

Ce polycopié 5 correspond au début du point III.

III. Les controverses théoriques sur la monnaie

La monnaie a fait l'objet de multiples débats théoriques depuis la naissance de l'économie politique, on pourrait presque dire que les réflexions économiques sont nées à partir de questions monétaires.

Dans l'analyse économique de la monnaie, il y a plusieurs débats qui se recoupent partiellement seulement :

- **La création de monnaie est-elle exogène ou endogène ?** C'est le débat qui oppose la théorie du multiplicateur de crédit à celle du diviseur de crédit. Chaque économiste qui s'intéresse aux questions monétaires doit se positionner sur l'adoption d'une conception exogène ou endogène de la monnaie. La conception exogène de la monnaie signifie que l'activité économique a une existence propre qu'elle n'est pas en soi monétaire et que la création monétaire au-delà des besoins de l'économie réelle peut plus ou moins la perturber : en d'autres termes, l'échange préexiste à la monnaie (cf le troc par exemple).

- **La monnaie joue-t-elle un rôle actif ou passif dans l'économie ?** c'est-à-dire la monnaie est-elle un voile qui n'a aucun impact sur les variables réelles comme la production et la consommation ou au contraire la monnaie est-elle active, ce qui signifie qu'elle a des effets réels ?

Le croisement de ces deux critères donne lieu à une distinction entre trois conceptions de la monnaie.

Repères dans les clivages théoriques sur la monnaie

	EFFETS NOMINAUX (monnaie voile) ¹	EFFETS RÉELS (monnaie active) ²	
	DICHOTOMIE FORTE	DICHOTOMIE FAIBLE	REFUS DE LA DICHOTOMIE
MONNAIE EXOGÈNE ³	Currency School Irving FISHER (1911) Nouvelle économie classique	Effets transitoires : Milton FRIEDMAN ⁶	Effets durables et nocifs : Friedrich HAYEK Jacques RUEFF
MONNAIE ENDOGÈNE ⁴	5	Banking School Knut WICKSELL (école de Stockholm) Tradition Keynésienne Joseph Aloïs SCHUMPETER	

Source : d'après Marc Bassoni et Alain Beitone, *Monnaie : théories et politiques*, Sirey, 1994.

1. La quantité de monnaie en circulation a pour seul effet de déterminer le niveau général des prix

2. La quantité de monnaie en circulation a également un effet sur le niveau de l'activité économique

3. C'est l'approche du multiplicateur de crédit

4. C'est l'approche du diviseur de crédit

5. Cette case est vide par définition car si la monnaie est endogène, elle est forcément active

6. Alfred MARSHALL et Arthur Cecil PIGOU ont développé cette approche avant Milton FRIEDMAN

De ces deux questions on débouche sur celle de savoir si l'émission de monnaie perturbe l'activité économique ou au contraire si elle peut stimuler la production.

Enfin, les théories économiques de la monnaie posent une question moins strictement économique : la monnaie repose-t-elle sur des facteurs institutionnels et sociaux ? C'est une question que nous avons évoquée à l'occasion du débat sur ce qu'est la monnaie pour les économistes : un actif qui permet le développement des échanges car il remplit trois fonctions économiques ou un actif dont la nature l'amène à remplir des fonctions économiques mais aussi politiques et sociales qui sont essentielles dans une économie capitaliste. Nous n'y reviendrons pas ici mais il faut garder cet élément en tête.

1. L'approche dichotomique

Cette approche considère que la monnaie est exogène et n'a pas d'effet sur le niveau de l'activité économique (c'est-à-dire pas d'effets réels).

Dans cette approche, la création de monnaie trouve son origine en dehors de l'activité économique :

- dans la quantité de métaux précieux en régime d'étalon métallique
- dans la base monétaire émise par la Banque centrale quand la monnaie perd cette référence et devient purement fiduciaire et scripturale.

Selon cette approche, un changement de la quantité de monnaie en circulation dans l'économie n'a pas d'impact sur l'activité économique mais modifie le niveau général des prix.

Nous allons voir dans ce point 1. les étapes de la construction de cette approche dichotomiste de la monnaie. Il y en a quatre :

1. À l'origine de l'approche dichotomiste de la monnaie : le débat entre Malestroit et Bodin au 16^{ème} siècle
2. Le débat entre la Banking school et la Currency school à la fin du 18^{ème} siècle en Angleterre
3. La théorie quantitative de la monnaie
4. La distinction entre une approche faible et une approche forte de la dichotomie de la monnaie.

1.1. La naissance de l'approche dichotomiste de la monnaie au 16^{ème} siècle

Cette approche repose sur la mise en évidence d'une relation entre monnaie et prix. L'idée de cette relation est assez ancienne puisque le jésuite espagnol Martin de Azpilcueta l'évoque en 1556 dans son *Comentario* où il commente l'effet sur les prix de l'afflux d'argent en Espagne suite à l'exploitation des mines de Potosi au Pérou. Mais on doit à Jean Bodin une présentation convaincante de cette relation 12 ans plus tard, à l'occasion d'une controverse devenue célèbre avec Monseigneur de Malestroit.

Le débat entre Malestroit et Bodin

- **Origine du débat**

Le 16^{ème} siècle est marqué par une hausse des prix qui affecte en premier l'Espagne, le pays le plus dynamique à l'époque, et s'étend ensuite aux autres pays européens. La situation contrastant avec la relative stabilité des prix aux siècles précédents, la chambre des comptes de Paris est chargée en 1563 de mener une étude qui sera conduite par le seigneur de Malestroit et publiée en 1566 sous le titre *Enquête sur le renchérissement de toutes choses et la dépréciation des monnaies*.

- **Analyse du Monseigneur de Malestroit**

Pour Malestroit, la hausse des prix s'explique par la réduction de la quantité d'argent (c'est-à-dire de métal précieux) contenue dans la monnaie de compte donc par une dévaluation de la livre tournois (monnaie alors utilisée en France). Il montre ainsi que la hausse des prix n'est qu'une illusion : la quantité de métal précieux permettant d'acquérir un bien n'a pas changé.

- **La réponse de Jean Bodin**

Il réfute la pertinence de l'analyse de Malestroit et développe à cette occasion l'ébauche d'un raisonnement quantitatif. Il explique en effet la hausse des prix par l'afflux de métal que reçoit d'abord l'Espagne et il montre comment cette hausse se répercute en France :

- 1^{er} mécanisme : l'afflux d'argent en Espagne conduit à la mise en circulation de davantage de monnaie, la demande de produits augmente alors que la production n'a pas augmenté ce qui fait augmenter les prix en Espagne
- 2^{ème} mécanisme : le supplément de métal dont dispose l'Espagne lui permet également d'importer des produits français et ce d'autant plus que le renchérissement des produits nationaux stimule les importations

- 3^{ème} mécanisme : l'augmentation des prix a fait augmenter les salaires ce qui attire sur le territoire espagnol une importante main-d'œuvre française, les salaires y étant selon Bodin « le triple de ce qu'ils font en France ».

Ainsi, le paiement des importations et le rapatriement des salaires des émigrés procurent un apport de métal à la France à l'origine d'une augmentation des prix sur le territoire.

L'analyse de Jean Bodin ne peut pas être considérée comme une formulation achevée de ce que sera plus tard la théorie quantitative de la monnaie. Jean Bodin se borne en effet à expliquer un phénomène historiquement daté, l'accroissement des prix en France, par un supplément de monnaie métallique. Mais il ne va pas jusqu'à généraliser ce raisonnement en prétendant que toute augmentation de la quantité de monnaie amène automatiquement une élévation des prix dans les mêmes proportions. Il considère en effet que d'autres éléments peuvent expliquer une augmentation des prix comme de mauvaises récoltes ou encore la baisse du degré de concurrence entre les entreprises. Cependant, il amorce dans sa critique adressée à l'analyse de Malestroit une réflexion que développeront des économistes pendant plusieurs siècles.

