Introduction à l'analyse des fluctuations macroéconomiques

Gilles Saint-Paul

November 7, 2017

1 Introduction

La macroéconomie émerge comme champ distinct de la microéconomie après la publication de la *Théorie générale* de Keynes. Le sous-emploi de masse observé pendant la grande dépression était incompatible avec l'idée selon laquelle les marchés s'équilibrent, qui fondait la microéconomie de l'époque. La théorie générale comprend un corpus d'idées qui, mises ensemble, prédisent qu'une situation de sous-emploi est possible et que le gouvernement peut y remédier grâce à la politique budgétaire. Ces idées fondatrices sont les suivantes:

- Le *rationnement de l'offre:* les producteurs ne peuvent pas réaliser leurs plans de production parce que la demande est insuffisante. Leur production est limitée par la demande, d'où une situation globale de sous-emploi des ressources.

- Le *paradoxe de l'épargne:* Une hausse de la propension à épargner réduit la demande globale de biens et services et donc la production et l'emploi.

- Le *multiplicateur:* Une hausse de la consommation publique a un effet induit sur la demande privée qui est positif. En conséquence, un euro injecté par le gouvernement augmente l'activité de plus d'un euro.

- Les *animal spirits:* Les fluctuations de l'optimisme des investisseur – justifiées ou pas – se traduisent par des fluctuations dans le PIB et l'emploi.

- La *préférence pour la liquidité:* Le taux d'intérêt nominal ne peut pas baisser pour que l'investissement rétablisse le plein emploi, parce qu'il est déterminé par l'allocation de portefeuille entre la liquidité (c'est-à-dire les encaisses monétaires) et les autres actifs.

L'histoire de la recherche en macroéconomie à la suite de la théorie générale peut se résumer en deux grandes phases. Dans une première phase, on s'attache à formaliser mathématiquement les idées de Keynes, ce qui permet de mettre au point les gros modèles macroéconomiques des administrations. Ces modèles sont ensuite utilisés pour formuler des prévisions et évaluer l'effet de différentes mesures de politique économique. A l'issue de cette première phase est mis au point le coeur de la macroéconomie: le modèle AS-AD et sa version en économie ouverte, appelée Mundell-Fleming. La seconde phase fait suite aux objections soulevées à l'encontre du modèle AS-AD par l' "école de Chicago", et notamment Friedman et Lucas. Ceux-ci font valoir l'absence de fondements microéconomiques du modèle AS-AD et montrent que cette absence rend problématique son utilisation pour évaluer la politique économique. En effet, les paramètres (ou coefficients) des équations utilisées ne sont pas stables mais dépendent en particulier des anticipations, lesquelles sont elles-mêmes affectées par la politique économique. Ainsi, une politique donnée ne peut pas être évaluée grâce à un modèle estimé sur données passées car ses coefficients seront différents du fait de la mise en place de la nouvelle politique. On ne peut comprendre comment ces coefficients changeront qu'en donnant des fondements microéconomiques aux modèles. La recherche s'est donc consacrée, depuis le milieu des années 70, à une "défense et illustration" du modèle AS-AD, en s'attachant à lui donner des fondements microéconomiques tout en préservant le coeur de sa structure. Cette évolution a débouché sur les modèles modernes, dits DSGE, qui, à partir d'une spécification complète du comportement des agents, obtiennent des relations agrégées interprétables comme des versions améliorées des équations du modèle AS-AD.

Quels fondements microéconomiques peuvent-ils sous-tendre un équilibre de sous-emploi? Toutes les réponses qui ont été proposées, de Barro et Grossman (1971) aux modèles DSGE, reposent sur l'idée de *rigidité nominale.* En effet, dans une situation de rationnement de l'offre, les producteurs auront tendance à baisser leur prix, ce qui signifie que l'équilibre de sous-emploi ne peut perdurer. Il doit donc exister un mécanisme empêchant une telle baisse des prix – cette idée était déjà présente chez Keynes, qui considérait les salaires nominaux comme rigides.

Par rigidité "nominale", on entend que le prix d'un bien, exprimé en euros, ne s'ajuste pas à la baisse lorsque l'offre est supérieure à la demande – du moins pas suffisamment pour rétablir l'équilibre. Cette notion est différente de celle de rigidité "réelle", qui exprime le fait que le prix relatif entre un bien et un autre bien ne peut pas s'ajuster pour réaliser l'équilibre de marché. Ainsi, la rigidité nominale des salaires signifie, par exemple, qu'un travailleur ne peut être payé moins que 10 euros de l'heure. En revanche, une rigidité réelle signifie qu'il ne peut être payé moins qu'une certaine fraction de l'indice des prix à la consommation. La différence entre les deux notions est importante. En effet, lorsque le sous-emploi provient d'une rigidité réelle (situation de chômage classique), une hausse de la demande ne parvient pas à rétablir le plein emploi.

Il est utile de garder à l'esprit les deux remarques suivantes, malgré leur caractère relativement subtil.

Première remarque: les rigidités nominales sont une condition nécessaire à la possibilité d'un équilibre de sous-emploi. Cela ne signifie pas pour autant qu'en absence de rigidités nominale, l'économie converge naturellement vers un équilibre de plein emploi. Certes, c'est ce que l'on suppose habituellement dans les modèles avec prix flexibles. Mais une baisse du niveau *général* des prix n'a pas d'effet allocatif, car elle n'affecte pas les prix *relatifs.* Si on s'attend à ce qu'une telle baisse se produise lorsque l'offre est en général supérieure à la demande, cette baisse ne tend pas en elle-même à rétablir l'équilibre, du moins en l'absence de monnaie. Une manière de voir la chose est de noter qu'une baisse du niveau général des prix rend les biens moins chers, mais qu'elle implique également une baisse proportionnelle des revenus; de sorte que le pouvoir d'achat n'augmente pas, et la demande de biens non plus. En présence de monnaie, les choses sont différentes: la baisse du niveau général des prix augmente la valeur réelle des encaisses monétaires (effet Pigou). Les consommateurs qui détiennent ces encaisses sont alors plus riches en termes *réels,* ce qui tend à accroître la consommation et donc la demande.

Deuxième remarque: même en présence de rigidités nominales des *prix*, il existe potentiellement un autre moyen pour rétablir le plein emploi: l'ajustement des *taux d'intérêts.* En effet, une baisse du taux d'intérêt augmente la consommation et l'investissement, et donc le niveau d'activité. Quel que soit le niveau d'activité courant, il lui est associé un taux d'intérêt réel d'équilibre. Ce taux d'intérêt réalise l'égalité entre épargne et investissement (ou de manière équivalente, l'équilibre sur le marché des fonds prêtables) *à niveau d'activité* donné. Plus ce niveau d'activité est faible, plus le taux d'intérêt réel d'équilibre correspondant est élevé. On peut imaginer que les marchés s'ajustent de manière à ce que le taux d'intérêt réel d'équilibre soit précisément celui qui rétablisse le plein emploi. Notons cependant que si l'économie est dans un équililbre de sous-emploi, l'équilibre sur le marché des fonds prêtables est réalisé; il n'y a donc aucune force qui ferait baisser le taux d'intérêt réel de façon à rétablir le plein emploi. Cependant, les hypothèses keynésiennes impliquent que même si cette force existait, elle serait inopérante. En d'autres termes, dans le modèle AS-AD, il existe un unique taux d'intérêt réel d'équilibre et donc un unique niveau d'activité. En effet, le taux d'intérêt *nominal* est déterminé par l'équilibre sur le marché de la monnaie, et, les prix étant rigides, cette détermination implique un unique taux réel d'équilibre, qui n'est pas en général celui qui réalise le plein emploi., Le taux d'intérêt nominal ne peut à la fois réaliser l'équilibre sur le marché de la monnaie et le plein emploi, et pour obtenir ce dernier, les prix devraient s'ajuster librement pour que le taux d'intérêt réel soit celui de plein emploi. Cependant – et c'est là le message du modèle IS-LM, les autorités monétaires peuvent généralement ajuster la quantité de monnaie de façon à ce que, étant donné les rigidités qui caractérisent la formation des prix, le taux d'intérêt nominal d'équilibre donne précisément le taux d'intérêt réel qui correspond au plein emploi.

2 Le modèle IS-LM ou la *Théorie générale* en deux équations

Le modèle ISLM (Hicks 1937) représente la première formalisation mathématique des idées de Keynes. Il repose sur deux conditions d'équilibre:

- L'équilibre sur le marché des biens, qui est un équilibre de sous-emploi
- L'équilibre sur le marché de la monnaie, qui est réalisé par l'ajustement du taux d'intérêt nominal.

2.1 L'équilibre sur le marché des biens: la courbe IS

La production est déterminée par la demande, qui est la somme de trois composantes: la consommation des ménages (C), l'investissement des entreprises (I) et les dépenses du gouvernement en biens et services (G). Soit Y le PIB. Cet équilibre s'écrit de la manière suivante:

$$Y = I + C + G. \tag{1}$$

Cette équation coincide avec l'identité fondamentale de la comptabilité nationale, qui stipule que la somme des dépenses est toujours égale à la somme des revenus. Mais dans le cadre d'analyse keynésien, cette identité a une interprétation comportementale: les producteurs ne peuvent vendre que la demande qui leur est adressée – en particulier, les prix ne s'ajustent pas à la baisse lorsque celle-ci est inférieure à leurs plans de production – et celle-ci dépend du comportement des consommateurs, des investisseurs, et du gouvernement. Ces comportements sont à leur tour représentés par une "fonction de consommation" et une "fonction d'investissement", tandis que G est généralement considéré comme exogène.

Une fonction de consommation "keynésienne" typique est

$$C = mY + c_0.$$

Elle suppose simplement que les agents consomment d'autant plus que leur revenu est élevé. Le paramètre m décrit l'effet d'un euro de revenu supplémentaire sur la consommation et on l'appelle traditionnellement *propension marginale à consommer.* On suppose traditionnellement que $0 < m < 1$, c'est à dire que les consommateurs épargnent une fraction constante et positive de leurs revenus supplémentaires.

Une fonction d'investissement "keynésienne" simplifiée repose sur l'idée qu'il existe un lien positif entre coût d'usage du capital et taux d'intérêt, de sorte que l'investissement est une fonction décroissante de ce dernier. Soit i le taux d'intérêt, alors

$$I = f(i),$$

avec $f' < 0$.

Avec ces deux hypothèses comportementales et le fait que G est exogène, la relation (1) détermine le niveau d'activité de manière unique en fonction du taux d'intérêt:

$$Y = \frac{f(i) + G + c_0}{1 - m}. \tag{2}$$

Cette relation entre taux d'intérêt et PIB est décroissante puisque $f' < 0$ et est appelée "courbe IS". De plus, elle rend compte d'un certain nombre d'intuitions keynésiennes. Ainsi,

- $dY/dG = 1/(1 - m) > 1$. Une hausse des dépenses publiques de 1 euro augmente le PIB de plus de 1 euro. C'est le multiplicateur keynésien. Il est égal à $1/(1 - m)$, et donc d'autant plus élevé que la propension marginale à consommer m est élevée.

- $dY/dc_0 > 0$. Lorsque c_0 baisse, les ménages tentent d'épargner plus à revenu donné, mais cela réduit la demande et donc le revenu national Y. Et même, à cause du multiplicateur, l'épargne nette des ménages est inchangée à l'équilibre. En effet, celle-ci est égale à

$$S = Y - C = (1 - m)Y - c_0 = f(i) + G.$$

L'effet de la baisse du revenu sur l'épargne compense exactement celui de la baisse de c_0. On retrouve là le paradoxe de l'épargne.

- Enfin, une hausse de l'optimisme des investisseurs se traduira par une hausse de la valeur de $f(i)$ pour toute valeur de i, ce qui augmente Y. L'équation IS rend donc compte des *animal spirits.* Cependant elle ne nous dit pas si ces anticipations optimistes sont fondées ou pas.

Il est intéressant de comparer cette détermination du PIB avec l'approche "classique". Celle-ci ne remet évidemment pas en cause l'équation (1) puisque c'est une identité mais elle considère qu'elle est satisfaite *au plein emploi* grâce à l'ajustement des prix *et* des taux d'intérêts. Soit Y^* le niveau du PIB de plein emploi. Alors, selon le modèle classique, le taux d'intérêt i s'ajuste de sorte que l'équation (2) est satisfaite *au plein emploi.* C'est à dire:

$$Y^* = \frac{f(i) + G + c_0}{1 - m}. \tag{3}$$

Selon l'approche classique, le taux d'intérêt s'ajuste pour réaliser (1) à travers l'*équilibre sur le marché des fonds prêtables.* En effet, l'équation (1) signifie non seulement que le revenu national est égal à la somme des dépenses, mais aussi, de façon concomitante, que *l'épargne est égale à l'investissement.* En effet, soit T le montant des impôts versés par les ménages à l'Etat. Leur revenu disponible est égal à $Y - T$, et leur épargne à $Y - T - C$. L'épargne de l'Etat est elle égale à $T - G$, de sorte que l'épargne nationale est égale à

$$S = Y - C - G,$$

qui doit effectivement être égal à I à l'équilibre. Considérons une situation telle que $S > I$. Alors l'offre de fonds prêtables en provenance des ménages est supérieure à la demande en provenance des entreprises, et les taux d'intérêt auront tendance à baisser, ce qui augmente I, i.e. rend plus intéressant d'emprunter. Ce processus d'ajustement continue jusqu'au point où l'offre de fonds est égale à la demande, c'est à dire que l'équation (1) est satisfaite. Cependant, comme on l'a mentionné dans l'introduction, ce raisonnement n'implique nullement que l'égalité entre S et I se réalise *au niveau de plein emploi.* Un équilibre de sous-emploi tel que $Y < Y^*$ et qui satisfait à l'équation (2) satisfait également, par construction, à (1). On a donc $S = I$ dans cet équilibre et il n'y a aucune force pour ramener i vers le niveau de plein emploi. En réalité l'approche classique considère que les prix et les taux d'intérêt s'ajustent *simultanément* de manière à ce que l'on obtient à la fois l'équilibre sur le marché des biens et celui sur le marché de la monnaie au niveau de plein emploi, c'est à dire à la fois $Y = I + C + G$ et $Y = Y^*$. Cela ne signifie donc pas que l'équation (3) est obtenue par le seul ajustement des taux d'intérêt, puisque celui-ci implique en réalité seulement (2), mais pas que l'on ait $Y = Y^*$.

2.2 L'équilibre sur le marché de la monnaie: la courbe LM

L'équation (2) ne nous dit pas quel sera le taux d'intérêt d'équilibre. Pour que le modèle soit clos, il doit comporter autant d'équations que d'inconnues, c'est à dire de variables endogènes. Sachant que C et I ne dépendent que de Y et i et que G est exogène, le système peut être réduit à deux variables endogènes: Y et i. Il nous manque donc une relation entre Y et i et celle-ci est donnée par l'équilibre sur le marché de la monnaie.

L'idée est que l'offre de monnaie est déterminée par la quantité de monnaie en circulation M, elle-même fixée par la banque centrale. La demande de monnaie est supposée être une fonction décroissante du taux d'intérêt i. L'idée fondamentale est ici que le taux d'intérêt est *le coût d'opportunité de la détention de monnaie.* En effet, la monnaie ne rapporte pas d'intérêt alors que les actifs financiers rapportent i. Détenir une partie de son portefeuille sous forme de monnaie suppose donc un manque à gagner par unité de temps égal à i multiplié par le montant de ces encaisses.

On supposera donc que la demande de monnaie est égale à $Yg(i)$, avec $g' < 0$. Cette demande est proportionnelle à Y parce que la monnaie est un moyen d'échange et que, toutes choses égales par ailleurs, la quantité de monnaie nécessaire pour effectuer un certain volume de transactions est naturellement proportionnelle à ce volume.

L'équation LM nous dit que le taux d'intérêt i réalise l'équilibre sur le marché de la monnaie:

$$M = Yg(i). \tag{4}$$

Cette équation définit une relation positive entre Y et i, appelée courbe LM. La détermination jointe de Y et de i est obtenue par l'intersection de la courbe IS avec la courbe LM, selon le célèbre "diagramme ISLM" (Figure 1).

Le modèle ISLM rend compte de l'idée fondamentale de Keynes selon laquelle la "préférence pour la liquidité" joue un rôle important dans le sous-emploi. Supposons en effet que $Y = Y^* > I + C + G$ mais qu'il y ait équilibre sur le marché de la monnaie. Dans une telle situation de déséquilibre, il y a excès d'offre sur le marché des fonds prêtables puisque $S > I$. On s'attendrait à ce que i baisse. Cependant, une baisse marginale de i incite les consommateurs à réallouer leurs actifs en faveur de la monnaie, ce qui n'est pas possible étant donné le niveau de l'offre de monnaie. Comme les taux d'intérêts ne baissent pas, le déséquilibre sur le marché des fonds prêtables ne se résorbe que parce que le revenu national Y baisse, les producteurs étant incapables d'écouler leur production au niveau de plein emploi.

2.3 L'approche moderne: IS-MP

Le modèle ISLM est encore enseigné dans la plupart des manuels modernes. Cependant, il remonte à une époque où l'on concevait le travail de la banque centrale comme consistant à fixer, ou du moins à cibler, la quantité de monnaie. De nos jours, les banques centrales ne prêtent plus guère attention à la quantité de monnaie et fixent directement les taux d'intérêt. On considère qu'elles poursuivent un objectif de stabilisation du PIB et de l'inflation, ce qui signifie qu'une politique plus contractionniste sera mise en place lorsque le PIB Y croît ou lorsque l'inflation s'accélère. A ce stade de notre discussion où le niveau des prix ne joue encore aucun rôle, on peut considérer que la politique monétaire fixe un taux d'intérêt d'autant plus grand que Y est élevé:

$$i = h(Y), h' > 0.$$

Cette relation qui décrit le comportement de la banque centrale dans la fixation de ses taux directeurs définit une relation positive entre i et Y et est donc formellement équivalente à la courbe LM. Le modèle tel que nous le présentons est donc fondamentalement inchangé, si ce n'est que la courbe LM qui s'interprétait comme une relation d'équilibre sur le marché de la monnaie est désormais remplacée par une courbe MP qui a les mêmes propriétés mais qui s'interprète comme, la règle de fixation du taux d'intérêt par la banque centrale.

2.4 Les effets de la politique budgétaire

Le modèle IS-LM/IS-MP est traditionnellement utilisé pour analyser les effets de la politique économique. Considérons tout d'abord le rôle de la politique budgétaire. Celle-ci est représentée par la variable G, le niveau des dépenses publiques. D'après (3), une hausse de G déplace la courbe IS vers le haut. La courbe LM est inchangée puisque G n'apparaît pas dans (4). Comme le montre la figure 2, cette hausse de G a pour effet de stimuler l'activité (Y augmente) et d'accroître le taux d'intérêt.

En différentiant (2) et (4), on peut calculer l'effet sur Y d'un acroissement marginal dG des dépenses publiques. La différentiation de (2) s'écrit

$$dY = \frac{dG + f'(i)di}{1 - m}, \tag{5}$$

et celle de (4):

$$dM = g(i)dY + Yg'(i)di. \tag{6}$$

Ces deux équations définissent un système linéaire dont les inconnues sont dY et di, et dans le cas qui nous préoccupe, on a $dG > 0$ et $dM = 0$, et l'on obtient

$$\begin{aligned} dY &= \frac{1}{1 - m + \frac{f'(i)g(i)}{g'(i)Y}} dG; \qquad (7) \\ di &= -\frac{g(i)}{(1-m)g'(i)Y + f'(i)g(i)} dG. \end{aligned}$$

Puisque $g' < 0$ et $f' < 0$, on a bien $dY/dG > 0$ et $di/dG > 0$, ce qui confirme l'analyse de la figure 2.

La hausse du PIB induite par la demande additionnelle en provenance de l'Etat est associée à une augmentation de la demande de monnaie. Les agents, ayant besoin de liquidités supplémentaires pour financer leurs transactions, vendent des obligations pour les convertir en monnaie. Ceci crée un excès de demande de monnaie et un excès d'offre de bons qui fait baisser le prix de ces derniers ou, ce qui est équivalent, fait augmenter les taux d'intérêt. Cette hausse des taux d'intérêt réduit l'investissement, ce qui tempère l'effet net du stimulus fiscal initial sur le PIB. C'est ce que l'on appelle l'*effet d'éviction* dit encore crowding out. D'après la formule (7), dY/dG est d'autant plus faible, et donc l'effet d'éviction d'autant plus élevé, que $|f'|$ est élevé et $|g'|$ est faible. Cela se comprend aisément: plus g' est faible en valeur absolue, moins la demande de monnaie est réactive au taux d'intérêt, et plus la hausse du taux d'intérêt nécessaire pour rétablir l'équilibre sur la marché monétaire est forte. Plus f' est élevé en valeur absolue, plus une hausse donnée du taux d'intérêt entraîne une chute importante de l'investissement.

Il est intéressant de considérer la valeur du multiplicateur dans deux cas d'école.

Le premier est celui de la *trappe à liquidité* (figure 3). Il s'agit d'une situation où la demande de monnaie est localement infiniment élastique au taux d'intérêt. Cela signifie que la courbe LM devient horizontale, car g' est très élevé – en effet, d'après (6), di/dY est très faible si g' est très élevé. Cette situation se produit, typiquement, lorsque le taux d'intérêt tombe à zéro. La monnaie étant plus liquide que les autres actifs, ces derniers doivent avoir un rendement supérieur ou égal à la première, qui est nul par définition. Le taux d'intérêt ne peut donc tomber au-dessous de zéro. Lorsqu'il atteint ce plancher, le coût d'opportunité de la monnaie est nul. Cela signifie qu'un euro supplémentaire détenu en espèces n'a plus aucune valeur en tant que moyen de paiement; si ce n'était pas le cas, il serait profitable de se le procurer en vendant des actifs, puisque le coût de cette opération est nul si $i = 0$. Le marché de la monnaie ne serait pas alors à l'équilibre. La monnaie est alors parfaitement substitut avec les autres actifs, car elle a perdu son avantage propre en termes de liquidité. C'est pourquoi l'élasticité de la demande de monnaie devient infinie: si le taux d'intérêt baisse même très faiblement, les agents liquident tous leurs actifs financiers pour ne détenir que de la monnaie.

Dans une trappe à liquidité, le terme $\frac{f'(i)g(i)}{g'(i)Y}$ dans (7) devient pratiquement nul et l'on trouve que $dY/dG = 1/(1-m)$. C'est la valeur du multiplicateur keynésien à taux d'intérêt inchangé. Il n'y a pas d'effet d'éviction des dépenses publiques sur l'investissement privé. La demande de monnaie commence à augmenter sous l'effet du supplément d'activité, mais il suffit d'une hausse négligeable des taux pour convaincre les agents de détenir des bons plutôt que de la monnaie et ramener le marché monétaire à l'équilibre.

