-Unis, nous allons revenir sur les théories macroéconomiques qui ont découlé du fameux livre de John Maynard Keynes³, *The General Theory of Employment, Interest and Money* parue en 1936. Parmi les théories développées, nous allons procéder à l'explication du modèle **IS-LM-FX** qui est une extension du modèle **IS-LM** "classique" aussi appelé **Hicks–Hansen model** du nom de leurs auteurs. Nous avons fait le choix du modèle IS-LM-FX car il permet de prendre en compte les échanges entre pays, en rajoutant au modèle le marché des devises. En effet, ce modèle fait l'hypothèse d'une économie ouverte ce qui a pour objectif d'être plus ajusté à la réalité économique des pays. À travers ce type de modélisation, nous pourrions mettre en évidence l'impact que peuvent avoir les taux d'intérêt sur le marché des devises affectant les taux de change et ceux-ci, sur la balance commerciale qui affecte à son tour la production du pays. Le modèle que nous allons étudier est une variante du modèle IS-LM, utilisé amplement dans l'étude d'une économie fermée et qui met en évidence la relation entre les taux d'intérêt et la production réelle sur le marché des biens et services ainsi que le marché monétaire. Dans la section suivante, nous expliquerons dans le détail le fonctionnement du modèle IS-LM-FX qui nous permettra de mettre en évidence les relations interdépendantes que peuvent avoir le taux d'intérêt et la production. Dans le but de simplifier la théorie macroéconomique, nous avons fait le choix de l'étudier à court terme. Or, l'hypothèse principale des modèles keynésiens à court terme est la rigidité des prix de telle sorte que ce soient les déplacements de la demande qui déterminent la production. Bien entendu, cela suppose que nous ne prenions pas en compte l'inflation⁴ dans le modèle suivant.

2.1 La demande en économie ouverte

Comme nous l'avons expliqué ci-dessus, à court terme, on suppose que le niveau des prix domestiques et étranger, \bar{P} et \bar{P}^* , reste fixe. Il n'existe donc pas d'inflation espérée ($\pi^e = 0$), ce qui signifie que les valeurs réelles et nominales peuvent être confondues. Nous supposons que les dépenses publiques, \bar{G} et les impôts nets, \bar{T} sont fixes et constants, mais peuvent être utilisés dans le cadre de politiques économiques. Enfin, comme notre principal intérêt est d'étudier les équilibres et fluctuations de l'économie domestique (États-Unis), nous allons considérer que la production étrangère \bar{Y}^* et son taux d'intérêt \bar{i}^* sont donnés et constants. Pour simplifier, nous supposons que la production et le revenu sont équivalents ce qui revient à supposer que le produit intérieur brut (PIB) soit égal au revenu national brut disponible (RNBD). Maintenant que les hypothèses sont formalisées, nous allons expliquer les déterminants de chaque composante affectant la production. Ces composantes sont aux nombres de quatre :

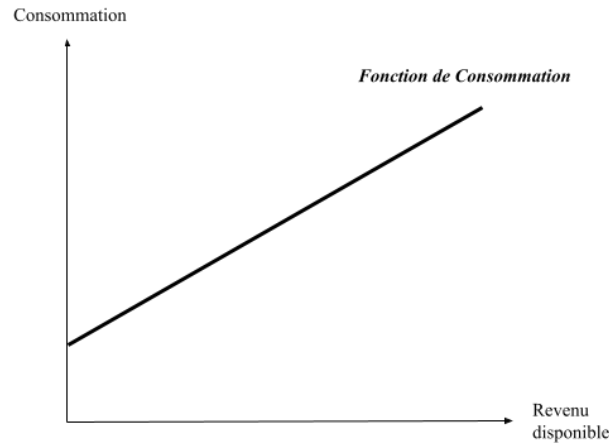
- **Consommation** : La consommation privée notée C dépend du revenu disponible Y_d c'est-à-dire du revenu total perçu moins les impôts \bar{T} . La 'fonction de consommation' peut donc s'écrire de la manière suivante :

$$C = C(Y - \bar{T}) \quad (1)$$

3. économiste britannique né en 1883 et mort en 1946. Il est le fondateur de la macroéconomie keynésienne et considéré comme l'un des plus influents théoriciens de l'économie du xxe siècle.

4. Hausse soutenue et autoentretenu du niveau général des prix

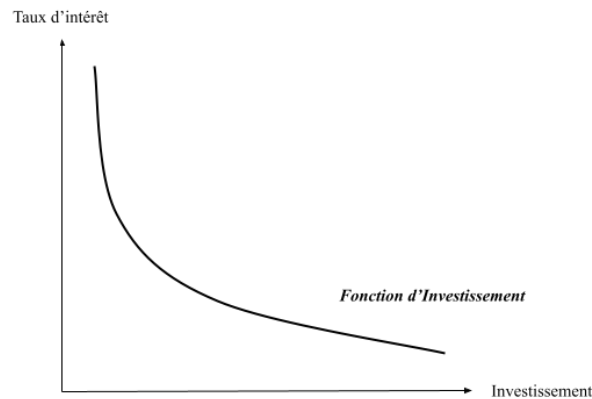
La consommation tend à croître lorsque le revenu disponible augmente c'est donc une fonction croissante du revenu disponible que nous avons représenté graphiquement ci-dessous. La pente de cette fonction de consommation représente la **Propension marginale à consommer** (PMC), qui nous indique quelle proportion pour chaque unité additionnelle de revenu disponible sera dédiée à la consommation.



- **Investissement** : En règle générale, les entreprises investissent dans des projets si et seulement si les rendements du projet sont supérieurs au coût de financement de leurs investissements. Le coût de financement est le **taux d'intérêt réel espéré** qui correspond au taux d'intérêt nominal moins le taux d'inflation espéré : $r^e = i - \pi^e$. Cependant, comme nous avons supposé que le taux d'inflation espéré est nul, le taux d'intérêt réel est égal au taux d'intérêt nominal : $r^e = i$. Quand le taux d'intérêt espéré diminue, on peut s'attendre à une hausse des projets d'investissement. Notre modèle suppose donc que l'investissement I est une fonction décroissante du taux d'intérêt réel. Qui peut s'écrire ainsi :

$$I = I(i) \quad (2)$$

Le graphique suivant représente cette fonction.



- **Gouvernement** : Dans ce modèle les fonctions du gouvernement sont très simples, il collecte les impôts T sur les ménages et réalise des dépenses publiques G en biens et services tels que les dépenses militaires, universitaires, en infrastructures et bien d'autres. Nous pouvons faire face à trois possibilités en ce qui concerne le budget gouvernemental : Le cas où les dépenses publiques coïncident parfaitement avec les recettes ce qui correspond à un budget équilibré $G = T$. Lorsque $G > T$ alors il y a un excédent budgétaire et enfin dans le cas contraire ou $G < T$ alors le gouvernement fait face à un déficit budgétaire. La politique fiscale consiste à gérer cet équilibre entre les recettes qui émanent des impôts et les dépenses nécessaires au bon fonctionnement d'un pays. Dans notre analyse nous allons simplement supposer que les impôts et les dépenses publiques sont fixés et donc exogènes :

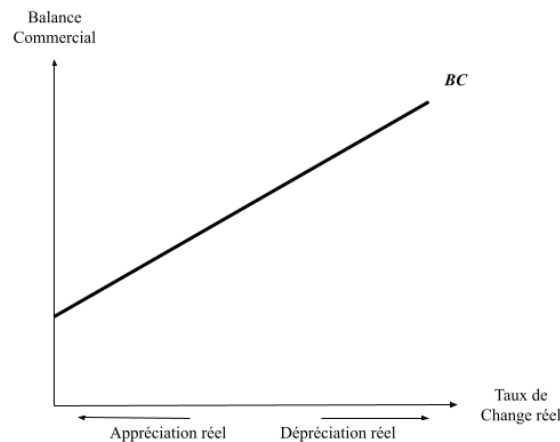
$$G = \bar{G} \quad (3)$$

$$T = \bar{T} \quad (4)$$

- **Balance Commerciale** : La balance commerciale est égale aux exportations moins les importations et mesure l'impact du commerce extérieur sur la demande de production domestique. Les principaux facteurs qui influencent la balance commerciale seront dans ce modèle le taux de change réel et les niveaux de revenu domestique ainsi qu'étranger. En ce qui concerne le taux de change réel, ce sont les prix des biens ou services à l'étranger comparé à ceux du pays domestique. Si le taux de change de l'économie domestique est E , et le niveau des prix \bar{P} et \bar{P}^* à l'étranger, alors le taux de change réel q est égal à $q = \frac{E\bar{P}^*}{\bar{P}}$. Nous pouvons aussi supposer que le niveau des revenus disponibles domestiques devrait augmenter les importations et donc diminuer la balance commerciale. De la même manière, une augmentation du revenu disponible pour le reste du monde augmentera les exportations provoquant une hausse de la balance commerciale. Si nous combinons les trois critères, la fonction de la balance commerciale peut s'écrire de la manière suivante :

$$BC = BC\left(\frac{E\bar{P}^*}{\bar{P}}, Y - \bar{T}, Y^* - \bar{T}^*\right) \quad (5)$$

La représentation graphique ci-dessous met en évidence la relation croissante entre la balance commerciale et le taux de change réel.



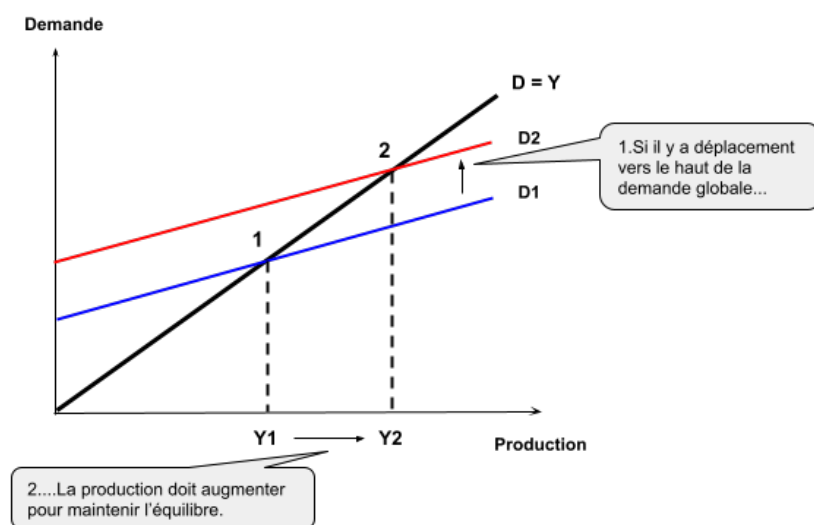
2.2 Équilibre sur le marché des biens et des devises

2.2.1 La croix keynésienne

Jusqu'ici nous avons analysé les déterminants de chacune des composantes de la demande. Nous allons maintenant les réunir pour montrer que le marché des biens et services est en équilibre lorsque la **demande globale** est égale à l'**offre globale**. L'offre globale de biens et services correspond à la production totale : $Offre = PIB = Y$. En ce qui concerne la demande globale, elle est la somme des différentes sources de demande auquel répond l'offre globale et s'écrit comme : $Demande = D = C + I + G + BC$. On peut remplacer chaque partie de la demande par l'équation qui lui correspond ce qui revient à dire que la demande est égale à :

$$D = C(Y - \bar{T}) + I = I(i) + \bar{G} + BC(\frac{E\bar{P}^*}{\bar{P}}, Y - \bar{T}, Y^* - \bar{T}^*) \quad (6)$$

Nous obtenons donc la condition d'équilibre sur le marché des biens et services lorsque la demande D est égale à l'offre Y à l'équilibre. Le graphique suivant montre l'exemple d'une hausse de la demande et son effet sur la production. L'équilibre initial sur le marché des biens correspondant au point 1, où l'offre et la demande s'ajuste avec une production égale à Y_1 . Une hausse de la demande D , pouvant venir de l'une ou plusieurs des composantes de la demande C , I , G ou BC , déplace la courbe de D_1 à D_2 . L'équilibre se déplace au point 2 où l'offre et la demande se réajuste à un niveau de production supérieur Y_2 .