Ce débat sur l'impact de la quantité de monnaie en circulation dans l'économie va être ravivé à la fin du 18^{ème} en Angleterre.

1.2. Le débat entre la *Banking school* et la *Currency school*

Ce débat constitue la deuxième étape de la construction d'une approche dichotomiste de la monnaie.

Aux origines du débat : la controverse bullioniste en Angleterre

L'origine du bullionisme : nous verrons dans le prochain chapitre que les économistes mercantilistes mesurent la richesse d'une Nation à la quantité de métaux précieux amassée (grâce au commerce extérieur notamment). Cela les a amenés à développer une analyse monétaire qualifiée, parfois à tort, de « bullioniste », le bullionisme désignant l'amour du lingot (bullion en anglais), dans laquelle la richesse et l'or sont confondus.

Une controverse bullioniste (« the bullionist controversy ») naît en Angleterre entre 1797 et 1821. Elle porte sur les conséquences de la suspension de la convertibilité-or illimitée et à prix fixe des billets émis par la Banque d'Angleterre. Dans un contexte de tensions très fortes entre la France et l'Angleterre, l'Angleterre fait face à un problème financier : comment financer l'effort de guerre ? Les billets sont convertibles en or mais cette conversion est régulièrement suspendue pour financer la guerre. Cela signifie l'établissement d'un cours forcé pour protéger les réserves d'or du pays contre la surémission de papier-monnaie du fait de l'effort de guerre. Ainsi, le 3 mai 1797 le *Bank Restriction Act* suspend la convertibilité-or des billets . Or l'Angleterre connaît à partir de 1808 une hausse des prix qui conduit la Chambre des Communes à demander un rapport sur « la cause du prix élevé des lingots d'or » : ce rapport est rendu en juin 1810 (il s'appelle « the Bullion report »), un de ses auteurs H. Thornton (1760-1815) préconise le retour de l'instrument de circulation à « son étalon véritable et légal ».

Ce rapport va susciter un débat qui oppose deux positions :

- Celle des bullionistes dont Ricardo (qui n'a pas participé au rapport mais a publié la même année, en 1810, un texte intitulé « Le cours élevé du lingot, preuve de la dépréciation des billets de banque ») qui dans la lignée du rapport de Thornton défendent le retour à l'étalon métallique.

Leur analyse est que c'est l'excès d'émission de monnaie inconvertible (billets) qui est à l'origine de la hausse des prix des marchandises, de la dépréciation du change et de la hausse du prix de l'or. Il est ainsi nécessaire de limiter absolument la circulation des billets et de revenir à l'étalon-or.

- Celle des anti-bullionistes, incarnés par T.Tooke, qui appartiendra plus tard au courant de la *banking school*.

Selon ce dernier et ses partisans, les causes de l'inflation ne tiennent pas à la suspension de la convertibilité-or de la livre mais à des causes réelles, c'est-à-dire à la guerre qui a désorganisé les échanges commerciaux et perturbé les récoltes. Tooke considère en fait que l'augmentation de la circulation de monnaie a suivi l'augmentation des prix. L'idée est que la monnaie est endogène c'est-à-dire que l'évolution de la quantité de monnaie en circulation est commandée par les besoins de l'économie réelle.

La préconisation du Bullion report sera repoussée par la Chambre des communes en 1810. Mais les bullionistes vont gagner en 1819 avec le principe d'un retour de la livre au pair imminent (c'est-à-dire à la convertibilité en or des billets et à la suppression du cours forcé) mais la reprise de la convertibilité-or pose des problèmes. En effet les stocks d'or diminuent car les banques demandent l'assistance de la *Bank of England* pour faire face aux retraits des déposants. De fait, le retour à la convertibilité-or des billets décidée en 1819 ne sera effective qu'en 1821. Ces difficultés ravivent la controverse sur la question de savoir s'il faut s'en tenir à une monnaie définie par un stock de métal ou lui donner une définition plus large.

La réponse à cette question va donner naissance à deux écoles dont les analyses s'opposent

- Dans la lignée de Ricardo, la *Currency school* (école de la circulation) défend l'idée que la monnaie est exogène et neutre.

Le billet est simplement un substitut commode à l'or, qui permet une meilleure circulation de la monnaie. Si une banque émet du papier monnaie, elle doit donc en posséder l'équivalent en métal dans ses coffres. En 1844 le *Peel's Act* (ou *Bank Charter Act*) consacre la victoire de la *Currency school* : la Banque d'Angleterre est scindée en deux départements :

- le département bancaire qui collecte les dépôts et escompte les effets de commerce
- le département d'émission qui met en circulation les billets et assure que les émissions sont couvertes à 100% par l'or.

On est donc dans une orthodoxie monétaire de la politique d'émission puisque les émissions monétaires (pièces et billets) sont dépendantes des réserves d'or, ce qui ouvre la voie au monopole d'émission qui sera accordé à la Banque d'Angleterre en 1826. On a ainsi une conception exogène de la monnaie qui est en outre considérée comme passive. La monnaie est ainsi exclue du champ de l'économie réelle. Seule exception, les *Bills of Pardon* émis par le Parlement anglais pour autoriser temporairement la *Bank of England* à émettre au-delà de ses réserves en or afin de financer l'État.

- À l'opposé la *Banking school* (ou *banking principle*) défend une définition large de la monnaie qui a deux caractéristiques :

- c'est une définition nominaliste de la monnaie : quelle que soit sa substance et sa forme (pièces ou billets de banque), la monnaie est une convention qui assure les trois fonctions habituellement attribuées aux pièces métalliques (unité de compte, intermédiaire des échanges et réserve de valeur)
- la monnaie est endogène c'est-à-dire doit être créée en fonction des besoins de l'économie réelle : il est nécessaire de pouvoir émettre de la monnaie pour ne pas limiter la croissance économique c'est-à-dire pour répondre aux besoins de financement des agents. Ainsi, quand la distribution de crédits ne fait que répondre à la demande des agents, elle n'est pas inflationniste et ce d'autant que les prix ne sont pas déterminés par la quantité de monnaie mais par les revenus des agents.

Le deux seuls points d'accord entre la *currency school* et la *banking school* portent sur :

- la nécessité d'un monopole d'émission des billets pour la banque d'Angleterre (parce que « the freedom of banking is freedom of swindling » (swindle : escroquer))
- la nécessité de limiter le droit d'émission des autres banques privées.

1.3. La théorie quantitative de la monnaie

Elle constitue la troisième étape de la construction de l'approche dichotomiste de la monnaie.

Nous l'avons vu avec le débat précédent, dans l'approche dichotomiste de la monnaie la création de monnaie trouve son origine en dehors de l'activité économique dans la quantité de métaux précieux en régime d'étalon métallique. Quand la monnaie perd cette référence et devient purement fiduciaire et scripturale, la création de monnaie se fait également en dehors de l'activité économique, dans le rôle de la Banque centrale (cf l'approche en terme de multiplicateur de crédit). Par ailleurs, l'approche dichotomiste telle qu'elle a été développée affirme que quand la quantité de monnaie change, cela n'a pas d'impact sur l'activité économique. Cette idée trouve son origine chez Jean Bodin en 1568, et elle sera formalisée par Irving Fisher en 1911 dans le cadre de la théorie quantitative de la monnaie.