Le second cas d'école est celui de la *théorie quantitative de la monnaie* (figure 4). Il s'agit d'une situation où la demande de monnaie est inélastique au taux d'intérêt ($g' = 0$), en d'autres termes sa vélocité $v = Y/M$ est constante. Dans une telle situation, la courbe LM est verticale: il existe une valeur unique du PIB compatible avec l'équilibre sur le marché monétaire. On a donc $dY/dG = 0$ – en effet, dans (7), le ratio $\frac{f'(i)g(i)}{g'(i)Y}$ devient infini. L'effet d'éviction est alors total: tant que l'investissement n'a pas suffisamment baissé pour ramener Y à son niveau initial, les taux d'intérêt continuent à monter sous l'effet de l'excès de demande de monnaie; l'équilibre n'est rétabli que lorsque la chute de l'investissement compense exactement la hausse des dépenses publiques.

2.5 Les effets de la politique monétaire

D'après (4), une hausse de la quantité de monnaie M déplace la courbe LM vers la droite: étant donné Y, la valeur de i qui réalise l'équilibre sur le marché monétaire est plus faible. De même, la courbe MP se déplacerait vers la droite, ce qui signifie simplement que la banque centrale pratique une politique de taux plus bas.

On voit (figure 5), que l'injection monétaire a un effet positif sur Y et négatif sur i. Pour absorber la monnaie supplémentaire, les taux doivent baisser; en effet les agents tentent de convertir ces encaisses excessives en actifs, ce qui fait monter le prix de ces derniers et donc baisser leur rendement. Cette baisse des taux rend plus rentable d'investir en capital physique, ce qui accroît I et donc Y. La consommation C augmente également sous l'effet de la hausse des revenus.

On peut résoudre le système (5)-(6) avec $dM > 0$ et $dG = 0$, et l'on obtient

$$\frac{dY}{dM} = \frac{1}{g(i) + \frac{Yg'(i)(1-m)}{f'(i)}}.$$

Cette formule nous montre que la politique monétaire est d'autant plus efficace que

-$|f'|$ est grand: l'investissement est plus élastique à i.

- $|g'|$ est faible: la demande de monnaie est plus inélastique à i.

En particulier, dans une trappe à liquidité, la politique monétaire devient inopérante ($dY/dM = 0$): les encaisses injectées dans l'économies sont détenues par les agents sans que les taux d'intérêt aient à baisser. Il n'y a pas lieu pour eux de les convertir en actif puisque ceux-ci sont parfaitement substituables avec la monnaie.

3 Le modèle AS-AD: la formation des prix

Le modèle ISLM ignore le rôle des prix. On peut l'interpréter comme une analyse de très court terme où le niveau général des prix est fixé par le passé. Mais dans la réalité, les prix changent. Il est donc naturel d'exiger du modèle qu'il puisse prévoir l'évolution du niveau général des prix. De plus, l'ajustement des prix joue un rôle essentiel dans notre compréhension du sous-emploi puisque ce dernier ne peut persister que si les prix sont "rigides".

Le modèle AS-AD est une extension du modèle IS-LM qui introduit une variable supplémentaire: le niveau général des prix. Son évolution est régie par une équation supplémentaire, dite courbe d'offre agrégée (AS), qui postule une relation positive entre les prix et le PIB. De plus, il est possible d'enrichir les courbes IS et LM en prenant en compte le rôle des prix, et l'on va montrer que si l'on élimine le taux d'intérêt entre ces deux relations, on trouve une relation négative entre Y et le niveau général des prix appelée courbe de demande agrégée (AD).

Ces dénominations proviennent d'une analogie avec l'équilibre partiel sur le marché d'un bien: la demande de tomates est une fonction décroissante du prix des tomates, l'offre de tomates est une fonction croissante de ce prix. Mais cette analogie est trompeuse, puisque le prix qui intervient ici n'est pas un prix relatif, mais le niveau général des prix qui ne devrait avoir a priori aucun rôle allocatif.

Historiquement, la courbe d'offre agrégée est la contrepartie théorique de la courbe de Phillips. L'économiste anglais A.W. Phillips avait en effet observé une corrélation négative entre taux de chômage et inflation salariale. On a également observé, par la suite, une corrélation négative entre taux de chômage et inflation du niveau général des prix. Par une sorte de glissement sémantique, cette propriété a été intégrée au modèle AS-AD comme équation structurelle qui rend compte du comportement de formation des prix, en l'absence a priori d'interprétation économique. C'est précisément la recherche d'une telle interprétation par des économistes tels que Phelps et Friedman qui a conduit à une profonde remise en question de l'appareil keynésien, notamment, comme on le verra, à travers une critique de la courbe de Phillips en tant que relation structurelle stable. A ce stade de notre analyse, contentons-nous d'interpréter la courbe d'offre agrégée comme rendant compte de l'idée générale selon laquelle les prix ont tendance à augmenter avec l'activité, parce que la hausse de celle-ci crée des "tensions" sur les marchés.

3.1 Le modèle IS-LM-AS

Reprenons notre modèle IS-LM et introduisons-y le niveau général des prix. Cette variable était absente de notre exposition du modèle IS-LM, mais elle doit maintenant être prise en compte dans les équations IS et LM. En particulier, la variable qui détermine l'investissement est le taux d'intérêt *réel,* i.e. la différence entre le taux d'intérêt nominalb et l'inflation anticipée. On doit donc réécrire l'équation IS comme suit:

$$\begin{aligned} Y &= C + I + G \\ &= mY + c_0 + f(i - \pi^e) + G, \end{aligned} \tag{8}$$

où Y, C, I, et G sont des variables *réelles,* c'est à dire exprimées en *volume*, i est le taux d'intérêt *nominal* et π^e désigne l'inflation anticipée, soit $1 + \pi^e = \frac{P'^e}{P}$, P étant le niveau courant des prix et P' leur niveau futur, et "e" signifiant "anticipé". le taux d'intérêt réel est $r = i - \pi^e$. Nous traiterons les anticipations d'inflation commes *données.* Dans un modèle complet, celles-ci devraient bien entendu être endogénéisées, ce qui soulève la question cruciale de leur formation. Pour l'instant, contentons-nous d'une exposition simplifiée et statique, et considérons donc π^e comme une variable exogène.

Quant à l'équation LM, elle doit aussi être récrite, pour faire apparaître le fait que la demande d'encaisses monétaires est une demande d'encaisses réelles. En effet, toutes choses égales par ailleurs, un doublement des prix induit un doublement de la valeur des transactions à financer et donc de celle des moyens d'échange que doivent détenir les agents. On peut donc récrire l'équilibre sur le marché de la monnaie de la manière suivante:

$$\frac{M}{P} = Yg(i). \tag{9}$$

Notons que c'est le taux d'intérêt nominal qui intervient et pas le taux réel. En effet, le coût d'opportunité de la monnaie provient du fait que contrairement aux actifs financier, son rendement *nominal* est nul par définition. La différence entre le rendement des actifs financiers i et celui de la monnaie est en quelque sorte une quantité réelle, puisqu'il s'agit d'un prix relatif. Mais comme le rendement de la monnaie est fixé à zéro en termes nominaux, cette différence coïncide avec le taux d'intérêt nominal i. Si, par un moyen quelconque, la banque centrale émettait un moyen de paiement dont le rendement est fixé à zéro en termes réels, alors ce serait le taux d'intérêt réel et non le taux d'intérêt nominal qui apparaîtrait dans la courbe LM.[1]

Enfin, la variable P est elle-même régie par la courbe d'offre agrégée AD:

$$P = h(Y), h' > 0.$$

3.2 La courbe de demande agrégée

Le modèle IS-LM-AS est un modèle de trois équations à trois inconnues: Y, P, et i. Il est souvent commode d'analyser l'équilibre sur un graphique, avec deux courbes qui s'intersectent. C'est pourquoi l'on réduit souvent ce modèle à un modèle à deux équations qui déterminent simultanément les prix et la production, le modèle AS-AD. Pour ce faire, on élimine i entre les relations IS

[1] Une thématique importante de la *Théorie générale* est qu'en période de déflation, la "préférence pour la liquidité" – l'équilibre sur le marché monétaire – empêche i de baisser; si $\pi^e < 0$, le taux d'intérêt réel d'équilibre a des chances d'être trop élevé par raport au niveau qui réalise le plein emploi. Une hausse de M peut remédier à cette situation, sauf en présence de trappe à liquidité. Dans une telle situation, Keynes préconise de faire baisser le rendement financier de la monnaie en introduisant de la monnaie estampillée, i.e. devenant caduque après une certaine date (ce qu'ont pratiqué les états successeurs de l'empire Austro-Hongrois lors de la transition entre la monnaie impériale et leur propre monnaie).

(8) et LM (9) et l'on obtient la courbe de demande agrégée (AD), une relation décroissante entre prix et quantités, que l'on peut exprimer algébriquement:

$$P = \frac{M}{Yf^{-1}((1-m)Y - c_0 - G)}.$$

Le mécanisme par lequel P réduit Y, toutes choses égales par ailleurs, montre à quel point l'interprétation de la courbe AD est différente de celle d'une courbe de demande traditionnelle. Une hausse de P réduit la valeur réelle des encaisses monétaires, ce qui incite les agents à vendre des bons afin de la reconstituer. Ces ventes d'actifs font baisser leur prix et donc augmenter leur rendement (le taux d'intérêt), ce qui réduit les incitations à investir. Cela a un effet négatif sur la demande globale et donc sur l'activité.

Lorsque le prix des tomates augmente, les agents réduisent leur consommation de tomate car il est plus intéressant de les remplacer par des poivrons. La demande de tomates diminue parce que leur prix *relatif* a augmenté. Ici, c'est le niveau général des prix qui augmente – donc une quantité purement *nominale* – et il réduit la demande *réelle,* par un mécanisme de transmission en équilibre général du marché monétaire vers le coût du capital, la demande d'investissement, et enfin la demande de biens. Et ce mécanisme n'opère que parce que la quantité nominale de monnaie est fixée, de sorte que sa quantité réelle est une fonction décroissante du prix. En l'absence de rigidité nominale, la dichotomie classique serait rétablie et la courbe de demande agrégée serait verticale[2].

La détermination jointe des prix et des quantités est représentée sur la figure 6. Une hausse des dépenses publiques ou de la masse monétaire déplace la courbe AD vers la droite et vers le haut: Y augmente et P aussi. La prise en compte des prix introduit un nouveau mécanisme d'éviction: la hausse de

[2] Dans le modèle IS-MP, la règle de politique monétaire ne dépend pas du niveau des prix, et la courbe de demande agrégée est verticale en l'absence d'autres rigidités nominales. Cependant, on considère que la règle de politique monétaire indexe le taux d'intérêt sur le taux d'inflation avec un coefficient supérieur à 1. Cela signifie que les autorités monétaires s'arrangent pour que le taux d'intérêt réel augmente lorsque l'inflation augmente. Si c'est le cas, alors la courbe AD – une relation négative entre PIB et *niveau des prix* – sera remplacée par une relation négative entre le PIB et le *taux d'inflation.*

P réduit la demande agrégée, relativement à l'analyse IS-LM où les prix sont considérés comme inchangés, à travers la contraction des encaisses réelles discutées plus haut. Notons aussi qu'une hausse des anticipations d'inflation a un effet expansionniste car elle déplace également la courbe AD vers la droite: c'est l'effet Mundell-Tobin. En effet, le taux d'intérêt nominal ne peut croître du même montant que les anticipations d'inflation. Si c'était le cas, Y serait inchangé et dans (9) la demande de monnaie serait inférieure à l'offre. Le taux nominal augmente nécessairement moins que les anticipations d'inflation[3].

3.3 Les chocs d'offre

Le modèle AS-AD permet aussi d'analyser des chocs d'offre. De manière un peu tautologique, ceux-ci sont définis comme déplacements de la courbe AS. En réalité, ils comprennent deux grandes catégories:

- Les chocs sur la capacité productive de l'économie, comme par exemple une hausse de la productivité ou du stock de capital,
- Les chocs sur la formation des prix, comme une hausse du pouvoir de monopole des syndicats ou des entreprises

Pour prendre en compte ces deux sources de chocs, on peut réécrire la courbe d'offre agrégée comme suit:

$$P = h(\frac{Y}{Y^*}), h' > 0, \tag{10}$$

où Y^* est le PIB potentiel, interprété comme le niveau du PIB en situation de plein emploi. L'indicateur de "tension" qui intervient dans cette relation est donc Y/Y^*; une hausse de Y^* à Y donné signifie donc que sous l'effet de la hausse de la productivité (par exemple), les "tensions dur les capacités de production" sont relâchées, ce qui rend l'économie moins inflationniste.

[3]Cet effet disparaît et est même inversé dans le modèle IS-MP, voir note précédente.

L'effet d'un choc d'offre sur l'économie est représenté sur la figure 7. Un choc d'offre adverse réduit l'activité Y tout en augmentant le niveau général des prix P.

Ces chocs ont été pris en compte dans les modèles dans les années 70 pour rendre compte des phénomènes de stagflation créés par les deux premiers chocs pétroliers. En réponse à de tels chocs d'offre, l'économie ne se déplace plus le long de la courbe de Phillips, puisque le chômage et l'inflation augmentent simultanément. Cela a constitué une surprise pour les économistes de l'époque qui considéraient la courbe de Phillips comme immuable, jusqu'au jour où ils ont compris que les chocs d'offre déplaçaient cette courbe (ou, de manière équivalente, la courbe d'offre agrégée).

Une hausse du prix des matières premières importées implique une baisse de Y^*; en effet, la quantité de valeur ajoutée produite par les facteurs de production domestiques (travail, capital) diminue car une plus grande quantité doit être soustraite du PIB pour payer les matériaux importés. Cela réduit mécaniquement la valeur ajoutée, mais aussi le niveau d'équilibre du PIB Y^*, car l'activité productive est moins rentale qu'avant[4]

3.4 Le modèle "classique"

Le modèle dit "classique" est un cas particulier du modèle AS-AD où la courbe d'offre agrégée est verticale à un niveau tel que $Y = Y^*$. (Figure 8). Le niveau général des prix s'ajuste pour maintenir l'économie au plein emploi. L'ajustement simultané de i et de P permet de réaliser l'équilibre sur le marché des fonds prêtables au niveau de production de plein-emploi, grâce au taux d'intérêt réel d'équilibre r^* qui garantit que la demande globale soit égale au PIB potentiel Y^*, tout en ayant le taux d'intérêt nominal i qui satisfait à l'équilibre sur le marché monétaire. Le mécanisme est le suivant: tant que $Y < Y^*$, le niveau général des prix baisse – c'est une conséquence du comportement de fixation des prix des entreprises tel que résumé par la

[4] Il en va de même d'une hausse du pouvoir de monopole des syndicats ou des entreprises: à cause du comportement malthusien du monopole, le PIB potentiel Y^* diminue.

courbe d'offre agrégée verticale. Cette baisse ne tend pas par elle-même à rétablir l'équilibre, puisque P n'apparaît pas dans la courbe IS; cependant, les taux d'intérêts baisse car les agents ajustent leurs encaisses réelles en essayant de convertir leur monnaie en bons; cette baisse des taux stimule la demande à travers l'investissement.

Les principales propriétés du modèle classique sont les suivantes:

1. La monnaie n'a pas d'effets réels

En effet, puisque $Y = Y^*$ à l'équilibre, le taux d'intérêt réel est celui qui réalise le plein emploi dans (8), soit

$$r^* = f^{-1}(Y^*(1-m) - G - c_0),$$

qui ne dépend clairement pas de M. Le taux nominal correspondant est $i^* = r^* + \pi^e$, d'où, d'après (9)

$$P = \frac{M}{Y^* g(i^*)}. \tag{11}$$

Une hausse de la quantité de monnaie se traduit par une hausse proportionnelle du niveau des prix et n'a pas d'effet sur les grandeurs réelles C, Y ou r.

2. Une hausse des dépenses publiques est sans effet sur Y; le taux réel r^* augmente, l'investissement chute, le niveau général des prix augmente.

Le mécanisme est le suivant: la hausse des dépenses publiques augmente la demande de biens et services; les entreprises s'ajustent en augmentant leurs prix. Les agents, ayant besoin de liquidités supplémentaires pour conduire leurs transactions, vendent des bons pour les convertir en monnaie, ce qui réduit leur prix et augmente leur rendement. L'investissement chute. Le mécanisme se poursuit jusqu'à ce que le niveau d'équilibre du PIB soit rétabli. Il y a donc effet d'éviction complet entre les dépenses publiques et l'investissement privé.

3. L'inflation est un phénomène purement monétaire

Considérons une économie à plusieurs périodes où l'équilibre au sein de chaque période t est décrit par le modèle ci-dessus. Supposons que la banque centrale fasse croître la masse monétaire à taux constant μ :

$$M_{t+1} = M_t(1+\mu).$$

Essayons de caractériser un état "stationnaire" où le taux d'inflation est constant et égal à π. Il est raisonnable de supposer que le long de cet état stationnaire, les anticipations d'inflation sont correctes, $\pi^e = \pi$ (en réalité la seule propriété dont nous avons besoin est que π^e soit également constant). D'après ce qui précède, $i^* = r^* + \pi$ est également constant au cours du temps. D'après (11), $P_t = kM_t$ où $k = \frac{1}{Y^* g(i^*)}$ est constant au cours du temps. On a donc nécessairement

$$\pi = \frac{P_{t+1} - P_t}{P_t} = \frac{M_{t+1} - M_t}{M_t} = \mu.$$

Le taux d'inflation est égal au taux de croissance de la masse monétaire: la cause ultime de l'inflation est la croissance de la quantité de monnaie.

Remarque: supposons que le PIB potentiel Y^* croisse à taux constant g,

$$Y_{t+1} = Y_t(1+g),$$

par exemple sous l'effet du progrès technique. Alors $P_t = k' \frac{M_t}{Y_t}$, où $k' = 1/g(i^*)$ est constant au cours du temps. On a

$$1 + \pi = \frac{P_{t+1}}{P_t} = \frac{M_{t+1}}{M_t} \frac{Y_t}{Y_{t+1}} = \frac{1+\mu}{1+g} \approx 1 + \mu - g,$$

d'où

$$\pi \approx \mu - g.$$

Le taux d'inflation est égal à la différence entre le taux de croissance de la masse monétaire et le taux de croissance du PIB réel, g.

3.5 La synthèse néo-classique

La synthèse néo-classique constitue le schéma intellectuel commun aux macroéconomistes "orthodoxes". Toute la recherche depuis la fin des années 1980 a consisté à préserver ce schéma, en l'enrichissant et en lui donnant des fondements rationnels, ce qui a permis de passer du cadre simplifié exposé ci-dessus aux modèles DSGE actuels.

La synthèse néo-classique tient en deux propositions:

- A court terme, l'économie se comporte de façon "keynésienne" parce que les prix sont rigides. Cette rigidité est prise en compte par une courbe d'offre agrégée de court terme dont la pente est positive et finie. Plus cette pente est faible, plus les prix sont rigides – plus une hausse de la demande agrégée est expansionniste, moins elle est inflationniste.

- A long terme, l'économie se comporte de façon "classique", car les prix sont flexibles. La courbe d'offre agrégée de long terme est verticale. Le niveau de PIB ne peut être différent de son niveau "naturel" Y^*.

Remarque: Le niveau naturel Y^* ne coïcinde pas forcément avec le plein emploi. C'est le niveau de production obtenu une fois l'ajustement des prix réalisés. Cet ajustement n'est pas nécessairement concurrentiel, et le niveau d'équilibre des prix relatifs n'est pas nécessairement le même que celui d'un équilibre walrasien. Y^* peut être inférieur au niveau de plein emploi à cause de rigidités réelles: pouvoir de monopole, pouvoir syndical, frictions sur le marché du travail, barrières à l'entrée, rigidités microéconomiques empêchant l'ajustement des salaires, etc.

L'effet d'une hausse des dépenses publiques est décrit sur la figure 9. l'économie se trouve initialement au point A. Ce point est un équilibre de long terme: il est situé sur la courbe d'offre agrégée de long terme, le PIB est égal à son niveau naturel, $Y = Y^*$. A court terme, le PIB augmente; les prix augmentent aussi mais l'effet d'éviction qu'ils engendrent n'est pas assez fort pour empêcher une expansion. L'économie se déplace le long de la courbe d'offre agrégée de court terme du point A au point B.

A long terme, les prix continuent à augmenter du fait que le PIB est supérieur à son taux naturel Y^*. Cela signifie que la courbe d'offre agrégée se déplace vers le haut: l'arbitrage entre inflation et activité devient moins favorable. Ces déplacements induisent un mouvement de l'économie du point B vers le point C, qui est son nouvel équilibre de long terme. Pour que le point C constitue un tel équilibre, la courbe de demande agrégée, la courbe d'offre agrégée de long terme et celle de court terme doivent nécessairement se couper au point C.

La politique budgétaire – et plus généralement les politiques de demande – n'ont donc qu'un effet de court terme sur l'activité. A long terme le PIB est ramené vers son niveau naturel sous l'effet de l'ajustement des prix.

Le modèle est jusqu'ici muet sur les mécanismes économiques sous-jacents aux déplacements de la courbe de demande agrégée. Une formule telle que (10) ne dit rien sur les raisons pour lesquelles la fonction $h()$ pourrait changer. Cette opacité résulte de l'absence de fondements microéconomiques pour la courbe d'offre agrégée. Une fois de tels fondements introduits, on verra que la fonction h doit être étendue pour prendre en compte les anticipations des agents, de telle sorte que Y ne pas être différent de Y^* si ces anticipations sont correctes. C'est donc l'ajustement de ces anticipations qui déplace la courbe AS pour ramener l'économie à son taux naturel.

Historiquement, ce n'est que sous l'effet des déplacements de la courbe de Phillips pendant les années 60 que l'on a pris consciences qu'une relation telle que (10) n'était pas stable. On considérait auparavant l'essence de la politique macroéconomique comme consistant pour les décideurs à choisir leur point préféré sur une relation décroissante stable entre inflation et chômage ("inflation-output tradeoff"). Ainsi, un gouvernement "de gauche" pour lequel le chômage est plus coûteux que l'inflation, choisirait un point tel que G sur la figure 10, tandis qu'un gouvernement "de droite" privilégiant la stabilité des prix opterait pour le point D.

Cette conception a été battue en brèche par la synthèse néoclassique puisqu'elle prédit qu'un tel choix n'est possible qu'à court terme. Ainsi, un gouvernment "de gauche" qui tenterait de faire baisser le chômage au-dessous de son taux naturel augmenterait les dépenses publiques pour amener l'économie à un point tel que B sur la figure 9, et à long terme ne parviendrait ainsi qu'à créer de l'inflation.

4 Critique du modèle IS-LM/AS-AD

L'instabilité de la courbe de Phillips dans les années 1960 a débouché sur une critique du modèle AS-AD qui s'est traduite par la révolution des anticipations rationnelles, laquelle a conduit à refonder la macroéconomie sur des bases microéconomiques explicites.