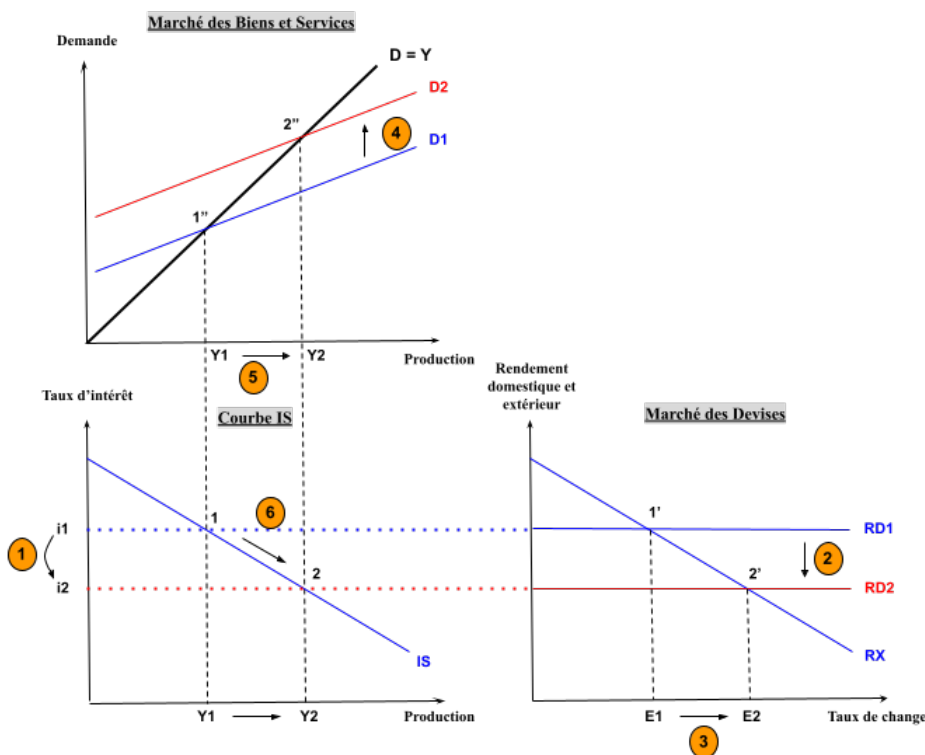
Une augmentation de la demande provoquant une hausse de la production, comme le représente le graphique précédent, peut provenir de politiques fiscales telles qu'une baisse de l'imposition \bar{T} qui augmente le revenu disponible et donc la consommation des ménages ou une hausse des dépenses publiques \bar{G} . Elle peut provenir d'une politique monétaire de réductions des taux d'intérêt i qui incite à l'investissement des entreprises. Enfin cette hausse peut provenir d'une dépréciation du taux de change E due à une réduction des prix domestiques P ou une hausse des prix à l'étranger P^* qui favorise les exportations et réduit les importations améliorant ainsi la balance commerciale.

2.2.2 la courbe IS

Maintenant que nous avons vu comment s'obtient l'équilibre sur le marché des biens et services, nous pouvons construire la courbe IS qui est l'une des composantes du modèle IS-LM-FX. La courbe IS montre toutes les combinaisons de la production, Y , et du taux d'intérêt, i , qui permettent l'équilibre simultané sur le marché des biens ainsi que celui des devises. Le schéma qui suit nous montre cette construction de la courbe IS avec les équilibres présents sur les deux marchés. En ce qui concerne le marché des devises, l'équilibre s'obtient lorsque le rendement espéré d'un dépôt bancaire dans le pays domestique, RD est égal au rendement espéré d'un dépôt bancaire à l'étranger RX . Cette égalité est connue sous le nom de **parité des taux d'intérêt non couverte** (PTINC) et peut s'écrire de la manière suivante :

$$\underbrace{i}_{\text{Tx d'intérêt domestique}} = \underbrace{i^*}_{\text{Tx d'intérêt étranger}} + \underbrace{\left(\frac{E^e}{E} - 1\right)}_{\text{Tx de dépréciation de la monnaie domestique}} \quad (7)$$

La partie droite de l'équation représente donc les rendements extérieurs attendus représentés par la courbe RX sur le marché des devises. Avec un taux d'intérêt étranger donné i^* et un taux de change espéré connu E^e , le rendement extérieur diminue au fur et à mesure que le taux de change E augmente (se déprécie). L'intuition découlant de ce résultat est que plus la monnaie étrangère coûte chère aujourd'hui, plus le rendement espéré d'un dépôt à l'étranger sera faible ceteris paribus. La courbe RX est donc décroissante. Le rendement domestique représenté par la ligne horizontale RD correspond, quant à lui, au taux d'intérêt domestique i .



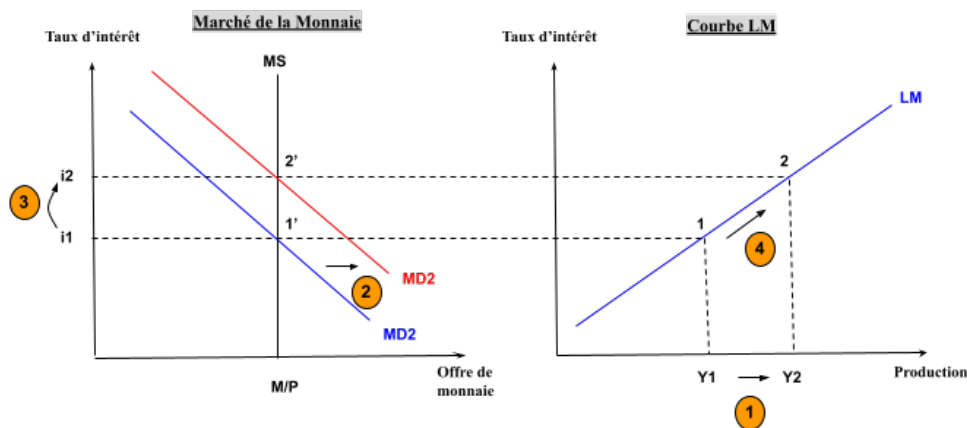
Dans notre cas, ce qui nous intéresse dans le graphique précédent est de montrer les liens qui existent entre le taux d'intérêt et la production en économie ouverte. En effet, une diminution du taux d'intérêt, provoque une diminution des rendements domestiques et donc une dépréciation du taux de change pour conserver la PTINC sur le marché des devises. Cette diminution du taux d'intérêt incite à l'investissement et la dépréciation du taux de change améliore la balance commerciale ce qui fait augmenter la demande globale sur le marché des biens et services. La production s'ajuste à la nouvelle demande passant de $Y1$ à $Y2$. La théorie économique nous indique donc que l'on peut s'attendre à une hausse de la production lorsque le taux d'intérêt diminue.

2.3 Équilibre sur le marché de la monnaie : la courbe LM

Nous allons rajouter maintenant le marché de la monnaie à notre modèle qui aboutira à la construction de la courbe LM, deuxième composante du modèle IS-LM-FX. Le marché de la monnaie est constitué de demandeurs de monnaie réel $L(i)Y$ qui demanderont de la monnaie à chaque fois qu'ils ont un besoin de liquidité pour la réalisation de transactions diverses. L'augmentation des transactions sera due à la hausse du revenu disponible Y_d qui découle du revenu national Y . Cependant, la demande de monnaie diminuera à chaque fois qu'augmentera le coût d'opportunité (le taux d'intérêt) de conserver la monnaie. En effet, la hausse du taux d'intérêt provoque une augmentation de la rentabilité des dépôts bancaires (la monnaie prend de la valeur par elle-même) et donc moins grande est la demande de monnaie fiduciaire. Pour répondre à cette demande de monnaie, la banque centrale s'occupe de l'offre de celle-ci et à court terme cette offre est constante $\frac{M}{P}$ et elle est représentée par une ligne verticale. L'équilibre sur le marché de la monnaie exige donc que la demande réelle de monnaie soit égale à l'offre réelle de monnaie de telle sorte que :

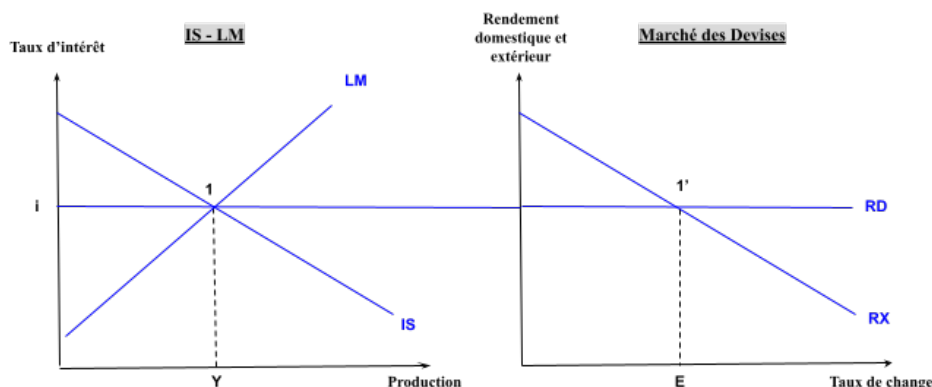
$$\frac{M}{P} = L(i)Y \quad (8)$$

Cette équation est à la base de la courbe LM : ceteris paribus, une hausse de la production provoque une hausse de la demande de monnaie et donc du taux d'intérêt. La courbe LM représente donc l'ensemble des valeurs (Y, i) qui garantit l'équilibre sur le marché de la monnaie avec une pente positive.