La théorie quantitative de la monnaie développée par Irving Fisher

Cette théorie part d'une égalité comptable qui est toujours vraie : sur une année, la valeur de la monnaie qui a circulé est égale à la valeur des marchandises achetées. Soit :

$M.V = P.T$ où M représente la quantité de monnaie qui circule dans l'économie, V sa vitesse de circulation (le nombre de fois où, en moyenne, une pièce ou un billet est utilisé), P est le niveau général des prix et T le nombre de transactions effectuées.

À partir de là, Fisher fait les hypothèses suivantes : la vitesse de circulation est constante et la monnaie est neutre (autrement dit M n'a pas d'influence sur T). Par conséquent, M n'a d'influence que sur P . Cela signifie concrètement que si la quantité de monnaie dans l'économie double, les prix, les salaires, soit toutes les grandeurs exprimées en monnaie doublent, mais les grandeurs réelles restent identiques, car les prix relatifs restent tous les mêmes.

Remarque : on trouve une formulation plus détaillée de cette équation dans laquelle on distingue la monnaie fiduciaire (M , dont la vitesse de circulation est V) et la monnaie scripturale (M' dont la vitesse de circulation est V') : $M.V + M'.V' = P.T$

Les hypothèses étant les mêmes (V et V' sont stables et M et M' n'ont d'impact que sur P), le résultat est le même : la masse monétaire doit progresser dans la même proportion que les échanges pour ne pas être inflationniste.

La reformulation de la théorie quantitative de la monnaie par Alfred Marshall

Marshall propose une reformulation de la théorie quantitative de la monnaie sous la forme dite d'équation de Cambridge qui consiste en une approche par le revenu :

$M = k.PY$ avec

- M qui est la demande d'encaisse des agents économiques (c'est-à-dire demande de monnaie des agents économiques)
- Y qui est le revenu réel des agents (il n'y a plus ni V ni T comme chez Fisher). Rappel : le revenu réel désigne le pouvoir d'achat du revenu nominal distribué dans l'économie
- P qui est le niveau général des prix

Ce que décrit Marshall avec cette équation est que la demande de monnaie (M) est dans une relation constante (k) avec le revenu nominal (PY) : les agents économiques désirent de la monnaie pour acheter des biens.

C'est une première formulation sous forme d'équation de la demande de monnaie : Marshall développe la théorie quantitative en faisant de la demande de monnaie une demande d'encaisse. « Cette demande de monnaie s'explique par la volonté des individus de détenir des sommes permettant de saisir des opportunités de transaction. La monnaie donne [en effet] à son détenteur la possibilité d'effectuer des dépenses [futurs]. La détention de monnaie implique toutefois un manque à gagner puisque la somme ainsi conservée aurait procurée un revenu si elle avait été investie. Ce n'est donc qu'une fraction de leurs avoirs que les agents conservent sous cette forme » (G. Jacoud, *La monnaie dans l'économie*, 3^{ème} édition, Nathan, 2003).

Cette équation permettra à Arthur Cecil Pigou de formuler une théorie des encaisses réelles qui fournit une autre explication au phénomène d'inflation : l'augmentation des encaisses (M) conduit les agents à élever leurs dépenses (se croyant plus riches, ils vont dépenser davantage), ce qui va conduire à une augmentation du niveau général des prix P. Ce mécanisme va se poursuivre, jusqu'à ce que les individus retrouvent la même proportion k entre leurs encaisses réelles M/P et leur revenu réel (en effet, on déduit de l'équation de Marshall $M = kPY$ que $M/P = kY$).

Les deux versions de la théorie quantitative de la monnaie (celle de Fisher et celle de Marshall et Pigou) aboutissent au même résultat : la détermination du niveau général des prix P est fonction de la quantité de monnaie en circulation M, avec une différence de raisonnement :

- chez Fisher ce résultat repose sur la stabilité, pour des raisons techniques, de la vitesse de circulation de la monnaie et la monnaie est neutre à court et à long terme
- chez Marshall et Pigou ce résultat repose sur la stabilité, pour des raisons psychologiques, des encaisses monétaires désirées par les agents économiques et la monnaie n'est pas neutre à court terme mais elle l'est à moyen terme.

Dans la théorie quantitative de la monnaie, l'analyse de la monnaie est dichotomique : une modification de M n'a aucun impact sur Y qui mesure le niveau de l'activité (le revenu national ou la valeur de la production réalisée). Cette analyse dichotomique est héritée des économistes classiques (nous le verrons dans le chapitre 3) et va être reprise à leur suite mais on distingue alors deux approches dans cette analyse.

1.4. La distinction entre une approche faible et une approche forte de la dichotomie de la monnaie

La question qui se pose est de savoir si la monnaie est neutre à court ou à long terme :

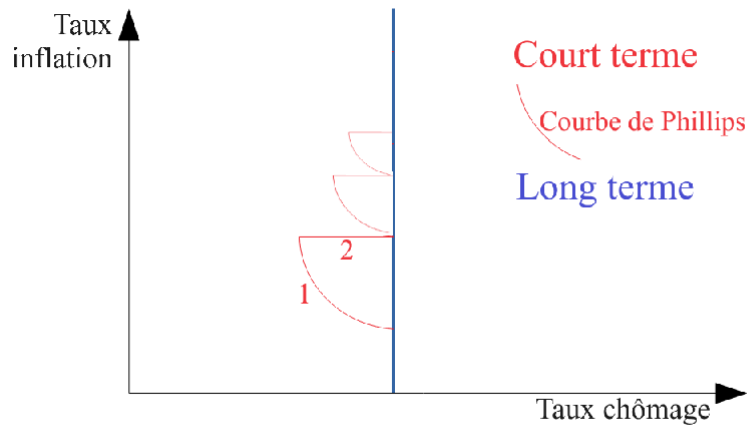
- **Dichotomie forte**

Elle est héritée de Jean-Baptiste Say (début 19^{ème} siècle) pour qui « la monnaie n'est qu'un voile ». Les agents produisent et vendent des biens et services pour acquérir d'autres biens et services. La monnaie n'est qu'un intermédiaire des échanges. Elle n'est pas désirée pour elle-même, juste pour ce qu'elle permet d'acheter. Donc, les produits s'achètent contre des produits, et il ne peut y avoir de déséquilibre entre offre et demande, grâce aux ajustements des prix : « l'offre crée sa propre demande » (c'est ce que Keynes appellera la « loi de Say »). On retrouve cette conception chez beaucoup d'économistes : Bodin, Smith, Ricardo, Fisher.

- **Dichotomie faible**

C'est une version affaiblie de l'approche dichotomique, elle est notamment portée par Milton Friedman. Selon lui, les agents font des anticipations sur l'évolution de leurs revenus nominaux et des prix mais sont victimes d'une illusion monétaire, car ils anticipent mieux les premiers que les seconds. Quand la quantité de monnaie augmente, tous les prix augmentent dans l'économie, y compris les salaires. Or les agents économiques voient d'abord l'augmentation de leurs salaires nominaux avant de voir celle des prix de ce qu'ils consomment (autrement dit, ils ne voient pas que leur pouvoir d'achat n'a pas changé). Ayant l'illusion que leur richesse a augmenté, les agents économiques vont augmenter leurs dépenses ce qui va stimuler la production de richesses et augmenter encore les prix de ce qu'ils consomment. Dans un deuxième temps, les agents économiques vont se rendre compte de la hausse des prix et corriger leurs anticipations. Ils reviendront alors à leur niveau de dépenses initial mettant ainsi fin à l'augmentation de la production.