Le modèle AS-AD manque de fondements microéconomiques. Notamment:

- Il ignore les aspects intertemporels du comportement des agents (pourquoi épargne–t-on? Pourquoi les entreprises investissent-elles? Pourquoi détient-on de la monnaie?)
- Il ignore le rôle des anticipations (comment mes croyances sur le futur influent-elles sur mon comportement présent?)
- Il ignore la contrainte budgétaire du gouvernement (comment les dépenses publiques sont-elles financées? Quelles sont les conséquences du fait que la dette publique doive être remboursée?)

La synthèse néoclassique est silencieuse sur l'articulation entre le court terme et le long-terme: à quelle vitesse la courbe d'offre de court-terme se déplace-t-elle? Sous l'effet de quelles forces?

A la suite de ces critiques, sont nées des controverses entre "nouveaux classiques" et "nouveaux keynésiens".

Pour les "nouveaux classiques", le modèle AS-AD n'est pas viable. On doit l'oublier et refonder la macroéconomie à partir de la théorie microéconomique de l'équilibre général.

Pour les nouveaux keynésiens, le modèle AS-AD doit être réparé. On doit lui trouver des fondements microéconomiques compatibles avec ses caractéristiques fondamentales: prépondérance de la demande dans la détermination de court terme de l'activité, inefficacité des fluctuations du point de vue du bien-être social, rôle du gouvernement dans la stabilisation de ces fluctuations, arbitrage entre inflation et chômage. Compte tenu de ses aspects, les fondations doivent nécessairement impliquer des *imperfections de marché.*

4.1 Les fondements

La littérature sur les fondements peut se répartir en deux grandes classes de modèles:

- Ceux qui s'intéressent à la demande agrégée, et qu'on peut répartir en deux sous-catégories. D'une part, les travaux qui s'attachent à déduire les différentes composantes de la demande agrégée: fonction de consommation, fonction d'investissement, stocks, demande de monnaie, etc, de comportements d'optimisation de la part des entreprises et des ménages. D'autre part, ceux qui se préoccupent de l'allocation des ressources en équilibre général lorsque les prix ne s'ajustent pas. Dans une telle situation, il y aura du rationnement et l'on veut savoir dans quelle mesure ce rationnement valide les intuitions keynésiennes.

- Ceux qui s'intéressent à la formation des prix, donc aux fondements de l'offre agrégée. On distingue entre les modèles d'équilibre où une courbe de Phillips existe malgré le caractère purement nominal de l'inflation – comme le fameux modèle de misperceptions de Lucas, et des modèles incorporant explicitement les rigidités nominales de prix, mais considérant ceux-ci comme endogènes et fixés optimalement par les agents, compte tenu de ces rigidités.

5 Les fondements de la demande agrégée

5.1 La théorie du revenu permanent

Considérons un consommateur vivant deux périodes, dont le revenu à la date t est y_t, pouvant librement prêter et emprunter à taux réel r entre ces deux dates. Soit c_t la consommation à la date t.La fonction d'uilité du consommateur est $u(c_1, c_2)$ et sa contrainte budgétaire peut s'écrire comme suit:

$$c_2 = y_2 + (1 + r)(y_1 - c_1). \tag{12}$$

Le plan optimal de consommation satisfait au programme suivant

$$\max_{c_1, c_2} u(c_1, (1 + r)[R - c_1]),$$

où

$$R = y_1 + \frac{y_2}{1 + r}$$

est la valeur actualisée des revenus du consommateur, dite encore *revenu permanent.*

Le plan optimal de consommation ne dépend donc du flux de revenu (y_1, y_2) qu'à travers le revenu permanent R. En l'absence d'information sur le revenu futur, celui-ci sera naturellement remplacé par sa valeur anticipée y_2^a. On aura donc

$$c_1 = f(R^a, r),$$

avec $f_1 > 0$ et

$$R^a = y_1 + \frac{y_2^a}{1 + r}.$$

Supposons maintenant les préférences *homothétiques.* Cela signifie que le ratio $k = c_2/c_1$ ne dépend que des prix (ici le taux réel r) et pas du revenu total R. Puisque $c_2 = (1 + r)(R - c_1)$ une hausse de R à r donné doit nécessairement se traduire par une augmentation proportionnelle de c_1 et de

c_2. En effet cette relation est équivalente à

$$c_1 = \frac{1+r}{1+r+k}R,$$

avec k indépendant de R. La forme de la fonction f est donc $f(R^a, r) = g(r)R^a$.

Cette relation peut se généraliser au cas d'un consommateur vivant un grand nombre de périodes. Il est commode de considérer ce nombre comme infini. Le revenu permanent à la date t est alors égal à

$$R_t^a = \sum_{i=0}^{+\infty} \frac{y_{t+i}^a}{(1+r)^i},$$

et la consommation à la date t est égale à

$$c_t = g(r)R_t^a.$$

Cette fonction de consommation a des propriété sensiblement différentes de la fonction de consommation keynésienne traditionnelle $C = mY + c_0$. En particulier:

1. La propension marginale a consommer n'est pas un paramètre structurel. Elle dépend non seulement du taux d'intérêt mais aussi et surtout du caractère permanent ou temporaire des changements de revenus envisagés. Considérons une hausse du revenu égale à Δy. Si cette hausse est permanente, et connue comme telle, elle induit une hausse du revenu permanent égale à

$$\begin{aligned} \Delta R_t^a &= \sum_{i=0}^{+\infty} \frac{\Delta y}{(1+r)^i} = \Delta y \sum_{i=0}^{+\infty} \frac{1}{(1+r)^i} \\ &= \Delta y \left[\frac{1}{1 - \frac{1}{1+r}} \right] = \frac{1+r}{r} \Delta y. \end{aligned}$$

La propension marginale à consommer une hausse permanente du revenu est donc égale à

$$m_P = g(r)\frac{1+r}{r}.$$

Considérons une hausse transitoire du revenu. Supposons que celui-ci n'augmente de Δy que pendant la période t. Alors on a désormais

$$\Delta R_t^a = \Delta y,$$

d'où l'on déduit que la propension marginale à consommer une hausse transitoire du revenu est

$$m_T = g(r).$$

Avec un taux d'intérêt de 5%, m_T est environ 20 fois inférieure à m_P.

2. Une fonction de consommation estimée sur données économétriques conduit à une mauvaise estimation du multiplicateur. Supposons que le revenu du consommateur soit soumis à des chocs aléatoires persistants, de sorte qu'il est représentable par un processus dit "AR1" (autorégressif d'ordre 1):

$$y_t = \rho y_{t-1} + \varepsilon_t. \tag{13}$$

Dans cette expression, ε_t est un choc indépendant et identiquement distribué. Pour éviter des dynamiques explosives, on supposera $\rho \in [0, 1)$. La réponse dynamique du revenu y à un choc ε décroît exponentiellement avec le temps, et est d'autant plus persistante que ρ est élevé. En effet,

$$\frac{dy_{t+i}}{d\varepsilon_t} = \rho \frac{dy_{t+i-1}}{d\varepsilon_t} = ... = \rho^i \frac{dy_t}{d\varepsilon_t} = \rho^i. \tag{14}$$

La figure 11 compare la réponse de y à un choc ε entre deux valeurs de ρ.

Comment un choc sur le revenu à la date t affecte-t-il le revenu permanent? Supposons que les consommateurs anticipent correctement la réponse de l'économie au choc, soit $\frac{dy_{t+i}^a}{d\varepsilon_t} = \frac{dy_{t+i}}{d\varepsilon_t}$. On doit avoir

$$\begin{aligned} \frac{dR_t^a}{d\varepsilon_t} &= \sum_{i=0}^{+\infty} \frac{1}{(1+r)^i} \frac{dy_{t+i}^a}{d\varepsilon_t} \\ &= \sum_{i=0}^{+\infty} \left[\frac{\rho}{1+r}\right]^i = \frac{1}{1 - \frac{\rho}{1+r}} = \frac{r}{1+r-\rho}. \end{aligned}$$

D'où

$$\frac{dc_t}{d\varepsilon_t} = \frac{rg(r)}{1+r-\rho} = m_D.$$

Dans un échantillon de données où le revenu est régi par le processus (13), un choc de revenu de un euro se traduit par une hausse de la consommation courante de m_D euros. Notons que m_D est d'autant plus élevé que les chocs de revenu sont persistants. Un économètre qui estimerait une fonction de consommation keynésienne sur un tel échantillon en conclurait donc que la propension marginale à consommer le revenu est m_D. Il traiterait cette quantité comme un paramètre structurel qui caractérise le comportement des consommateurs. Supposons maintenant que sur la base d'une telle estimation, le gouvernement mette en place des mesures de stimulations fiscale *temporaires* qui augmentent le revenu uniquement à la date t. Ignorons pour simplifier les effets d'éviction. Le gouvernement s'attend à ce que sa politique ait un effet multiplicateur sur le PIB égal à $\frac{1}{1-m_D}$. En réalité la propension marginale à consommer sera plus faible car égale à m_T, le multiplicateur égal à $\frac{1}{1-m_T}$ sera également plus faible. La propension à consommer observée dans les données reflète la persistance moyenne des chocs dans l'échantillon, persistance égale à ρ; cette grandeur n'est pas pertinente pour analyser l'impact d'un choc sur le revenu artificiellement créé par les autorités budgétaires et dont la persistance sera a priori différente de ρ.

3. La politique budgétaire peut avoir des effets contra-cycliques sur la consommation. L'idée est la suivante: une hausse des déficits publics a un effet stimulant sur l'activité à court terme, mais elle implique une hausse de la dette publique et donc des impôts futurs. Ces hausses d'impôts anticipées peuvent en principe réduire les revenus futurs au point que le revenu permanent baisse. Si c'est la cas, la consommation baissera. Pour une économie dans une situation fiscale critique, une hausse des déficits peut avoir un effet important sur la probabilité de défaut souverain. On sait que de tels épisodes peuvent avoir un effet très néfaste sur l'économie. Dans une telle situation, il n'est pas à exclure que la consommation courante chute plus que la quantité d'euros injectés dans l'économie par la politique budgétaire (i.e. tout se passe comme si l'on avait $dC/dg < -1$) et les déficits peuvent être contractionnistes. Ces effets passent entièrement par la réaction des anticipations

et ne peuvent pas être pris en compte par une fonction de consommation keynésienne traditionnelle.

5.2 La théorie du q de Tobin

La théorie du q de Tobin est la contrepartie pour l'investissement de la théorie du revenu permanent pour la consommation. Elle prédit que l'investissement courant est affecté de manière cruciale par les perspectives de profitabilité future. De plus, ces perspectives sont refletées par la *valeur boursière de l'entreprise.* L'idée fondamentale est que l'acte d'investir consiste à transformer du capital non installé (par exemple des câbles électriques achetés chez un grossiste) en capital installé (ces mêmes câbles encastrés dans les parois et connectés au circuit d'alimentation d'une usine). La valeur du capital installé n'est autre que la valeur boursière de l'entreprise considérée.

Considérons une entreprise qui dispose d'un capital égal à k_t à la date t. Ce capital permet de produire et génère un flux de profits réels $\pi_t(k_t)$ qui sont reversés aux actionnaires de l'entreprise. La contribution au revenu permanent des actionnaires de ces profits est donc simplement égale à

$$V_t(k_t) = \sum_{i=0}^{+\infty} \frac{\pi_{t+i}(k_{t+i})}{(1+r)^i}, \tag{15}$$

où, comme dans la section précédente, r est le taux d'intérêt réel entre deux périodes consécutives, supposé constant. Le revenu permanent de ces actionnaires est inchangé si à la place de ce flux de dividendes on leur versait immédiatement une somme égale à $V(k_t)$. $V(k_t)$ est donc le prix auquel ils sont prêts à vendre leurs actions, c'est donc la valeur boursière de l'entreprise. Si celle-ci était supérieure à $V(k_t)$ tout le monde vendrait et le prix chuterait, s'il lui était inférieur il serait profitable pour tout le monde d'acheter et le prix augmenterait. La valeur du capital installé de l'entreprise ne peut donc être différente de V.

Le stock de capital suit l'équation d'évolution suivante:

$$k_t = (1-\delta)k_{t-1} + I_t,$$

où I_t est l'investissement à la date t et δ le taux de dépréciation du capital.

Pour augmenter son capital d'une unité à t, l'entreprise doit donc augmenter I_t d'une unité. On suppose que le prix d'achat des biens d'investissements est p_{Kt}, et que par ailleurs il existe des coûts convexes d'installation du capital, ceux-ci d'autant plus grands que l'investissement est élevé relativement à la taille de l'entreprise. Le coût total de l'investissement est donc

$$C_t(I_t) = p_{Kt} I_t (1 + c\left(\frac{I_t}{k_{t-1}}\right)).$$

Cette formule signifie que pour installer chaque unité d'investissement, on doit payer un coût égal à $c(I/k_{t-1})$ unités de bien d'investissement. On suppose la fonction de coût unitaire d'installation $c()$ telle que $c, c', c'' > 0$ pour $I > 0$ avec $c(0) = c'(0) = 0$ et $\lim_{x\to\infty} c(x) = \lim_{x\to\infty} c'(x) = +\infty$.

L'entreprise choisit I de manière à maximiser la valeur du capital installé nette des coûts d'investissement soit, à la date t :

$$\max V_t(k_t) - C_t(I_t).$$

La condition d'optimalité est

$$p_{Kt}\left(1 + c\left(\frac{I_t}{k_{t-1}}\right) + \frac{I_t}{k_{t-1}} c'\left(\frac{I_t}{k_{t-1}}\right)\right) = V'(k_t). \tag{16}$$

Soit h la fonction définie par $h(x) = 1 + c(x) + xc'(x)$. Par hypothèse, $h'(x) = 2c'(x) + xc''(x) > 0$, $h(0) = 1$, et $\lim_{x\to\infty} h(x) = +\infty$. $h()$ est donc une bijection de $[0, +\infty)$ vers $[1, +\infty)$, donc inversible sur ce domaine. Soit $g = h^{-1}$, l'on a

$$I_t = k_{t-1} g\left(\frac{V_t'}{p_{Kt}}\right).$$

L'investissement est une fonction croissante du ratio entre son effet marginal sur la valeur de l'entreprise et son prix d'achat. L'entreprise investira tant que, et d'autant plus que la valeur marginale de l'investissement est élevée relativement à son prix d'achat. Le ratio $q = \frac{V'}{p_K}$ est appelé *q marginal.* Notons que $g(1) = 0$. Si le q marginal est égal à 1, l'investissement est nul, le capital installé n'acquérant pas de valeur supplémentaire par rapport à son prix d'achat. Plus q est élevé, plus il est rentable d'investir, et plus l'entreprise est prête à consentir des coûts d'installation élevés pour convertir les biens d'investissement en capital installé. C'est là le sens de l'équation (16).

Il est facile de calculer V'. L'effet d'une unité supplémentaire de capital à la date t sur le capital à la date $t+i$ se calcule de façon formellement comparable à (14):

$$\frac{dk_{t+i}}{dk_t} = (1-\delta)\frac{dk_{t+i-1}}{dk_t} = ... = (1-\delta)^i \frac{dk_t}{dk_t} = (1-\delta)^i. \tag{17}$$

Il ne reste ensuite qu'à différentier (15):

$$V_t'(k_t) = \sum_{i=0}^{+\infty} \left(\frac{1-\delta}{1+r}\right)^i \pi_{t+i}'(k_{t+i}). \tag{18}$$

Notons que V', et donc q, décroît avec r. Le fait que I ne dépende que de q n'évacue pas la dépendance traditionnelle de I par rapport au taux d'intérêt réel. Une hausse de ce dernier réduit q mécaniquement en réduisant la valeur présente des profits marginaux futurs. La formule précédente implique aussi (comme pour la théorie du revenu permanent) que les anticipations sur la profitabilité future ont un impact sur l'investissement courant. Ainsi, si π_{t+i}' baisse à la date $t+i$, peut-être parce que l'on anticipe une récession ou une hausse du prix des matières premières, q_t baisse et l'investissment baisse dès la date t. Réduire son investissement aujourd'hui en anticipant les baisses futures de profitabilité du capital permet de lisser les coûts d'installation de ce dernier, de la même manière qure, dans la théorie du revenu permanent, une réduction du revenu à une période donnée se traduit par une baisse de la consommation à toutes les périodes, car le consommateur désire lisser son profil de consommation au cours du temps.

Enfin, si les rendements sont constants, alors les profits sont proportionnels au capital de l'entreprise. La fonction π est donc linéaire en k : $\pi_t(k_t) = \pi_t k_t$. On a alors $\pi_t'(k_t) = \pi_t$. π' ne dépend pas de k, et d'après (18),V' non plus. La valeur boursière du capital installé est donc proportionnelle à celui-ci:

$$V_t(K_t) = \omega_t K_t, V_t'(K_t) = \omega_t.$$

On a alors

$$q = \frac{\omega_t}{p_{Kt}} = \frac{\omega_t K_t}{p_{Kt} K_t} = \frac{V(K_t)}{p_{Kt} K_t}.$$

Le q marginal coïncide avec le *q moyen de Tobin,* défini comme le ratio entre la valeur boursière du capital productif de l'entreprise $V(K_t)$ et le coût de remplacement de ce capital, $p_{Kt}K_t$.

5.3 La demande de monnaie: la contrainte *cash in advance* et le modèle de Baumol-Tobin

Si les fondements de la courbe IS tels que nous venons de les analyser impliquent des fonctions de consommation et d'investissement fort différentes de celles supposées par l'approche keynésienne traditionnelle, en revanche la courbe LM est peu différente de ce que prédit un modèle théorique fondé sur l'optimisation. C'est le principal enseignement du modèle de demande de monnaie de Baumol-Tobin.

Pour que la demande de monnaie soit non nulle, il est nécessaire que les agents aient un motif pour la détenir, en dépuit de son caractère intrinsèquement inutile. On formalise simplement l'idée que la monnaie est un moyen d'échange en supposant que les agents sont soumis à une contrainte dite *cash in advance:* Ils ne peuvent conduire une transaction que s'ils disposent d'une quantité de monnaie égale au même montant. Cette contrainte est cependant insuffisante pour éviter que la demande de monnaie soit nulle, parce que les agents pourraient se procurer les liquidités nécessaires dans l'instant qui précède les transactions à effectuer; ils ne détiendraient de la monnaie que pendant une durée infinitésimale, de sorte que leur moyenne d'encaisses serait nulle. Pour éviter cette difficulté, on suppose qu'il faut payer un coût fixe chaque fois que l'on retire de l'argent (le fameux *shoeleather cost* de se rendre à la banque ou au distributeur automatique).

Ainsi, Baumol et Tobin considèrent un consommateur qui doit financer un flux d'achats égal à T par unité de temps, est soumis à une contrainte de cash-in-advance et doit payer un coût fixe b à chaque retrait d'argent. Ce consommateur doit décider de la fréquence optimale de ses trajets à la banque. S'il se rend plus fréquemment à la banque, il peut laisser un solde en moyenne plus élevé sur son compte, ce qui lui rapporte plus d'intérêts; en revanche, il doit payer le coût b plus souvent.

Soit C la quantité retirée à la banque à chaque trajet. Celle-ci est épuisée au bout d'un temps $t = C/T$. Le nombre de déplacements par période est donc égal à

$$n = \frac{1}{t} = \frac{T}{C}.$$

La figure 12 décrit l'évolution des encaisses monétaires détenues par l'agent au cours du temps. Entre deux trajets, cette quantité baisse linéairement de C à 0. En moyenne, la quantité détenue est donc égale à $C/2$. Le coût d'opportunité financier, par unité de temps, de la stratégie de l'agent est égal aux intérêts non perçus sur cette quantité, soit $iC/2$, où i est le taux d'intérêt nominal. Le coût des déplacements à la banque par unité de temps est lui égal à $bn = bT/C$. Le consommateur choisit donc C de façon à résoudre le programme suivant:*et*

$$\min_C L(C) = \frac{iC}{2} + b\frac{T}{C}.$$

La valeur optimale de C est déterminée par la condition du premier ordre, $L'(C) = 0$, soit encore

$$\frac{i}{2} - \frac{bT}{C^2} = 0.$$

La demande de monnaie M^d est égale à $C/2$, d'où

$$M^d = \sqrt{\frac{bT}{2i}}.$$

Le modèle prédit une élasticité au taux d'intérêt bien spécifique et égale à -1/2. Cette grandeur n'est pas incompatible avec celles obtenue dans la littérature empirique (voir par exemple Hoffman et Rasche, 1991).

Les grandeurs b et T sont spécifiées en termes nominaux. Soit P le niveau général des prix, β le coût réel des déplacements à la banque et Y le volume réel des transactions, interprété comme le PIB. Alors $\beta = b/P$ et $Y = T/P$. La demande de monnaie peut être réécrite comme

$$\frac{M^d}{P} = \sqrt{\frac{\beta Y}{2i}} \tag{19}$$

Comme dans l'équation (9), la demande d'encaisses nominales est proportionnelle au niveau général des prix. En revanche, (19) semble prédire une élasticité-revenu de la demande de monnaie égale à 1/2, tandis qu'on a supposé dans (9) une élasticité unitaire. Notons que cette différence ne change rien à l'analyse du modèle IS-LM-AS. Elle est dû au fait qu'une augmentation du volume des transactions à β inchangé se traduirait, en l'absence d'un réajustement de la fréquence n, par une hausse proportionnelle de la monnaie détenue et donc du coût financier correspondant, alors que le coût des déplacements serait lui inchangé. Il est préférable de ne pas augmenter ses encaisses proportionnellement et de se rendre plus souvent à la banque – ceci permet d'obtenir un meilleur équilibre entre les coûts d'opportunité financiers de détention d'encaisses et les *shoeleather costs*. Cependant, il est raisonnable de supposer que dans la plupart des cas, une hausse de Y se traduira par une hausse proportionnelle de β. Par exemple, un hausse de la productivité affectera proportionnellement la production et les salaires. Si β est le coût d'opportunité du temps consacré aux trajets (c'est à dire les revenus qui auraient été perçus si ce temps avait été consacré au travail), alors il est lui-même proportionnel au salaire. On s'attend donc à ce qu'une hausse de la productivité augmente β et Y dans les mêmes proportions, ainsi donc que la quantité $\sqrt{\beta Y}$ et la demande de monnaie. La formulation (9) n'est donc pas incompatible avec le modèle de Baumol-Tobin.

6 L'équilibre à prix fixes

Puisque la rigidité des prix joue un rôle important dans l'équilibre de sous-emploi, il est naturel d'étudier un modèle d'équilibre général à prix fixes, c'est à dire considérés comme exogènes. Ces prix diffèreront en général des prix d'équilibre walrasien. Sur chaque marché, l'offre sera différente de la demande et l'allocation des ressources soumise à un processus de *rationnement.* On suppose généralement que le côté "long" du marché est rationné. Si, sur un marché donné, l'offre est supérieure à la demande, la quantité échangée est égale à la quantité demandée et les producteurs ne peuvent pas produire autant qu'ils le veulent. Si c'est la demande qui est supérieure à l'offre, ce sont les consommateurs qui ne peuvent pas acheter autant du bien qu'ils le désireraient.