2.4 Un modèle à court terme d'une économie ouverte : IS-LM-FX

À partir de maintenant, nous avons toutes les conditions pour représenter l'équilibre d'une économie ouverte sur le marché des biens, de la monnaie et des devises à la fois. La figure suivante combine la courbe IS, la courbe LM et le marché des devises FX. Le marché des biens et des devises sont en équilibre si et seulement si l'économie se retrouve le long de la courbe IS. Le marché de la monnaie est en équilibre si et seulement si l'économie se retrouve le long de la courbe LM. À partir de ce constat, on peut en déduire que les trois marchés se retrouveront en équilibre si et seulement si l'économie se retrouve au point 1, l'unique point d'intersection entre les courbes IS et LM. Pour le marché des devises l'équilibre se situe au niveau du point 1', ou les rendements extérieurs, RX , sont égaux aux rendements domestiques, eux-mêmes égaux au taux d'intérêt domestique, $RD = i$. Les courbes IS et LM sont construites dans une unique figure qui détermine les valeurs de la production et du taux d'intérêt définissant l'unique point d'équilibre à court terme. L'ajout du diagramme du marché des devises permet de résumer les conditions sur les trois marchés conjointement : biens, monnaie, devises.



Le diagramme IS-LM-FX sert à évaluer l'impact à court terme de différentes politiques économiques ainsi que les conséquences sur la production et le taux d'intérêt d'un choc exogène d'offre ou de demande sur l'un des trois marchés. Ce diagramme est donc parfait pour comprendre quels sont les liens sous-jacents entre le taux d'intérêt i et la production Y à court terme d'une économie ouverte comme le sont les États-Unis.

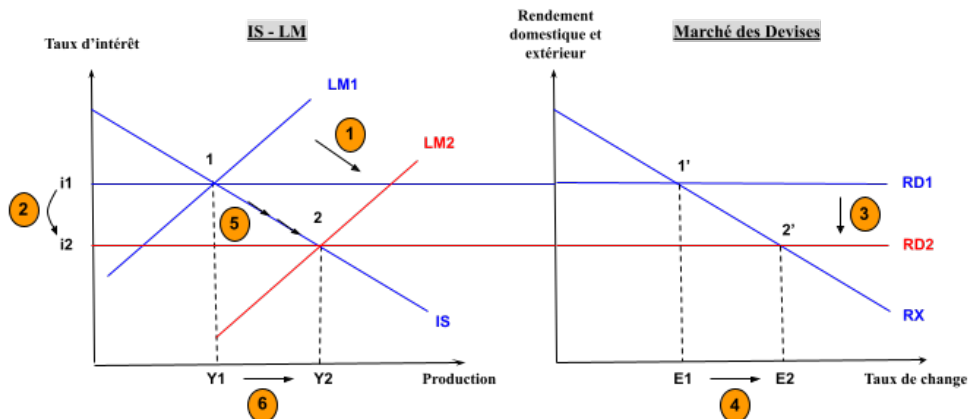
Avant de passer aux modélisations de type VAR, nous allons étudier comment le modèle IS-LM-FX, explique l'effet d'une politique monétaire expansionniste sur la production à court terme et avec un taux de change flottant⁵. Autrement dit, qu'est-ce que la théorie macroéconomique va prédire pour la quantité de production suite à une politique monétaire de réduction des taux d'intérêt. Les résultats pourront être vérifiés par la suite lorsque nous étudierons les liens entre le taux d'intérêt et la production des États-Unis avec des données réelles.

5. Un taux de change flottant correspond à une devise dont le cours est déterminé par des facteurs d'offre et de demande relatifs à d'autres devises. Un taux de change flottant diffère d'un taux de change fixe, qui est entièrement déterminé par le gouvernement de la devise en question.

2.5 Politiques monétaires

Dans un régime de taux de change flottant, le taux d'intérêt et le taux de change peuvent varier librement pour maintenir l'équilibre sur les marchés de la monnaie et des devises. Ainsi, la politique gouvernementale peut déplacer n'importe quelle courbe (IS ou LM ou les deux à la fois) pour atteindre ses objectifs de niveau de production potentielle⁶ par exemple.

Le graphique suivant montre les déplacements que provoque une politique monétaire expansionniste sur les différents marchés. Comme nous pouvons le constater, une politique monétaire expansionniste réduit le taux d'intérêt et donc diminue les rendements domestiques sur le marché des devises. Cette réduction des rendements domestiques entraîne une dépréciation du taux de change afin de maintenir l'égalité avec les rendements espérés étrangers. La dépréciation du taux de change améliore la balance commerciale et la réduction du taux d'intérêt incite à l'investissement des entreprises ce qui provoque une hausse de la demande globale sur le marché des biens et services et donc nécessairement un ajustement de la production globale. En résumé l'expansion monétaire déplace la courbe LM vers la droite, et la réduction des taux d'intérêt fait augmenter la production et déprécier le taux de change.



En règle générale, la possibilité de manipuler les courbes IS et LM donne au gouvernement la capacité de mettre en place des politiques de stabilisation de l'économie afin de faire face aux perturbations et ainsi de maintenir le niveau de production au plein emploi. Il est cependant compliqué de trouver les politiques adéquates à chaque situation d'autant plus qu'une fois les politiques prises, il faut un certain temps pour que celles-ci soient effectives, ce qui peut les rendre inefficaces et parfois même contre-productives. Il est temps de passer aux modélisations VAR mais avant cela nous reprendrons la méthodologie économétrique sous-jacente à ce type de modèle.

6. Désigne le niveau de produit intérieur brut le plus élevé soutenable à long terme pour une économie donnée.

3 Méthodologie économétrique

3.1 Modèle VAR(1)

Dans cette partie, nous nous intéressons à la méthodologie employée. En particulier, nous détaillons brièvement les modèles Vector Auto Regressive (VAR) et Structural Vector Error Correction Model (VECM). Nous passons également en revue certains tests statistiques utilisés ainsi que quelques notions de calcul matriciel qui nous seront utiles par la suite.

intéressons-nous dans un premier temps au modèle VAR, le modèle VECM étant une extension de ce dernier, la formalisation des deux modèles est similaire sur beaucoup de points. Nous reviendrons ensuite sur ce qui différencie un VAR d'un VECM. Le modèle VAR permet de modéliser des interactions dynamiques entre plusieurs grandeurs. Chaque variable considérée dans le modèle est fonction des autres, ce qui permet de modéliser simultanément l'effet de chaque variable sur les autres. Le modèle VAR peut s'écrire de plusieurs façons, nous présentons ci-dessous les écritures qui nous seront utiles ⁷ :

$$\begin{cases} y_{1,t} = \phi_{11}y_{1,t-1} + \phi_{12}y_{2,t-1} + \epsilon_{1,t} \\ y_{2,t} = \phi_{21}y_{1,t-1} + \phi_{22}y_{2,t-1} + \epsilon_{2,t} \end{cases} \quad (9)$$

Est l'écriture sous forme de système d'équations.

$$\begin{pmatrix} y_{1,t} \\ y_{2,t} \end{pmatrix} = \begin{bmatrix} \phi_{11} & \phi_{12} \\ \phi_{21} & \phi_{22} \end{bmatrix} \begin{pmatrix} y_{1,t-1} \\ y_{2,t-1} \end{pmatrix} + \begin{pmatrix} \epsilon_{1,t} \\ \epsilon_{2,t} \end{pmatrix} \quad (10)$$

Est l'écriture matricielle du système.

$$y_t = \Phi_1 y_{t-1} + \epsilon_t \quad (11)$$

Est l'écriture vectorielle du système.

Avec $y_t = \begin{pmatrix} y_{1,t} \\ y_{2,t} \end{pmatrix}$, $\Phi_1 = \begin{bmatrix} \phi_{11} & \phi_{12} \\ \phi_{21} & \phi_{22} \end{bmatrix}$ et $\epsilon_t = \begin{pmatrix} \epsilon_{1,t} \\ \epsilon_{2,t} \end{pmatrix}$

Nous pouvons constater que chaque variable apparaît dans chaque équation. Autrement dit, $y_{1,t}$ est déterminée à l'aide d'une combinaison linéaire de ses valeurs passées et des valeurs passées de $y_{2,t}$. Il en va de même pour $y_{2,t}$. Dans cet exemple précis, nous utilisons 2 variables et un retard d'une période. Nous sommes dans le cadre d'un VAR(1), 1 indiquant l'ordre du processus. Le modèle VAR est généralisable à davantage de variables et davantage de retards, on parle alors d'un VAR(p), p représentant le nombre de valeurs passées utilisées.

La matrice Φ_1 permet de mettre en évidence les relations dynamiques entre les variables, si ϕ_{12} et ϕ_{21} sont nuls, les variables ne sont pas corrélées et chaque série suit un processus AR(1) ⁸. Si ces deux coefficients sont non nuls, les variables ont un impact, l'une sur l'autre. Si seulement l'un des deux coefficients est nul, les corrélations sont unidirectionnelles, une variable suit un AR(1) et ne dépend que d'elle-même, tandis que l'autre dépend de la première et d'elle-même. Les coefficients sur la diagonale de Φ_1 sont les coefficients associés au processus AR(p) de chaque variable.

7. Nous nous plaçons dans un cadre avec 2 variables dynamiquement liées et un VAR(1)

8. Puisque qu'elles ne dépendent que de leurs propres valeurs retardées

3.2 Conditions de Stationnarité

La stationnarité est une propriété souhaitable lors de l'estimation de modèles de séries temporelles. La stationnarité assure que la distribution de la variable étudiée ne varie pas dans le temps, que sa moyenne inconditionnelle, ainsi que sa variance, sont constantes. Nous voyons ici les conditions qui font d'un processus VAR(p), un processus stationnaire. En réalité, nous nous intéressons ici au fait que VAR(p) soit **faiblement** stationnaire. À l'inverse d'un processus **strictement** stationnaire, où la distribution de la variable est constante, il suffit que la variance et la moyenne inconditionnelles soient constantes pour qu'un processus soit qualifié de faiblement stationnaire. Nous avons donc deux conditions pour atteindre cette stationnarité faible.

Si l'on reprend la forme vectorielle du VAR(1) : $y_t = \Phi_1 y_{t-1} + \epsilon_t$:

1. $E[y_t] = \mu$, avec μ scalaire
2. $Cov(y_t) = E[(y_t - \mu)(y_t - \mu)^T] = \Sigma_y$, une matrice de variance-covariance constante

Si la série débute à la marque $t = v$, avec pour valeur y_v , il est possible d'écrire la suite par récurrence suivante :

Démonstration.

$$\begin{aligned}
 y_t &= \Phi_1 y_{t-1} + \epsilon_t \\
 &= \Phi_1 (\Phi_1 y_{t-2} + \epsilon_{t-1}) + \epsilon_t \\
 &= \Phi_1^2 y_{t-2} + \Phi_1 \epsilon_{t-1} + \epsilon_t \\
 &= \Phi_1^3 y_{t-3} + \Phi_1^2 \epsilon_{t-2} + \Phi_1 \epsilon_{t-1} + \epsilon_t \\
 &= \dots \\
 &= \Phi_1^{t-v} y_v + \sum_{i=0}^{t-1} \Phi_1^i \epsilon_{t-i}
 \end{aligned}$$

□

Pour que les deux conditions citées plus haut soient respectées, et que VAR(1) soit stationnaire, il faut que y_t soit indépendant de y_v . Autrement dit ; on veut $\Phi_1^{t-v} \xrightarrow[v \rightarrow -\infty]{} 0$.