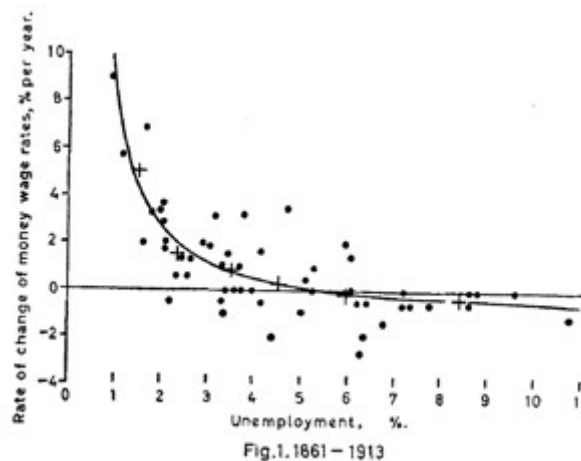
Illustration graphique du raisonnement de M. Friedman



Remarque

La thèse que Milton Friedman défend avec ce raisonnement est que la courbe de Phillips existe bien à court terme mais qu'elle disparaît à long terme (elle devient une droite verticale). Friedman explique en effet :

- Pourquoi à court terme il existe une corrélation négative entre le niveau de chômage et le niveau de l'inflation : c'est en raison du fait que les agents économiques sont victimes d'une illusion monétaire. Cette corrélation négative avait été mise en évidence par William Phillips grâce à deux séries statistiques (l'une portant sur l'évolution des salaires, l'autre sur celle du taux de chômage) au Royaume-Uni de 1861 à 1913. Ces données montraient que plus le taux de chômage était faible et plus le taux de variation des salaires était élevé et inversement :



Or comme le salaire constitue un élément central du coût de production des entreprises, on peut considérer que son évolution reflète également celle des prix de vente des entreprises et donc celle du niveau général des prix : quand le niveau du salaire augmente, les entreprises augmentent leurs prix (cela leur permet de garder un taux de profit identique). Comme toutes les entreprises fonctionnent ainsi, c'est l'ensemble des prix d'une économie qui va augmenter quand les entreprises ajusteront leurs prix suite à l'augmentation des salaires, provoquant ainsi de l'inflation, c'est-à-dire une augmentation du niveau général des prix. On doit à l'économiste canadien Richard Lipsey cette réinterprétation de la courbe de Phillips qui montre une corrélation négative entre inflation et chômage et que l'on appelle aujourd'hui la courbe de Phillips.

- L'analyse de Friedman montre aussi pourquoi à long terme il existe un niveau de chômage en dessous duquel on ne peut pas descendre sauf temporairement et au prix d'une inflation plus élevée. Ce taux de chômage est le NAIRU (*Non-Accelerating Inflation Rate of Unemployment*) qui

dépend des rigidités sur le marché du travail (chômage structurel). Le retour à moyen terme à ce niveau de chômage structurel après une tentative de le réduire grâce à une politique monétaire expansionniste vient du fait que les agents économiques font des anticipations adaptatives : ils mettent un certain temps à corriger leur erreur dans l'interprétation de l'augmentation de leurs salaires (qui ne venait pas de la croissance économique mais de l'inflation générée par la politique monétaire).

Cette représentation, selon laquelle la monnaie est active à court terme mais neutre à long terme, sera critiquée par les nouveaux classiques, Robert Lucas en tête. Selon lui, les agents font des anticipations rationnelles (et non des anticipations adaptatives comme le suppose Friedman) : ils anticipent à partir de toutes les données économiques disponibles et de leur parfaite connaissance du fonctionnement de l'économie, et donc ne se trompent pas sur l'évolution des prix. D'où un **retour à la dichotomie forte**.

2. Les analyses non dichotomiques

Ce sont des approches qui mettent en évidence des effets réels de la monnaie : une augmentation de la masse monétaire modifie la structure productive

Mais **au sein de ces approches non dichotomiques il y a deux courants opposés** :

Les analyses intégrationnistes (K. Wicksell et JM. Keynes) selon lesquelles la monnaie est endogène : l'offre de monnaie (la création monétaire) doit répondre à une demande de monnaie des agents économiques et peut anticiper sur la production de richesses supplémentaires dans le futur. Autrement dit, ces analyses considèrent qu'une augmentation de la masse monétaire peut avoir un effet positif sur l'activité économique ce qui conduit à défendre une approche endogène de la monnaie (dans une version plus atténuée pour Wicksell). La monnaie doit pouvoir être créée pour permettre aux individus d'anticiper sur leur richesse future comme dans le mécanisme du diviseur de crédit : c'est la masse monétaire qui doit déterminer la base monétaire.

Les analyses exogènes (R. Cantillon, F. Hayek) selon lesquelles la création monétaire ne peut pas stimuler l'activité économique de manière durable et est à l'origine de crises. Autrement dit, ces analyses qui considèrent, à l'inverse des précédentes, qu'une variation de la quantité de monnaie désorganise l'activité économique et que la monnaie doit donc être exogène. Plus précisément, la monnaie doit être créée *ex nihilo* par la Banque centrale comme dans le mécanisme du multiplicateur de crédit : c'est la base monétaire qui doit déterminer la masse monétaire.

2.1 L'approche intégrationniste

Selon cette approche la monnaie a des effets réels et est endogène.

À l'opposé de l'approche dichotomiste, des auteurs soutiennent que la monnaie doit être pensée en lien avec l'activité économique, et qu'elle a une influence sur celle-ci.

Nous l'avons vu, cette approche est défendue par la **Banking school** au milieu du 19^{ème} siècle, contre le *Currency principle* défendu par les économistes classiques, héritiers de Ricardo dans un contexte de développement du papier monnaie. Pour les partisans du *Currency principle*, il faut une stricte proportionnalité entre la quantité de métaux précieux et les émissions de billets (*Peel's Act*, 1844) : la monnaie est donc exogène. Les auteurs de la *Banking school*, dont Thomas Tooke, distinguent la monnaie métallique, dont la quantité est exogène, et la monnaie de crédit. La deuxième est créée en fonction des besoins des agents économiques, et est détruite lorsqu'elle est remboursée. L'important est donc moins la quantité de monnaie que la qualité de la monnaie : si elle finance des projets qui produisent des richesses, non seulement le crédit sera remboursé, mais l'activité économique progressera.

L'analyse de Knut Wicksell

L'analyse de K. Wicksell a inspiré celle de F. Hayek mais K. Wicksell considérait, contrairement à F. Hayek, que la monnaie est par nature endogène parce que la création monétaire permet aux agents économique d'anticiper sur leur richesse future.

L'interaction entre la sphère monétaire et la sphère réelle est mise en évidence par Wicksell dans *Interest and Prices*, publié en 1898. Elle repose sur une analyse des fluctuations économiques qui souligne l'existence de deux taux d'intérêt dans l'économie :

- **Dans la sphère réelle, la rentabilité du capital productif détermine un taux d'intérêt normal** (ou naturel ou encore réel)

Ce taux d'intérêt normal constitue le rendement espéré d'un capital productif nouvellement installé. Il correspond au taux d'intérêt qui égalise l'offre de fonds prêtables (l'épargne) et la demande de fonds nécessaires pour financer un achat de capital productif (l'investissement). Cette analyse s'inspire de la définition du capital productif comme détour de production : on renonce à consommer aujourd'hui pour consommer plus demain en consacrant une partie du revenu national à l'achat de biens productifs (que Wicksell appelle des biens capitaux).

Ainsi, si tous les fonds disponibles dans l'économie étaient prêtés à ce taux d'intérêt naturel, le montant de l'investissement financé serait limité par l'épargne désirée.

Mais dans une économie monétaire où il y a des banques, cela ne se produit pas parce que les banques peuvent faire de la création monétaire c'est-à-dire prêter davantage que l'épargne qui leur a été déposée. Il y existe donc dans une économie monétaire un deuxième taux d'intérêt :

- **Dans la sphère monétaire-financière, les banques prêtent à un taux d'intérêt qui est le taux du crédit**
- **Pour Wicksell, les fluctuations de l'activité économiques s'expliquent par la différence entre ces deux taux d'intérêt**

- Si le taux naturel est supérieur au taux d'intérêt bancaire, l'investissement va augmenter car il est plus rentable d'utiliser son épargne pour acheter des biens de production plutôt que de le placer auprès des banques, qui rémunèrent cette épargne au taux bancaire.