Une intuition keynésienne de base est que "la crise nourrit la crise": les travailleurs, se trouvant au chômage, voient leurs revenus diminuer, ce qui réduit leur demande de biens de consommation, aggravant les difficultés des entreprises, les conduisant à des licenciements supplémentaires, etc. Ce raisonnement suppose une boucle de feed-back entre le rationnement sur le marché du travail (les chômeurs ne trouvent pas d'emploi parce que l'offre de travail est supérieure à la demande) et le rationnement sur le macrhé des biens (les entreprises n'écoulent pas leurs produits parce que l'offre de biens est supérieure à la demande). Ce mécanisme est un mécanisme d'équilibre général au sens où il fait intervenir des interactiond entre différents marchés: le marché des biens et le marché du travail. On attend d'un modèle d'équilibre général à prix fixes qu'il puisse prédire la possibilité d'un tel régime de sous-emploi keynésien.

La première difficulté conceptuelle, lorsque l'on considère un équilibre avec rationnement, est que celui-ci pâtit d'une indétermination fondamentale. En effet, de par la loi de Walras, la somme de tous les revenus est égale à la valeur totale de toutes les quantités produites. Si ces quantités sont contraintes par la demande, on peut, partant d'un équilibre donné, réduire le revenu total d'une quantité arbitraire (par exemple 5 Euros) et construire un autre équilibre où la demande globale est également réduite de 5 euros, ce qui valide la croyance selon laquelle les revenus sont plus faibles. En d'autres termes, le revenu et la demande globale sont indéterminés, parce que l'équilibre implique seulement qu'ils doivent être égaux entre eux, et ne détermine pas leur niveau. Cette indétermination fondamentale est sans doute le fondement de l'intuition kéynesienne selon laquelle l'équilibre macroéconomique est instable et qu'on peut se coordonner sur un équilibre de sous-emploi. Mais ceci est peu opérationnel parce qu'on ne sait pas quel équilibre choisir parmi l''infinité d'équilibres de sous-emploi ainsi produits (dans le cas walrasien, on suppose par hypothèse que les agents ne font face à aucune contrainte de rationnement et que les prix relatifs sont compatibles avec l'égalité entre l'offre et la demande sur tous les marchés; mais dans un équilibre à prix fixe les contraintes de rationnement des producteurs sont endogènes et déterminées par le revenu des consommateurs, qu'ils perçoivent par ailleurs dans leurs activités de production, d'où l'indétermination). Dans le modèle IS-LM on considère l'équilibre à une date donnée en s'abstenant de formaliser le futur et en supposant la propension marginale à consommer le revenu courant inférieure à 1. Une baisse du revenu *courant* de 5 euros engendre donc une baisse de la demande *courante* inférieure à 5 euros; l'anticipation d'une telle baisse n'est donc pas autoréalisatrice et l'équilibre est donc unique. Mais l'indétermination que nous venons de mettre en lumière réapparaît dès lors que l'on considère un modèle bouclé où les consommateurs, sujets à leur contrainte budgétaire intertemporelle, allouent leur demande optimalement sur la totalité des biens et des périodes. Une façon de briser cette indétermination est d'introduire la monnaie. Contrairement aux autres biens, la quantité de monnaie ne peut pas baisser sous l'effet d'une baisse de la demande de monnaie parce que celle-ci n'est pas produite par les agents privés mais par la banque centrale, en une quantité fixe et exogène. En conséquence, parmi tous les équilibres de sous-emploi, il n'en existe qu'un qui est compatible avec la détention par les agents d'une quantité de monnaie exactement égale à celle choisie par la banque centrale. Dans le modèle que l'on va voir, la demande de monnaie provient simplement du fait que celle-ci entre dans la fonction d'utilité. Une hausse de revenu est répartie entre les agents entre une hausse de la demande de biens de consommation et une hausse de la demande de monnaie, ce qui implique une propension marginale à consommer le revenu inférieure à 1 comme dans le modèle IS-LM, rétablissant l'unicité de l'équilibre.

6.1 Le modèle (Barro-Grossman, 1971)

Il existe un consommateur représentatif dont la fonction d'utilité est

$$U(C, M) = \theta \ln C + (1 - \theta) \ln \frac{M}{p}.$$

C est la quantité consommée d'un unique bien final, M est la quantité de monnaie détenue et p est le niveau général des prix. Il est doté d'une unité de travail qu'il offre sur le marché de façon inélastique. Il est propriétaire de l'entreprise représentative, dont les profits lui sont reversés. Cette entreprise a une fonction de production

$$Y = L^{\alpha},$$

où L est l'emploi et Y est la production, avec $0 < \alpha < 1$. Le profit de l'entreprise est donc égal à

$$\pi = pL^{\alpha} - wL,$$

où w est le salaire.

Le gouvernement émet la monnaie M; pour se la procurer, les ménages doivent transférer au gouvernement des ressources de valeur égale (le seigneuriage). Le gouvernement utilise ses revenus de la façon suivante: une partie pG est utilisée pour financer les dépenses publiques, dont le niveau réel est G. Le reste, $S = M - pG$ est reversé aux ménages[5]. Le revenu du ménage représen-

[5] Si $pG > M$, on peut considérer que le gouvernement prélève une taxe forfaitaire sur les ménages égale à $pG - M$.

tatif est donc

$$R = S + \pi + wL = M + p(Y - G).$$

On peut alors calculer la demande de bien et de monnaie du consommateur, qui résoud le programme suivant:

$$\max_{C,M} U(C, M)$$

$$s.c. \; pC + M \leq R.$$

La solution est

$$\begin{aligned} C &= \theta \frac{R}{p} = \tilde{C}(R, p), \\ M &= (1 - \theta)R = \tilde{M}(R), \end{aligned}$$

Les fonctions $\tilde{C}$ et $\tilde{M}$ sont des fonctions de demande "notionelles" au sens où le consommateur ignore la possibilité de rationnement sur les marchés des biens et de la monnaie. En revanche on considère R comme le revenu réel de l'individu, et si l'offre de travail de celui-ci est contrainte, R sera inférieur à ce qu'il serait en l'absence de cette contrainte. En remplaçant R par sa valeur, on peut réécrire

$$\begin{aligned} \tilde{C} &= \frac{\theta}{p}(M + p(Y - G)) \qquad (20) \\ \tilde{M} &= (1 - \theta)\left(M + p(Y - G)\right). \end{aligned}$$

6.2 L'équilibre walrasien

Il est facile de de calculer l'équilibre walrasien de cette économie. Cela nous servira de base de comparaison pour étudier les équilibres à prix fixes.

La demande de travail est déterminée par la maximisation du profit des entreprises:

$$\max_{L} \pi = pL^{\alpha} - wL.$$

La condition du premier ordre est

$$\alpha p L^{\alpha-1} - w = 0.$$

D'où la demande de travail, fonction décroissante du salaire réel w/p :

$$L^d(\frac{w}{p}) = \left(\frac{w}{\alpha p}\right)^{-\frac{1}{1-\alpha}}. \tag{21}$$

L'offre de travail est $L^s = 1$. L'équilibre sur le marché du travail détermine donc le salaire réel de façon unique:

$$w = \alpha p.$$

L'équilibre sur le marché des biens s'écrit

$$C + G = Y.$$

La demande de biens est déterminée par (20), et puisque $L = 1$ on a également $Y = 1$. Donc

$$C = 1 - G = \tilde{C} = \frac{\theta}{p}(M + p(1-G)),$$

ce qui permet de calculer le prix d'équilibre

$$p = \frac{\theta M}{(1-G)(1-\theta)}, \tag{22}$$

d'où évidemment

$$w = \frac{\alpha\theta M}{(1-G)(1-\theta)}. \tag{23}$$

Les propriétés du modèle classique sont vérifiées. En particulier:

- Monnaie et dépenses publiques sont sans effet sur l'activité: $\frac{dY}{DM} = \frac{dC}{dM} = \frac{dY}{dG} = 0$
- Il y a effet d'éviction à 100% entre dépenses publiques et consommation privée: $\frac{dC}{dG} = -1$
- La monnaie augmente proportionnellement les prix: $\frac{M}{p}\frac{dp}{dM} = \frac{M}{w}\frac{dw}{dM} = 1$

- Les dépenses publiques augmentent le niveau des prix, car la demande de monnaie est proportionnelle à la consommation et celle-ci baisse: $\frac{dp}{dG} > 0$. Lorsque G augmente, le revenu disponible des agents diminue et ils tentent de répartir cette baisse entre baisse de la consommation et baisse de leurs encaisses réelles. L'effet d'impact de cette tentative d'ajustement est que la demande de biens est supérieure à l'offre, tandis que la demande de monnaie est inférieure à l'offre. Il en résulte des pressions inflationnistes qui conduisent à une hausse des prix. Il en va de même dans le modèle AS-AD lorsque la courbe AS est verticale, sous l'effet d'un déplacement de la courbe AD.

6.3 L'équilibre à prix fixes: le régime de chômage keynésien

Supposons maintenant que les prix soient fixes et exogènes, et donc diffèrent en général des niveaux walrasiens définis par (22) et (23). L'allocation des ressources se fera par le rationnement, et l'on suppose que le côté "long" du marché est rationné. Pour caractériser l'équilibre, il faut donc distinguer entre quatre régimes, selon qu'il y a excès d'offre ou de demande sur le marché des biens et le marché du travail[6]. Un régime appelé "de chômage keynésien" est un régime où il y a excès d'offre sur le marché des biens et sur le marché du travail. Essayons de caractériser un tel régime.

D'après (21), l'entreprise désire employer $L^* = L^d(w/p)$ personnes, ce qui lui permet de produire

$$\begin{aligned} Y^* &= L^d(w/p)^\alpha \\ &= \left(\frac{w}{\alpha p}\right)^{-\frac{\alpha}{1-\alpha}} \end{aligned}$$

unités de bien. Cependant, elle ne peut pas produire cette quantité car la demande est insuffisante. Elle produira donc à la place une quantité

$$Y = \tilde{C} + G.$$

[6] La nature du rationnement sur le marché de la monnaie est déterminée résiduellement. S'il y a excès de demande sur le marché des biens, il y a excès d'offre sur le marché de la monnaie, et vice-versa.

En substituant (20), on peut calculer Y :

$$Y = G + \frac{\theta}{1-\theta}\frac{M}{p}.$$

Le membre de droite de cette équation est une courbe de demande agrégée et en régime keynésien la production est égale à la demande agrégée.

Ce régime satisfait aux propriétés traditionnelles du modèle keynésien. En particulier

- Les dépenses publiques augmentent le PIB, avec un multiplicateur égal à 1: $\frac{dY}{dG} = 1$. Le multiplicateur est ici égal à un car une hausse des dépenses publiques, dans ce modèle statique, réduit d'autant le revenu disponible des agents (à travers la baisse du transfert S); il s'agit là d'un *balanced budget multiplier* dont on sait qu'il est égal à 1 (théorème de Haavelmo).

- De même $dY/dM > 0$, $dC/dM > 0$: la politique monétaire stimule l'économie car les consommateurs voudront dépenser les liquidités supplémentaires.

- Enfin $dY/dp < 0$: une hausse des prix est contractionniste car les agents tenteront de reconstituer en partie leurs encaisses, ce qui réduit la demande (effet Pigou).

Le régime keynesien ne prévaut que s'il y a effectivement excès d'offre sur les deux marchés. Cela implique d'une part que $Y < Y^*$, soit

$$G + \frac{\theta}{1-\theta}\frac{M}{p} < \left(\frac{w}{\alpha p}\right)^{-\frac{\alpha}{1-\alpha}}. \tag{24}$$

D'autre part, il doit y avoir excès d'offre sur le marché du travail. Le consommateur aimerait travailler une quantité $L = 1$. On doit donc avoir $L < 1$, ou, ce qui est équivalent , $Y = L^\alpha < 1^\alpha = 1$, soit

$$G + \frac{\theta}{1-\theta}\frac{M}{p} < 1. \tag{25}$$

Les conditions (24) et (25) définissent l'ensemble des prix (p, w) tels que l'équilibre avec rationnement correspondant se situe dans le régime keynésien.

6.4 Le régime de chômage classique

On appelle chômage classique un régime avec excès d'offre sur le marché du travail mais excès de demande sur le marché des biens. Ceci signifie que $\tilde{C} > Y - G$ et $L = L^* < 1 - L$ coïncide avec L^* puisque les entreprises ne sont contraintes sur aucun des deux marchés où elles interviennent.

Il est alors immédiat de calculer le PIB

$$Y = Y^* = \left(\frac{w}{\alpha p}\right)^{-\frac{\alpha}{1-\alpha}}. \qquad (26)$$

La consommation est calculée résiduellement comme $C = Y^* - G$.

Dans le régime classique,

- Les politiques de demande sont inopérantes: $dY/dG = dY/dM = 0$.
- Les politiques "structurelles" qui visent à une réduction du coût réel du travail résorbent le chômage: $dY/d(w/p) < 0$, $dL/d(w/p) < 0$.

Pour que l'économie soit dans ce régime, on doit avoir $L^* < 1$, soit

$$\frac{w}{\alpha p} > 1. \qquad (27)$$

Cette inégalité valide l'idée que le chômage classique est dû à des salaires réels excessifs.

Par ailleurs, on doit avoir $\tilde{C} > C$, soit $\frac{\theta}{p}(M + p(Y^* - G)) > Y^* - G$, ce qui est équivalent à

$$G + \frac{\theta}{1-\theta}\frac{M}{p} > \left(\frac{w}{\alpha p}\right)^{-\frac{\alpha}{1-\alpha}}. \qquad (28)$$

Cela signifie que la demande agrégée doit être supérieure à l'offre walrasienne, la condition inverse de (24) – si l'économie est sur sa courbe de demande agrégée en régime keynésien, elle se situe au-dessous de celle-ci en régime de chômage classique. De même, on peut définir la relation (26) comme courbe d'offre agrégée; l'économie se situe sur cette courbe en chômage classique mais en-deçà de cette courbe en régime keynésien.

Les conditions (27) et (28) définissent l'ensemble des prix (p, w) correspondant au chômage classique.

6.5 Le régime d'inflation réprimée

Il s'agit d'un régime avec excès de demande sur les deux marchés. On doit donc avoir $1 < L^*$ d'où

$$\frac{w}{\alpha p} < 1. \tag{29}$$

La production est celle de plein emploi:

$$Y = 1,$$

d'où $\tilde{C} = \frac{\theta}{p}(M + p(1 - G))$; l'excès de demande sur le marché des biens s'écrit $\tilde{C} + G > 1$, d'où

$$G + \frac{\theta}{1-\theta}\frac{M}{p} > 1. \tag{30}$$

Le régime d'inflation réprimée a les mêmes propriétés que l'équilibre walrasien, avec cette différence qu'une hausse de la quantité de monnaie ou des dépenses publiques n'est pas inflationniste mais se traduit par une rationnement accru sur le marché des biens, c'est à dire par un écart plus important entre le membre de gauche et le membre de droite de l'équation (30). On a donc $dY/dM = dY/dG = 0$, $dC/dG = 0$. Les politiques de prix sont ici également inopérantes: $dY/dp = dY/dw = 0$. A la marge, les prix ne sont pas allocatifs car les quantités sont contraintes par l'offre de travail (si cependant celle-ci était élastique au salaire réel une hausse de ce dernier augmenterait le niveau d'activité).

6.6 Synthèse

Le régime d'équilibre de l'économie peut être obtenu au moyen du diagramme de phase de la figure 13 dans le plan $(p, w/p)$. Les frontières entre les trois régimes sont définies par

$$w/p = \alpha,$$

qui, d'après (27) et (29), définit la frontière entre chômage classique et inflation réprimée;

$$G+\frac{\theta}{1-\theta}\frac{M}{p}=1,$$

qui définit la frontière entre chômage keynésien et inflation réprimée d'après (25) et (30), et correspond à une valeur unique de p, p_0; et enfin par

$$G+\frac{\theta}{1-\theta}\frac{M}{p}=\left(\frac{w}{\alpha p}\right)^{-\frac{\alpha}{1-\alpha}},$$

l'équation de la frontière entre les deux régimes de chômage d'après (24) et (28). Cette dernière équation définit une relation croissante entre p et w/p.

Il est facile de vérifier que l'intersection entre ces trois frontières – le point E sur la figure 13 – correspond à l'équilibre walrasien.

Remarquons également que l'économie peut se trouver en situation de régime keynésien alors même que $w/p > \alpha$, c'est à dire que les salaires réels sont trop élevés (point A sur la figure 13). Dans une telle situation les entreprises aimeraient produire à un niveau inférieur à celui de plein emploi, mais la demande agrégée est encore plus faible. Les politiques structurelle sont sans effet, les politiques de demande réduisent le chômage mais sans parvenir à l'éliminer complètement. Si l'économie se trouve au point A, une fois l'économie sortie du régime keynésien par des politiques de demande, elle se trouve en chômage classique et les politiques d'offre doivent prendre le relais.

Pour conclure, notons que le quatrième régime n'est pas possible. Supposons en effet qu'il y ait excès d'offre sur le marché des biens et excès de demande sur le marché du travail. Cela signifie que les entreprises sont contraintes sur les deux marchés. Cela n'est pas possible car la fonction de production lie l'offre notionelle sur le marché des biens (resp. la demande notionelle sur le marché du travail) à la quantité transactée sur le marché du travail (resp. sur le marché des biens). En excès de demande de travail, l'entreprise embauchera $L^s = 1$ travailleurs, et son offre de biens sera égale à $Y^s = 1$. On ne peut alors avoir $Y^s > Y^d = \tilde{C}+G$, car alors la demande de travail de l'entreprise serait égale à $L^d = Y^{d\frac{1}{\alpha}} < 1$, ce qui contredit l'hypothèse d'excès de demande sur le marché du travail.

7 Les fondemements de la courbe de Phillips

La courbe de Phillips, ou son équivalent théorique la courbe d'offre agrégée, stipule une relation positive entre l'activité et l'inflation (ou le niveau général des prix). Cette relation est en contradiction avec le principe selon lequel seuls les prix réels jouent un rôle dans les décisions de production. Ainsi, soit Y l'output et $q_i = p_i/\bar{p}$ le prix relatif du bien i par rapport à un numéraire quelconque dont les prix est noté $\bar{p}$. Supposons que Y et les q_i constituent un équilibre de l'offre agrégée, c'est à dire que les plans de production optimaux des entreprises étant donné les q_i impliquent un PIB réel égal à Y. Une hausse du niveau général des prix d'un facteur λ, connue des agents et correctement prise en compte par eux, laisse un tel équilibre inchangé puisque $q_i = \frac{\lambda p_i}{\lambda \bar{p}}$. L'offre agrégée ne dépend donc pas du niveau général des prix, *pour autant que celui-ci soit correctement anticipé.*

Comme on va le voir, tous les fondements de l'offre agrégée reposent sur une mauvaise évaluation par les agents du niveau général des prix au moment de prendre leurs décisions de production, c'est à dire sur des différences entre l'inflation anticipée et l'inflation réalisée. Mentionnons deux des principaux mécanismes sur lesquels la littérature s'est penchée.

7.1 Fixation à l'avance des salaires nominaux

La manière la plus simple d'obtenir une courbe d'offre agrégée est de supposer que les salaires sont fixés une période à l'avance. Soit ω le salaire réel anticipé d'équilibre; nous pouvons ignorer les mécanismes qui déterminent la valeur de ω. On peut considérer, par exemple, que le salaire nominal w est fixé par les syndicats un an à l'avance, et que ceux-ci essayent d'obtenir pour leurs membres un pouvoir d'achat ω. Mais l'analyse qui suit ne repose pas sur cette interprétation particulière, mais uniquement sur le fait qu'ω existe. Nous traitons ω comme constant pour simplifier l'analyse, mais celui-ci peut varier au cours du temps sous l'effet de chocs d'offre: gains de productivité, variations dans le pouvoir de négociation des syndicats, etc. Ces chocs déplacent la courbe d'offre agrégée, mais ce qui nous intéresse ici est le fait qu'elle ait une pente positive.

Le salaire nominal fixé à la date t pour la date $t+1$ est donc

$$w_{t+1} = \omega p^a_{t+1}, \tag{31}$$

où p^a_{t+1} est le niveau des prix anticipé à la date t pour la date $t+1$.

A la date $t+1$, les entreprises embauchent librement. La demande de travail est une fonction décroissante du salaire réel, et il en va de même de la production:

$$Y_{t+1} = f(\frac{w_{t+1}}{p_{t+1}}), f' < 0.$$

Comme w_{t+1} est prédéterminé, ceci définit bien une relation croissante entre le PIB et le niveau général des prix. Ce dernier est allocatif parce que tous les prix n'augmentent pas simultanément: les salaires restes fixés au niveau choisi à la date t. Par ailleurs, les salaires nominaux reflétant les anticipations d'inflation à travers (31), la position de la courbe d'offre agrégée dépend des anticipations d'inflation. On a

$$\frac{w_{t+1}}{p_{t+1}} = \frac{\omega p^a_{t+1}}{p_{t+1}} = \omega \frac{1+\pi^a_{t+1}}{1+\pi_{t+1}}$$

où $\pi_{t+1} = \frac{p_{t+1}}{p_t} - 1$ est le taux d'inflation et $\pi^a_{t+1} = \frac{p^a_{t+1}}{p_t} - 1$ le taux d'inflation anticipée.

Notons qu'une hausse anticipée de l'inflation n'a aucun effet sur la production: on a alors simplement

$$Y = Y^* = f(\omega).$$

Ce ne sont que les hausses non anticipées qui ont un effet, parce que les salaires ne pouvant s'ajuster instantanément à la hausse pour compenser cette surprise inflationniste, le coût réel du travail baisse ce qui favorise les embauches.

7.2 Confusion entre prix relatifs et niveau général des prix ("misperceptions")

Une idée assez différente, mais formellement assez similaire à celle de la rigidité des salaires nominaux, est celle de "misperceptions" due à Friedman. L'hypothèse est que les agents économiques, au moment de faire leur choix, observent le prix du bien qu'ils produisent mais pas les autres prix. Leur information ne leur permet donc pas de distinguer entre une hausse du prix relatif de leur bien et une hausse du niveau général des prix. Lorsqu'ils observent des prix plus élevés, ils considèrent que cela résulte d'un choc de demande qui leur est favorable, avec une certaine probabilité; dans ce cas, il est profitable de produire plus: les chocs inflationnistes engendrent donc une hausse de l'activité. On voit là aussi que c'est la différence entre l'inflation réalisée et l'inflation anticipée (c'est à dire l'évaluation du niveau général des prix, non observé, faite par les agents lorsqu'ils prennent leurs décisions de production) qui stimule l'économie. Si par exemple un producteur de chaussure observe que le prix des chaussures augmente de 8 % alors qu'il anticipe une inflation de 8 %, il conclura que le prix relatif des chaussures est inchangé et n'aura pas intérêt à accroître sa production. S'il observe une hausse de 9 %, il attribuera en moyenne un point de cette hausse à un choc de demande qui augmente le prix relatif des chaussures de 1 %, et accroîtra sa production en conséquence. Si l'inflation est de 9 % tous les producteurs augmenteront leur production, croyant chacun à tort bénéficier d'un choc de demande relatif qui leur est favorable. Ce n'est que parce que l'inflation est non observée au moment des décisions de production et supérieure à l'anticipation des agents que ceux-ci peuvent tous se tromper simultanément, et dans le même sens.