Or si les valeurs propres de la matrice Φ_1^{t-v} , λ_j^{t-v} sont égales à 0 Φ_1^{t-v} est une matrice nulle, ce qui implique que y_v n'a aucun impact sur y_t . Cette dernière condition requiert à son tour que toutes les valeurs propres de Φ_1 soient inférieures à 1 en valeur absolue.

3.3 Estimation

L'estimation d'un modèle VAR peut se faire de différentes manières :

- Generalized Least Squares (GLS)
- Ordinary Least Squares (OLS)
- Maximum Likelihood Estimation (MLE)

Nous présentons brièvement chacune d'entre elles. L'estimation par GLS d'un modèle VAR(p) se déroule comme suit :

Nous partons d'un modèle VAR(p) standard, à k équations tel que ;

$$k, t > 2, tk, t > 1, t \left\{ \begin{array}{l} y_{1,t} = \phi_{11}^1 y_{1,t-1} + \phi_{12}^1 y_{2,t-1} + \dots + \phi_{kp}^1 y_{k,t-p} + \epsilon_{1,t} \\ y_{2,t} = \phi_{11}^2 y_{1,t-1} + \phi_{12}^2 y_{2,t-1} + \dots + \phi_{kp}^2 y_{k,t-p} + \epsilon_{2,t} \\ \dots \\ y_{k,t} = \phi_{11}^k y_{1,t-1} + \phi_{12}^k y_{2,t-1} + \dots + \phi_{kp}^k y_{k,t-p} + \epsilon_{k,t} \end{array} \right. \quad (12)$$

Chaque ligne contient donc autant d'éléments que le système contient de variables (k) et de retards (p) pour un total de kxp éléments. Dans un VAR(1) à 2 variables, les lignes contiendraient 2 éléments auto-régressifs, la variable n°1 retardée et la variable n°2 retardée. Dans un VAR(2) à deux variables chaque ligne contiendra la variable n°1 deux fois, une pour chaque retard et la variable n°2 deux fois. Dans un VAR(p) à k variables

chaque ligne contient donc kxp éléments auto-régressifs. Si l'on pose : $\begin{pmatrix} y_{1,1} & \dots & y_{1,k} \\ \dots & \dots & \dots \\ y_{T,1} & \dots & y_{T,k} \end{pmatrix}$,

$Y_{t-1} = \begin{pmatrix} y_{t,1} & \dots & y_{t,kp} \\ \dots & \dots & \dots \\ y_{T,11} & \dots & y_{T,kp} \end{pmatrix}$, $B_1 = \begin{pmatrix} \phi_{1,1} & \dots & \phi_{1,k} \\ \dots & \dots & \dots \\ \phi_{pxk,1} & \dots & \phi_{pxk,k} \end{pmatrix}$, $U_t = \begin{pmatrix} \epsilon_{1,1} & \dots & \epsilon_{1,k} \\ \dots & \dots & \dots \\ \epsilon_{T,1} & \dots & \epsilon_{T,k} \end{pmatrix}$ nous pouvons récrire le système ci-dessus comme :

$$Y_t = Y_{t-1}B_1 + U_t \quad (13)$$

On peut ensuite vectoriser l'équation obtenue, $y = (I_k \otimes Y_{t-1})\beta + u$ où y est un vecteur $((T \cdot k) \times 1)$, $(I_k \otimes X)$ est une matrice $(kTx(k^2p + k))$, B est un vecteur $((k^2p + k) \times 1)$ et u est un vecteur $((T \cdot k) \times 1)$. Dans le système précédent, chaque équation représentant une variable au temps t , $t = 1, \dots, T$. La différence avec $y = (I_k \otimes Y_{t-1})\beta + u$ est que cette écriture n'est pas l'écriture du modèle VAR(p) mais l'ensemble des données sous forme de matrices, puis transformées en vecteurs. C'est à partir de ces vecteurs que les estimations par moindres carrés peut s'effectuer. On peut désormais dériver l'estimateur GLS \hat{B} de B :

Démonstration.

$$\begin{aligned}
SS(\hat{\beta}) &= u'(\Sigma_\epsilon \otimes I_t)^{-1}u \\
&= [y - (I_k \otimes X)\hat{\beta}]'(\Sigma_\epsilon^{-1} \otimes I_t)[y - (I_k \otimes X)\hat{\beta}] \\
&= [y' - \hat{\beta}'(I_k \otimes X')](\Sigma_\epsilon^{-1} \otimes I_t)[y - (I_k \otimes X)\hat{\beta}] \\
&= y'(\Sigma_\epsilon^{-1} \otimes I_t)y - 2\hat{\beta}'(\Sigma \otimes X')y + \hat{\beta}'(\Sigma_\epsilon^{-1} \otimes X'X)\hat{\beta}
\end{aligned}$$

$$\frac{\partial SS(\hat{\beta})}{\partial \hat{\beta}} = -2(\Sigma_\epsilon^{-1} \otimes X')y + 2(\Sigma_\epsilon^{-1} \otimes X'X)\hat{\beta} = 0$$

$$\begin{aligned}
\hat{\beta} &= (\Sigma_\epsilon^{-1} \otimes X'X)^{-1}(\Sigma_\epsilon^{-1} \otimes X')y \\
&= [\Sigma_\epsilon \otimes X'X]^{-1}(\Sigma_\epsilon^{-1} \otimes X')y \\
&= [I_k \otimes (X'X)^{-1}X']y \\
\hat{\beta} &= vec((X'X)^{-1}(X'Y))
\end{aligned}$$

$$\hat{B} = (X'X)^{-1}(X'Y)$$

□

Σ , matrice de variance covariance des résidus est estimée par : $\hat{\Sigma} = \frac{1}{T-p-(kp)} \sum_{t=p+1}^T \hat{U}'\hat{U}$. L'estimateur GLS est le meilleur estimateur si $\Sigma \neq I$. Autrement dit, si les résidus sont corrélés entre eux. Si les résidus sont non corrélés, $\Sigma \neq I$ et l'estimation GLS est identique à l'estimation OLS. Les résidus sont en réalité souvent corrélés, et dans ce cas de figure l'estimation OLS donne des estimateurs non-biaisés mais pas de variance minimale.

L'estimation peut également se faire par maximum de vraisemblance (MLE). Si l'on suppose que les résidus sont distribués normalement, la fonction de vraisemblance (de probabilité conditionnelle) des données peut s'écrire :

Démonstration.

$$\begin{aligned}
\mathcal{L}(y|X, \beta, \Sigma) &= \prod_{t=p+1}^T p(y|x, \beta, \Sigma) \\
\mathcal{L}(y|X, \beta, \Sigma) &= \prod_{t=p+1}^T p(u|x, \beta, \Sigma) \\
\mathcal{L}(y|X, \beta, \Sigma) &= \frac{1}{2\pi^{\frac{k}{2}}} |\Sigma|^{\frac{1}{2}} e^{(-\frac{1}{2}u'\Sigma^{-1}u)} \\
&\propto |\Sigma|^{\frac{-(T-p)}{2}} e^{(-\frac{1}{2} \sum_{t=p+1}^T trace(u'\Sigma^{-1}u))} \\
\log(\mathcal{L}) &= \frac{-(T-p)}{2} \log(|\Sigma|) - \frac{1}{2} \sum_{t=p+1}^T trace(u'\Sigma^{-1}u) \\
\text{Or ; } \sum_{t=p+1}^T trace(u'\Sigma^{-1}u) &= SS(\beta) \\
\log(\mathcal{L}) &= \frac{-(T-p)}{2} \log(|\Sigma|) - \frac{1}{2} SS(\beta)
\end{aligned}$$

□

En dérivant pour trouver β qui maximise la Log-vraisemblance, $\frac{\partial \log(\mathcal{L})(\beta)}{\partial \beta}$, le terme $-\frac{(T-p)}{2} \log(|\Sigma|)$ va disparaître et le problème revient à minimiser $SS(\beta)$. Autrement dit, l'estimation MLE est identique à l'estimation OLS pour le VAR(p).

3.4 Sélection du nombre de retards

Pour estimer un modèle VAR, il est nécessaire de savoir combien de retards seront inclus. Pour nous aider à choisir le nombre de retards optimal, nous pouvons nous appuyer sur les critères d'information. Ces critères permettent de mettre en valeur la qualité d'ajustement du modèle relative à sa complexité. En effet, le nombre de paramètres à estimer est fonction du nombre de retards ainsi que du nombre de variables (équations) à estimer. Ce nombre peut vite devenir très important et pour conserver un modèle pertinent, nous devons certes nous attarder sur sa qualité d'ajustement, mais également nous assurer qu'il n'est pas trop complexe. C'est dans cette optique que l'utilisation des critères d'information est pertinente ici. Pour comprendre comment ces critères mesurent la qualité du modèle, il est important de détailler la façon dont chacun d'entre eux est calculé. Nous utiliserons trois critères différents pour sélectionner le nombre de retards optimal :

1. Akaike Information Criterion, $AIC = 2k - 2 \log(\mathcal{L})$
2. Bayesian Information Criterion, $BIC = 2 \log(T)k - 2 \log(\mathcal{L})$
3. Hannan-Quinn Information Criterion, $HQ = 2 \log(\log(T))k - 2 \log(\mathcal{L})$

La pénalisation apportée au nombre de paramètres (complexité du modèle) est représentée par le terme qui multiplie k . Dès lors que $2 \log(n)k$ (resp. $2 \log(\log(T))k$) est plus grand que $2k$, la pénalisation apportée par BIC (resp. HQ) est plus sévère que celle de AIC. C'est à dire pour $T > 4$ (resp. $T > 15$). Par construction, BIC et HQ seront plus grands que AIC si le nombre d'observations est suffisant. La valeur pour laquelle les critères sont les meilleurs sont les valeurs les plus petits possibles. Puisque l'on retranche à la complexité du modèle sa qualité d'ajustement⁹, les valeurs faibles indiquent que le modèle est simple, et bien ajusté. Des valeurs élevées, montrent que le modèle est complexe et mal ajusté. Nous choisirons donc le nombre de retards qui donne les critères les plus faibles.