Par conséquent la demande de biens capitaux augmente ce qui fait augmenter leurs prix. La production de biens d'équipement augmente et les besoins des entreprises qui les fabriquent en facteurs de production (travail, terres et matières premières) augmentent également ce qui conduit à une augmentation de leurs prix. L'augmentation des salaires va alors conduire à une augmentation de la demande en biens de consommation faisant son tour augmenter le prix de ces biens. Il y aura par conséquent de l'inflation.

Si on considère que des taux d'intérêt bancaires bas sont la conséquence d'une offre de monnaie plus importante, on retrouve le lien établi par la théorie quantitative de la monnaie : l'augmentation de la quantité de monnaie en circulation dans l'économie se traduit par de l'inflation. Mais ce qui change fondamentalement de l'analyse de Fisher est que l'augmentation de l'offre de monnaie a été à l'origine d'une expansion puisqu'au fond elle a modifié les prix relatifs c'est-à-dire le prix des biens d'équipement en biens de consommation. L'expansion va cesser quand l'augmentation de la production de biens de consommation permise par l'augmentation de l'utilisation de biens capitaux ne sera pas écoulee par une augmentation de la demande de biens de consommation suffisante. En effet, les salaires augmentent mais les prix des biens de consommation aussi de sorte qu'à un moment donné l'augmentation du pouvoir d'achat va ralentir puis cesser.

- Wicksell montrera qu'à l'inverse, un taux d'intérêt bancaire supérieur au taux réel entraîne une dépression, de la déflation et du chômage.

La conclusion de Wicksell est qu'un système monétaire efficace doit avoir pour objectif de contrôler le taux d'intérêt bancaire de manière à le faire coïncider avec le taux naturel pour obtenir la stabilité des prix.

Synthèse de l'analyse de Wicksell

En l'absence de possibilité de création monétaire par les banques, le taux d'intérêt est déterminé de façon à égaliser la demande et l'offre d'épargne : ceci est le taux naturel. Mais dans une économie monétaire, le taux monétaire est différent du taux naturel, qui est un taux d'équilibre, parce que la demande et l'offre d'épargne ne sont pas confrontées dans leur expression naturelle mais dans leur expression monétaire, c'est-à-dire en monnaie, dont la quantité disponible peut être arbitrairement modifiée par les banques (autrement dit, par leurs prêts, la quantité de fonds prêtables n'est pas limitée par l'épargne). Ainsi les banques peuvent jouer un rôle perturbateur ; elles sont à l'origine de déplacements de l'équilibre entre offre et demande et de processus cumulatifs.

L'approche keynésienne de la monnaie

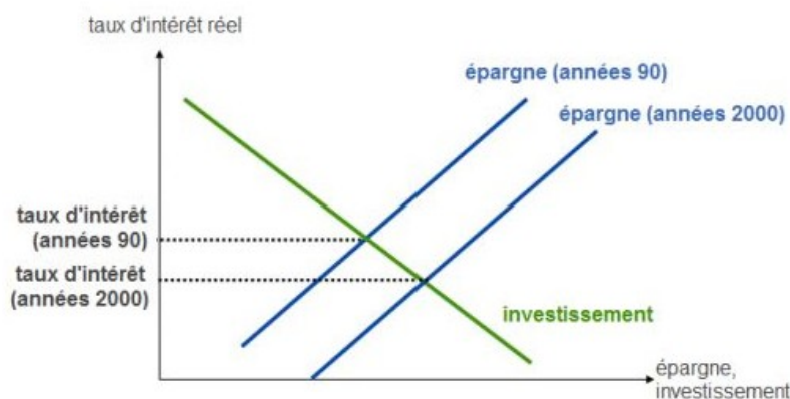
L'approche intégrationniste de la *Banking School* est approfondie par Keynes dans la *Théorie générale de l'emploi, de l'intérêt et de la monnaie* qui paraît en 1936. Selon lui, la monnaie n'est pas seulement détenue pour effectuer des transactions, mais surtout parce qu'elle est liquide. Elle est donc une réserve de valeur immédiatement disponible pour de futurs achats (motif de spéculation), et désirée pour elle-même. L'offre dans une économie (la production de marchandises et les revenus distribués à l'occasion de cette production) ne crée donc pas sa propre demande (une demande en biens de consommation et d'investissement) comme le pensaient les classiques dont Jean-Baptiste Say, car une partie du revenu distribué à l'occasion de la production de richesses peut être gardé sous forme d'encaisses monétaires, épargnée et non dépensée en biens de consommation ou d'investissement. La monnaie et la « préférence pour la liquidité » ont donc une influence sur l'activité économique.

- **Hypothèse centrale : le taux d'intérêt est le prix du marché de la monnaie** (et non celui du marché des capitaux)

Le premier élément que Keynes va remettre en cause est un élément fondamental de l'analyse des économistes classiques (un élément qui sera repris par tous les économistes néoclassiques dont Hayek), il concerne l'arbitrage entre consommation et épargne :

- Pour les économistes classiques et néoclassiques, cet arbitrage dépend du taux d'intérêt, ce qui signifie qu'épargne et investissement (soit capacités et besoins de financement) s'ajustent grâce à une variation du taux d'intérêt :

Prenons un exemple concret : au cours des années 2000, l'épargne a fortement augmenté sur un marché des capitaux devenu, comme vous le savez, mondial :



Cette augmentation s'est traduite par un déplacement de la courbe d'épargne vers la droite : à niveau de taux d'intérêt réel donné, les agents économiques épargnaient davantage.

Mais comme dans le même temps les comportements d'investissement des agents à besoin de financement ne se sont pas modifiés, il y a eu un excès de capacité de financement (S) par rapport aux besoins (I) sur le marché mondial.

Ce déséquilibre (excès d'offre par rapport à la demande) s'est traduit par une baisse du taux d'intérêt qui a conduit (déplacement sur les courbes) à une baisse de l'épargne (qui est une fonction croissante du taux d'intérêt) et à une augmentation de l'investissement (qui est une fonction décroissante du taux d'intérêt) jusqu'à un nouvel équilibre.

- Dans le cadre d'une analyse keynésienne, ce raisonnement est erroné. En effet, selon Keynes, l'arbitrage entre consommation et épargne, et donc le niveau de l'épargne, ne dépend pas du taux d'intérêt mais du niveau de revenu : plus le revenu est élevé, plus la propension à consommer est faible et donc plus le taux d'épargne est élevé (c'est selon lui une « loi psychologique fondamentale »).

En revanche, une fois le niveau de l'épargne déterminé, les individus arbitrent entre détenir cette épargne sous forme liquide ou la placer en achetant des titres sur le marché des capitaux (ou marchés financiers). Et cet arbitrage dépend du taux d'intérêt parce que le cours des titres sur les marchés financiers est négativement corrélé avec le taux d'intérêt. À ce stade de la présentation il nous faut expliquer deux éléments du raisonnement keynésien :

1. Pourquoi quand le taux d'intérêt diminue le cours des titres sur les marchés financiers augmente
2. Pourquoi quand le taux d'intérêt varie, la demande de monnaie varie

- **Le cours des titres est négativement corrélé au taux d'intérêt**

L'effet balançoire explique la corrélation négative entre l'évolution du cours des obligations et celle du taux d'intérêt

Le premier élément s'appelle l'effet balançoire et s'illustre avec l'évolution d'une obligation sur le marché secondaire. Partons d'un exemple :

- Soit une obligation vendue aujourd'hui 100 euros à un taux fixe de 5 % par an et avec une échéance de 10 ans. Elle rapporte donc 5 euros par an (son rendement est de 5%) pendant 10 ans
- Supposons maintenant qu'un an plus tard le taux d'intérêt augmente et passe à 8 %, les nouvelles obligations émises auront donc un rendement de 8 %
- À quel prix l'obligation émise à un taux de 5 % il y a un an va-t-elle pouvoir être revendue sur le marché financier secondaire ?