7.3 Les anticipations adaptatives et l'hypothèse accélérationniste

Ces fondements de la courbe de Phillips valident l'*hypothèse du taux naturel,* que nous formulerons ainsi:

HYPOTHESE DU TAUX NATUREL – *Si les anticipations d'inflation des agents sont correctes, alors le PIB (ou le chômage) est égal à son taux naturel. Ce taux ne dépend ni des prix ni de la demande agrégée.*

En d'autres termes, la courbe d'offre agrégée est verticale. Sa position peut fluctuer sous l'effets de chocs d'offre; le niveau d'équilibre du PIB est défini de façon unique par a position et le fait qu'elle est verticale – il est donc indépendant des prix et de la demande agrégée. La relation positive observée entre activité et inflation est en réalité due à une relation positive entre le PIB et la *surprise inflationniste,* c'est à dire la différence entre inflation et inflation anticipée:

$$Y = f(\pi - \pi^e). \tag{32}$$

Dès lors que des chocs inflationnistes affectent π plus que π^e, on observera (sous l'effet des chocs de demande) une corrélation positive entre Y et π[7]. Il n'en reste pas moins que les fluctuations anticipées de l'inflation sont sans effet puisqu'alors

$$Y = Y^* = f(0). \tag{33}$$

En d'autres termes (figure 14), l'équation (**??**) définit une courbe d'offre agrégée de court terme, à anticipations données, dont la position dépend de π^e, et qui se déplace vers le haut lorsque π^e augmente. La relation (33) définit une courbe d'offre agrégée de long terme, c'est à dire une relation qui décrit l'effet asymptotique sur Y d'un changement permanent de $\pi-$ il est en effet naturel de supposer qu'un tel changement sera à long terme entièrement pris en compte dans les anticipations des agents.

Une fois l'hypothèse du taux naturel énoncée, pour en savoir plus sur les co-mouvements entre inflation et activité, il est nécessaire de se faire une idée de la formation des anticipations. Jusqu'à ce que les anticipations rationnelles s'imposent, on supposait les anticipations *adaptatives*, c'est à

[7] Comme on l'a vu plus haut, les chocs d'offre, eux, génèrent une corrélation négative entre output et inflation.

dire que les agents corrigent graduellement leurs erreurs:

$$\pi_t^e = \pi_{t-1}^e + \alpha(\pi_{t-1} - \pi_{t-1}^e). \tag{34}$$

Cette formule signifie que, si par exemple $\pi_{t-1} > \pi_{t-1}^e$, les agents révisent leurs anticipations à la hausse en corrigeant une fraction α de l'erreur commise. Prenons une approximation linéaire de (32),

$$y_t = \pi_t - \pi_t^e, \tag{35}$$

où $y_t = \ln Y_t$ et le coefficient de y sur $\pi - \pi^e$ (resp. la constante) est normalisé à 1 (resp. 0) pour simplifier. Alors le taux naturel y^* est égal à zéro en logarithmes et y_t s'interprète naturellement comme "l'écart de production" ("output gap"). Supposons que pendant plusieurs périodes consécutives, $y = \bar{y} > 0$, soit un écart de production positif. Alors d'après (34), $\pi_{t+1}^e = \pi_t^e + \alpha\bar{y} > \pi_t^e$. Les anticipations d'inflations s'accélèrent. De plus, d'après (35), $\pi_t = \pi_t^e + \bar{y}$; la différence entre inflation anticipée et inflation réalisée étant constante, cette dernière s'accélère également. On en déduit que l'inflation s'accélère tant que l'écart de production est positif; s'il était négatif, on montrerait pareillement que l'inflation décélère. Cela conduit à formuler l'*hypothèse accélérationniste*:

HYPOTHESE ACCELERATIONNISTE – *L'inflation s'accélère (resp. décélère) lorsque le PIB est supérieur (resp. inférieur) au taux naturel.*

Cette formulation est relativement vague mais nous pouvons lui donner un contenu plus précis. Pour cela, dérivons une courbe de Phillips "dynamique", c'est à dire une relation entre y et les valeurs contemporaines et passées de π. En substituant dans (34) cette même expression écrite à la date précédente, on obtient

$$\pi_t^e = \alpha\pi_{t-1} + (1-\alpha)\left[(1-\alpha)\pi_{t-2}^e + \alpha\pi_{t-2}\right]. \tag{36}$$

En itérant cette opération, on obtient

$$\pi_t^e = \alpha \sum_{i=0}^{+\infty} (1-\alpha)^i \pi_{t-1-i}. \tag{37}$$

L'inflation anticipée est donc une moyenne pondérée des niveaux passés de l'inflation, à coefficients exponentiellement décroissants lorsque l'on va plus loin dans le passé. C'est bien une moyenne car les coefficients sont positifs et leur somme est égale à 1:

$$\alpha \sum_{i=0}^{+\infty} (1-\alpha)^i = \alpha \frac{1}{1-(1-\alpha)} = 1.$$

D'après (35), si les anticipations sont adaptatives, alors une courbe de Phillips spécifiée comme

$$y_t = \sum_{i=0}^{+\infty} a_i \pi_{t-i} \tag{38}$$

sera telle que

$$\sum_{i=0}^{+\infty} a_i = 0. \tag{39}$$

Cela a plusieurs conséquences. D'une part, on peut vérifier que $y = 0$ si π est constant au cours du temps. En état stationnaire, le PIB est égal à son taux naturel. D'autre part, sachant que $a_0 > 0$, pour avoir $y_t > 0$ on doit avoir $\pi_t > \bar{\pi}_{t-1} = \sum_{i=1}^{+\infty} \left(-\frac{a_i}{a_0}\right) \pi_{t-i}$. Or cette dernière quantité est une mesure de l'inflation moyenne passée, puisque, comme $\sum_{i=0}^{+\infty} a_i = 0$, on a nécessairement $\sum_{i=1}^{+\infty} \left(-\frac{a_i}{a_0}\right) = 1$. Si l'écart de production est positif, alors l'inflation s'accélère, au sens où l'inflation courante est supérieure à la moyenne de l'inflation passée telle que mesurée par $\bar{\pi}_{t-1}$.

Nous pouvons donc redéfinir l'hypothèse accélérationniste de façon plus opérationnelle:

HYPOTHESE ACCELERATIONNISTE 2 – *Si la courbe de Phillips obéit à la spécification (38), alors la somme des coefficients $\{a_i\}$ de cette relation est nulle.*

7.4 Désinflation optimale et accomodation

Une question importante est celle du choix optimal entre inflation et chômage. L'approche simpliste représentée sur la figure 10 ignore le fait que la position de la courbe de Phillips elle-même dépend des anticipations d'inflation, et que celle-ci est verticale à long-terme. Plus le gouvernement choisit un taux d'inflation élevé (et donc un chômage faible, plus les anticipations d'inflation seront élevées par la suite, ce qui déteriore l'ensemble de choix du gouvernement.

Supposons par exemple que le gouvernement maximise

$$u(y, \pi) = -\pi^2 - \beta y^2. \tag{40}$$

Le gouvernement essaie de ramener l'écart de production ainsi que l'inflation à zéro. Plus ces variables s'éloignent de ces points idéaux, plus son utilité est faible. Plus le paramètre β est élevé, plus le poids sur y est élevé et plus gouvernement est prêt à déstabiliser le niveau des prix en échange d'une meilleure stabilisation du niveau d'activité.

Un gouvernement complètement myope (avec une valeur de ρ extrêmement élevée) maximise seulement son utilité à la date t, $u(y_t, \pi_t)$, sous la contrainte (35). Comme π_t^e est l'inflation anticipée à la date $t-1$ pour la date t, cette variable ne dépend pas des actions du gouvernement décidées à la date t; le gouvernement considère donc π_t^e comme donné, et résoud le problème suivant

$$\max_{\pi_t} -\pi_t^2 - \beta(\pi_t - \pi_t^e)^2. \tag{41}$$

La solution est

$$\pi_t = \frac{\beta}{1+\beta}\pi_t^e$$

On constate que

- Le gouvernement *accommode* partiellement les anticipations d'inflation. Plus celles-ci sont élevées, plus le gouvernement choisira un taux d'inflation effectivement élevé. En effet, une hausse des anticipations d'inflation rend un taux d'inflation donné plus contractionniste, puisque c'est la différence entre inflation réalisée et inflation anticipée qui compte dans la courbe de πPhillips. Si $\pi_t^e > 0$, le gouvernement poursuit une politique de désinflation, puisque $\pi < \pi^e$, mais accomode partiellement les anticipations afin de limiter les effets contractionnistes de sa politique.

- La fraction de l'inflation anticipée qui est accomodée, $\frac{\beta}{1+\beta}$, est d'autant plus élevée que le poids des préférences du gouvernement sur la stabilisation du PIB, β, est élevé. Un gouvernement qui ne stabilise que le PIB ($\beta \to \infty$) accommode à 100 % les anticipations d'inflation ($\frac{\beta}{1+\beta} \approx 1$) et ne met en place aucune désinflation. Un gouvernement qui ne stabilise que les prix ($\beta = 0$) n'accomode pas l'inflation, mettant en oeuvre une inflation nulle quelles que soient les conséquences de cette politique sur le niveau d'activité.

Supposons maintenant un gouvernement qui ne se préoccupe que du long terme. Pour simplifier, supposons qu'à la date t il choisisse un taux d'inflation constant égal à π, qu'il laisse l'économie évoluer vers son état stationnaire, et qu'il maximise le membre de droite de (40) où y est la valeur de long terme de l'écart de production. Il est facile de voir à partir de (37) que $\pi^e \to \pi$. Cela implique qu'à long terme $y = y^* = 0$, et que notre gouvernement choisira π de manière à maximiser $-\pi^2$, soit $\pi = 0$.

Ceci montre que moins un gouvernement est myope, moins il accomode l'inflation, car plus il prend en compte l'effet adverse de cette accomodation sur les anticipations d'inflation future. A la limite, un gouvernement qui ne se préoccupe que du long terme internalise le fait qu'à long terme la courbe de Phillips est verticale, et qu'il ne peut mieux faire que choisir un taux d'inflation nul.

7.5 Test économétrique de l'hypothèse du taux naturel

Historiquement, on a testé économétriquement l'hypothèse du taux naturel en estimant une équation telle que (38) et en testant si (39) est valide. Les études de l'époque trouvent qu'en réalité $\sum_{i=0}^{+\infty} a_i > 0$. Elles en concluent que l'hypothèse du taux naturel doit être rejetée et qu'il existe un arbitrage de long terme entre inflation et chômage. En effet, d'après (38) une hausse permanente de l'inflation d'une quantité $\Delta\pi$ se traduit à long terme par une hausse de l'activité $\Delta y = \left(\sum_{i=0}^{+\infty} a_i\right) \Delta\pi > 0$. En d'autres termes, la courbe d'offre agrégée de long terme n'est pas verticale.

En fait, ces tests ne rejettent pas l'hypothèse du taux naturel, mais plutôt l'hypothèse accélérationniste. Or, on a vu que celle-ci est valide si l''hypothèse du taux naturel est vérifiée *et* si les anticipations sont adaptatives. Le fait que l'hypothèse accélérationniste soit rejetée n'implique donc pas nécessairement que celle du taux naturel le soit. Une autre possibilité est que les anticipations ne soient pas adaptatives. C'est cette observation qui a conduit à la formulation de l'hypothèse des *anticipations rationnelles.*

8 La critique des anticipations rationnelles

8.1 Retour sur l'hypothèse accélérationniste

Le point de départ des anticipations rationnelles est le suivant: il n'est pas généralement optimal de former ses anticipations de façon adaptative.

Considérons un exemple simple. Supposons que l'inflation π_t soit déterminée par un processus AR1 similaire à celui qui a été introduit dans notre analyse du revenu permanent:

$$\pi_t = \rho \pi_{t-1} + \eta_t, \tag{42}$$

où ρ est un paramètre et η_t un choc exogène d'inflation, i.i.d. et de moyenne nulle. A la date $t-1$ les agents observent le taux d'inflation courant et veulent former leurs anticipations pour le taux d'inflation de la date t sur la base de cette information. Il est raisonnable de supposer qu'ils veulent minimiser la valeur moyenne de l'erreur quadratique commise. Soit π_t^e leur anticipation d'inflation; la valeur optimale de cette anticipation est donc celle qui résoud

$$\min_{\pi_t^e} E\left[(\pi_t - \pi_t^e)^2 \mid I_{t-1}\right],$$

où E désigne l'espérance mathématique et I_{t-1} l'information connue des agents lorsqu'ils forment leurs anticipations, c'est à dire à la date $t-1$. Cette information se résume essentiellement au taux d'inflation de la date $t-1$, π_{t-1}. Les taux d'inflation des périodes précédentes, π_{t-2}, π_{t-3}, etc, sont également connus à $t-1$, mais d'après (42) ils n'apportent aucune information utile pour prévoir π_t.

La condition du premier ordre de ce problème est

$$\begin{aligned} 0 &= \frac{\partial}{\partial \pi_t^e} E\left[(\pi_t - \pi_t^e)^2 \mid I_{t-1}\right] \\ &= E\left[\frac{\partial}{\partial \pi_t^e}(\pi_t - \pi_t^e)^2 \mid I_{t-1}\right] \\ &= 2\left(E\left[\pi_t \mid I_{t-1}\right] - \pi_t^e\right). \end{aligned}$$

Puisque $E[\pi_t \mid I_{t-1}] = \rho\pi_{t-1}$ on doit donc avoir

$$\pi_t^e = \rho\pi_{t-1}. \tag{43}$$

Supposons maintenant l'hypothèse du taux naturel correcte, et qu'on puisse écrire la courbe d'offre agrégée sous la forme

$$y_t = a(\pi_t - \pi_t^e). \tag{44}$$

Alors, en remplaçant dans (44) π_t^e par sa valeur définie par (43), on trouve

$$y_t = a\pi_t - \rho a\pi_{t-1}. \tag{45}$$

Cette relation est l'équivalent de (38) et la somme des coefficients est clairement égale à $a(1-\rho)$, qui est strictement positif dès lors que $\rho < 1$. La courbe de Phillips estimée n'*est donc pas accélérationniste.* D'après cette relation estimée, une hausse unitaire de l'inflation, permanente, augmente le PIB d'une quantité a à court terme et d'une quantité $a(1-\rho)$ à long terme. Un économètre ayant estimé une relation telle que (45) recommenderait donc une telle hausse permanente de l'inflation à un décideur politique qui serait prêt à la tolérer pour réduire le chômage à long terme.

Supposons maintenant que cette décision soit mise en oeuvre. Cela signifie qu'à l'évolution autonome de l'inflation décrite par le processus (42), se superpose désormais une composante de politique économique. Pour simplifier, supposons qu'à chaque période cette composante augmente l'inflation d'un montant fixe égal à $\Delta\pi$, relativement au niveau qu'elle atteindrait en l'absence d'intervention. Cela signifie que l'économie se trouve dans un nouveau régime où l'inflation obéit désormais, à la place de (42), au processus suivant:

$$\pi_t = \rho\pi_{t-1} + \eta_t + \Delta\pi. \tag{46}$$

On a désormais $E[\pi_t \mid I_{t-1}] = \rho\pi_{t-1} + \Delta\pi$, et donc

$$\pi_t^e = \rho\pi_{t-1} + \Delta\pi.$$

Le processus de formation des anticipations a changé car il reflète désormais le nouveau régime de politique économique. En particulier, d'après (44), le niveau du PIB à la date t est égal à>0

$$y_t = a\eta_t.$$

La politique suivie est donc inopérante, contrairement à ce que l'on aurait anticipé sur la base de la forme réduite (45). Cette inefficacité de la politique économique est due au fait que l'inflation ainsi générée est parfaitement anticipée. Or, d'après l'hypothèse du taux naturel, l'inflation anticipée n'a pas d'effet sur l'activité. La relation (45) n'est valide que tant que les anticipations obéissent à (43), or cette relation n'est plus valide au sein du nouveau régime de politique économique. Il était donc erronné de s'appuyer sur une telle relation pour évalue l'effet de la politique considérée.

8.2 L'hypothèse des anticipations rationnelles

Lorsque les anticipations sont adaptatives, les agents corrigent mécaniquement et partiellement leurs erreurs de prévision. Au contraire, lorsque celles-ci sont rationnelles, ils utilisent de façon pertinente toute l'information dont ils disposent. Cette approche a un certain nombre de conséquences importantes.

Tout d'abord, l'anticipation d'une variable quelconque X est égale à son espérance mathématique, comme on l'a montré dans l'exemple précédent. Cela signifie que pour former ses anticipations, les agents doivent prendre en compte une distribution de probabilité pour les variables concernées. Par hypothèse, on suppose qu'ils utilisent la distribution de probabilité qui est correcte, c'est à dire celle qui prévaut à l'équilibre considéré.

Comme en général, les anticipations des agents déterminent leur comportement, et donc les valeurs d'équilibre des variables endogènes, cela signifie que les anticipations et l'équilibre sont déterminés simultanément. Un *équilibre à anticipations rationnelles* est une trajectoire de l'économie telle que (i) la distribution de probabilité des variables endogènes est déterminée par les équations de comportement et simultanément (ii) les anticipations qui interviennent dans ces équations sont égales à l'espérance mathématique de la variable considérée, conditionnellement à l'information dont dispose les agents.

Tout se passe comme si les agents qui peuplent l'économie utilisaient, de la même manière que les macroéconomistes qui travaillent pour des administrations, un modèle pour former leurs prédictions, et si ce modèle se trouvait être le *vrai* modèle de l'économie. [En réalité, malgré l'utilité de cette métaphore, dans bien des cas il n'est pas nécessaire de connaître le vrai modèle pour prédire une variable dès lors que celle-ci est observable; il suffit de connaître la distribution de probabilité des observables, c'est à dire la forme réduite du modèle].

8.3 Calculer un équilibre à anticipations rationnelles: l'exemple de Muth (1960)

L'exemple suivant, emprunté à Muth (1960) montre comment calculer un équilibre à anticipations rationnelles.

Il s'agit d'un modèle d'équilibre partiel d'un marché où les décisions de production ont lieu avant la réalisation de l'équilibre. Soit y la quantité produite et p le prix; notons p^e le prix anticipé par les producteurs au moment de leur décision, alors l'équation d'offre est

$$y = \gamma p^e + u, \tag{47}$$

où u est un choc d'offre de moyenne nulle ($Eu = 0$) et $\gamma > 0$.

L'équation de demande est donnée par

$$y = -\beta p, \tag{48}$$

où $\beta > 0$.

Il est élémentaire de calculer le prix d'équilibre *à anticipations données:*

$$p = -\frac{\gamma}{\beta} p^e - \frac{1}{\beta} u. \tag{49}$$

Selon l'hypothèse des anticipations rationnelles, $p^e = Ep$. En prenant les espérances mathématiques des deux côtés de l'équation précédente, on trouve donc

$$p^e = -\frac{\gamma}{\beta} p^e, \tag{50}$$

d'où

$$p^e = 0.$$

Contrairement aux anticipations adaptatives qui ne dépendent que du passé, il y a ici une authentique *détermination d'équilibre* des anticipations. On peut interpréter p^e dans le membre de gauche de l'équation (49) comme l'anticipation de prix de la *moyenne* des producteurs (en effet, l'équation (47) détermine l'offre globale sur le marché). Le membre de gauche de (50) s'interprète lui comme l'anticipation optimale d'un producteur quelconque. Tous les producteurs ayant la même information, cette anticipation est la même pour tous et doit donc coïncider avec sa valeur moyenne. L'équation (50) implique qu'il n'existe qu'une seule valeur de p^e qui satisfait à ces conditions d'équilibre, soit $p^e = 0$.

Pour conclure que leur meilleure anticipation de prix est $p^e = 0$, les agents qui peuplent cette économie doivent se tenir le raisonnement qui précède. Ils n'observent pas l'anticipation de prix des autres agents, mais sachant que les autres agents sont identique à eux-mêmes, ils comprennent que cette anticipation doit nécessairement satisfaire à (50), et qu'il n'y a pas d'autre possibilité que $p^e = 0$. Si en moyenne tout le monde anticipait une autre valeur, alors il serait rationnel pour chaque agent d'anticiper autre chose que cette valeur, ce qui est une contradiction.

Une fois calculé p^e, il est trivial de calculer le prix d'équilibre comme fonction de la réalisation du choc u. D'après (49), on a en effet

$$p = -\frac{1}{\beta} u.$$

L'analyse qui précède illustre un principe assez général de résolution des équilibres à anticipations rationnelles: la méthode en deux temps. Dans un premier temps, on prend les espérances mathématiques des relations structurelles, ce qui permet de calculer les valeurs d'équilibre de ces espérances – et donc des anticipations. C'est ce qu'on fait dans (50). Dans un second temps, on résoud le modèle de manière classique en utilisant les valeurs d'équilibre des anticipations que l'on a calculées dans la première étape, ce qui revient à les traiter comme des variables exogènes.

Il est instructif d'appliquer cette méthode au cas où le choc u n'est pas de moyenne nulle: $Eu \neq 0$. Eu s'interprète naturellement comme l'anticipation de u tandis que la différence $u - Eu$ est la composante non anticipée de u. En appliquant l'opérateur espérance mathématique E à (49), on trouve maintenant:

$$Ep = p^e = -\frac{Eu}{\beta + \gamma}. \tag{51}$$

Il suffit ensuite de substituer ceci dans (49) et l'on trouve

$$\begin{aligned} p &= \frac{\gamma}{\beta(\beta + \gamma)} Eu - \frac{1}{\beta} u \\ &= -\frac{1}{\beta}(u - Eu) - \frac{1}{\beta + \gamma} Eu. \end{aligned}$$

Cette formule implique notamment qu'un choc d'offre non anticipé a un effet plus élevé en valeur absolue sur le prix d'équilibre qu'un choc d'offre anticipé. En effet, dans le premier cas l'élasticité est $-\frac{1}{\beta}$ et dans le second $-\frac{1}{\beta+\gamma}$, or $\frac{1}{\beta+\gamma} < \frac{1}{\beta}$. Cela est dû au fait que si un choc $u > 0$ est anticipé, $Ep < 0$ et d'après (47) les producteurs réduisent leur offre, anticipant un prix plus faible. Cette réponse de l'offre a pour effet de mitiger la baisse des prix.