3.5 Diagnostic des Résidus

L'efficacité des modèles VAR reposent sur certaines hypothèses, notamment une selon laquelle les résidus ne sont pas autocorrélés et une autre selon laquelle ils suivent une loi normale. Nous avons vu plus tôt l'impact de ces hypothèses sur les estimations du modèle. Si elles sont vérifiées, les estimateurs sont les meilleurs estimateurs possibles. Si elles ne sont pas vérifiées, les estimateurs ne seront pas de variance minimale au mieux, au pire, biaisés. Pour vérifier la validité de ces hypothèses, des tests statistiques ont été spécifiés. Le test de Jarque-Bera de normalité des résidus vérifie que les moments d'ordre 3 et 4 (skewness et kurtosis) sont bien ceux d'une loi normale (0 et 3). Ce test revient donc à vérifier :

- H_0 : les résidus sont normalement distribués.
- H_a : les résidus ne sont pas normalement distribués.

9. Mesurée par la vraisemblance.

La statistique du test $JB = \frac{T-k}{6}(S^2 + \frac{(K-3)^2}{4})$ ¹⁰ suit une loi du χ^2 à deux degrés de liberté sous l'hypothèse H_0 . Une statistique JB supérieure à la valeur critique (p-value inférieure au seuil de confiance établi) signifie ainsi que l'on rejette l'hypothèse H_0 de normalité des résidus.

Le test Q de Ljung-Box vérifie l'auto-corrélation des résidus :

- H_0 : les résidus sont indépendamment distribués.
- H_a : les résidus ne sont pas indépendamment distribués.

La statistique du test $Q = T(T+2) \sum_{k=1}^h \frac{\rho_k^2}{T-k}$ ¹¹ suit une loi du χ^2 à h degrés de liberté sous l'hypothèse H_0 . Une statistique Q supérieure à la valeur critique (p-value inférieure au seuil de confiance établi) signifie ainsi que l'on rejette l'hypothèse H_0 d'absence d'autocorrélation dans les résidus jusqu'à l'ordre h.

3.6 Vector Error Correction Model

Les modèles VAR supposent que les variables sont stationnaires, ou I(0). Il est en revanche relativement commun que les séries macro-économiques ne soient pas stationnaires. Les grandeurs comme le PIB, l'indice des prix à la consommation, les salaires, etc... sont des grandeurs qui s'accroissent avec le temps il est alors très probable que ces séries exhibent une tendance et ne soient pas stationnaires en moyenne. Le modèle VAR serait, dans ce cas, mal spécifié. Il existe cependant un modèle capable de modéliser des séries I(1) sous certaines conditions. Le Vector Error Correction Model (VECM) prend plusieurs variables I(1) cointégrées et modélise leurs relations dynamiques à l'aide d'un modèle VAR où les effets sont décomposés en effets de long terme (relation de cointégration), effets de courts termes de retour à l'équilibre et effets autorégressifs. On voit bien comment, avec un VAR standard les effets associés à la relation de cointégration, de long terme, seraient à tort considérés comme les effets des variables entre elles.

Si nos variables sont I(1), et cointégrées, nous estimerons donc un VECM pour prendre en compte cette relation de long terme. Si les données sont I(0) alors nous pourrions estimer un VAR standard. Si les variables sont I(1) mais ne sont pas cointégrées alors nous pourrions estimer un VAR sur les données différenciées. il faut donc déterminer quel schéma suivent nos données afin de choisir la spécification la plus appropriée. Vérifier si les données sont I(0) ou I(1) est relativement aisé, il existe nombre de tests statistiques qui permettent de mettre en évidence la présence d'une racine unitaire dans les processus générateurs de données. Comme nous l'avons vu, les racines doivent se trouver en dehors du cercle unitaire pour qu'un processus soit stable, si nous sommes en mesure de prouver que ces racines se trouvent sur le cercle unitaire, nous apportons la preuve que le processus n'est pas stationnaire. Ici, nous utiliserons trois tests différents pour obtenir des résultats plus robustes.

1. Augmented Dickey-Fuller (ADF) : le test ADF repose sur un test de significativité de coefficient, on pose le modèle $\Delta y_t = \delta y_{t-1} + u_t$ ¹². Tester pour la présence d'une racine unitaire revient à tester $\delta = 0$. On calcule donc la statistique $DF = \frac{\hat{\delta}}{\sigma^2(\hat{\delta})}$ laquelle nous permet de vérifier $H_0, \delta = 0$ qui stipule que la série possède une racine unitaire.

10. Avec k le nombre de paramètres, K le kurtosis de l'échantillon et S son coefficient d'asymétrie.

11. Avec ρ_k le coefficient de corrélation au lag k et h le nombre de lags testés.

12. Δ est l'opérateur de différenciation de sorte à ce que $\Delta y_t = y_t - y_{t-1}$.

2. Kwiatkowski–Phillips–Schmidt–Shin (KPSS) : le test KPSS scinde la série en un trend déterministe, une random walk et un bruit blanc. On utilise ensuite un test du multiplicateur de Lagrange pour vérifier que la variance de la partie marche aléatoire est nulle, ce qui revient à tester H_0 la série est stationnaire.
3. le test de Elliot Rothenberg and Stock, appelé ADF-GLS est une version modifiée du test ADF qui retire à la série ses constante et tendance de long terme avant d'effectuer un test ADF classique sur la série de-trended.

Ces trois tests nous permettent donc de différencier une série $I(1)$ d'une série $I(0)$. Ils ne nous permettent pas de vérifier qu'une relation de cointégration (une ou plusieurs) existe entre nos variables. Pour ce faire nous employons la procédure de Johansen. Partons d'un VAR(1) avec 2 variables :

$$\begin{pmatrix} y_{1,t} \\ y_{2,t} \end{pmatrix} = \begin{pmatrix} \beta_{10} \\ \beta_{20} \end{pmatrix} + \begin{bmatrix} \beta_{11} & \alpha_{11} \\ \alpha_{21} & \beta_{21} \end{bmatrix} \begin{pmatrix} y_{1,t-1} \\ y_{2,t-1} \end{pmatrix} + \begin{pmatrix} \epsilon_{1,t} \\ \epsilon_{2,t} \end{pmatrix} \quad (14)$$

En différenciant les variables, nous obtenons :

$$\begin{pmatrix} y_{1,t} - y_{1,t-1} \\ y_{2,t} - y_{2,t-1} \end{pmatrix} = \begin{pmatrix} \beta_{10} \\ \beta_{20} \end{pmatrix} + \begin{bmatrix} \beta_{11} - 1 & \alpha_{11} \\ \alpha_{21} & \beta_{21} - 1 \end{bmatrix} \begin{pmatrix} y_{1,t-1} \\ y_{2,t-1} \end{pmatrix} + \begin{pmatrix} \epsilon_{1,t} \\ \epsilon_{2,t} \end{pmatrix} \quad (15)$$

On pose $\pi_{11} = \beta_{11} - 1$, $\pi_{12} = \alpha_{11}$, $\pi_{21} = \alpha_{21}$, $\pi_{22} = \beta_{21} - 1$ pour obtenir :

$$\begin{pmatrix} \Delta y_{1,t} \\ \Delta y_{2,t} \end{pmatrix} = \begin{pmatrix} \beta_{10} \\ \beta_{20} \end{pmatrix} + \begin{bmatrix} \pi_{11} & \pi_{12} \\ \pi_{21} & \pi_{22} \end{bmatrix} \begin{pmatrix} y_{1,t-1} \\ y_{2,t-1} \end{pmatrix} + \begin{pmatrix} \epsilon_{1,t} \\ \epsilon_{2,t} \end{pmatrix} \quad (16)$$

La procédure décrite par Johansen consiste à vérifier le rang de la matrice Π . Si une combinaison linéaire existe entre deux variables, et est $I(0)$, alors ces deux variables sont cointégrées et le rang de Π sera inférieur au nombre d'équations. En vérifiant que le rang de Π est égal au nombre de variables, nous vérifions en réalité qu'aucune combinaison linéaire n'existe entre les variables. Si cette combinaison existe et se révèle être stationnaire, alors nous pourrions dire qu'une relation de cointégration existe entre nos variables. Nous pouvons donc tester $H_0, rang(\Pi) = 0$ contre $H_a, rang(\Pi) > 0$. Si H_0 est rejetée, il existe au moins 1 relation de cointégration. On peut ensuite tester $rang(\Pi) = 1$ contre $H_a, rang(\Pi) > 1$, le rejet de H_0 signifiera alors qu'il existe au moins deux relations de cointégration. Le test continue ainsi jusqu'à trouver le nombre total de relations dans le modèle. Deux méthodes permettent d'obtenir ces résultats, le test de trace et le test de valeurs propres. Si la procédure de Johansen détecte une ou plusieurs relations de cointégration, il est judicieux d'estimer un VECM.

Nous supposons jusqu'à présent un VAR(1), nous allons désormais partir d'un VAR(p). Avec plus d'un retard, il est plus simple de voir comment le VECM combine la relation de long terme pris en compte par la relation de cointégration, et les ajustements de court terme pris en compte par la partie autorégressive sur les variables différenciées. Nous nous plaçons donc dans le cas d'un VAR(p) tel que :

$$Y_t = \phi_1 Y_{t-1} + \dots + \phi_p Y_{t-p} + \epsilon_t \quad (17)$$

Après différenciation des variables, on obtient :

$$\Delta Y_t = \Pi Y_{t-1} + \sum_{i=i+1}^{p-1} \Gamma_i \Delta Y_{t-i} + \epsilon_t \quad (18)$$

Avec $\Gamma_i = -\sum_{i=i+1}^p \Phi_i$. Si $\Pi = 0$ alors $\Delta Y_t = \sum_{i=i+1}^{p-1} \Gamma_i \Delta Y_{t-i} + \epsilon_t$ qui équivaut à un VAR(p) standard, sur les variables différenciées. Si Π est de rang plein (k , le nombre d'équations) alors les séries sont stationnaires, ce cas de figure ne devrait pas se présenter dès lors que les séries sont I(1). Si $0 < \text{rang}(\Pi) < k$ alors il existe $\text{rang}(\Pi) = m$ relations de cointégration, nous nous retrouvons dans le cadre rencontré plus tôt où il convient d'estimer un VECM.