L'obligation rémunérée à 5 % ne sera achetée sur le marché secondaire par un agent qui si elle procure un rendement de 8 % ce qui permet de déduire son nouveau cours : il sera de 62,50 euros car 5 euros représentent 8 % de 62,50 euros ($5 \times 100 / 8$).

- À l'inverse si le taux d'intérêt avait baissé et était passé à 4 %, à quel prix l'obligation émise à un taux de 5 % il y a un an va-t-elle pouvoir être revendue sur le marché financier secondaire ?

Une rémunération de 5 % est cette fois-ci été plus attractive que la rémunération des nouvelles obligations émises d'où une forte demande amenant à augmenter son cours jusqu'au point où son rendement devienne le même que celui des nouvelles obligations soit 125 euros (5 euros représentent 4 % de 125 euros)

On le voit avec cette exemple, le cours des obligations sur le marché secondaire augmente quand le taux d'intérêt diminue et inversement. Il en va de même pour le cours des actions. Comment l'expliquer ?

L'effet de levier explique la corrélation négative entre l'évolution du cours des actions et le taux d'intérêt

Faible du taux d'intérêt à RE constante (avec $i < RE$) \Rightarrow Fort effet de levier \Rightarrow Forte RF \Rightarrow Forte demande pour les actions sur le marché financier \Rightarrow Cours des actions élevé sur le marché secondaire

Comme vous le savez, plus le taux d'intérêt est faible (à rentabilité économique constante), plus les entreprises peuvent bénéficier d'un effet de levier qui augmente leur rentabilité financière, qui est la

seule qui intéresse les actionnaires. Or plus la rentabilité financière des entreprises est forte, plus il y aura de personnes prêtes à racheter des actions sur le marché secondaire ce qui fera augmenter leur cours.

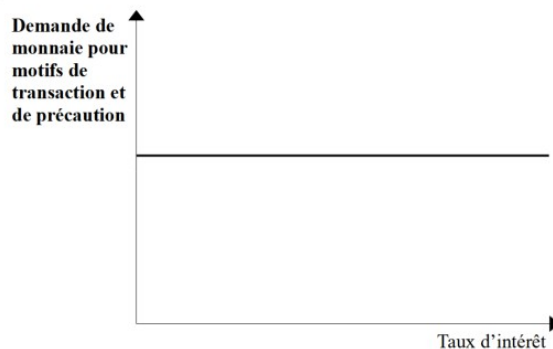
À l'inverse plus le taux d'intérêt est élevé moins la rentabilité financière sera forte et plus la demande d'actions sur le marché secondaire sera faible ce qui fera baisser leurs cours.

On peut compléter cette analyse par une explication toute simple : quand le taux d'intérêt augmente, les entreprises cotées en bourse qui doivent s'endetter devront verser un montant d'intérêts plus élevé à leurs banques ce qui réduira leurs profits et donc les dividendes qu'elles pourront verser à leurs actionnaires. Or si les dividendes diminuent, les actions de ces entreprises seront moins attractives sur le marché secondaire car moins bien rémunérées, ce qui fera baisser leurs cours.

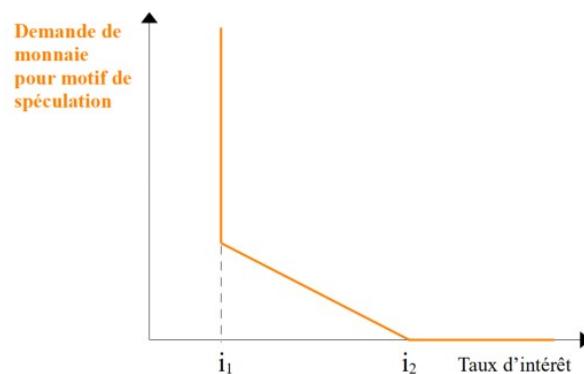
Maintenant que vous savez pourquoi le cours des titres sur les marchés financiers secondaires varie en sens inverse du taux d'intérêt, voyons l'impact que cela a sur la demande de monnaie.

- **Pour Keynes la demande de monnaie a trois motifs**

- Deux motifs qui dépendent du revenu : il s'agit du motif de transaction (un motif déjà présent chez A. Marshall et ses prédécesseurs) et du motif de précaution. Cette demande est donc indépendante du niveau du taux d'intérêt :



- Un motif qui dépend du taux d'intérêt : le motif de spéculation. Cette demande de monnaie vient du fait que les agents à capacité de financement désirent garder une partie de leur épargne sous forme liquide pour profiter de futures opportunités de placement sur les marchés financiers. Cette demande est décroissante du niveau du taux d'intérêt :



Explication du graphique

- Quand le taux d'intérêt est très faible et considéré comme minimum par les agents qui ont de l'épargne, cela signifie que le cours des titres est à son maximum ce qui fait qu'il ne pourra plus augmenter. Les agents économiques gardent alors toute leur épargne sous forme liquide.

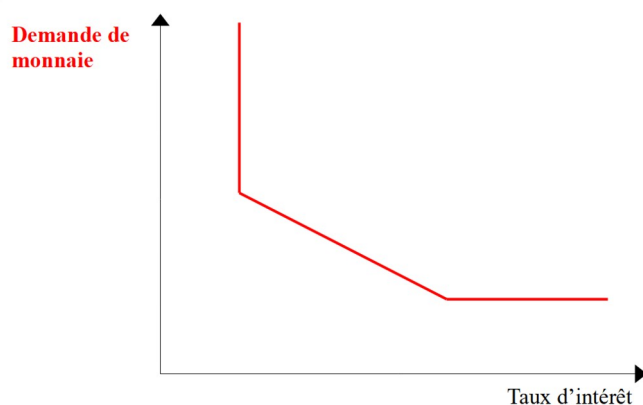
=> i_1 est le taux d'intérêt en dessous duquel les agents économiques désirent détenir toute leur épargne sous forme liquide pour profiter de futures opportunités de placement (les cours des titres

financiers sont tellement élevés que les individus pensent qu'il faut attendre qu'ils baissent et donc que le taux d'intérêt augmente)

- Puis, à mesure que le taux d'intérêt augmente, le cours des titres va diminuer ce qui incitera de plus en plus d'individus à renoncer à garder leur épargne sous forme liquide et ce jusqu'à ce que le taux d'intérêt ait atteint un tel niveau que la demande de monnaie pour motif de spéculation soit nulle.

=> i_2 est le taux d'intérêt à partir duquel la demande de monnaie pour spéculation est nulle car les agents économiques ont placé toute leur épargne en titres financiers

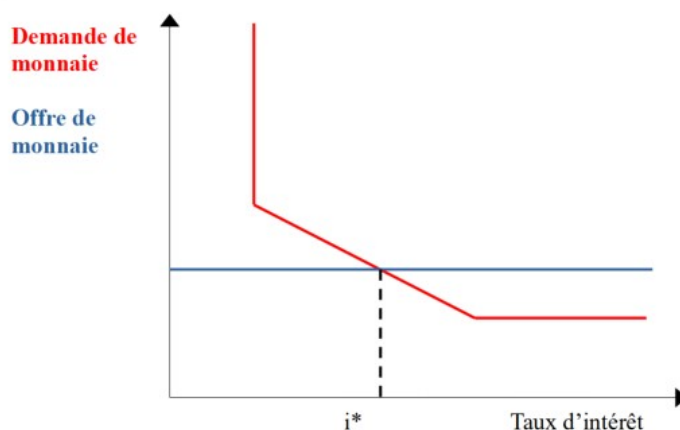
Ces trois motifs additionnés permettent de représenter la demande de monnaie qui est une fonction décroissante du taux d'intérêt



Poursuivons le raisonnement mené par Keynes, c'est-à-dire expliquons comment se forme le taux d'intérêt dans l'économie.