Remarque: cette structure forme la base du modèle "Cobweb" d'oscillations déterministes. Si les agents anticipent un prix égal au prix passé, alors $p_t^e = p_{t-1}$ et l'offre à t est égale à $\gamma p_{t-1} + u_t$. Le prix d'équilibre à t vaut $p_t = -\frac{\gamma}{\beta} p_{t-1} - \frac{1}{\beta} u_t$. La dynamique des prix est donc oscillatoire et explosive si $\gamma/\beta > 1$. En anticipations rationnelles ces oscillations disparaissent, les agents formant leurs anticipations sur leur connaissance de la structure future de l'équilibre et pas à partir de prix passés non pertinents.

8.3.1 Exercices

Exercice 1 – Suppons que la structure que nous venons d'analyser se répète à chaque période t, et que les décisions de production pour la date t se font sur la base de l'information disponible à la date $t-1$, soit $p_t^e = E_{t-1}p_t$. Supposons que le choc u_t suive un processus AR1

$$u_t = \rho u_{t-1} + \varepsilon_t,$$

avec ε_t iid et $E\varepsilon_t = 0$.

1. Montrer que

$$p_t^e = -\frac{1}{\beta+\gamma}\rho u_{t-1}$$

2. Montrer que

$$u_t = -\beta p_t - \gamma p_t^e$$

3. En conclure que

$$p_t^e = \frac{\beta}{\beta+\gamma}\rho p_{t-1} + \frac{\gamma}{\beta+\gamma}\rho p_{t-1}^e.$$

4. Pour quelle valeur de ρ le processus de révision des anticipations coïncide-t-il avec les anticipations adaptatives?

5. Montrer que la fraction des erreurs corrigée par les agents dépend des paramètre structurels des équations d'offre et de demande. Montrer que les anticipations sont révisées plus rapidement lorsque β/γ est plus élevé. Pouvez-vous expliquer pourquoi?

Exercice 2 – On considère une version à deux périodes du modèle de Muth. On suppose qu'il existe des spéculateurs qui peuvent acheter le produit à la date $t=1$ pour le revendre à la date $t=2$. La quantité achetée est d'autant plus grande que les gains de la spéculation sont élevés, i.e. que $p_2^e - p_1$ est élevé. Ainsi, à la date $t=1$ l'offre est déterminée par (47):

$$y_1 = \gamma p_1^e + u_1,$$

en revanche la demande comporte une composante spéculative:

$$y_1 = -\beta p_1 + \alpha(p_2^e - p_1),$$

avec $\alpha > 0$. A la date $t = 2$ la demande est déterminée par (48),

$$y_2 = -\beta p_2, \tag{52}$$

mais l'offre reflète la revente des stocks spéculatifs:

$$y_2 = \gamma p_2^e + u_2 + \alpha(p_2^e - p_1).$$

On suppose u iid avec $Eu_1 = Eu_2 = 0$. p_2^e est le prix anticipé à $t = 2$ conditionnellement à l'information disponible à $t = 1$: $p_2^e = E_1 p_2$. p_1^e est le prix anticipé à la date $t = 1$ avant la réalisation des chocs et de l'équilibre à cette date, soit en l'absence de toute information.

1. Montrer que l'équilibre en période 2 implique que

$$p_2^e = \frac{\alpha p_1}{\alpha + \beta + \gamma}.$$

Expliquer pourquoi $dp_2^e/dp_1 > 0$.

2. Calculer la valeur d'équilibre de p_2 en fonction de p_1 et de u_2.

3. Montrer que le prix d'équilibre à la date $t = 1$ est tel que

$$-kp_1 = \gamma p_1^e + u_1,$$

avec

$$k = \frac{\alpha(2\beta + \gamma) + \beta(\beta + \gamma)}{\alpha + \beta + \gamma}.$$

4. En déduire que $p_1^e = 0$ et donc que $p_1 = -u_1/k$.

5. Calculer $dk/d\alpha$ et en conclure que la volatilité de p_1 décroît avec l'intensité de la spéculation. Expliquer.

9 Le modèle de misperceptions de Lucas (1973)

On se souvient de l'idée de Friedman selon laquelle l'existence d'une courbe de Phillips serait due à une confusion faite par les agents, au moment de planifier leur production, entre hausse du niveau général des prix et hausse du prix relatif de leur produit. Le modèle de Lucas (1973) donne des fondations théoriques à cette idée et montre qu'elle peut impliquer une courbe d'offre du type

$$y = \alpha(p - p^e),$$

cohérente avec l'hypothèse du taux naturel. De plus, le modèle de Lucas implique que le paramètre α n'est pas structurel mais dépend en particulier de la volatilité de la politique monétaire – ce qui implique en particulier qu'un changement de régime de politique monétaire a un effet sur la pente de la courbe de Phillips. Enfin, Lucas emprunte à Muth l'hypothèse d'anticipations rationnelles, $p^e = Ep$, pour en conclure qu'une politique de stabilisation systématique est sans effet sur l'activité. Seules les suprises, la partie non anticipée de l'inflation égale à $p - Ep$, ont un effet sur l'activité. La politique monétaire n'a donc donc d'intérêt que si la banque centrale dispose d'une information supérieure au secteur privé. Dans ce cas, elle peut, sur la base de cette information, créer une suprise inflationniste qui a des effets réels sur l'activité, tout en ramenant y et/ou p vers des valeurs-cibles désirées. Si la banque centrale a la même information que le secteur privé, elle ne peut affecter y qu'en rajoutant du bruit non corrélé avec son information, ce qui réduit le bien être car cela augmente la volatilité de y en lui ajoutant une composante aléatoire. Enfin, même dans le cas où la banque centrale a un avantage informationnel sur les agents privés, elle peut toujours rendre son information publique, ce qui aura tendance à réduire l'écart entre p et p^e et à ramener y vers son niveau naturel.[8]

[8] Les autorités ne peuvent espérer mieux que de stabiliser y autour de son taux naturel. Sous l'hypothèse du taux naturel, toute tentative de maintenir y à un niveau supérieur à ce taux (par exemple) conduira à une spirale inflationniste.

9.1 Le modèle d'inférence normal

Le modèle de Lucas repose sur le modèle d'inférence normal, que nousn présentons brièvement ici. Soit une variable aléatoire $x \sim N(\bar{x}, \sigma_x)$: x est distribuée normalement avec une moyenne $\bar{x}$ et une variance σ_x^2. Supposons que l'on observe un signal bruité de x,

$$y = x + \varepsilon,$$

où $\varepsilon \sim N(0, \sigma_\varepsilon)$ est le bruit. Alors on peut montrer en utilisant la loi de Bayes que la meilleure inférence possible sur x est

$$E(x \mid y) = \theta\bar{x} + (1 - \theta)y,$$

où

$$\theta = \frac{\sigma_\varepsilon^2}{\sigma_\varepsilon^2 + \sigma_x^2}.$$

Cette espérance conditionnelle est donc une moyenne pondérée entre la réalisation du signal y et la moyenne inconditionnelle de $x, \bar{x}$. Le poids sur cette dernière est d'autant plus grand que σ_ε est grand: plus le signal est bruité, moins on lui accorde d'importance. Dans le cas limite où le rapport signal/bruit est nul, i.e. $\sigma_\varepsilon/\sigma_x \to \infty$, on a $\theta = 1$: le signal est complétement non informatif et la moyenne de x postérieure au signal est la même qu'en l'absence de signal. Si au contraire le signal est très peu bruité, $\sigma_\varepsilon/\sigma_x \to 0$, on a au contraire $\theta = 0$; l'inférence sur x coïncide avec la réalisation du signal.

9.2 La courbe d'offre de Lucas

Lucas considère un continuum de producteurs individuels, produisant des biens indexés par i. Le prix du bien i à la date t, à l'équilibre, est donné par (en logarithmes)

$$p_{it} = p_t + z_{it},$$

où p_t est le niveau général des prix et z_{it} le prix relatif d'équilibre de ce bien. Ce prix dépend de chocs qui affectentdemande du bien correspondant. Ainsi, un bien i bénéficiant d'une demande élevée aura, à l'équilibre, une valeur z_{it} élevée.

Le niveau général des prix p_{it} fluctue sous l'influence de chocs inflationnistes, et pour un bien donné i, le prix relatif z_{it} fluctue sous l'effet des chocs spécifiques de demande qui frappent ce secteur. On suppose qu'au moment de prendre leurs décisions de production, les producteurs de bien i observent le prix p_i, mais n'observent ni le niveau général des prix ni la valeur du prix relatif z_{it}. Cependant, les agents connaissent le vrai modèle de l'économie; en particulier, ils connaissent les distributions d'équilibre des variables p et z, ce qui leur permet de former des inférences sur la réalisation de celles-ci, à partir de l'observation de leur prix propre p_i.

Nous allons supposer que $z_{it} \sim N(0, \sigma_z)$. Cela signifie que dans chaque secteur i, la valeur d'équilibre de z est tirée selon cette loi. En principe, z est endogène puisque c'est un prix relatif d'équilibre. Si le modèle était spécifié complètement, il y aurait une équation d'offre et une équation de demande pour le bien i; ces équations serait affectées par des chocs exogènes, et l'on devrait calculer la valeur d'équilibre de z en fonction de ces chocs. Mais ici "l'action" est focalisée sur le comportement d'offre des producteurs et leurs misperceptions; comme les décisions de ceux-ci ne dépendent que des prix d'équilibre et pas des forces sous-jacentes qui mènent à ces prix, on ne perd guère à traiter z comme une variable exogène. Il est toujours possible de construire artificiellement une courbe de demande pour le bien i avec une structure idoine des chocs de demande, qui garantit qu'à l'équilibre sur le marché i le prix relatif suit bien une loi normale de moyenne nulle et de variance σ_z^2.

Nous supposons également que $p_t \sim N(\bar{p}_t, \sigma_p)$. Le prix à la date t suit une distribution normale qui est connue des agents. Cependant, cette distribution est ici endogène et il nous faudra ultérieurement calculer les valeurs d'équilibre de $\bar{p}_t$ et σ_p. Pour des raisons de commodité technique, nous voulons utiliser le modèle d'inférence normal et nous supposons *a priori* qu'à l'équilibre p suit une loi normale. Ce n'est *pas* une hypothèse *du modèle* puisque p est endogène, c'est une hypothèse *de travail.* Elle nous permet de résoudre notre modèle et donc en particulier de calculer la distribution d'équilibre de p. Il faudra s'assurer qu'elle est bien normale.

Dans un secteur i quelconque, l'offre de biens est une fonction croissante du prix relatif d'équilibre anticipé:

$$y_{it} = \gamma E(z_{it} \mid p_{it}).$$

Notons qu'il n'y a aucune illusion monétaire. Les agents ne se préoccupent que de grandeurs réelles et c'est le prix relatif qui détermine l'offre. Cependant, celui-ci n'est pas observé et l'inférence est polluée par les fluctuations du niveau général des prix, une grandeur purement nominale.

L'équation qui précède se réécrit ainsi:

$$\begin{aligned} y_{it} &= \gamma E(p_{it} - p_t \mid p_{it}) \\ &= \gamma \left(p_{it} - E(p_t \mid p_{it})\right). \end{aligned}$$

D'après le modèle d'inférence normal, on a

$$E(p_t \mid p_{it}) = \theta \bar{p}_t + (1-\theta) p_{it}$$

avec

$$\theta = \frac{\sigma_z^2}{\sigma_z^2 + \sigma_p^2}. \tag{53}$$

D'où $y_{it} = \gamma\theta(p_{it} - \bar{p}_t)$, relation au niveau sectoriel que nous pouvons agréger entre secteurs en appliquant la loi des grands nombres, ce qui nous donne la courbe d'offre agrégée suivante

$$y_t = \gamma\theta(p_t - \bar{p}_t). \tag{54}$$

Ici, deux remarques s'imposent.

Notons tout d'abord à nouveau que le prix d'équilibre p est endogène. Il dépend en général des paramètres du modèle, et en particulier de la pente de la courbe d'offre $\gamma\theta$. Cela implique que (53) détermine θ de manière implicite. A ce stade de notre résolution, nous ne connaissons pas la valeur de $\theta-$ pas plus que nous ne sommes certains que p suive effectivement une loi normale.

Deuxièmement, $\bar{p}_t$ est la moyenne de la distribution *a priori* de p_t avant observation par les producteurs de leur prix propre p_{it}, mais conditionnellement à toute l'information dont ils disposent au début de la date t. En d'autres termes

$$\bar{p}_t = E_{t-1} p_t.$$

On supposera que toutes les variables antérieures à la date t sont observées lorsque les agents résolvent le problème d'inférence discuté plus haut. C'est le cas en particulier du niveau général des prix p_{t-1} : si le niveau général des prix courant n'est pas observé lorsque les agents forment leurs plans de production, il est connu la période suivante.

Enfin, soit $\pi_t = p_t - p_{t-1}$ le taux d'inflation entre $t-1$ et t. On a $E_{t-1}\pi_t = E_{t-1}p_t - E_{t-1}p_{t-1} = E_{t-1}p_t - p_{t-1}$. Donc, d'après (54)

$$y_t = \theta\gamma(\pi_t - E_{t-1}\pi_t).$$

C'est la fameuse "courbe d'offre de Lucas", qui implique que seule la composante non anticipée de l'inflation a des effets réels.

9.3 Le bouclage du modèle: la courbe de demande agrégée

Le modèle de Lucas est bouclé grâce à une courbe de demande agrégée. Les fluctuations de y et π sont dues à des déplacements de cette courbe AD sous l'effet de chocs sur la quantité de monnaie.

On suppose une courbe de demande agrégée fondée sur la théorie quantitative:

$$p_t + y_t = m_t. \tag{55}$$

La quantité de monnaie suit une marche aléatoire

$$m_t = m_{t-1} + \varepsilon_t.$$

On suppose ε_t iid et $\varepsilon_t \sim N(0, \sigma_\varepsilon^2)$. En éliminant y entre (54) et (55), on se ramène à une seule équation qui détermine p (ou plutôt le processus suivi par p) en fonction de m :

$$\gamma\theta(p_t - E_{t-1}p_t) + p_t = m_t. \tag{56}$$

Pour le résoudre, on utilise la même méthode en deux temps que dans le modèle de Muth.

1. On applique l'opérateur d'anticipations E_{t-1} de part et d'autre de l'équation (56):

$$E_{t-1}p_t = E_{t-1}m_t = m_{t-1}. \tag{57}$$

2. On remplace $E_{t-1}p_t$ par sa valeur d'équilibre obtenue dans (57), dans (56), ce qui permet de calculer la valeur d'équilibre de p :

$$p_t = \frac{1}{1+\gamma\theta}\left(m_t + \gamma\theta m_{t-1}\right). \tag{58}$$

On peut enfin utiliser la courbe AD (55) pour calculer également y :

$$\begin{aligned} y_t &= \frac{\gamma\theta}{1+\gamma\theta}(m_t - m_{t-1}) \qquad (59) \\ &= \frac{\gamma\theta}{1+\gamma\theta}\varepsilon_t. \end{aligned}$$

Cette formule confirme que seule les surprises (ε_t) ont un effet réel. La politique monétaire anticipée n'a pas d'effets.

Remarquons que puisque m est distribué normalement, p l'est aussi d'après (58). Notre hypothèse de travail selon laquelle p est normale est donc cohérente. De plus, la variance de p_t conditionnellement à l'information disponible à $t-1$, σ_p^2 est égale à

$$\begin{aligned} \sigma_p^2 &= E_{t-1}(p_t - E_{t-1}p_t)^2 \\ &= E_{t-1}(p_t - m_{t-1})^2 \\ &= \frac{\sigma_\varepsilon^2}{(1+\gamma\theta)^2}. \qquad (60) \end{aligned}$$

Les équations (53) et (60) déterminent simultanément les valeurs d'équilibre de θ et σ_p^2. On montrera dans un exercice que $d\theta/d\sigma_\varepsilon^2 < 0$. Cela signifie que plus la volatilité des chocs monétaires est élevée, plus l'arbitrage inflation/chômage à court terme est défavorable (i.e. plus la courbe d'offre agrégée de court terme est proche de la verticale). En effet, plus σ_ε^2 est élevé, plus il est rationnel pour les producteurs d'interpréter les fluctuations de leur prix propre comme résultant d'un choc monétaire plutôt que d'un choc de demande idiosyncratique, et plus la réponse de leur quantité produite à ces fluctuations sera faible.

9.4 Evidence empirique

Le modèle de Lucas implique une prédiction simple: la courbe de Phillips est d'autant plus pentue que la volatilité de l'inflation est élevée (Figure 14). Cette prédiction est à peu près valide sur les données analysées par Lucas dans son article de 1973.

Barro (1977) estime une régression du PIB sur composantes anticipées et non anticipées de la monnaie, c'est à dire l'équivalent empirique de l'équation (59). Ses résultats sont cohérents avec l'article de Lucas: l'hypothèse que la composante anticipée (qui serait m_{t-1} ici) est sans effet n'est pas rejetée; l'hypothèse que la composante non anticipée (ici $m_t - m_{t-1} = \varepsilon_t$) est sans effet est rejetée.

9.5 Exercices

Exercice 1 – On considère un modèle du type suivant:

$$x = f(y); \tag{61}$$

$$y = g(x). \tag{62}$$

Le première équation décrit le choix par les agents d'une variable x comme fonction d'une grandeur agrégée y. La seconde décrit la manière dont le choix de x affecte la valeur d'équilibre de y. Un équilibre est un couple (x, y) qui est solution de (61)-(62).

Supposons qu'un tel équilibre existe. Notons-le (x^*, y^*). Nous voulons savoir si cet équilibre est "stable" par tâtonnement. Le processus de tâtonnement est défini comme suit: à l'étape k, les agents choisissent x en fonction de la valeur de y observée à l'étape précédente, y_{k-1}. Ceci définit le choit de x à l'étape k, x_k :

$$x_k = f(y_{k-1}).$$

Ce comportement détermine la nouvelle valeur d'équilibre de y à l'étape k

$$y_k = g(x_k).$$

L'équilibre est stable s'il existe un voisinage de (x^*, y^*) tel, que si les valeurs initiales de x et y, x_0 et y_0, se trouvent dans ce voisinage, alors la suite $\{(x_k, y_k)\}$ ainsi définie converge vers (x^*, y^*).

1. Montrer que si (x_{k-1}, y_{k-1}) sont assez proche de x^*, y^*, alors au premier ordre:

$$y_k - y^* = g'(x^*)f'(y^*)(y_{k-1} - y^*).$$

2. En déduire que si $-1 < g'(x^*)f'(y^*) < 1$, l'équilibre est stable.

3. Appliquons ce schéma au modèle de Lucas. Supposons que x soit θ, qui est proportionnel à la réponse des producteurs à l'inflation non anticipée. Supposons que y soit σ_p^2, la variance des prix à l'équilibre. Quelles sont alors les fonctions $f()$ et $g()$?

4. Supposons que $g'f' < 1$, ce qui garantit la stabilité de l'équilibre. Montrer qu'alors $d\theta/d\sigma_\varepsilon^2 < 0$.

Bibliographie

Hoffman, Dennis et Robert H. Rasche (1991), "Long-run Income and Interest Elasticities of Money Demand in the United States", *Review of Economics and Statistics*