Il est dès lors possible d'obtenir l'équation de la relation de cointégration à partir de la matrice Π . Cette matrice est décomposée en deux matrices ($k \times m$) telles que ; $\Pi = \alpha \beta'$, le VECM devient alors :

$$\begin{pmatrix} \Delta y_{1,t} \\ \dots \\ \Delta y_{k,t} \end{pmatrix} = \begin{pmatrix} \alpha_1 \\ \dots \\ \alpha_k \end{pmatrix} \begin{bmatrix} \beta_1 \\ \dots \\ \beta_k \end{bmatrix}' \begin{pmatrix} y_{1,t-1} \\ \dots \\ y_{k,t-1} \end{pmatrix} + \sum_{i=i+1}^{p-1} \Gamma_i \Delta Y_{t-i} + \begin{pmatrix} \epsilon_{1,t} \\ \dots \\ \epsilon_{k,t} \end{pmatrix} \quad (19)$$

$$\begin{pmatrix} \Delta y_{1,t} \\ \dots \\ \Delta y_{k,t} \end{pmatrix} = \begin{pmatrix} \alpha_1 \\ \dots \\ \alpha_k \end{pmatrix} [\beta_1 \dots \beta_k] \begin{pmatrix} y_{1,t-1} \\ \dots \\ y_{k,t-1} \end{pmatrix} + \sum_{i=i+1}^{p-1} \Gamma_i \Delta Y_{t-i} + \begin{pmatrix} \epsilon_{1,t} \\ \dots \\ \epsilon_{k,t} \end{pmatrix} \quad (20)$$

β et Y_{t-1} sont donc ($m \times k$) et ($k \times 1$), en les multipliant, on obtient :

$$\begin{pmatrix} \Delta y_{1,t} \\ \dots \\ \Delta y_{k,t} \end{pmatrix} = \begin{pmatrix} \alpha_1 \\ \dots \\ \alpha_k \end{pmatrix} + (\beta_1 y_{1,t-1} + \dots + \beta_k y_{k,t-1}) + \sum_{i=i+1}^{p-1} \Gamma_i \Delta Y_{t-i} + \begin{pmatrix} \epsilon_{1,t} \\ \dots \\ \epsilon_{k,t} \end{pmatrix} \quad (21)$$

L'expression entre parenthèses représente l'équation de la relation de long terme et les coefficients de α représentent les vitesses de retour à cette relation de long terme. Les coefficients de Γ représentent le processus auto-régressif entre les variables différenciées et leur valeurs passées.

3.7 Impulse Response Functions (IRF)

Dans le cas où nous avons des séries I(1) cointégrées, le modèle VECM permet de séparer les effets de long terme, de retour à l'équilibre de long terme et d'autocorrélation. En revanche le modèle tel qu'il est ne permet pas de mettre en évidence les effets d'un choc d'une variable sur une autre. Nous faisons l'hypothèse que les termes d'erreurs ne sont pas corrélés dans le temps, en revanche ils peuvent être corrélés entre eux. La réponse d'une variable à un choc exogène (sur ϵ_t) peut être contaminée par les autres termes d'erreur et l'impact obtenu ne mesure les effets ceteris paribus, ce qui est particulièrement intéressant lorsque l'on s'intéresse à l'effet d'une politique macro-économique sur les variables d'intérêt. Afin d'identifier précisément les effets des chocs exogènes, nous devons donc nous débarrasser de la corrélation des résidus et rendre la matrice de variance-covariance de ces derniers, Σ , diagonale.

Cette matrice peut s'écrire sous la forme d'un produit de PP' où P est une matrice triangulaire de la taille de Σ , en introduisant la variable $\eta_t = P^{-1}\epsilon_t$ on obtient un bruit

blanc de variance = 1 ($E[\eta_t] = P^{-1}E[\epsilon_t] = 0$ et $Cov(\eta_t) = P^{-1}\Sigma(P')^{-1} = I$). L'intérêt de cette manipulation est qu'elle rend les chocs orthogonaux entre eux, il devient alors possible d'obtenir des effets ceteris paribus. Ces effets correspondent aux dérivées partielles de la variable répondante aux chocs, k , par rapport aux chocs sur une autre variable, f , $\frac{\partial Y_{k,t}}{\partial \eta_{f,t-h}} = \Psi_{kf}^{(h)}$, h étant l'horizon, $k \neq f$. L'orthogonalisation des chocs suppose en contrepartie que la matrice Ψ des effets des chocs sur les variables est triangulaire. Dans le cas d'une matrice triangulaire inférieure, cela revient à faire l'hypothèse que le choc de la première variable n'a d'effet que sur la première variable, qu'un choc sur la deuxième variable n'impacte que sur la première et la deuxième variable ainsi de suite. Un choc sur n^{ieme} variable n'a d'effet que sur les n premières variables. Puisque les valeurs au-dessus de la diagonale de Ψ sont composées de zéros, les chocs ne peuvent pas se diffuser sur les variables suivantes, uniquement sur les variables précédentes. A l'inverse si Ψ est une matrice triangulaire supérieure, les chocs n'auront d'effet que sur les variables précédentes. L'ordre des variables est donc particulièrement important, il reflète les hypothèses que nous faisons quant à la diffusion des chocs sur les variables. D'autres méthodes permettent d'orthogonaliser les chocs pour obtenir des résidus η bruit blanc sans avoir à choisir l'ordre des variables. Si l'on reprend la matrice P précédente, nous avons $Cov(\eta_t) = P^{-1}\Sigma(P')^{-1} = I$, reformulé autrement ; $\Sigma = P^{-1}Cov(\eta_t)(P')^{-1} = P^{-1}(P')^{-1}$ car $Cov(\eta_t) = I$. Au lieu de rendre cette décomposition possible en supposant que P est triangulaire, on peut voir $\Sigma = P^{-1}(P')^{-1}$ comme un système d'équations et retrouver la matrice P . En revanche Σ ne possède en réalité que $\frac{k(k+1)}{2}$ éléments (l'élément de la i^{ieme} ligne, j^{ieme} colonne étant identique au terme de la j^{ieme} ligne, i^{ieme} colonne). Le système d'équations possède donc $(k \times k)$ inconnues (les $(k \times k)$ éléments de P) et $\frac{k(k+1)}{2}$ équations. Il faut donc $\frac{k(k-1)}{2}$ restrictions sur P pour pouvoir résoudre le système. Ces restrictions peuvent forcer les éléments de P à zéros, rendre 2 ou plusieurs éléments de P égaux ou proportionnels (de sorte à ce que si l'un des éléments est identifié, les autres le soient aussi de manière mécanique) ou selon la théorie économique. La décomposition de Σ effectuée, nous pourrions calculer $\frac{\partial Y_{k,t}}{\partial \eta_{f,t-h}} = \Psi_{kf}^{(h)}$ et obtenir les effets ceteris paribus des chocs sur les variables.

Nous avons vu la spécification du modèle VAR, les conditions pour qu'un tel processus soit stationnaire, comment ce modèle peut être estimé, ainsi que les diagnostics possibles des résidus. Nous nous sommes ensuite intéressés au modèle VECM, extension du modèle VAR avec plusieurs variables cointégrées et comment ce modèle permet de décomposer les relations entre les variables en une composante de long terme et une composante de court terme. Enfin nous avons passé en revue les fonctions de réponse aux chocs, comment les obtenir et quelles informations en tirer. Nous allons désormais estimer ces modèles sur nos données.

4 Résultats

4.1 Ordre d'Intégration des Variables

Dans cette section nous présentons les principaux résultats obtenus en suivant la méthodologie ci-dessus. Nous commençons par importer les données afin de visualiser nos deux séries. les graphiques suivants illustrent les évolutions du GDP et des taux d'intérêt au fil du temps.

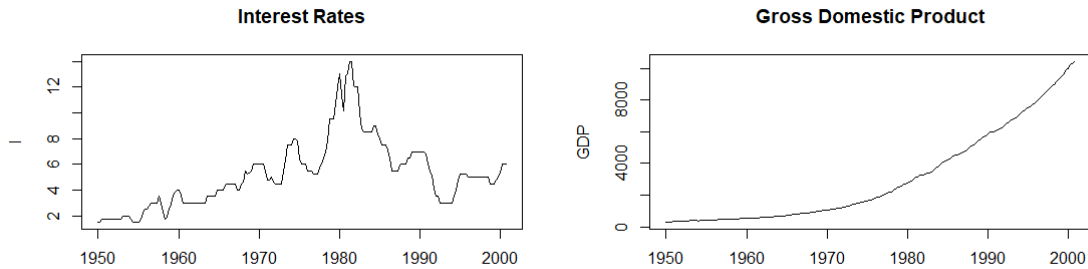


FIGURE 1 – Séries Brutes

Visuellement, nous pouvons soupçonner que les séries ne sont pas stationnaires. Le GDP en particulier semble clairement suivre une tendance positive. Les taux d'intérêt semblent également suivre une tendance et n'est pas stationnaire en variance. Ces observations ne sont pas suffisantes pour déterminer si nos séries sont stationnaires. Nous employons donc les tests ADF, KPSS et ADF-GLS, avec un seuil de significativité de 5%. Les p-values des test ADF et ADF-GLS sont supérieures à 0.05 et celle de KPSS inférieure à 0.05. ADF et ADF-GLS ne peuvent donc pas rejeter l'hypothèse H_0 de présence d'une racine unitaire, tandis que KPSS rejette l'hypothèse H_0 d'absence de racine unitaire. Les séries ne sont donc pas stationnaires et les tests confirment nos observations. Si les séries s'avèrent être $I(1)$ ¹³, il sera possible de vérifier si elles sont cointégrées, auquel cas une estimation VECM serait plus pertinente qu'une estimation VAR.

Dans un premier temps, nous précédon à une première différenciation des nos variables, pour déterminer si les variables sont $I(1)$ ou $I(p)$, $p > 1$. Si les variables sont $I(1)$, après avoir été différenciées, elles devraient être stationnaires en moyenne. Les graphiques ci-dessous représentent les deux séries différenciées.

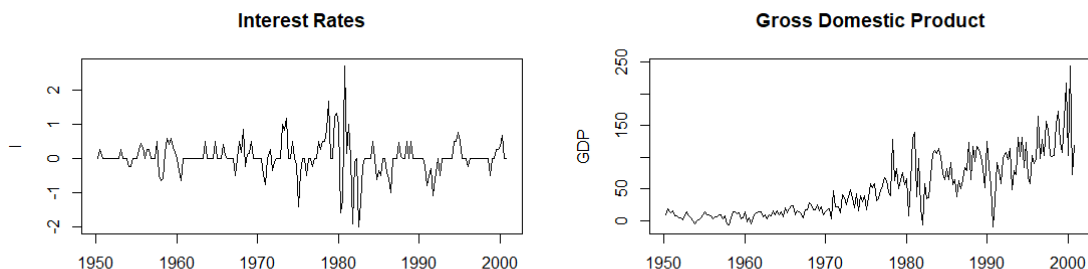


FIGURE 2 – Séries Différenciées

13. Intégrées d'ordre 1

Aucune des deux séries ne semble stationnaire en variance, en revanche le taux d'intérêt semble être stationnaire en moyenne. Le GDP exhibe toujours une légère tendance. Nous utilisons les mêmes tests que précédemment pour déterminer si ces nouvelles séries différenciées sont stationnaires. Nous noterons ces séries ΔGDP et ΔI . Pour la série ΔI , les 3 tests s'accordent à considérer la série comme stationnaire. Pour la série ΔGDP , ADF suggère la stationnarité, ADF-GLS et KPSS suggèrent la non-stationnarité. Il est possible que la non-stationnarité soit engendrée par la variance de la série qui semble augmenter avec le temps et non pas à la tendance. Afin de vérifier cette hypothèse, nous passons les séries en logarithme. La transformation logarithmique permet de réduire les distances entre les valeurs extrêmes, et permet, dans certains cas de stabiliser la variance d'un processus. Nous présentons donc les séries $\Delta \log(GDP)$ et $\Delta \log(I)$ ainsi que les résultats des tests de stationnarité correspondants.