- **La confrontation entre offre et demande de monnaie fixe le taux d'intérêt**

C'est ensuite la rencontre entre cette demande de monnaie et l'offre de monnaie, qui est indépendante du niveau des taux d'intérêt puisqu'elle relève des décisions de la Banque centrale, qui va fixer le taux d'intérêt dans l'économie :



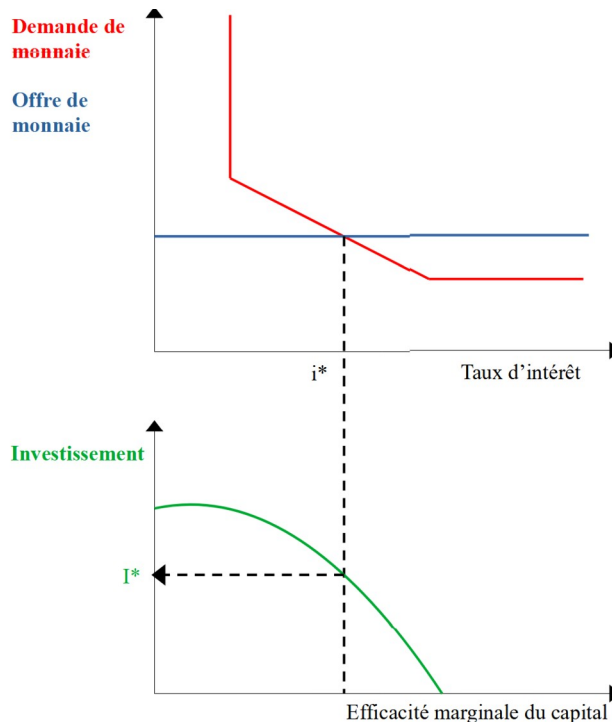
Maintenant que nous savons comment Keynes montre que le taux d'intérêt se fixe sur le marché de la monnaie, voyons comment Keynes fait le lien entre la sphère monétaire et la sphère réelle (celle de la production).

- **Le taux d'intérêt détermine le niveau de l'investissement**

C'est ensuite le taux d'intérêt qui détermine l'investissement et donc le niveau de l'activité économique. En effet dans l'analyse keynésienne la variation de l'investissement est le principal moteur de la variation de la demande globale et donc de la production.

Expliquons le lien que fait Keynes entre taux d'intérêt et investissement. Pour Keynes, les entreprises décident de leur niveau d'investissement en comparant ce que leur coûte le fait d'investir (le montant aujourd'hui de l'investissement) à ce que cela leur rapporte (c'est-à-dire des recettes supplémentaires dans le futur). Keynes explique que la valeur aujourd'hui de ces recettes futures détermine l'efficacité marginale du capital. L'efficacité marginale du capital est le taux d'intérêt qui égalise le coût de l'investissement aujourd'hui à la valeur aujourd'hui des recettes qu'il permettra de faire dans le futur : $Emc = i$ tel que $I_t = \sum_{h=t+1}^{t+T} \frac{R_h}{(1+i)^h}$.

L'Emc est ainsi le taux d'intérêt qui justifie un certain niveau d'investissement, ce qui permet de comprendre le lien entre taux d'intérêt et niveau de l'investissement, soit le lien entre la sphère monétaire et la sphère réelle :



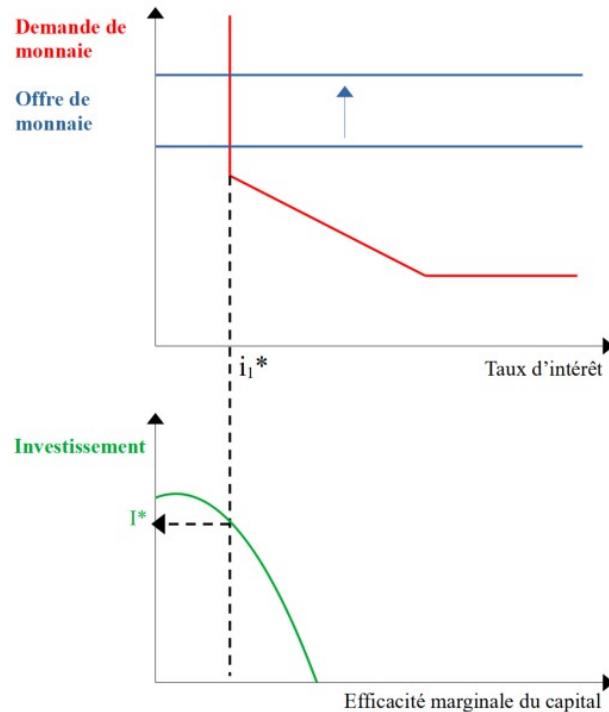
Le second graphique montre comment se détermine le niveau de l'investissement : le niveau d'investissement choisi par les agents économique est celui pour lequel le taux d'intérêt dans l'économie (qui résulte de la confrontation entre offre et demande de monnaie, cf 1^{er} graphique) est égal à l'efficacité marginale du capital.

On voit sur ce graphique que l'Emc est une fonction décroissante de l'investissement. En effet, plus il y a d'investissement au niveau macroéconomique, plus la production est forte et donc plus la concurrence entre les entreprises est importante et cela fait baisser les recettes supplémentaires attendues de l'investissement et donc l'Emc.

Cette analyse keynésienne permet de montrer l'efficacité d'une politique monétaire expansionniste dans un contexte de chômage dû à une insuffisance de la demande globale (un chômage conjoncturel) : la création de monnaie supplémentaire permet de baisser le niveau du taux d'intérêt ce qui relance l'investissement augmentant ainsi le niveau de la demande globale.

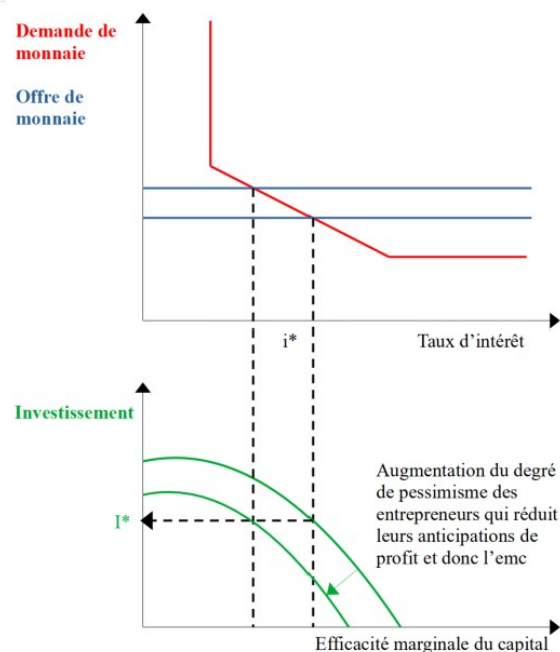
Mais Keynes croyait peu à cette solution pour lutter contre le chômage en cas de crise pour deux raisons :

1. La politique monétaire est sans effet quand elle intervient dans une situation où les individus souhaitent garder toute leur épargne sous forme liquide (trappe à liquidité) :