Figure 1: Le diagramme IS-LM

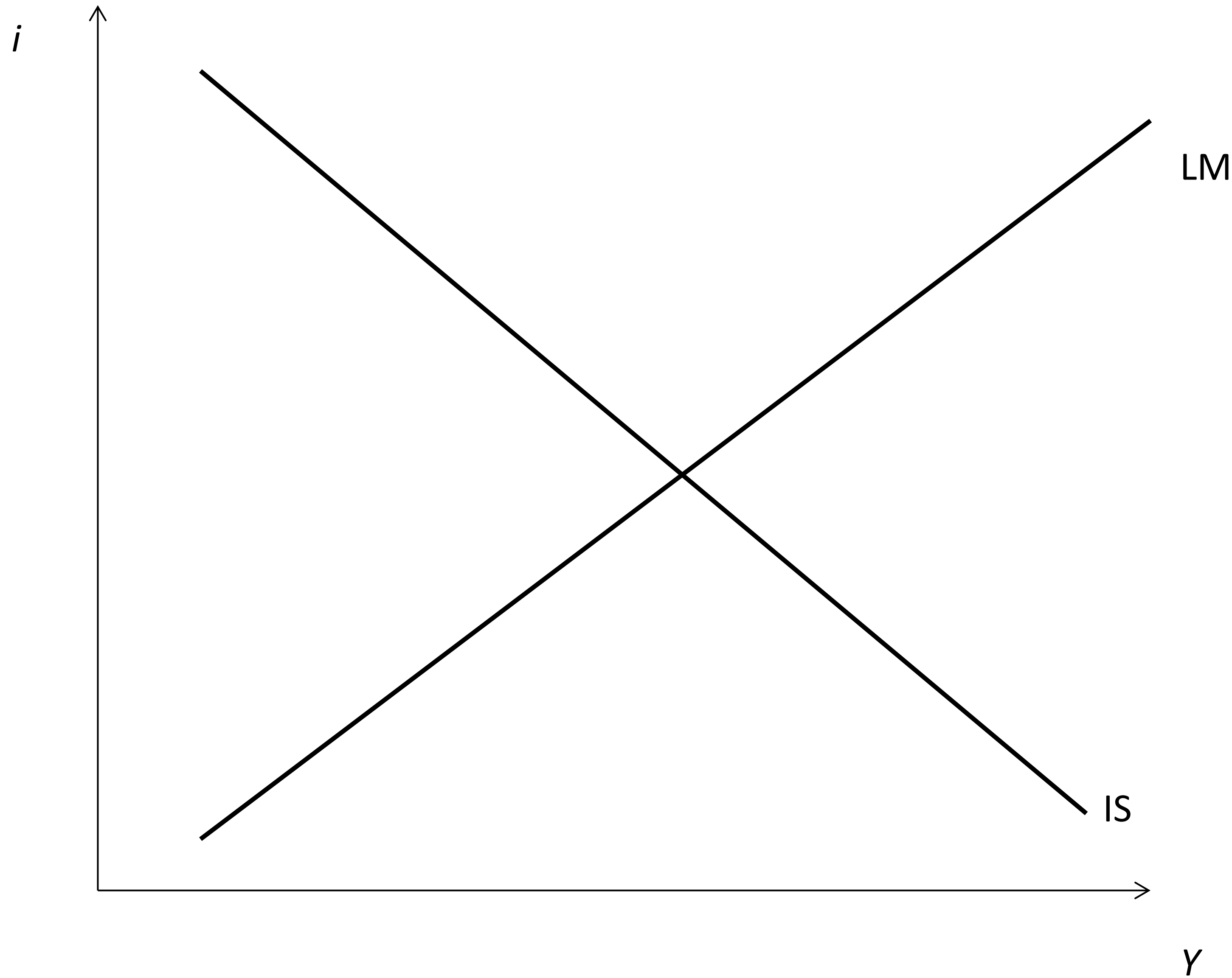

Figure 2: Impact d'une hausse des dépenses publiques

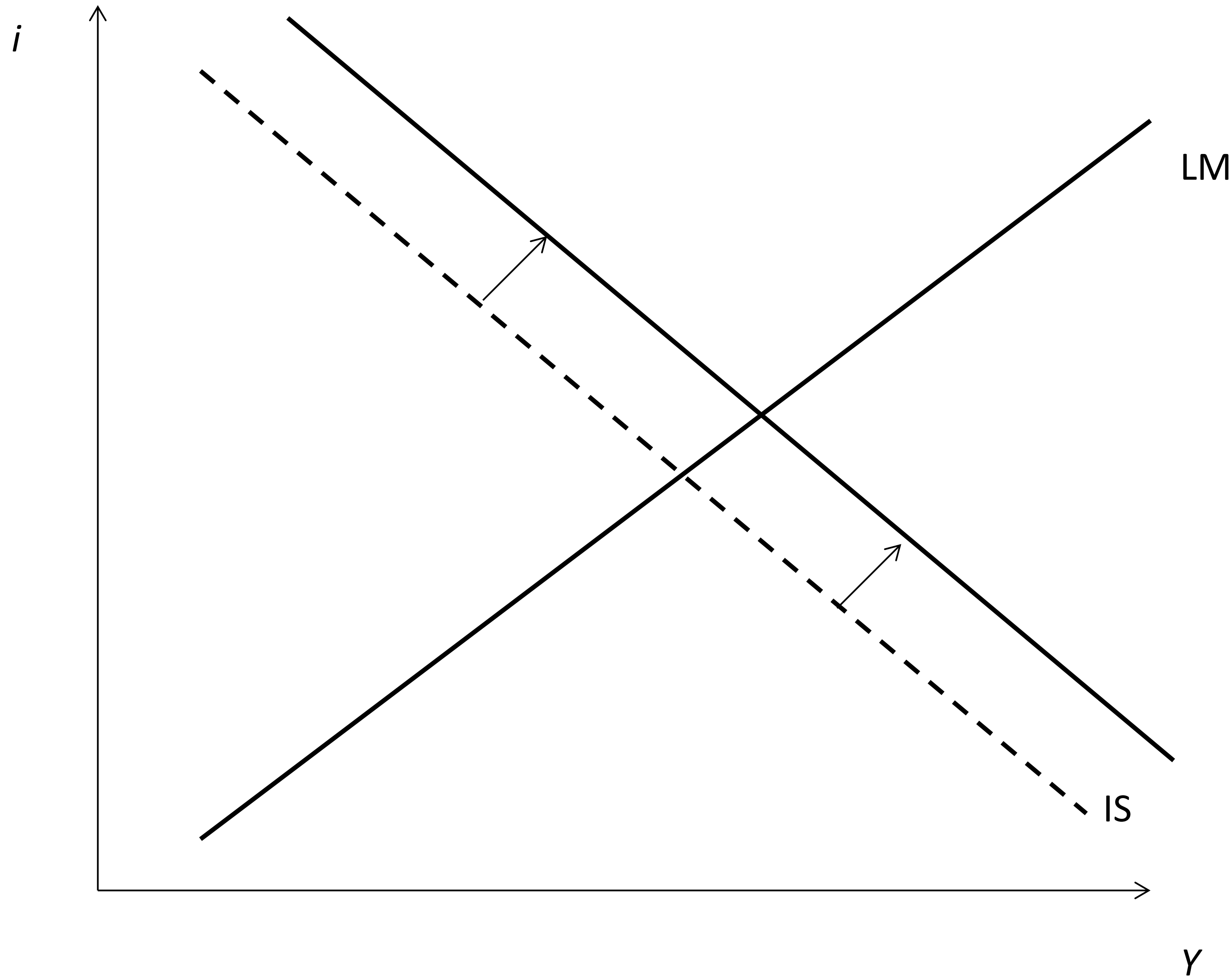

Figure 3: Le cas de la trappe à liquidité

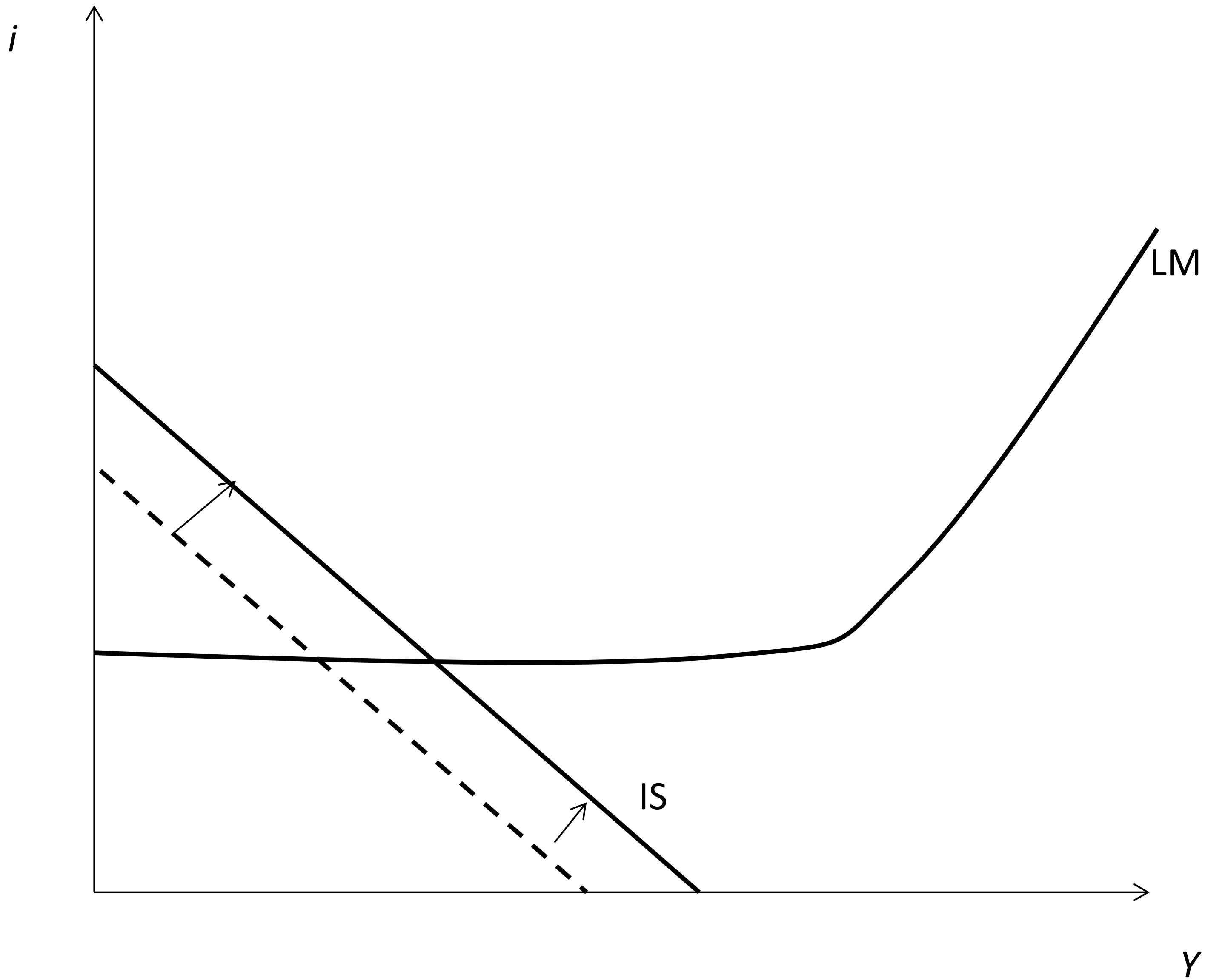

Figure 4: Le cas de la théorie quantitative de la monnaie

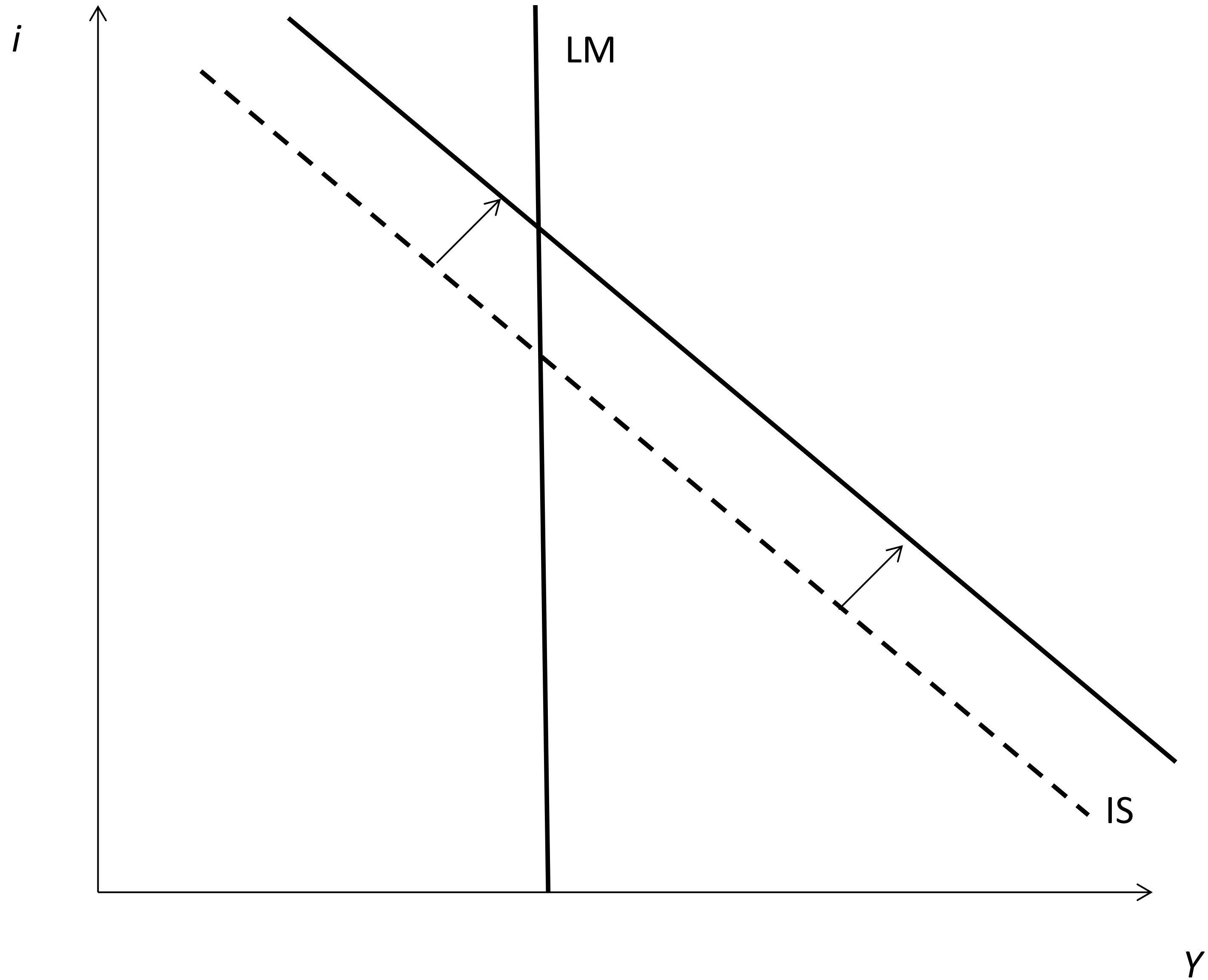

Figure 5: Impact d'une hausse de la masse monétaire

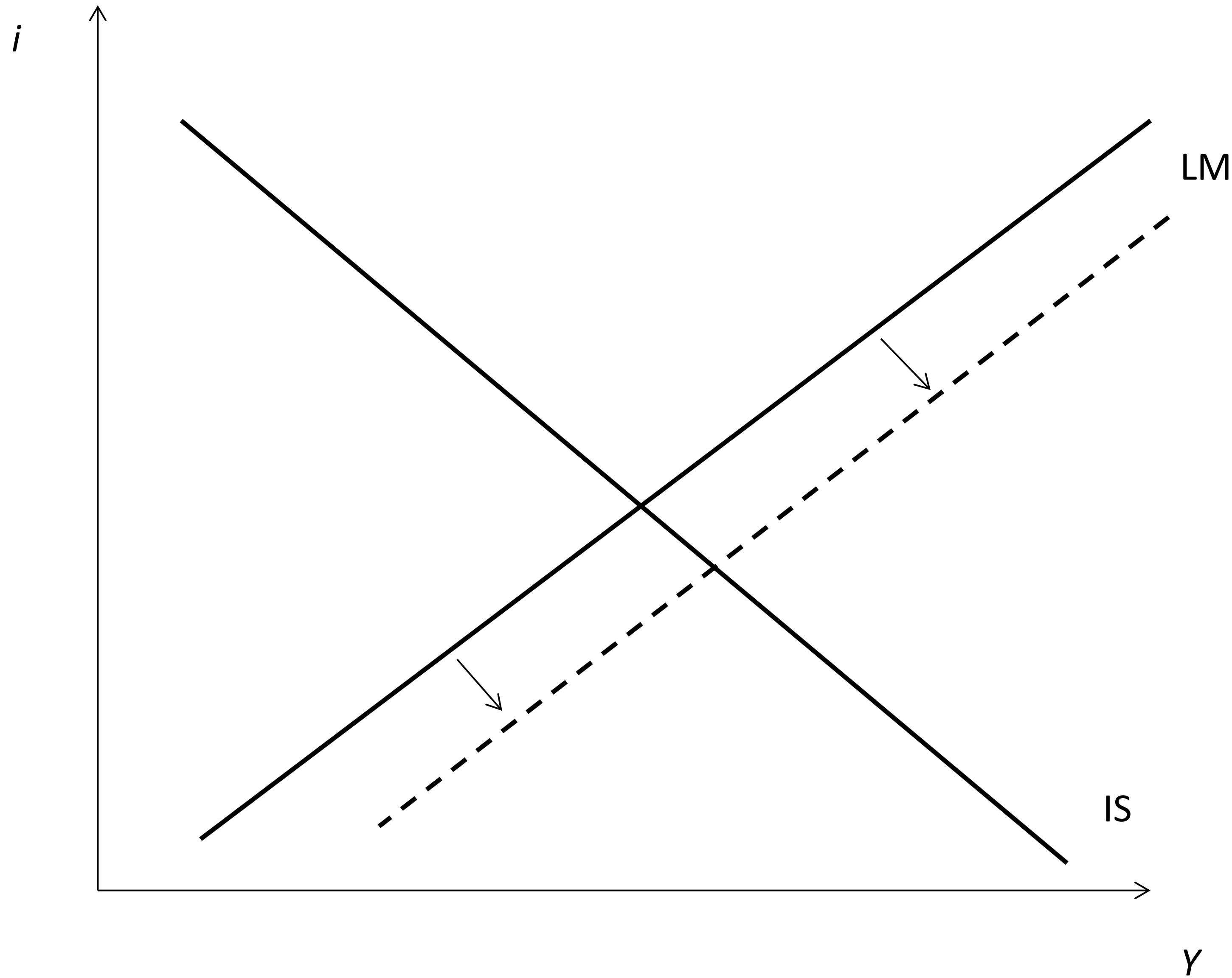

Figure 6: Le diagramme AS-AD

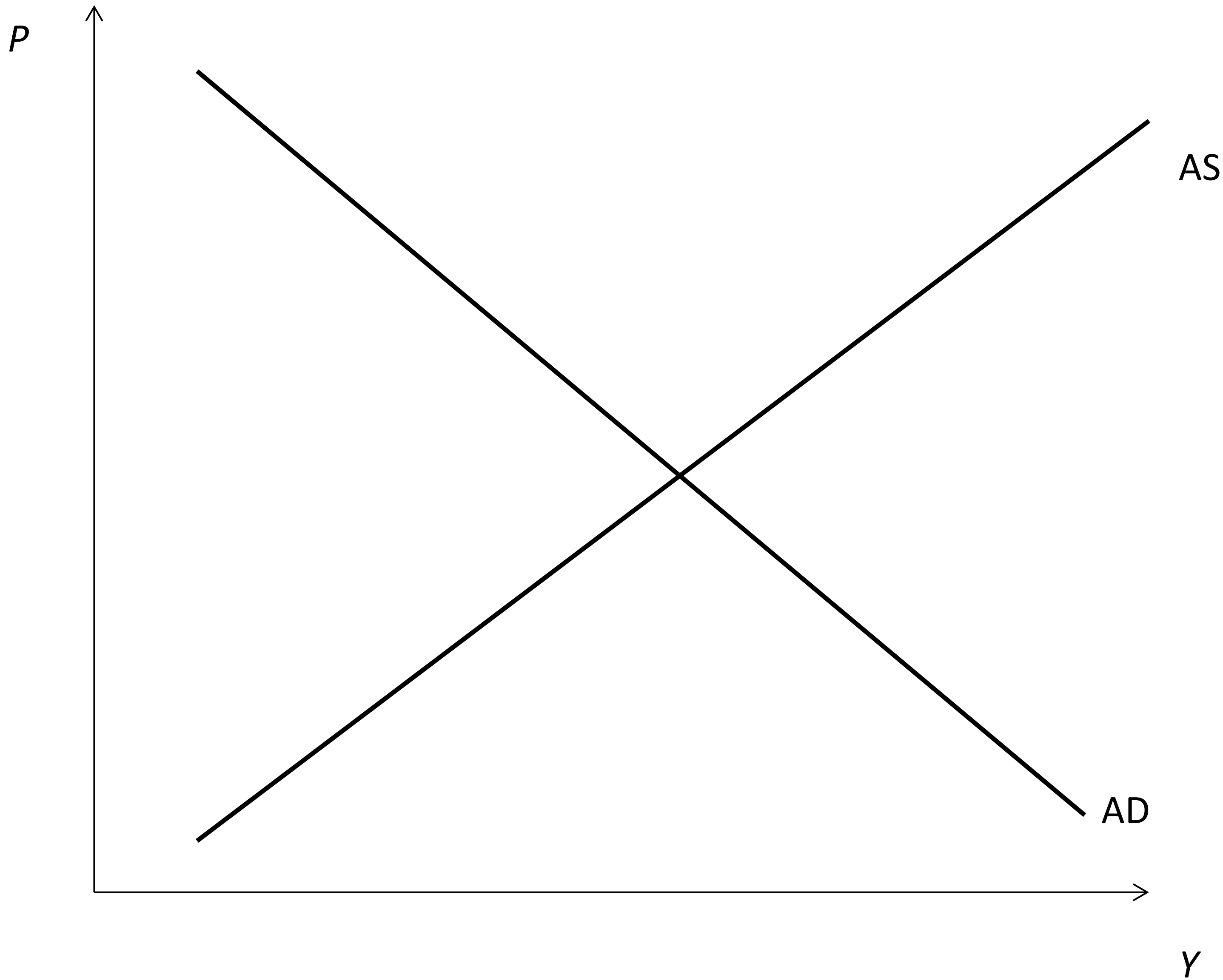

Figure 7: Impact d'un choc d'offre

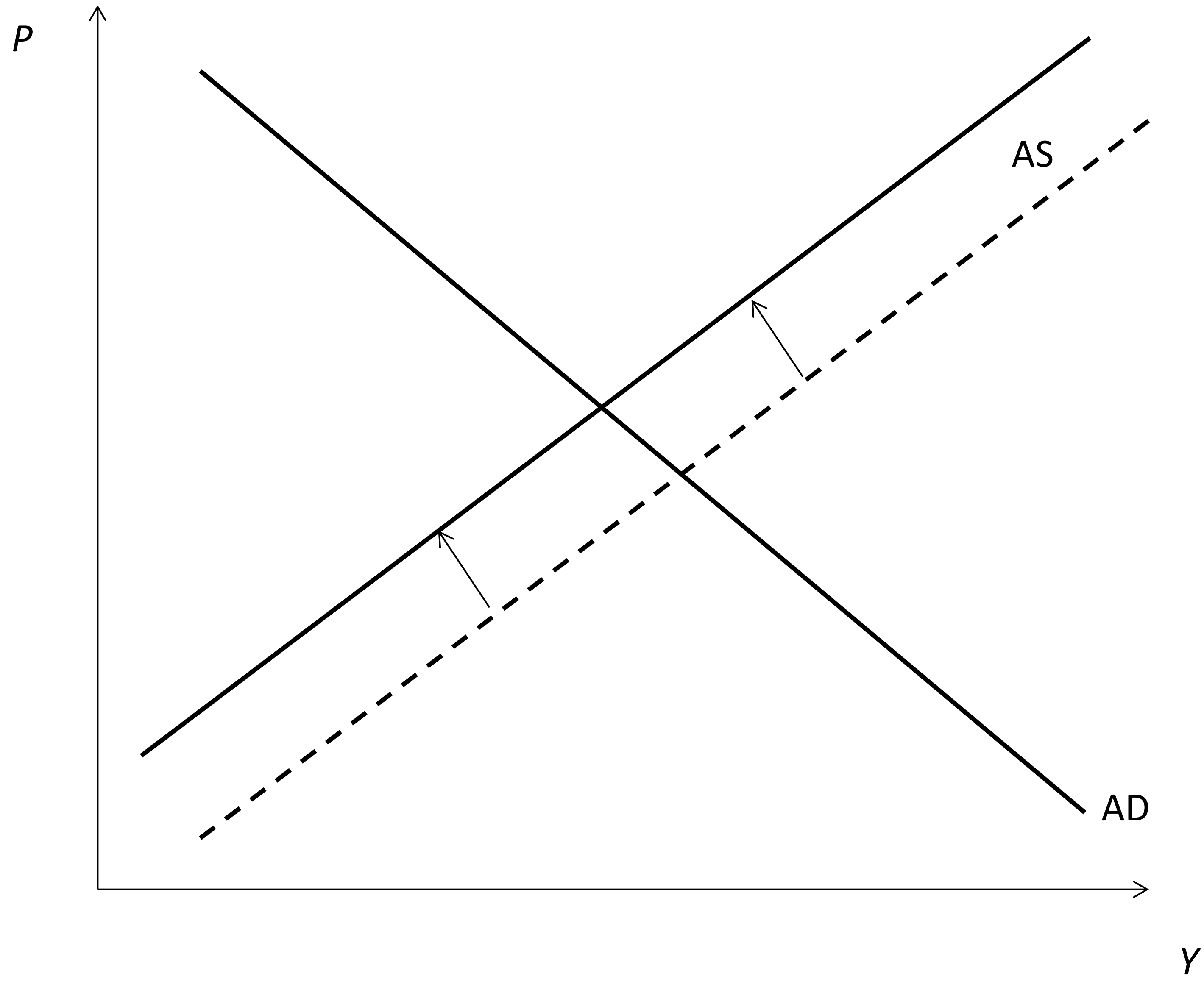

Figure 8: Le modèle “classique”