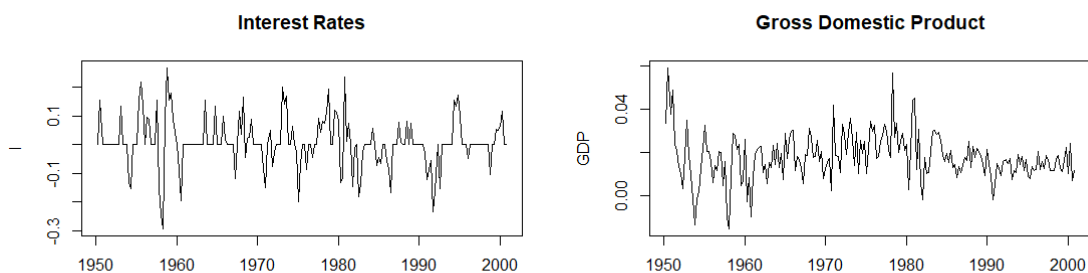


FIGURE 3 – Log-Séries Différenciées

Visuellement, les deux séries paraissent stationnaires, en tendance et en variance. Les tests confirment cette hypothèse pour les deux séries. ADF et ADF-GLS rejettent l'hypothèse H_0 de présence d'une racine unitaire dans les processus et KPSS ne rejette pas l'hypothèse H_0 de stationnarité des variables. $\log(GDP)$ et $\log(I)$ sont bel et bien $I(1)$. Si elles s'avèrent être cointégrées, nous estimerons leurs relations dynamiques à l'aide d'un VECM, si aucune relation de cointégration n'existe, nous estimerons un VAR sur les séries différenciées $\Delta \log(GDP)$ et $\Delta \log(I)$. Afin d'appliquer la procédure de Johansen pour déterminer si une relation de cointégration existe, nous devons au préalable choisir le nombre de retard optimal. Nous partons donc des séries $I(1)$ $\log(GDP)$ et $\log(I)$ puis nous estimons séquentiellement des modèles VAR(p) $p = 1, \dots, 12$. Nous calculons les critères d'information pour chaque nombre de retards considéré. Le nombre retard maximum que nous testons est 12, si les critères suggèrent que 12 est le nombre de retards optimal, nous pourrions effectuer cette procédure une nouvelle fois en testant $p = 12, \dots, 20$. Il est cependant peu probable que 12 retards soient nécessaires, empiriquement, entre 2 et 4 retards selon la fréquence des données sont suffisants.

Nous utilisons les trois critères décrits dans la section précédente (AIC, BIC et HQ). Les résultats obtenus sont unanimes, tous les critères choisissent 3 comme nombre de retards optimal. Le fait que les trois critères s'accordent sur une seule valeur nous donne confiance en la robustesse de la procédure et le reste de ces travaux seront donc menés en considérant $p = 3$ dans les modèles. Nous pouvons désormais utiliser la procédure de Johansen pour vérifier si une relation de cointégration existe entre nos 2 variables. Autrement dit, nous vérifions si une relation de long terme existe entre le GDP et le taux d'intérêt. Il est fort probable que oui, les séries de grandeurs économiques sont souvent cointégrées.

4.2 Procédure de Johansen

Il est possible d'inclure à la relation de cointégration, une constante ou une moyenne temporelle. Il est également possible de n'inclure aucun de ces deux phénomènes. La théorie économique peut parfois aider à choisir si oui ou non il est judicieux d'inclure ces subtilités dans la relation de cointégration. Dans notre cas en revanche, nous préférons estimer trois VECM, un pour chacun des cas de figure et comparer leurs critères d'information. Lors de l'estimation de la procédure de Johansen, nous choisirons alors la spécification de la relation de cointégration qui donne les meilleurs résultats. A noter que bien que nous estimions des VECM, nous ne sommes pas certains qu'une relation de cointégration existe. Ces modèles intermédiaires ont pour seul et unique but de mettre en évidence la meilleure spécification de la relation de cointégration. Les résultats sont présentés ci-dessous.

Spécification	AIC	BIC
None	-2923.595	-2874.12
Const	-2932.713	-2883.239
Trend	-2928.058	-2871.987

TABLE 1 – Tests de spécification de la relation de long terme

Aux vues des résultats, nous décidons d'inclure une constante dans la relation de cointégration entre **GDP** et **I**. Nous pouvons désormais lancer la procédure de Johansen pour déterminer si cette relation existe bel et bien. Pour plus de robustesse nous utiliserons à la fois le test de trace et le test de valeurs propres. Les résultats sont résumés ci-dessous :

TRACE				EIGEN			
	test	5%	1%		test	5%	1%
$r \leq 1$	3.43	9.24	12.97	$r \leq 1$	3.43	9.24	12.97
$r = 0$	53.64	19.96	24.60	$r = 0$	50.21	16.67	20.20

TABLE 2 – Tests de trace et de valeurs propres

Nous pouvons constater que les deux tests s'accordent sur les résultats. Pour le test de la trace nous avons $53.64 > 24.60$ et pour le test des valeurs propres ; $50.21 > 20.20$ qui signifient qu'il existe au moins une relation de cointégration et ce avec un seuil de confiance de 99%. Puisque nous n'utilisons que deux variables, nous savons d'ores et déjà qu'il n'existera pas plus d'une relation de cointégration. Autrement dit, la procédure de Johansen confirme que nos deux variables sont cointégrées. Nous choisissons donc d'estimer un modèle VECM. Comme nous l'avons détaillé plus tôt, l'inclusion d'une relation de long terme dans le modèle VECM(3)¹⁴ conduit à la nécessité de poser certaines restrictions sur les matrices de coefficients α et β . Il est bien sur recommandé de choisir minutieusement ces restrictions pour conserver un modèle le plus pertinent possible. Dans un second temps, nous passerons en revue les restrictions qui nous semblent justifiées. Pour le moment, afin d'estimer un VECM, nous choisissons des restrictions dites "à la Johansen" où le premier coefficient de la matrice β est ramené à un, les autres coefficients (dans notre cas l'autre unique coefficient) est proportionnellement modifié. Cette restriction par défaut est justifiée plus tard.

14. Nous rappelons que nous choisissons $p = 3$.

4.3 Diagnostiques des résidus

Une fois ces restrictions mises en place, le modèle VECM peut être estimé, et les fonctions de réponses pourront être obtenues. Nous estimons donc le modèle et nous reportons les résultats ci-dessous. Nous commençons par vérifier que les résidus sont non autocorrélés et distribués normalement.

Test	p-value
Ljung-Box	0.096
Jarque-Bera	< 0.01

TABLE 3 – Diagnostique des résidus

Nous rejetons l'hypothèse H_0 de normalité des résidus au seuil de 1% pour le test de Jarque-Bera. Le test de Ljung-Box ne rejette pas l'hypothèse H_0 d'absence d'autocorrélation des résidus au seuil de 5% mais cette hypothèse est rejetée au seuil de 10%. Ces résultats semblent indiquer que les résidus du modèle VECM tel qu'il est spécifié ne permettent pas de prouver que les coefficients ne sont pas biaisés ou efficaces. Dans le graphique ci-dessous, nous représentons la distribution des résidus face à celle d'une loi normale de moyenne et variance proportionnelles.

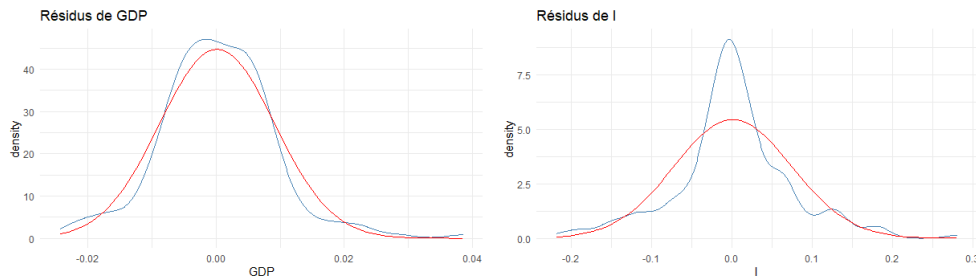


FIGURE 4 – Distribution des résidus

On constate que les résidus de GDP ¹⁵ paraissent normalement distribués tandis que ceux de I sont clairement leptokurtiques. Les tests d'hypothèses de Jarques-Bera et Ljung-Box sont conduits sur les deux séries de résidus simultanément. On peut imaginer comment une combinaison des deux distributions ci-dessus peut amener à conclure à la non-normalité des résidus.

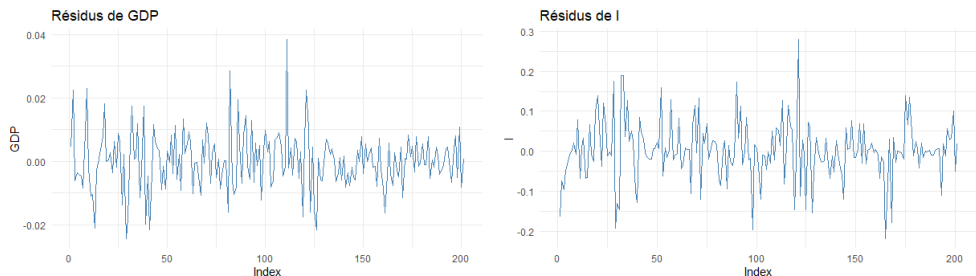


FIGURE 5 – Séries des résidus

15. La distribution des résidus apparaît en bleu et la loi normale en rouge.

L'autocorrélation est plus difficile à identifier visuellement sur les graphiques. Le test de Ljung-Box confirme cependant la présence d'autocorrélation avec un seuil de 10%. Nos résidus ne respectent pas les hypothèses requises. Nous tentons alors d'inclure une tendance temporelle dans la relation de cointégration, cette spécification étant la deuxième meilleure que nous ayons obtenue. Un nouveau modèle est estimé, et les mêmes diagnostics sont appliqués aux résidus de ce nouveau modèle.

Test	p-value
Ljung-Box	0.12
Jarque-Bera	< 0.01

TABLE 4 – Diagnostic des résidus

Les résultats ne sont pas drastiquement différents selon si la relation de cointégration est estimée avec une constante ou avec une tendance temporelle. Avec la seconde spécification, nous ne pouvons plus rejeter l'hypothèse d'absence d'auto-corrélation des résidus au seuil de risque de 10%. Nous préférons donc l'ancienne spécification et nous incluons une constante dans la relation de long terme. En revanche, les résidus ne sont toujours pas distribués normalement. Nous rappelons que les séries sont déjà log transformées, nous n'avons donc pas la possibilité d'utiliser cette transformation pour tenter d'obtenir des résidus normalement distribués. Le biais introduit par la non-normalité des résidus pose un problème d'interprétabilité des coefficients. En revanche il est possible que le modèle reste relativement performant en terme de pouvoir prédictif. Si tel est le cas, le VECM estimé peut-être utilisé dans une optique prédictive mais les fonctions de réponses, et les coefficients ne pourront pas être interprétés comme des relations macro-économiques entre nos variables. Ou tout du moins, ces relations seront biaisées. Les interprétations qui suivent ne sont donc pas à considérer comme étant parfaitement justes.