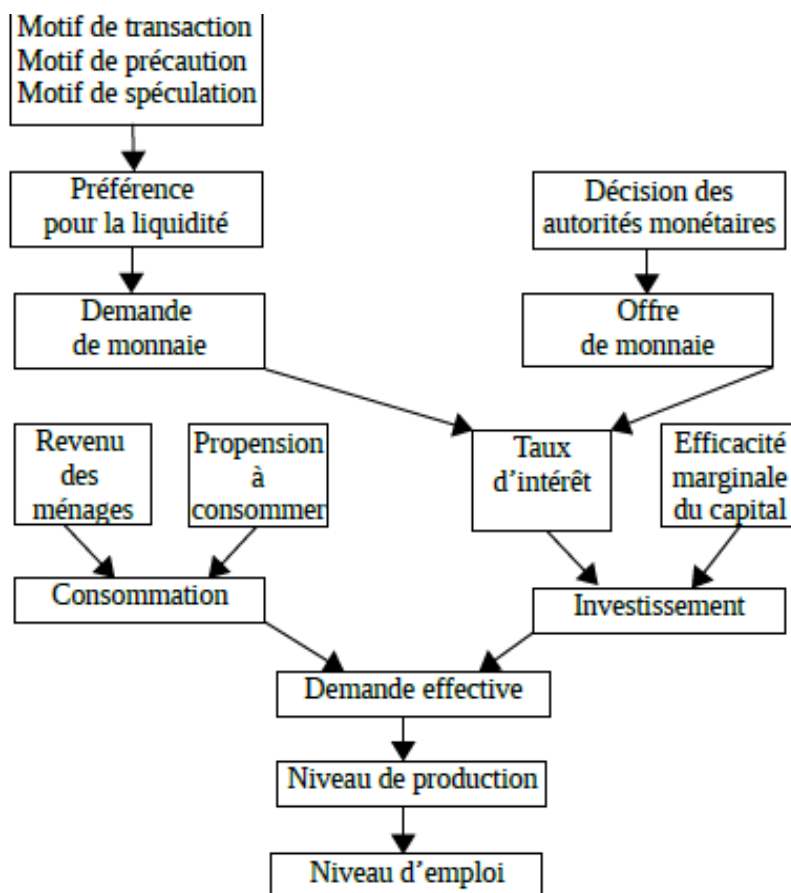


Si le niveau du taux d'intérêt qui égalise offre et demande de monnaie dans l'économie est i_1 , une politique monétaire expansionniste ne permettra pas de le diminuer davantage car l'économie est en situation de trappe à liquidité

2. Une crise se traduit par une augmentation du pessimisme des entrepreneurs qui déplace la courbe d'Emc vers le bas ce qui rend là encore une politique monétaire expansionniste inefficace pour relancer la demande globale via l'investissement :



Synthèse de l'analyse keynésienne



Gilles Jacoud, *La monnaie dans l'économie*, Nathan, 1994.

2.2. Une monnaie exogène et active

Des auteurs cherchent à concilier une conception exogène de la création monétaire et l'idée que la monnaie joue un rôle actif sur l'économie. Nous allons présenter ici l'origine de cette analyse avec les écrits de Richard Cantillon et développer l'analyse de Friedrich August von Hayek.

L'analyse originale de Richard Cantillon

On trouve l'origine de cette idée de monnaie exogène et active chez Richard Cantillon, un économiste irlandais inspiré par les physiocrates, qui faisait profession de banquier à Paris. Il est notamment l'auteur d'un *Essai sur la nature du commerce en général* paru en 1755. Son analyse repose sur l'idée que la monnaie n'est pas introduite de manière uniforme dans l'économie : elle augmente les richesses de ceux qui la reçoivent, et donc les prix de ce qu'ils achètent, déformant la production en faveur de ces productions. Mais si Cantillon pose la base d'un raisonnement qui permet de montrer que l'augmentation de la quantité de monnaie affecte de manière différenciée les prix dans une économie et modifie ainsi les prix relatifs, il ne voit pas que cela peut avoir un impact sur l'offre. En effet, même s'il évoque le fait que l'augmentation de la demande pour une marchandise, consécutive à l'augmentation de la quantité de monnaie, est de nature à stimuler sa production, il n'en déduit pas que cette augmentation de l'offre risque de contrer la tendance à la hausse des prix.

Cependant son analyse, aussi incomplète soit-elle, a donné naissance à ce qu'on appelle l'« effet Cantillon » qui est un mécanisme intéressant pour comprendre :

1. Les mécanismes à l'œuvre au cours de la crise des subprimes

La création monétaire pilotée par la FED entre 2001 et mi 2004 a fait augmenter les prix de l'immobilier mais pas l'inflation. Cela est dû à deux phénomènes : le développement du crédit immobilier et l'absence de prise en compte dans l'indice des prix à la consommation de l'évolution des prix de l'immobilier.

2. La montée des inégalités aux États-Unis après la crise des subprimes

Le professeur Guido Hulsmann établit un lien entre la création monétaire et les inégalités. Il souligne que les salaires dans la finance new-yorkaise sont très élevés. Or New York est le point d'entrée de la création monétaire de la Banque centrale américaine puisque c'est sur ce marché financier qu'elle a acheté massivement des titres dans le cadre de sa politique monétaire non conventionnelle. La création monétaire a par ce mécanisme augmenté le patrimoine des plus riches en étant à l'origine de l'augmentation du cours des actifs financiers qu'ils détenaient.

3. L'analyse des premiers effets de la politique budgétaire de relance de D. Trump en 2020, une politique budgétaire soutenue par la politique monétaire de la FED

Article à lire sur cahierdeprepa : « Les banques centrales ressuscitent l'effet Cantillon »

L'analyse de Friedrich August von Hayek

Dans son ouvrage *Prix et production* (1931) F.A. Hayek considère, à la suite des néoclassiques autrichiens, que le capital est un détour de production : au lieu de produire directement, on produit des biens qui servent à produire d'autres biens, ce qui allonge le processus de production. Or il montre que cet allongement dépend du niveau des taux d'intérêt car, pour le réaliser, il faut emprunter des fonds.

Si la quantité de monnaie augmente dans l'économie (et ce de manière exogène, c'est-à-dire sans que cela ne soit justifié par une augmentation de la production), cela entraînera une abondance de fonds prêtables et donc une baisse du taux d'intérêt qui incite à investir, c'est-à-dire à allonger les détours de production. Mais cet allongement n'est pas justifié par un surcroît d'épargne (c'est-à-dire par le fait que la consommation a diminué), et ne peut être maintenu qu'en accroissant de manière indéfinie la masse monétaire. Les prix de certains biens augmentent (ceux qui sont achetés pour produire) contrairement à d'autres, ce qui déforme aussi la structure de production.

Dès que la Banque centrale interrompt ses injections de monnaie, il y aura un ajustement très fort : la hausse du taux d'intérêt va provoquer une réduction brutale du détour de production et une crise de surproduction.

Selon les partisans d'une conception active et exogène, la monnaie déstructure l'économie (car elle modifie les prix relatifs des biens) si on la manipule, c'est-à-dire si on augmente ou baisse sa quantité au-delà de ce que l'activité économique nécessite (autrement dit lorsque la masse monétaire augmente différemment de la production). Hayek va même jusqu'à l'idée de la neutraliser, par exemple en ne changeant plus du tout sa quantité dans l'économie. Il a ainsi été un temps partisan de l'étalon or, mais s'est ensuite rallié à l'idée que les monnaies devaient être mises en concurrence et privatisées pour laisser le marché sélectionner les plus aptes à s'adapter à l'activité économique, c'est-à-dire les plus « neutres ».