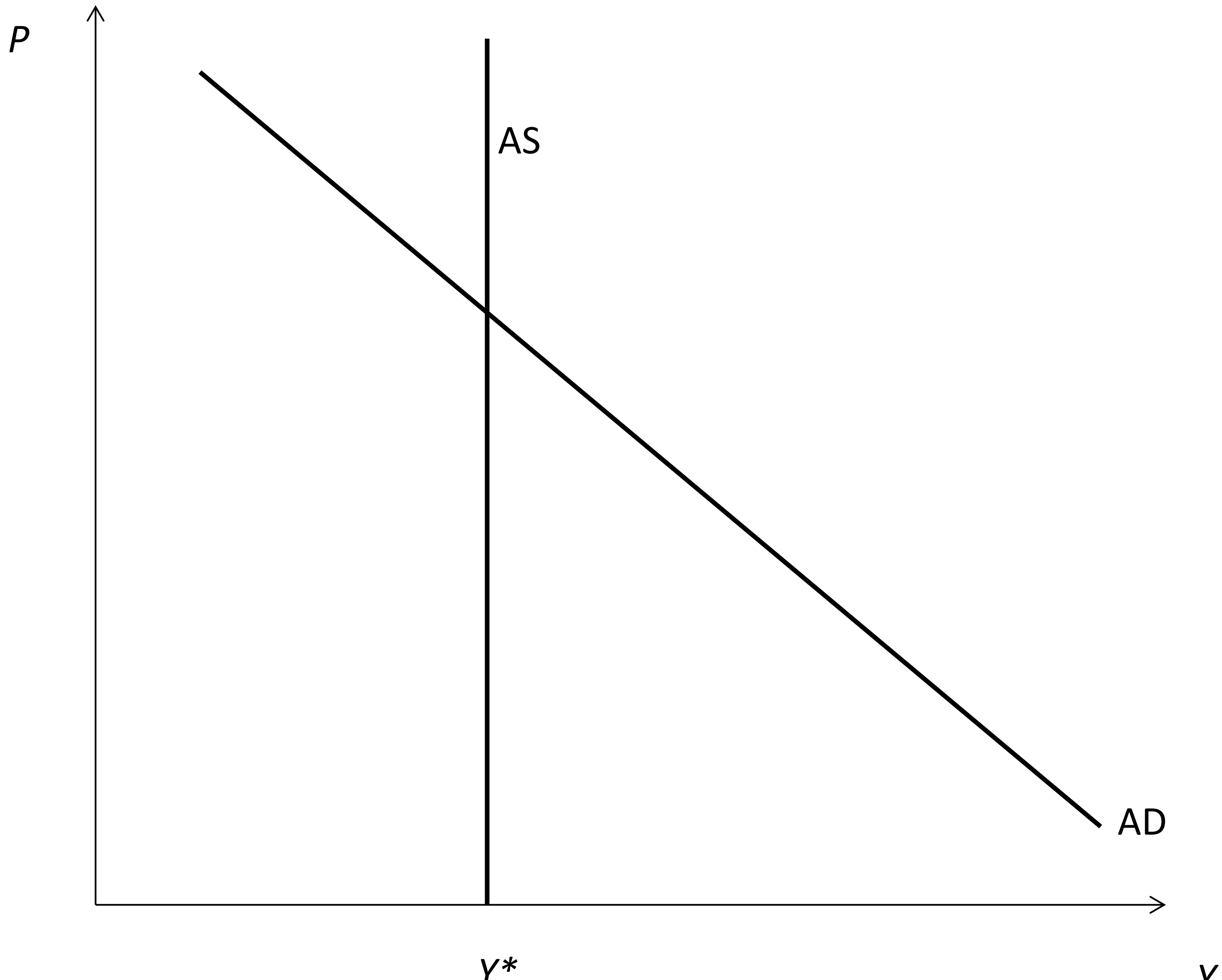

Figure 9: Effet à court et long terme d'un choc de demande

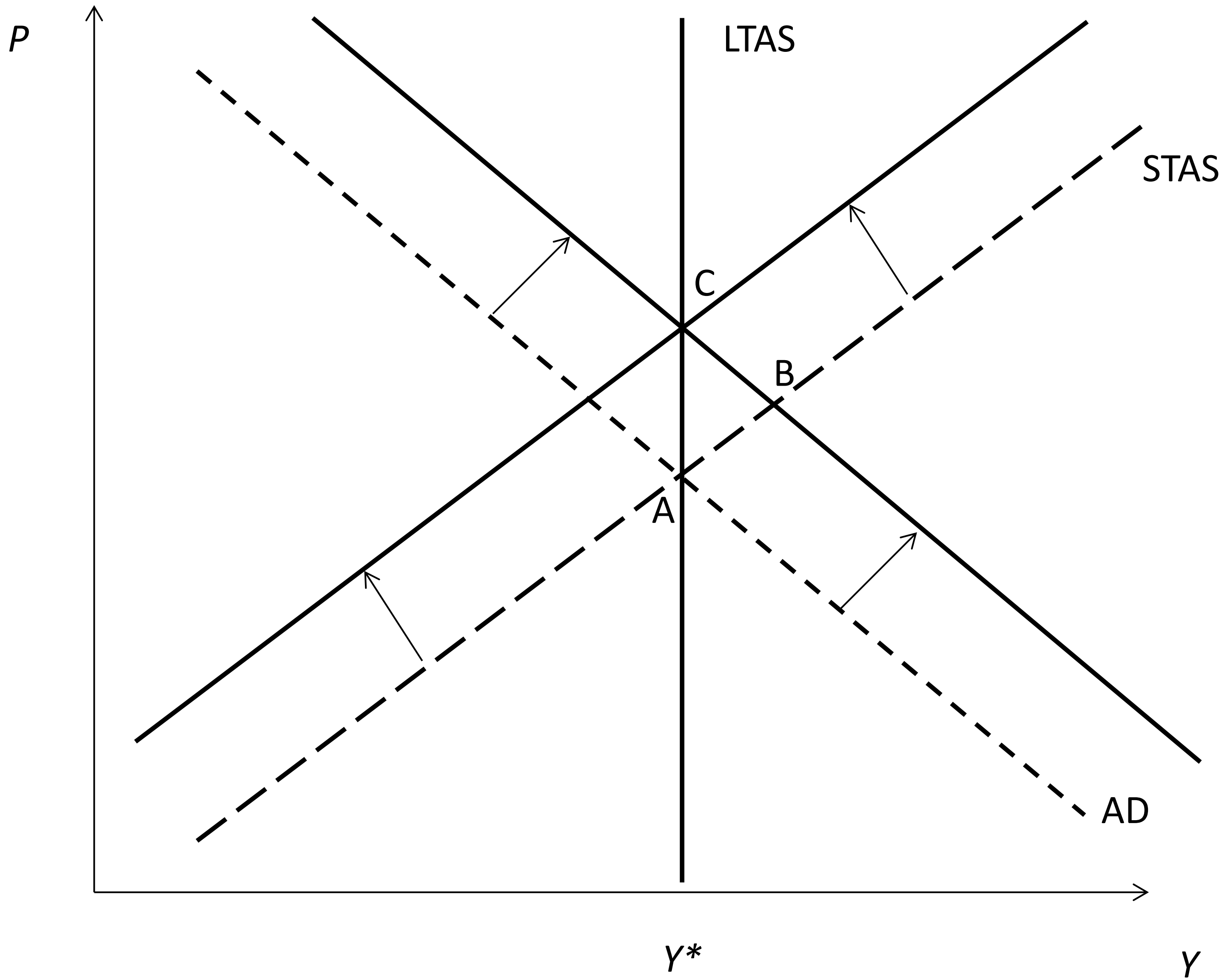

Figure 10: La conception traditionnelle de la politique économique

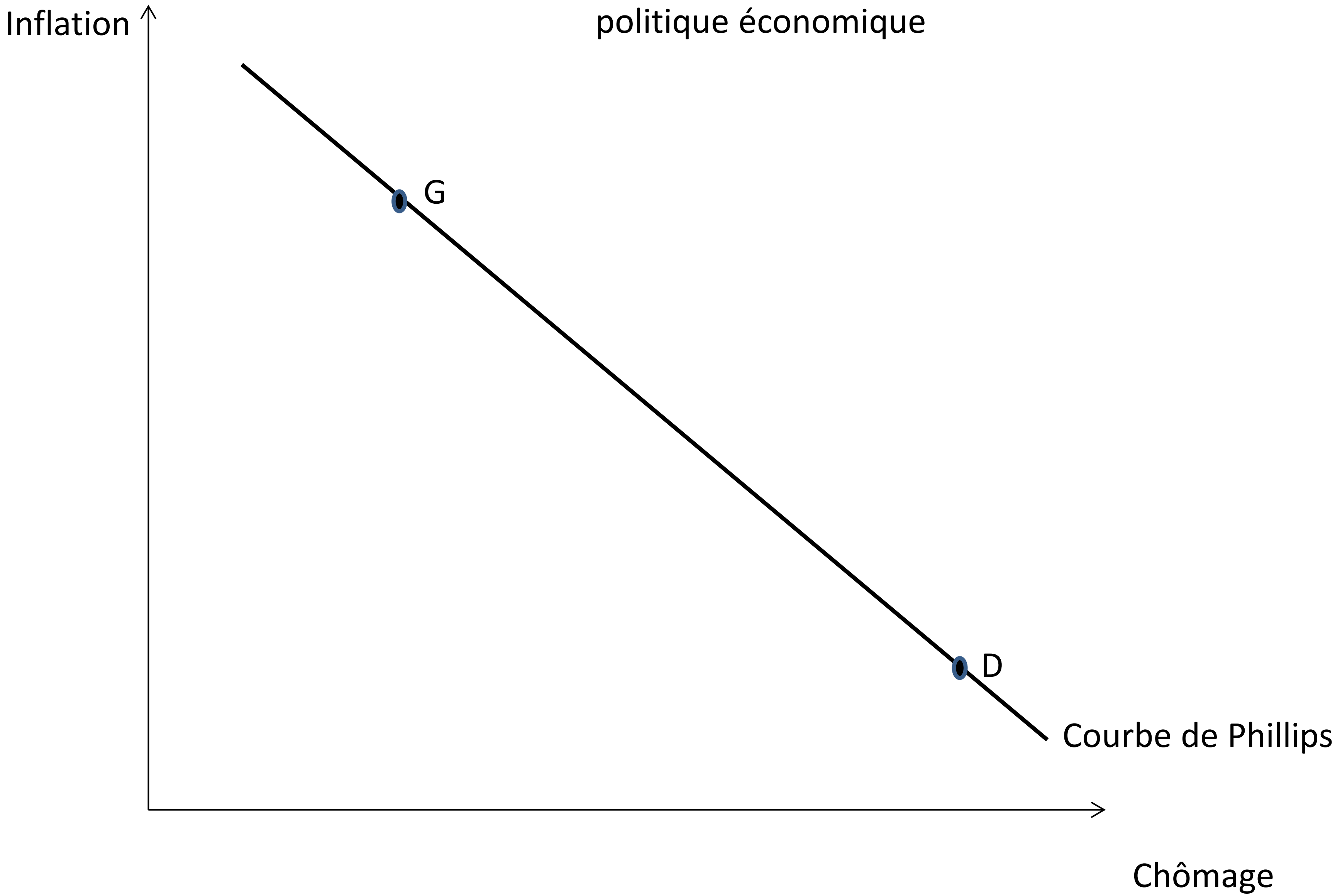

Figure 11: Le diagramme AS-AD

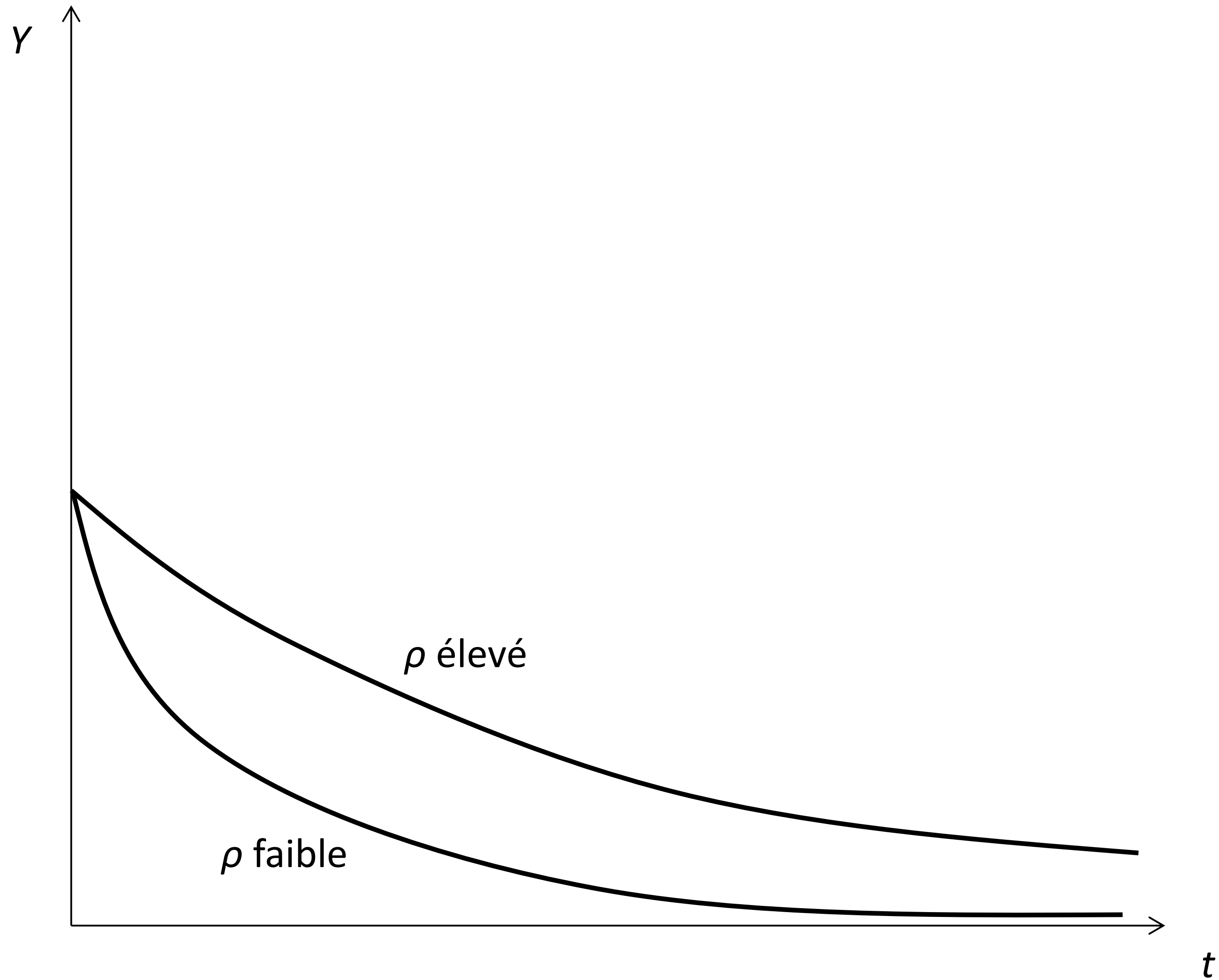

Figure 12: Evolution des encaisses dans le modèle de Baumol-Tobin

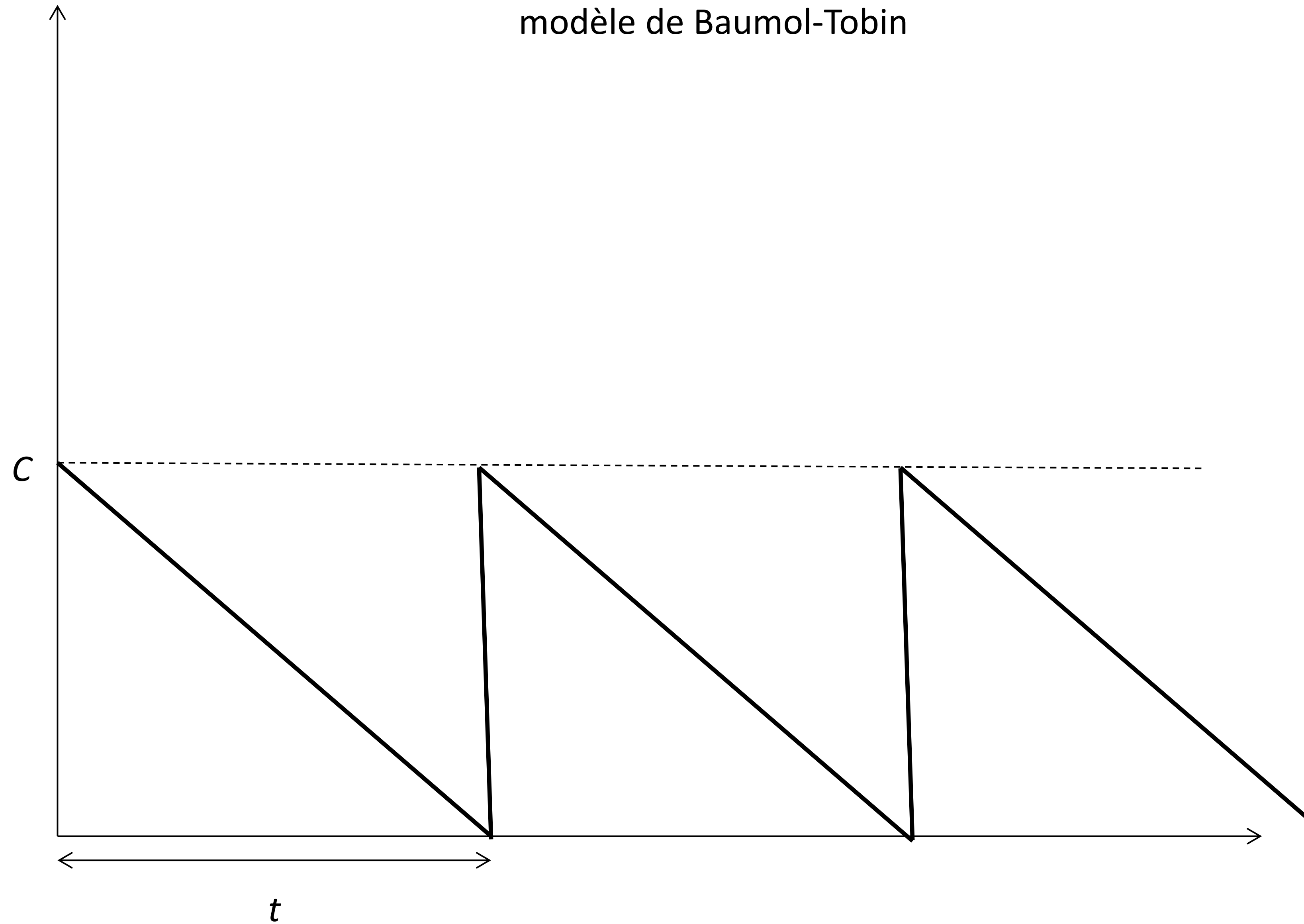

Figure 13: Détermination du régime en prix fixes

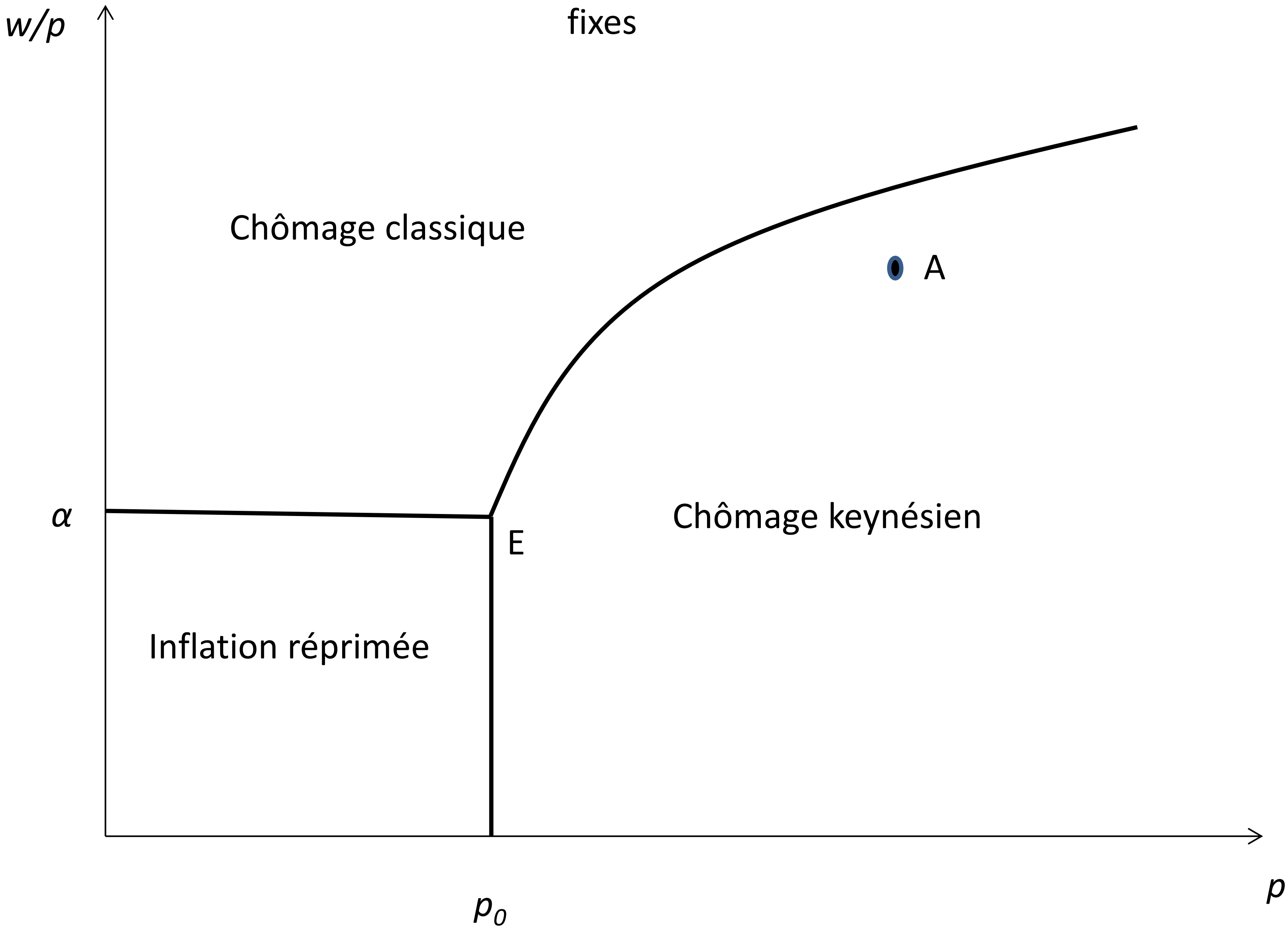

Figure 14: La pente de la courbe de Phillips dépend de la volatilité de l'inflation