Plus tôt, nous avons évoqué la possibilité de choisir différentes spécifications des matrices Π de coefficients de cointégration et P d'orthogonalisation des chocs. Ces différentes spécifications correspondent à différentes hypothèses que nous faisons sur la relation de long terme d'une part, et sur la diffusion des chocs dans le processus d'autre part. Pour la matrice Π nous utilisons les restrictions de Johansen, qui force le coefficient de la première variable de β à 1. Cette restriction revient à faire l'hypothèse que la variable dépendante dans cette relation de long terme est le PIB, cette variable est expliquée par celle dont le coefficient n'est pas supposé a priori, le taux d'intérêt dans notre cas. Une restriction de ce type nous paraît justifiée dans la mesure où il est raisonnable de supposer que le PIB d'un pays réagit plus aux évolutions du taux d'intérêt que le taux d'intérêt ne réagit aux évolutions du PIB d'un pays. Quant à la matrice P , nous utilisons une décomposition de Cholesky avec **GDP** en première position, ce qui suppose qu'à court terme, les chocs sur le PIB ne se répercutent que sur le PIB de la période suivante et pas sur le taux directeur de la période suivante. Les chocs sur le taux d'intérêt en revanche, impactent simultanément **GDP** et **I** à la période suivante. Cette restriction, bien qu'étant celle par défaut, nous paraît faire sens pour les mêmes raisons évoquées plus haut ; il est raisonnable de supposer que ce soit le taux d'intérêt qui impacte le PIB et non pas l'inverse.

4.4 Interprétation des Résultats

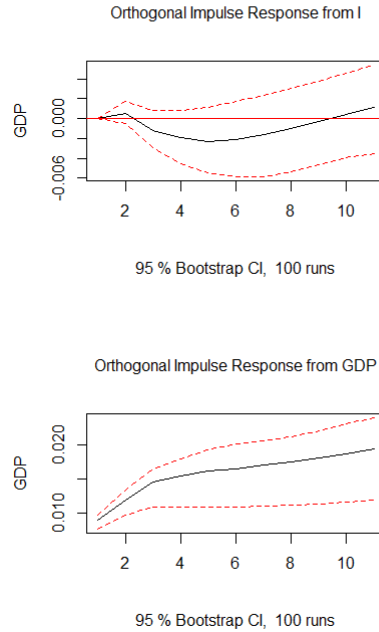


FIGURE 6 – Fonctions de réponses

La réponse du PIB après un choc de un écart-type sur le taux d'intérêt est positive. Autrement dit quand le taux d'intérêt augmente de un écart-type, le PIB augmente, puis au terme de deux trimestres, la relation devient négative ; l'augmentation du taux d'intérêt diminue le PIB, conformément à la théorie issue du modèle IS-LM-FX. Plus l'horizon considéré est lointain, plus cette relation négative s'accroît. A noter qu'en présence d'un VAR standard, ces effets positifs ou négatifs sont censés s'estomper dans le temps et retomber à 0 à partir d'un certain horizon. Ici, puisque nos deux variables sont $I(1)$, les fonctions de réponses aux chocs ne sont pas repoussées vers 0, c'est la raison pour laquelle les deux fonctions, se comportent de cette manière.

Le tableau suivant donne les écarts-types et les p-values pour ces coefficients :

ΔGDP_t				ΔI_t			
Coeff	Valeur	Sd	p-value	Coeff	Valeur	Sd	p-value
α_{GDP}	-0.001	0.0002	9.78e-07	α_I	0.006	0.0017	0.0002
$\gamma_{\Delta GDP_{t-1}}$	0.32	0.07	6.52e-06	$\gamma_{\Delta GDP_{t-1}}$	2.16	0.56	0.0001
$\gamma_{\Delta GDP_{t-2}}$	0.23	0.07	0.001	$\gamma_{\Delta GDP_{t-2}}$	0.62	0.59	0.3
$\gamma_{\Delta I_{t-1}}$	0.007	0.009	0.42	$\gamma_{\Delta I_{t-1}}$	0.29	0.07	9.99e-05
$\gamma_{\Delta I_{t-2}}$	-0.03	0.008	0.0008	$\gamma_{\Delta I_{t-2}}$	-0.08	0.07	0.24
R^2	0.81			R^2	0.26		

TABLE 5 – Estimations

Nous avons constaté que la fonction de réponse du PIB avait la forme d'une courbe en J, positive pour les 2 premiers retards, puis négative par la suite. Nous remarquons désormais que le coefficient associé au taux d'intérêt retardé d'une période n'est pas significatif dans l'équation du PIB. Seul le coefficient de corrélation avec I retardée de deux périodes est significatif. Il est possible que cet effet de courbe en J soit dû à ce coefficient non-significatif qui est inclus dans le modèle. Le taux d'intérêt n'a pas réellement d'impact sur le PIB avec un retard d'un trimestre mais commence à avoir un effet négatif (-0.029) à partir de deux trimestres de retard. Pourtant cet effet faiblement positif au premier retard (0.007) est inclus dans le modèle, entraînant une relation positive entre le taux d'intérêt et le PIB pour les quelques premiers retards. Les variations du PIB suivent bien un processus auto régressif d'ordre 2, on peut remarquer que les coefficients associés au PIB retardé sont significatifs dans l'équation du PIB. Les variations du taux d'intérêt semblent en revanche suivre un AR(1), seul le coefficient à la variable retardée une fois est significatif. D'une manière générale, nous pouvons constater que le PIB d'un pays et le taux d'intérêt évoluent bien de manière simultanés et sont cointégrés.

La qualité d'ajustement des modèles est tout à fait raisonnable, avec un R^2 de 0.81 pour l'équation du PIB et 0.26 pour l'équation du taux d'intérêt. Notre modélisation des dynamiques entre ces deux variables permet d'avantage les mécaniques du PIB que celles des taux d'intérêts. Ce résultat ne nous surprend pas, d'une part l'équation du PIB contient davantage de coefficients significatifs, ce qui signifie que nos variables permettent d'expliquer le PIB tandis qu'elles ne permettent pas d'expliquer le taux d'intérêt. Nous pouvions donc anticiper que notre modèle ait un meilleur pouvoir explicatif sur le PIB que sur le taux d'intérêt. Nous rappelons que ces coefficients sont biaisés devraient être considérés en conséquence.

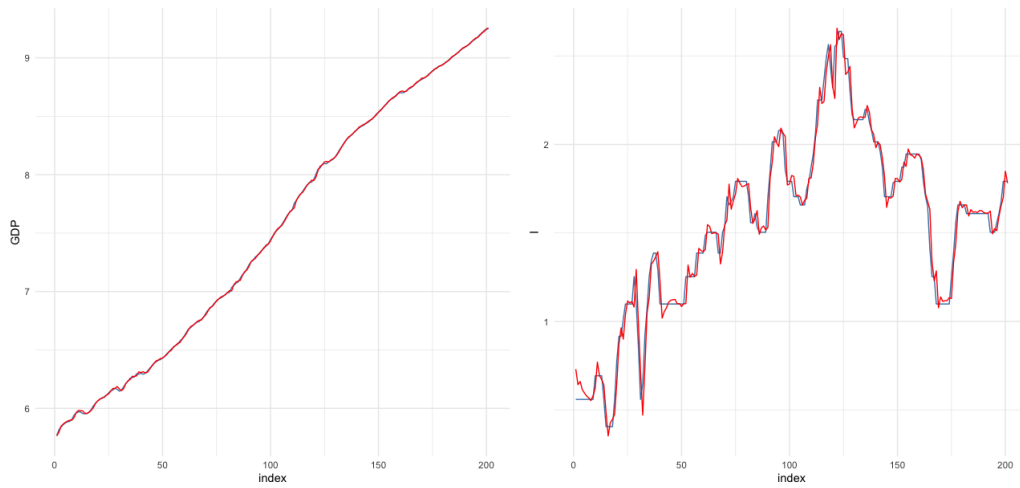


FIGURE 7 – Prédications contre valeurs réelles

On voit sur la Figure 7 - que la série prédite (en rouge) pour les deux variables sont proches de la série réelle (en bleu). Nous pouvons voir d'avantage de variabilité dans les prédictions du taux d'intérêt, qui explique probablement le R^2 inférieur à celui de la série **GDP**.

5 Conclusions

Notre objectif était de modéliser les relations dynamiques qui existent entre le PIB et le taux d'intérêt aux États-Unis. Nos hypothèses quant à ces relations sont qu'elles peuvent expliquer par les modèles IS-LM-FX. Pour modéliser ces relations, nous avons opté pour une spécification type VAR, qui permet de représenter un ensemble d'équations autorégressives, où chaque équation a pour but d'expliquer une variable grâce aux autres variables retardées. Nous avons vu que le modèle VAR(p) est un modèle qui doit être estimé sur des données stationnaires. Après avoir constaté que nos deux variables sont $I(1)$, nous décidons d'utiliser le modèle VECM qui permet de modéliser des variables cointégrées. Pour vérifier qu'une combinaison linéaire stationnaire de nos deux variables existe, nous utilisons la procédure de Johansen qui indique qu'il existe bien une relation de long terme qui doit être prise en compte. Les restrictions nécessaires sont posées de sorte à pouvoir estimer les paramètres du modèle. Ces restrictions supposent que le taux d'intérêt a un impact plus important sur le PIB que le PIB n'a d'impact sur le taux d'intérêt. Après estimation du modèle, nous montrons que les deux variables semblent être dictées par des processus autorégressifs. Une hausse d'un écart type du taux directeur entraîne une légère hausse du PIB sur 2 trimestres puis sa diminution sur les horizons suivants. Cette relation positive entre le taux directeur et le PIB n'est pas significative tandis que la relation négative est significative à 1%. Les résidus du modèle ont été détectés comme étant auto-corrélés et ne suivant pas une distribution normale, il est donc judicieux de rappeler que les coefficients sont probablement biaisés, en revanche, la qualité d'ajustement est relativement bonne et le modèle peut être utilisé pour faire de la prédiction. Le modèle confirme une réponse négative du PIB aux chocs positifs sur le taux d'intérêt ce qui confirme la théorie du modèle IS-LM-FX. Une extension de ce modèle consisterait à incorporer toutes les autres variables du modèle IS-LM-FX pour vérifier si toutes les relations supposées par la macro-économie tiennent et sont statistiquement significatives.